THE Ergonomics society

Contemporary Ergonomics

2008

Editor:

Philip D. Bust

Contemporary Ergonomics 2008

Contemporary Ergonomics 2008

Editor

Philip D. Bust

Loughborough University

Taylor & Francis

The Ergonomics society

Do I Really Need A Limo? Fit For Purpose Ergonomics/Human Factors Workshop
A. Jackson & J. Wilkinson

Initial Testing of an Inspectors' Tool for the Assessment of Repetitive Tasks (ART)
J. Ferreira, M. Birtles & L. Stanley

Development of an Inspectors' Assessment Tool for Repetitive Tasks of the Upper Limbs
J. Ferreira, M. Gray, L. Stanley & D. Riley

Baggage Handling In Narrow-Bodied Aircraft: Further Risk Identification
S. Tapley, D. Riley & L. Oxley

Improving Railway Safety Critical Communications
C. Dickinson

The Epidemiology of Slips, Trips, and Falls in the U.S. Army
M. Canham-Chervak, K.G. Hauret & B.H. Jones

Ergonomics Issues In A Jet Car Crash
M. Gray

Typeset by Charon Tec Ltd (A Macmillan Company), Chennai, India
Printed and bound in Great Britain by Antony Rowe Ltd (A CPI-group Company), Chippenham, Wiltshire

ISBN 978-0-415-46575-5

Contents

Preface XVII

AGEING POPULATION

Understanding workplace design for older workers: A case study 3
P. Buckle, V. Woods, O. Oztug & D. Stubbs

Evaluation of the effectiveness of telecare for independent living in Surrey 8
J. Anderson & K. Horton

The effects of users' age, nationality and computer experience on icon recognition 14
H-F. Wang & S-Y. Yu

APPLICATIONS OF ERGONOMICS

Problems associated with police equipment carriage with body armour, including driving 23
D. Stubbs, G. David, V. Woods & S. Beards

To queue or not to queue – where are the biscuits? 29
E. Chandler & S. Hull

Teaching ergonomics at a distance: A virtual workplace 34
M. Richardson & A. Baird

Influence of vibration on workload while reading and writing on Indian trains 40
R. Hancock, N.J. Mansfield, V.K. Goel & R. Narayanamoorthy

COMPLEX SYSTEMS

Using organisational systems engineering to improve the bid proposal process 49
G.A.L. Kennedy, E-M. Molloy, C.E. Siemieniuch, M.A. Sinclair & R.J. Clayton

A new decision-making systems framework 55
E-M. Molloy, C. Siemieniuch & M. Sinclair

Managing human factors trade-offs in armoured fighting vehicle projects 61
R. Cummings

DESIGN

Ergonomics aspects of video conferencing 69
N.C.C.M. Moes, A. Kooijman, E. van Doorn & N. Fain

Avoiding project failure: Putting people first 75
L. Barrett & M. McNeill

The dream economy – design, emotion and the user experience 81
P.W. Jordan

The cultural imagery used in icon design 86
H-F. Wang & K-L. Wang

English, Slovenian, Swedish: Capturing user requirements on large projects 92
D. Golightly, A. Stedmon, M. Pettitt, S. Cobb, S. Sharples, H. Patel, J. Hollowood, K. Moncrieff, M. D'Cruz & J. Wilson

ERGONOMICS 4 SCHOOLS

Diagnosing failures in organisational behaviour using influence diagrams 101
M. Tainsh

A post occupancy evaluation of five primary schools built to coventry's model brief 106
M. Newman, A. Woodcock & P. Dunham

A child centred approach to the design of e-learning material for the ICT curriculum 112
R.A. Hassana & A. Woodcock

ERGONOMICS AND SECURITY

Posture, position & biometrics: Guidelines for self-service technology 121
J. Ward, C. Riley & G. Johnson

A human-centered approach for improving airport security screening 127
V. Vassiliades, P. Evans, L. Kaufman, J.W. Chan & D. Downes

Keeping track of a suspect on CCTV: How hard can it be?! 133
S. Harris, A. Stedmon, S. Sharples & J. Wilson

Tracking a suspicious person using CCTV: But what do we mean by suspicious behaviour? 139
A. Stedmon, S. Harris, A. Carse, S. Sharples & J. Wilson

Exploring essential skills of CCTV operators: The role of sensitivity to non-verbal cues 145
A. Blechko, I. Darker & A. Gale

HAZARDOUS INDUSTRIES

Do I really need a limo? Fit for purpose ergonomics/human factors workshop 153
A. Jackson & J. Wilkinson

Current trends in human reliability assessment 158
B. Kirwan & H. Gibson

Nuclear Action Reliability Assessment (NARA), further development of a data-based HRA tool 164
W.H. Gibson, B. Kirwan, R. Kennedy, J. Edmunds & I. Umbers

Recurring factors in reportable events of nuclear power plants 170
M. Richter & T-O. Solisch

HEALTH AND WELL BEING OF CONSTRUCTION WORKERS

Recommendations for ergonomic design of workplaces and machinery in construction 179
A. Piedrabuena, A. Ferreras, P. Castelló, C. García, A. Oltra, J. Gálvez, L. Tortosa & Á. Page

Evaluating technology and research knowledge gaps for prevention of falls from scaffolds 185
H. Hsiao

The importance of power tool consumables on vibration exposure 191
L. Notini, N.J. Mansfield & D.J. Edwards

Implementing a health management toolkit in the construction industry 197
P. Bust, A.G.F. Gibb & C. Brace

Survey of heavy equipment operators whole-body vibration profiles 203
G. Newell, N.J. Mansfield & L. Notini

Maintenance workers and asbestos: Understanding influences on worker behaviour 209
C. Tyers & S. O'Regan

Employing foreign workers in Sweden construction sector 215
R.K. Ahmad & B. Hansson

Ageing workforce in construction – equipment use and the prevention of early retirement 221
J. Leaviss, A.G.F. Gibb & P.D. Bust

Cost effectiveness of safety initiatives in the Hong Kong construction industry 227
S. Rowlinson & B. Yip

HCI

Human-computer interaction: The elephant in the corner 239
H. David

The cultural issues in icon design 246
H-F. Wang

Charles babbage and the "facility of using" 252
R.B. Stammers

INCLUSIVE DESIGN

Reasonable adjustments: Collecting user experiences for a "solutions database" 261
N. Geddes & V. Haines

A framework for estimating exclusion from work tasks during inclusive design 267
D. Nery, P. Langdon & J. Clarkson

Categorising design methods: How designers view the roles of user methods in design 273
J. Goodman-Deane, P. Langdon, S. Clarke & P.J. Clarkson

Operationalising analytical inclusive design evaluation 279
P. Langdon, U. Persad & J. Clarkson

The assistive eye: An inclusive design approach 285
E. Moellenbach, A. Gale & J.P. Hansen

Calibrating capability loss simulators to population data 291
S. Waller, P. Langdon, C. Cardoso & P.J. Clarkson

Measuring prior experience in learned product interactions 297
P. Langdon, T. Lewis & J. Clarkson

Tactile paving a necessary intervention, but does it suit everyone? 303
M. Faruk, M. Ormerod, R. Newton, H. MacLennan & M.Y. Abbas

Time is marching on. What really determines the time it takes us all to get from A to B? 309
H. MacLennan, M. Ormerod, R. Newton, M. Faruk & M.Y. Abbas

Pedestrians: Just a simple flow? 315
M.Y. Abbas, M. Ormerod, R. Newton, M. Faruk & H. MacLennan

Pattern and trend of pedestrian casualties in the UK: A case of misunderstood ergonomics? 321
M.Y. Abbas, M. Ormerod, R. Newton, H. MacLennan & M. Faruk

Say what you see? The effects of being able to see what's happening on vocabulary design for speech input 327
A. Stedmon & M. Aujla

Inclusive communication – the other breakthrough 333
R. Chesters, R. Baines & J. Mitchell

Inclusion – blending disabled user and professional insights 338
R. Baines, R. Chesters & J. Mitchell

21 'building blocks' for 21st century inclusion 343
J. Mitchell, R. Baines & R. Chesters

MEDICAL ERGONOMICS

Task envelopes: The missing link 351
S. Hignett & P. Waterson

Reducing errors in surgical care: Theory and evidence 357
K. Catchpole, A. Mishra, S. Kreckler & P. McCulloch

Radiologist performance in the 21st century: The impact of workstation type and MSD's 363
S. Taylor-Phillips, A.G. Gale & M. Wallis

Human reliability assessment in healthcare – where next? 369
M. Lyons

Anytime, anywhere mammographic interpretation training 375
Y. Chen, A. Gale & H. Scott

Head banging: The role of ergonomics in the interpretation of neuroradiological images 381
L.H.K. Cooper & A. Gale

METHODS AND TOOLS

Laptops in the lecture theatre: An ergonomic focus on the critical issues 389
R. Benedyk & M. Hadjisimou

Development of software simulation to enable HF assessment of some ASMGCS's features 395
K.Y. Lim, D. Tanuj, R. Yung, I. Koh, S.Z.Y. Ow & D.R.C. Han

Assessment of physical activity at workplaces 400
B. Weber, I. Hermanns, R.P. Ellegast & J. Kleinert

A follow-up study using participatory ergonomics to reduce WMSD's 406
J. Devereux & R. Manson

A collaborative design approach to preventing work-related musculoskeletal disorders 411
H.K.G. Punchihewa & D.E. Gyi

Crossmodal congruency benefits for tactile and visual signalling 417
P.A. Hancock, L.T.C. James & L. Merlo

OCCUPATIONAL HEALTH AND SAFETY

Management of work-relevant upper limb disorders 425
A.K. Burton, N.A.S. Kendall, B.G. Pearce, L.N. Birrell & L.C. Bainbridge

Glove effects on hand performance capabilities in a light assembly task 431
I. Dianat & C.M. Haslegrave

Long-term study of musculoskeletal problems during a reorganisation in Royal Mail 436
C. Parsons

Teaching about musculoskeletal disorders – are we barking up the wrong tree? 441
A. Baird

Evaluation and introduction of electrically powered equipment for mail delivery 447
A. Woodcock & C. Parsons

Initial testing of an inspectors' tool for the Assessment of Repetitive Tasks (ART) 453
J. Ferreira, M. Birtles & L. Stanley

Development of an inspectors' assessment tool for repetitive tasks of the upper limbs 459
J. Ferreira, M. Gray, L. Stanley & D. Riley

Baggage handling in narrow-bodied aircraft: Further risk identification 465
S. Tapley, D. Riley & L. Oxley

Ergonomics review of the Highways Agency control rooms 471
N. Heaton, J. Evans, J. McKenzie & M. Trigg

PHYSICAL ERGONOMICS

The effects of a mid-morning chocolate snack on mood and glucose tolerance 479
N. Morris

Reduced physiological strain in firefighters using a practical cooling strategy 485
D. Barr, W. Gregson & T. Reilly

Methods and applications of body composition analysis 491
T. Reilly & L. Sutton

Virtual electronic game playing by children can be active 496
L. Straker, C. Pollock, J. Piek, R. Abbott & J. Coleman

Military load carriage injuries and discomfort:
A questionnaire survey 502
S.A. Birrell & R.A. Haslam

RAIL ERGONOMICS

Improving railway safety critical communications 511
C. Dickinson

Systems engineering meets user centred design:
In cab signalling displays 517
M. Lowten & G. Roberts

Comparison of different types of reports in the investigation of
Signals Passed at Danger (SPADs) 523
*J. Hutchings, B. Ryan, M. Nolan-McSweeney, E. Lowe &
W. Weatherill*

Human systems integration in rail: Developing the interface for
integration 529
J. McGovern Narkevicius

The influences on safety behaviours of rail maintenance staff 534
P. Murphy

Fatigue management in rail: The role of ergonomics 540
F. Kenvyn

Applications of handheld computers for the rail industry 546
Y. Dadashi, J.R. Wilson, S. Sharples & T. Clarke

Structured observations of automation use 552
N. Balfe, J.R. Wilson, S. Sharples & T. Clarke

Network rail signaller's workload toolkit 558
E. Lowe & L. Pickup

Use of visual scenarios to understand and improve
engineering work on the UK railway 564
A. Schock, B. Ryan, J.R. Wilson, T. Clarke & P. Dingwall

Use of off-workstation displays in rail control applications 570
M. Carey

Improving productivity in rail maintenance 577
T. Farrington-Darby, G. King & T. Clarke

Fingers on your buzzers please 583
A. Scott

Integration is more than a human factors issue 589
F. Bird

SAFETY AT SEA

Too tired to start? Travel to the vessel and subsequent fatigue at work among seafarers 597
E.J.K. Wadsworth, P.H. Allen & A.P. Smith

A diary study of fatigue offshore 602
A. Smith

Seafarers' fatigue: Conclusions and the way forward 607
A.P. Smith, P.H. Allen & E.J.K. Wadsworth

Designing hyperbaric decompression chambers 613
E. Mackie, A. Woodcock & C. Johnson

SLIPS, TRIPS AND FALLS SYMPOSIUM

An evaluation of the different styles of footwear worn at the time of a stair fall 621
V. Halford & H.H. Cohen

Reflectivity as an indicator of slip risk 626
P. Lehane & R. Clark

Accidents on English dwelling stairs are directly related to going size 632
M. Wright & M. Roys

Behavioural safety – reduction of slips and trip accidents in a station environment 638
I. Rowe

Kinematic and tribological analysis of slip phenomena during walking 644
T. Yamaguchi, S. Hatanaka & K. Hokkirigawa

Impact of spreading flour and salt on floor slipperiness 651
F. Quirion & P. Poirier

A procedure to prepare fouled and worn quarry tiles 657
F. Quirion, A. Massicotte, P. Poirier & N. Hadjev

Perception and cognition during walking while concurrently using a cellular phone 663
M. Kuzel, M. Heller, R. Gray, S. Di Jorio & S. Straughn

Effect of shoe type on kinematics of stair negotiation in women 669
M. Heller, S. Di Jorio, M. Kuzel, L. Ciccarelli & M. Carhart

To slip or not to slip: A comparison of matched slip and non-slip trials under similar conditions 675
R.W. McGorry, C-C. Chang & A. DiDomenico

The minimum cost of non-fatal fall accidents to the NHS in England 681
M. Roys & M. Wright

Optimum friction level for walking over a floor 687
T. Zamora, E. Alcántara, F. Payá, J. Portolés & E. Algora

Multidisciplinary research to prevent Slip, Trip, and Fall (STF) incidents among hospital workers 693
J.W. Collins, J.L. Bell, R. Grönqvist, T.K. Courtney, D.A. Lombardi, G.S. Sorock, W-R. Chang, L. Wolf, S. Chiou, B. Evanoff, H.M. Wellman, M. Matz & A. Nelson

Slips and/or trips at the menopause 699
J.C. Davies, G.J. Kemp & C. Bowes

The epidemiology of slips, trips, and falls in the U.S. Army 705
M. Canham-Chervak, K.G. Hauret & B.H. Jones

A preliminary study on slip potentials of stepping barefoot on slippery floors 710
H. Nagata, M. Kato, H. Watanabe, Y. Inoue & I-J. Kim

Modeling shoe-floor-contaminant friction applied to a pin-on-disk apparatus 717
K.E. Beschorner, M. Lovell, C.F. Higgs III & M. Redfern

Comparison of site-based friction measurement and slip potential 723
K. Hallas, R. Houlihan, R. Shaw & S. Thorpe

Nature of the shoe wear: Its uniqueness, complexity and effects on slip resistance properties 728
I-J. Kim & H. Nagata

Linear regression models of floor surface parameters on friction at shoe-floor interface 735
W-R. Chang, S. Matz, R. Grönqvist & M. Hirvonen

TRANSPORT

Combined posture and vibration investigations among occupational drivers 743
N. Raffler, I. Hermanns, R. Ellegast & S. Fischer

Safe and fuel efficient driving: Defining the benchmarks 749
M.S. Young, S.A. Birrell, N.A. Stanton, N. Thorpe & M. Fowkes

Detection of memory lapse during flight: A study of back-up systems and feedback 755
K. Enflo

Ergonomics issues in a jet car crash 761
M. Gray

Drivers and driving: A segmentation based on behaviours and attitudes 767
P.W. Jordan, F. Chen & A. Lindgren

Different riders on different motorcycles: A quasi-experimental approach to rider comfort 773
A. Stedmon, S. Hargreaves, M. Fitz-Gibbons & C. Wood

Profiling a new user group: Anthropometric data for motorcyclists 779
A. Stedmon, S. Robertson, T. Yates & B. Fitzpatrick

Spatial differences in sensible and latent heat losses under a bicycle helmet 785
G. De Bruyne, J-M. Aerts, G.V. Perre, J. Goffin, I. Verpoest & D. Berckmans

Author Index 791

Preface

Contemporary Ergonomics 2008 is the proceedings of the Annual Conference of the Ergonomics Society, held in April 2008, for the second consecutive year, at the Jubilee Campus of the University of Nottingham, UK. The conference is a major international event for ergonomists and human factors specialists, and attracts contributions from around the world.

Papers are chosen by a selection panel from abstracts submitted in the autumn of the previous year and the selected papers are published in *Contemporary Ergonomics*. Papers are submitted as camera ready copy prior to the conference. Each author is responsible for the presentation of their paper. Details of the submission procedure may be obtained from the Ergonomics Society.

The Ergonomics Society is the professional body for ergonomists and human factors specialists based in the United Kingdom. It also attracts members throughout the world and is affiliated to the International Ergonomics Association. It provides recognition of competence of its members through its Professional Register. For further details contact:

The Ergonomics Society
Elms Court
Elms Grove
Loughborough
Leicestershire
LE11 1RG
UK

Tel: (+44) 1509 234 904
Fax: (+44) 1509 235 666

Email: ergsoc@ergonomics.org.uk
Web page: http://www.ergonomics.org.uk

AGEING POPULATION

UNDERSTANDING WORKPLACE DESIGN FOR OLDER WORKERS: A CASE STUDY

Peter Buckle, Valerie Woods, Ozhan Oztug & David Stubbs

Robens Centre for Public Health, Faculty of Health and Medical Sciences
University of Surrey, Guildford GU2 7TE

The ageing workforce presents a significant challenge to ergonomists. Whilst our knowledge of changes in physical and cognitive capacities with age has been researched, there has been little attention paid to the perceptions of the older workforce regarding their experience of ageing in the modern workplace. This case study has used qualitative methods to gain new insights into the motivations, health issues, coping capacities and self perceptions of ageing workers in a food processing industry.

Introduction

Few workplaces or work organizations have been designed with the needs of the 65 year old (and older) worker in mind Huppert 2003, Woods and Buckle 2002). Despite the "Inclusive Design" research agenda there is little knowledge of what this workforce sector requires. There is also little known on what help organizations need to accommodate older workers.

This case study focused on identifying organizational issues that older workers believe need to be addressed for a healthy working environment that meets their abilities, capacities and expectations. The study is one of four workplace case studies undertaken as part of the research councils' Strategic Promotion of Ageing Research Capacity (SPARC) initiative.

Methods

The study took place at the manufacturing centre of a food processing organization. Focus group interviews were undertaken with workers from age groups of 40–49, 50–59, 60+ years and a sample of recently retired workers. In addition, two occupational health professionals and a representative of the human resource management team were interviewed. In total 14 workers participated; these were mainly from the shop floor but a number of office workers also took part in the study. "Older workers" refers to the participants who were aged 50 years and above and "younger workers" to those aged 40–49 years. Although both male and female workers took part in the focus groups all participants are referred to as male in

the following summary to preserve anonymity. All the interviews were transcribed verbatim and coded line by line using qualitative data analysis software QSR NUD*IST (Richards and Richards, 1991). Codes were arranged into coherent themes using the approach advanced by Taylor-Powell and Renner (2003).

Results

The issues raised during the interviews were categorized under the following main themes: "perceptions of organizational culture", "motivations to leave or continue working", "health issues", "perceptions of own capacity", and "coping with job demands". A brief summary of the issues raised and examples of what participants said are presented below.

Perceptions of organizational culture: This category comprised the issues that the participants perceived to be important with respect to their age and the organizations' culture. As they aged, some participants reported that their career opportunities were limited and that their accumulated knowledge was superseded "I'm thinking more of a permanent fixture thing and I personally don't think that industry now rewards the older person with the knowledge, it's almost like we get to the burn-out stage and then we are put out to grass."

There was a consensus that limited career opportunities affected their motivation for personal development; however it appeared that some were happy with their existing knowledge, that they had accumulated over the long years of service, and did not wish for further development. One of the participants expressed his resistance for progress as follows: "If I'd wanted to, I think I should be able to be given the opportunity of doing something, but I feel that it's not there. They may contradict me and say 'Yes, it is.' But it's not that apparent. It's definitely not apparent to me, but as I say my choice is that I wouldn't want to at this stage."

Similar concerns were raised by the younger workers. They believed that as workers aged they were likely to be perceived as less adaptable in dynamic work settings in comparison with younger workers. One of the workers expressed this as follows: "I think part of people's perception is as you're getting older that they can't think of you in the same way as the younger person coming into the business and that you're not as flexible."

Motivations to leave or continue working: The participants expressed a range of factors as motivators for continuing to work. "Finance" was the most frequently reported motivator for older workers to stay and continue working. Overall both younger and older workers reported that they were quite satisfied with their earnings. One financial reason mentioned by an older worker for continuing to work was better investment in pension plans: "As I said, 18 years ago, I took out an investment plan. If you remember, 18 years ago insurance companies weren't allowed to give you advice like they do nowadays. If hindsight had prevailed, I could have been retired now, but my investment plan takes me until I'm 65."

Similar views were expressed by a younger worker: "I said 60 [previously thought age for retirement] but you just don't know, do you? You're on a pension, you see

the changes being made now, we could be here until we're 65 because monetary wise you could be worse off if you don't. You just don't know."

Some of the older participants reported social reasons for continuing to work. One of them thought that coming to work provided a structure to his life. Another reported that he was enjoying coming to work as he had good social networks there: "It's the discipline isn't it really? Getting up every day, knowing that you've got to do certain things . . ."

The participants also expressed some factors as motivators to leave the job. One of the older participants, a shift worker, said having regular sleep was a motivator to leave the job: "But what I have also noticed as well is that on some occasions, some of the blokes that you see now, that you haven't seen for two or three years, they look younger now than when they worked here because . . . They're getting sleep. They're getting sleep; they're getting regular sleep . . ."

Another older worker thought that the attitudes of the management had changed and that it was more difficult to communicate problems to the senior management. He expressed this as a motivator for leaving the job earlier than the normal retirement age of 65: "I must admit, if you'd have asked me 10 years ago whether I wanted to retire, I'd have said no. I'd have been quite happy working as long as I could, but it's the atmosphere, I think, in the factory that's changed. I wouldn't want to stay on."

Some of the younger workers expressed health status as a reason for possible early retirement. One of them thought that working for many years might lead to ill-health that might result in premature death: "My dad took early retirement when he was 59 and he always said to me he'd lost a lot of money with his ABCs and different things, but he wanted to be at home with my mum. Good job because he died when he was 67. If he'd have stopped when he was 65 he'd have only had two years . . ."

There was a consensus among workers that the shift patterns were difficult to follow: "Some of the things that are killing me off with age . . . pressures, I don't take them as I used to, and the hours. That is what I find difficult as I've got older, the pressure of the work and the physical hours of the shift work."

One of the younger workers suggested that shift work was a source of stress as it made working conditions difficult and was one of the reasons people were choosing to retire early and changing to less stressful jobs: "But you look at the people that have actually retired from . . . . . that took early retirement from . . . . . and got other jobs and are enjoying themselves because they've got different jobs without any stress related to them . . . . People have gone. They want to get away from this after they've been in it for such a long time."

Health issues: The older workers perceived regular health checks as a positive aspect of working as an older worker. One of them perceived this as being cared for. There was a consensus among all the workers that the Occupational Health Department was looking after older workers well: "Certainly the Health & Safety departments look after the older worker. They monitor our health, I think more frequently than others . . ." and "As I said, the only positive I see [about getting older in this organisation] is the occupational health, as far as the fitness thing goes, and the physio."

One of the participants on the other hand raised a concern regarding the environmental conditions and believed that older workers might become more sensitive to environmental exposures such as heat, dust and noise: "The working environment at the factory as a whole is quite good. I mean, it [dust, noise, heat] can get terrible out on the Process." (Note: "Process" here refers to the food processing area of the organization.) When asked if these factors got worse as they got older "Oh yes . . . Oh yes. I certainly find now that things like that would aggravate me far more now than they would have done. It can make you feel unwell. I mean, you can probably hear that I've got a problem with my throat anyway, which the dust aggravates. It does aggravate it. I should wear a dust mask, but unfortunately with that, these hair nets, you feel as though you're in a helmet and it's hot out there. It's dirty, it's wet."

Perceptions of own capacity: Two issues emerged under perceptions of older workers' own capacity. These were cognitive ability and patience. Some workers thought that as they aged their cognitive ability to learn and recall information was declining. One of them expressed this as follows: "The other thing I find wrong is that as you get older, although you might get there eventually, it takes a bit longer to take things in. You can't be as quick. Twenty years ago I was quite sharp in the brain and now it takes a bit longer." Another older worker suggested that as he was getting older he perceived himself as becoming less patient: "I just find that at my age now, whoever said 'With age you get patience' is a liar. You don't get patient. You get less patient." In contrast, one of the younger workers thought that as individuals grow older they become more mature in their views: "I don't know. I think you get a more rounded view on things, don't you? I think when you're young you tend to just hone in on specifics and treat them as the centre of the earth at that particular time."

Coping with job demands: The results demonstrate that there might be a trade-off among some of the older workers and their younger colleagues. A number of the older workers said they found the physical aspect of their jobs difficult to carry out and sometimes they asked for help from their younger colleagues: "It's very difficult on the process because the lines that we run – I'm on the process – you've got to do every job on that line to be able to . . . be an active part . . . It just doesn't work any other way. I'm lucky that the two people I work with will take a fair bit off my shoulders on the heavy stuff."

A number of the younger workers confirmed that they sometimes helped their older colleagues in carrying out physically demanding tasks. One of them expressed this as "carrying older people": "You definitely feel that you're carrying the older people? Yes, not with the technical bit. The fellow I work with, he's as technically advanced as anyone and he's 62, but the physical nature of it, there comes to a point where people find it more difficult to do the job that we're doing." "Like we say, we carry the older people and it happens a lot on the night work."

Discussion

The ageing workforce is an inevitable consequence of demographic trends and current pension policies. The challenge this presents to ergonomists can only be met

if our knowledge of the difficulties faced by the older workforce are understood and acknowledged. This study has gained insights into the perception of work by older members of the workforce and, importantly, has provided insights into teamwork issues, physical demands and cognitive capacities that will need to be considered in future work system design.

References

Huppert, F. 2003 Designing for older users. In: J. Clarkson, R. Coleman, S. Keates and C. Lebbon (Eds) Inclusive Design: Design for the Whole Population. London: Springer Verlag.

Platman, K. and Taylor, P. 2004 Themed section on age, employment and policy, Social Policy and Society, 3, 2, 143–200.

Richards, T. and Richards, L. 1991 The NUDIST System, Qualitative Sociology, 14, 289–306.

Taylor, P. 2002 New Policies for Older Workers. Joseph Rowntree Foundation: The Policy Press. ISBN 1 86134 463 5.

Taylor-Powell, E. and Renner, M. 2003 *Analyzing qualitative data.* University of Wisconsin, Cooperative Extension. Available: http://learningstore.uwex.edu/pdf/G3658-12.PDF

Woods, V. and Buckle, P. 2002 Work, Inequality and Musculoskeletal Health. Norwich: HSE Books, CRR 421/2002. ISBN 0 7176 2312 2.

EVALUATION OF THE EFFECTIVENESS OF TELECARE FOR INDEPENDENT LIVING IN SURREY

Janet Anderson & Khim Horton

Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey GU2 7TE

The aim of this research was to evaluate the introduction of telecare in Surrey from the perspectives of users and the staff involved in the implementation of telecare. Users of telecare were surveyed about their experiences with the equipment and its benefits. Staff completed an online questionnaire assessing their experiences, including the provision of training, the resources provided for telecare, whether their working patterns had changed and whether telecare was beneficial for their patients. Users had a positive attitude to telecare and felt that it enhanced their confidence and independence. However, some people reported difficulties operating the equipment and understanding how it worked. Staff also had positive attitudes to telecare, but reported that training was not always adequate, there was a lack of resources to support the increased workload and there was a need for more clarity in assessment and referral criteria. In conclusion, there is a need for careful evaluation of equipment before procurement to assess its fit with the capacities of the users. Although this study showed that telecare can provide benefits, there is still a need for a strategic and sustainable approach to telecare that involves the whole health and social care system.

Introduction

Emphasis on independence and choice has led to developments in the provision of care and support for older people in the UK. A number of factors have raised the demand for telecare including issues about how the needs and expectation of older people are going to be met by public services (Alaszewski & Cappello 2006). Brownsell and Bradley (2003) proposed that telecare could be positioned as “a form of intermediate care” where people could be discharged from hospital and be provided in their own home with a telecare system. The use of telecare has been reported in a variety of forms and recent applications of telecare cover activities such as remote consultations in specialities ranging from dermatology to psychiatry, the transmission of electrocardiograms and radiological images, and education for health professionals (Currell *et al*. 1997). The benefits are likely to include a reduction of the cost of care provision, the prevention of unnecessary admission to institutional care (Bayer *et al*. 2005 p. 4), maintenance of health, preservation of independence of older people and a reduced burden of care on their relatives (Bowes 2006).

The Department of Health released significant funding over two years from 2006 to provide telecare for 160,000 people in the UK to help them remain independent at home. Surrey County Council received funding under the Preventative Technologies Grant to implement telecare in Surrey. Telecare includes a wide range of assistive devices such as community alarms, activity lifestyle monitors, vital signs monitoring for people with chronic conditions and medication reminder systems. The introduction of telecare is motivated by pressures on the healthcare system to adopt proactive, preventative models of care, to shift care from the acute care sector to primary care and to support earlier discharge from hospital. The successful integration of telecare into assessment and care delivery processes is complex. It requires a change in the culture of health and social care, new ways of delivering services, extensive staff training and processes to manage the supply of equipment and provide round the clock response services. The aims of this research were to evaluate the implementation of telecare in Surrey: In particular, the research aimed to:

1. Assess users' satisfaction with the telecare equipment and service.
2. Assess the experiences of staff involved in implementing telecare and their opinions of the benefits of telecare and the challenges involved in implementing telecare.

Method

Participants

The participants were 65 people who had received telecare equipment under the Preventative Technology Grant and 56 staff who were involved in providing telecare services or referring people for telecare services. In the telecare users group there were 21 males, 36 females and 8 people who did not specify their gender. Their ages ranged from 40 to 95 years, with a median age of 80 years. In the staff group there were 9 males and 47 females. Twenty staff worked in the healthcare sector, 19 in social care, 18 in boroughs and districts and 3 could not be classified (respondents could nominate more than one employment sector).

Questionnaire

Questionnaires were designed to elicit opinions about and experiences with telecare. The questionnaires were sent to all users who had received equipment under the Preventative Technologies Grant. The staff questionnaire was administered via an online questionnaire. In addition a hard copy of the questionnaire was mailed to potential respondents who did not have email or internet access. Questionnaires were reviewed for relevance and appropriateness by the stakeholders and approved by the University of Surrey ethics committee. The questionnaires contained a mix of question formats. There were open ended questions, Likert rating scales and closed questions that required a yes/no answer. Open ended questions were analyzed thematically and the responses were classified and coded so that group data could be reported.

Results

Telecare users

Table 1 shows the type of telecare equipment had been received by the user group.

Users were asked an open ended question about how they found out about telecare. Although social services was a key source of information for 29% of the sample, almost the same number of people (24%) found out about telecare from family or friends. Smaller numbers of people found out about it from healthcare professionals, hospitals and the council.

Using five point Likert scales people were asked to rate how easy the equipment was to use, whether they felt more reassured about their safety, whether they felt more independent, more safe, whether they had improved quality of life, how confident they were using the equipment and how satisfied they were with the equipment. For all these questions a rating of 5 was positive and a low rating indicated a negative response. Figure 1 shows that participants reported positive experiences of telecare with mean scores over 4 for all items.

Reported problems with the equipment fell into four categories: oversensitivity which led to the generation of unnecessary alarms, problems associated with the users, which were often difficulties associated with ageing such as memory or hearing loss, problems generated by a difficulty with the phone line and problems associated with the power supply.

Table 1. Type of telecare equipment installed.

Equipment installed	Number	Equipment installed	Number
Bed sensor	4	External door release	41
Pendant alarm	64	Pill dispenser	3
Smoke alarm	5	Door contact	1
Carer buzzer	2	Wristcare acitivity monitor	1

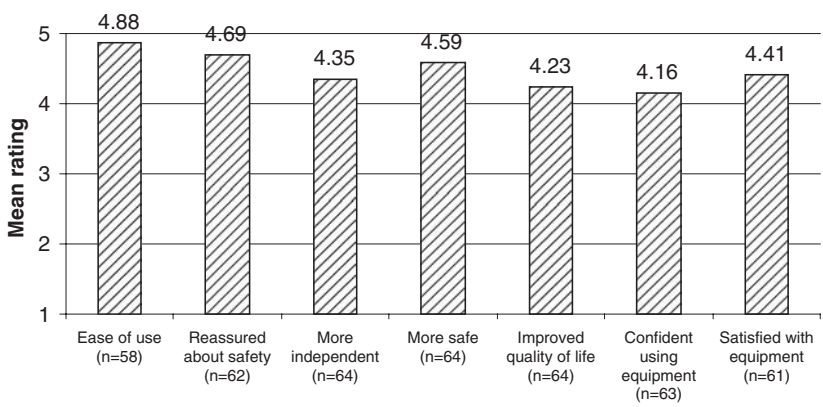


Figure 1. User's experiences of telecare equipment.

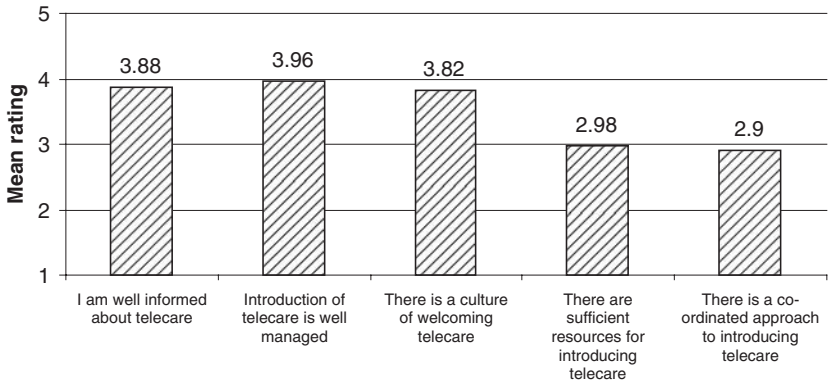


Figure 2. Staff perceptions of implementation of telecare.

Staff

Statements about the implementation of telecare included in the survey are shown in Figure 2 along with the mean rating of agreement for each statement. Fifty one people answered these questions. Staff were slightly positive about how well informed they felt, the introduction of telecare being well managed and the culture in their organisation relating to telecare. They were neutral about whether there were sufficient resources for telecare and whether there was a coordinated approach to telecare.

Forty two percent of the staff stated that the training they received was sufficient, but 38% said it was not sufficient and 20% were not sure. Twelve percent reported receiving very little or no training and a further ten percent received only informal discussion or written information. Suggestions for future training included one to one sessions, equipment demonstrations with hands on experience, explanations of how the equipment works, the costs, how to order the equipment, and technical sessions from suppliers. Higher level training issues were identified including how to match the equipment to the people who would benefit from it the most, including case studies and examples of how telecare had benefited patients and considerations of professional accountability, risk assessment and management of telecare use. These responses indicate some high level concerns about how telecare is incorporated into traditional care delivery processes.

On average staff were neutral about whether there were clear criteria for deciding who should receive telecare and whether telecare was incorporated into care/discharge plans. Staff satisfaction with telecare was marginally positive but staff did not agree that telecare resulted in better care outcomes for their clients, as shown in Figure 3. Respondents were asked an open ended question: “In what ways if any does telecare change your relationships with clients?” Forty five people responded to this question and some people identified more than one issue. Many respondents indicated that telecare had a positive impact on their relationships with clients due to their being able to offer the clients more choices, better support and continuity of care. Negative effects on client relationships resulted when

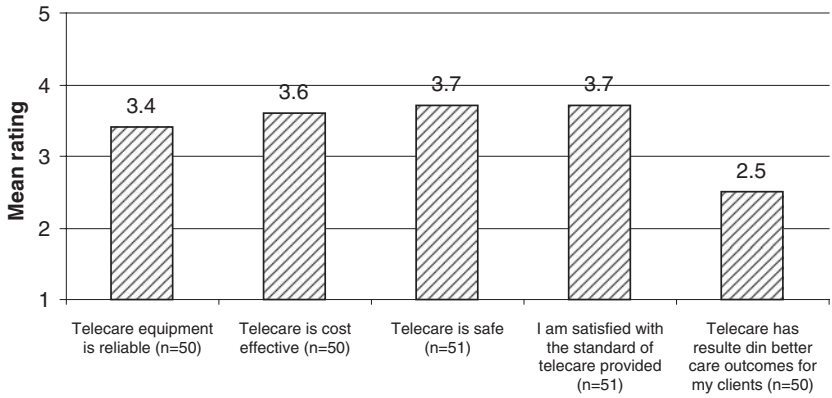


Figure 3. Staff satisfaction with telecare.

the client felt marginalised and telecare was seen as intrusive. Some respondents also indicated that if clients were not satisfied with the equipment or were worried about the costs their frustration and disappointment would impact on the relationship. Interestingly, many people felt that telecare did not change their relationships with clients.

Another open ended question asked respondents whether telecare changed their responsibilities at work. All respondents answered this question with some people identifying more than one issue. Although many respondents said that telecare had not changed the way they work this may be because of the degree to which they have been involved with providing telecare. Others reported that telecare represented a big increase in workload, resulting in less patient time. The benefits of decreased workload, cost savings and enhanced partnership working were mentioned by only a few people. Interestingly, three people saw telecare as increasing their professional responsibility because it involved relying on technology rather than personal care and therefore they had to be sure that the technology was appropriate and set up correctly. Six people specifically mentioned that telecare meant they had more options to offer people.

Conclusions

Although users and staff in general displayed positive attitudes towards telecare and the benefits it could bring, there were some particular areas of concern that were identified. These included the provision of training, publicising the telecare service both for staff and the general public, and improving the referral process for telecare provision. Some concerns were also expressed about the models of care that patients expect to receive and staff expect to provide and how telecare challenges these expectations. The potential of telecare to reduce workload and reduce the need for the provision of personal care does not appear to be as high as has been expected. The telecare technology is more likely to change the way that

care is provided, introducing new tasks and responsibilities to replace old ones. Telecare has the potential to be of great benefit to older people but the introduction of a new service is challenging and there are clearly issues of implementation and strategy that need to be addressed in order to realize the benefits.

References

Alaszewski A. & Cappello R. 2006 *Piloting Telecare in Kent County Council: The Key Lessons Final Report*. Centre for Health Services Studies, University of Kent.

Bayer S., Barlow J. & Curry R. (2005) *Assessing the impact of a care innovation: telecare*. Tanaka Business School Discussion Papers: TBS/DP05/38 London: Tanaka Business School, Imperial College.

Bowes A.M. (2006) Evaluating technology for dementia care. In *Evaluation in Dementia Care* (A. Innes and L. McCabe eds), Jessica Kingsley, London, pp. 87–105.

Brownsell S. & Bradley D. (2003) *Assistive Technology and Telecare: Forging solutions for independent living*. Edited by J. Porteus. The Policy Press, Bristol.

Currell R., Urquhart C., Wainwright P. & Lewis R. (2000) Telemedicine versus face to face patient care: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Review*: CD002098 (2).

THE EFFECTS OF USERS' AGE, NATIONALITY AND COMPUTER EXPERIENCE ON ICON RECOGNITION

Hsiu-Feng Wang & Shu-Yin Yu

Department of Commercial Design, Ming Chuan University, Gwei-Shan, Taoyuan County, Taiwan

The aim of this research is to explore the relationship between users' age, computer experience and nationality and their ability to identify computer icons. The experiment involved 40 British and 40 Taiwanese participants divided into groups, and used a questionnaire to evaluate the effects of age, nationality and computer experience. It took the form of a recognition test using 20 icons on a computer screen. Results suggested that computer experience and age were the main factors that helped people to identify the computer icons shown.

Introduction

Many improvements have taken place in human computer interface design since punched cards and binary codes. Although problems with usability are still occurring, a growing number of interface guidelines and rules are being developed. The International Organization for Standardization (ISO), for example, provides information on ergonomic requirements for the use of visual display terminals for office tasks (ISO 9241) as well as information on how to improve ergonomics at work, to enhance effectiveness and efficiency (ISO 13407). Other user interface design guidelines are also available (Martel and Mavrommati, 2001). The interface designers, however, cannot use them effectively because these guidelines are not developing according to potential users' needs.

Preece (1994) states that "a highly successful approach in interface design is to capitalize on users' existing knowledge and to use metaphors". Thus, it follows that since life experiences, common everyday objects and meanings attached to objects often differ from country to country (and culture to culture), interface designers should use metaphors and objects from the target computer users' environment. Evers (1998) supports this argument stating that "when localising a computer interface it is likely that the metaphors used need to be re-evaluated and perhaps replaced to make the interface intuitive to its users".

Another issue when designing interfaces for users is that many nations have aged societies and that elderly people have become a substantial consumer group within the world of computing. The elderly have many problems with using electronic

products as well as with computer interfaces (Roma *et al.*, 2001). To allow the elderly to effectively operate computers, designers have to keep users' needs in mind, knowing their physiological and psychological abilities for using computer interfaces.

Many interface designers exclude potential users unnecessarily. Such exclusion often arises because of a misunderstanding between designers' perceptions of the wants and needs of the users and their actual wants and needs. Sometimes the misunderstanding is the result of the designers being unaware of either the age or nationality of the target users. Therefore, this paper explores the relationship between users' age, computer experience and nationality and their ability to identify computer icons.

Method

Participants

The experiment involved 40 Taiwanese citizens and 40 British citizens and took the form of a recognition test using icons on a computer screen. Participants were asked to match the icons shown to referents written on a list. The term *referent* is used in this paper to describe the name given to an icon by its program designer.

Prior to the commencement of the experiment, all participants completed a questionnaire. The questionnaire collected personal details, age and data relating to computer experience. The experimenter communicated with the Taiwanese participants in Mandarin Chinese and with the British participants in English. The questionnaire showed that there were no participants who were not familiar with computers, or aged below 30, therefore we divided participants into two groups:

Group A: Young participants who were familiar with computers (participants were aged 20 to 29 years, $M = 23.8$, $SD = 2.7$)

Group B: Elderly participants who were not familiar with computers (participants were aged 60 to 68 years, $M = 65.2$, $SD = 3.5$)

Materials

The recognition test used 20 icons shown in an environment that emulated the widely used word processing package 'Microsoft Word'. Icon presentation was achieved by drawing the icons to be used in a computer graphics program and then importing them into Microsoft PowerPoint to be displayed on a computer screen. Forty labels were also provided on a sheet of paper. The icons employed in the experiment are shown in Table 1.

This experiment uses the term 'label' to refer to a possible name for an icon. A label can be either a referent (the name given to an icon by its program designer) or a dummy label. In this experiment, the term 'dummy label' refers to a label which is not an icon's referent. Dummy labels were used in the experiment to increase the number of alternative answers from which a participant could pick. The labels used in this experiment were written as a list in Chinese for Taiwanese

Table 1. Testing icons.

Referents	Icons	Referents	Icons	Referents	Icons	Referents	Icons
1. Notepad		6. Clock		11. Terminal		16. Keyboard	
2. Text editor		7. Address book		12. Browser		17. Mouse	
3. Calculator		8. Media player		13. Mail		18. File manager	
4. Calendar		9. CD player		14. Settings		19. Installation	
5. Trash can		10. Volume level		15. Display		20. System	

participants and English for English participants. The order of the labels on the list was randomised for each participant. In total, 40 labels were shown in the experiment, of which 20 were referents and 20 were dummy labels.

Procedure

Prior to the tests, participants who had never used a computer before were shown Microsoft Word and were given a brief but succinct explanation of what the program was for and how it worked. This was done not to introduce the participant to Microsoft Word *per se* but to give an overview of how computer interface tools enable computer users to carry out tasks.

Each participant tested was individually shown a mix of 20 different icons. The icons were shown on a computer screen one after the other in a random order. Participants were asked to match each icon shown with a label from a list of 40 on a sheet of paper. Participants were told that there was no restriction on how often a label could be used and that they should take as much time as they needed to respond to each icon shown. The experimenter recorded all comments made during the process.

Results

Table 2 shows the number of correct labels chosen by each group with respect to Taiwanese and British participants. A total of 400 icon viewings were performed in the recognition test to each group (20 different icon designs were shown to all 20 participants in each group).

Chi-square was used to compare frequencies of correct responses with respect to computer experience, age and nationality. The results of these tests are given Table 2.

Table 2. The number of correct labels chosen for each icon.

Icons	Taiwanese Group A	Taiwanese Group B	British Group A	British Group B
1. Notepad	17	6	18	8
2. Text editor	16	7	15	9
3. Calculator	20	14	20	15
4. Calendar	18	8	18	12
5. Trash can	20	5	18	13
6. Clock	20	15	17	14
7. Address book	15	5	16	7
8. Media player	12	7	13	9
9. CD player	15	13	16	13
10. Volume level	20	12	18	12
11. Terminal	10	4	12	2
12. Browser	20	7	19	13
13. Mail	17	7	17	10
14. Settings	12	3	10	4
15. Display	15	5	18	6
16. Keyboard	20	12	18	12
17. Mouse	18	13	15	12
18. File manager	15	3	16	2
19. Installation	18	5	17	4
20. System	16	2	15	1
Total	334	153	326	178

Table 3. The number of correct responses according to Taiwanese and British participants.

Group A and Group B	Taiwanese participants	British participants
Correct labels chosen	487	503

Nationality

A comparison of the frequencies of correct responses between Taiwanese and British computer users is insignificant ($\chi^2 = 0.26$, 1df, $p > 0.05$; see Table 3). In other words, Taiwanese and British computer users recognized icons equally well.

When nationality was looked at separately, the results showed that young participants who were familiar with computers (Group A) selected a similar number of correct labels for icons regardless of nationality. (334 and 326 in Table 4, $\chi^2 = 0.1$, 1df, $p > 0.05$.)

This result is similar to that for elderly participants who were not familiar with computers (Group B) who demonstrated that nationality is not the main factor in the task of icon recognition (153 and 178 in Table 4, $\chi^2 = 1.89$, 1df, $p > 0.05$).

Table 4. The number of correct responses in Group A and Group B.

Group A and Group B	Group A	Group B
Taiwanese participants	334	153
British participants	326	178
All icons	660	331

Computer experience and age

Combining the results of both nationalities (Taiwanese and British), the total number of correct labels chosen for the icons shown in the icon recognition test was 660 for Group A, compared with 331 for Group B (see Table 4). A comparison of the frequencies of correct responses between elderly participants who were not familiar with computers and young participants who were familiar with computers is significant ($\chi^2 = 109.22$, 1df, $p < 0.05$; see Table 4). In other words, a participant's computer familiarity and age played a significant role in the understanding of the icons shown in the tests.

To summarise, the results seem to indicate that computer experience and age assisted participants in the task of icon recognition. Nationality did not seem to assist participants to recognise icons.

Confusion analysis

Confusion can occur with the recognition of computer icons for a variety of reasons. For example, metaphors can be too indirect or imagery may be too culturally based. Confusion analysis was thus performed to discover the likely causes of confusion in the recognition test. The results show that in Group B (elderly participants who were not familiar with computers), the trash can icon returned twice as many correct responses for British participants as for Taiwanese participants (the icon returned 5 correct responses from amongst the Taiwanese members, and 13 correct responses from amongst the British, of Group B, see Table 2). A similar pattern can be seen for the browser icon, where Taiwanese participants returned 7 correct responses and British participants returned 13 correct responses for the same icon. It seems that participants of Group B in general linked the images portrayed by icons with real-life objects, because of their limited computer experience. Thus, if the object used in the icon is not familiar to the participants, they will find it difficult to recognise. For example, in Taiwan there is no object in everyday life which looks like the trash can shown in the icon.

Conclusion

Results suggested that computer experience and age were the main factors that helped people to identify the computer icons shown. These results are consistent

with those of earlier study (Gillan *et al*., 1995) in that computer experience and age have an impact on icon recognition. Nationality was seen to positively influence those participants not familiar with computers who recognized some icons because of their similarity to local features.

The information gained in this research raises the need to take into account national differences when designing icons, especially for the elderly who were not familiar with computer. It may also broaden the scope of the HCI research community because it signals the importance of paying attention to the needs of all users. Our evaluations demonstrate that designers need to consider the effect of ageing beyond just capabilities and anthropometrics, to include experience, skills and aesthetics.

References

Evers, V. 1998, Cross-cultural understanding of metaphors in interface design. In C. Ess, and F. Sudweeks (eds.) *Proceedings CATAC'98, Cultural Attitudes towards Technology and Communication*, (Science Museum, London), 1–10

Gillan, D. J., Fogas, B. S., Aberasturi, S. and Richard, S. 1995, Cognitive ability and computer experience influence interpretation of computer metaphors. In *Proceedings of the Human factors society 39th annual meeting*, (The Human Factors Society, Santa Monica), 243-247

ISO 13407, 1999, *Human-centred design processes for interactive systems.* Geneva: International Organization for Standardization

ISO 9241, 1998, *Ergonomic requirements for office work with visual display terminals (VDTs).* Geneva: International Organization for Standardization

Martel, A. and Mavrommati, I. 2001, Design principles. In K. Bauman, B. Thomas, (eds.) *User interface design for User interface design for electronic appliances*, (John Wiley & Sons, New York), 77–107

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. and Carey, T. 1994, *Human-computer interaction*. (Addison-Wesley, Harlow)

Roma, M. D., de Ridder, H. and Bouma, H. 2001, Technology generation and age in using layered user interfaces. *Gerontechnology* , **1**, 25–40

APPLICATIONS OF ERGONOMICS

PROBLEMS ASSOCIATED WITH POLICE EQUIPMENT CARRIAGE WITH BODY ARMOUR, INCLUDING DRIVING

D. Stubbs*, G. David*, V. Woods* & S. Beards**

**Robens Centre for Public Health, PGMS, Faculty of Health and Medical Science, University of Surrey, GU2 7TE*
***Occupational Health (OH), Welfare and Safety Team, Southern Support and Training HQ, Hamble, Hampshire, SO31 4TS*

This paper presents an investigation of the difficulties that police officers have with the items of equipment that are carried, on either their Utility Belt (UB) or Body Armour (BA). Additionally particular concern associated with equipment carriage worn on either BA or UB and driving was assessed. An ergonomics approach to equipment carriage problems was adopted that included a survey of the relevant literature, extensive officer interviews and the development of usability criteria. The analysis of the data identified a number of primary usability issues. These will be discussed and the paper will conclude by briefly describing the recommendations made.

Introduction

In recent years there has been increasing ergonomics interest in the problems faced by police officers with respect to equipment carriage and how it affects them and their performance. From the literature some of the equipment design and carriages issues relating to poor usability, performance and comfort are identified and summarised including:

- Design and carriage issues:
  - Size, bulk, weight and lack of flexibility,
  - Amount and placement of load.
- Performance issues:
  - Stability, running and difficult arrests,
  - Movement and mobility restrictions.
- Interaction with other uniform:
  - Incompatibility with high visibility jacket.
- Comfort issues:
  - Seated (desk or car): digs into lower back, ‘cuts into throat’, bruising from equipment,
  - Weather conditions: thermal stress, ventilation particular problem in summer.
- Fit:
  - Equipment to suit wearer in terms of size, fit and weight,
  - Gender: females less body and muscle mass and smaller waists.

- Public perception:
  - Recognisable and smart.
- Care:
  - Ease of maintenance and serviceable.

(Buchanan and Hooper 2005; Douse, 2006; Hooper, 2003; Edmonds and Lawson, 2001 and Soldo, 2003)

With respect to equipment carriage and driving additional important issues were highlighted including:

- Performance:
  - Access and egress,
  - Reduces ability to control the car.
- Car seat design:
  - Type of vehicle and shape of patrol car seat,
  - Inadequate lumbar support.
- Posture and discomfort:
  - BA rides up whilst seated causing discomfort/choking.
- Interaction between equipment and car:
  - Seat belt fastening and function,
  - Any equipment at rear of body will interfere with backrest.

(Barge, 2003; Brown et al, 1998; Hooper, 1999 and Lomas and Haslegrave 1998)

Design solutions

There is little indication in the literature on how the problems of equipment carriage whilst driving (in multi-user vehicles) should be addressed. Design, training and maintenance considerations, and issues for review identified in the literature included:

- fully adjustable seats and steering columns as each user requires a unique driving position in order to drive in a comfortable and safe environment,
- customised seats, e.g. with section cut out to accommodate equipment carriage, wrap-round seats,
- information on car set-up for safety and comfort, e.g. placing hard objects on lumbar spines should be avoided,
- soft pouch to be placed over lumbar spine,
- review items (and their design) stored on UB,
- review time spent in police vehicles,
- remove equipment while driving.

Information provision and assessment

Finally a number of papers indicate the importance of dynamic risk assessment (Spencer 2005), provision of information and advice to officers, e.g. increase awareness on how to set up vehicle for both comfort and safety (Barge, 2003) and the importance of a user participatory approach in design (Morris et al, 2004).

Officer interviews: Unaffected and affected officers

In order to gain a greater understanding of how equipment carriage might be influencing previously unaffected and affected officers in the performance of their duties a number of Police Stations were contacted by the Force Occupational Health Department (OHD) to ask whether the researcher team could have access to the station to interview some of their officers. For affected officers who had previously reported problems, they were contacted directly by OHD to seek permission to be interviewed. All interviews with the officers were pre-arranged, confidentiality was assured and the interviews were tape recorded. In total 40 officers were interviewed (20 affected and 20 unaffected).

Summary analysis: All officers

The following provides a summary of the some of the salient points arising from the analysis of the reports from the 20 'unaffected' and 20 'affected' officers.

All officers interviewed appreciate the need for BA and would not like to be without it.

The points raised have been summarised under four main headings although there is some overlap between them, and alternative/conflicting views are sometimes apparent.

Task related

- The added weight of UB/BA restricts activities such as running/chasing and also the performance of more mundane tasks, e.g. felt restricted when changing car wheel.
- BA restricts movement when restraining/grappling with suspects/prisoners.
- Heat load and sweating during summer even for moderate levels of activity can reduce capacity to run/chase.
- Prolonged standing wearing BA may cause back ache, as a result of the increased weight and changed static postures.

Clothing (size and fitting)

- UB may be loose when running or walking – this chastens/rubs on hips and makes them sore and may be exacerbated by a small waist or hips.
- High Visibility clothing, worn over BA/UB restricts access to equipment (e.g. to cuffs when needed urgently); the thermal load may also be increased.
- A good fit in one dimension for BA does not guarantee a good fit in all dimensions; e.g. if it fits at the waist, then it may be restrictive around chest.
- There is a tendency to keep BA on at the station because of fitting time.

Environmental demands

- BA imposes a high heat load in warm weather and shirt worn beneath BA can become soaked with sweat and cause embarrassment.

- When cold, BA provides an additional layer of insulation around the torso, but leaves the arms exposed.
- The BA carrier is not waterproof: it can become saturated and the contents of pockets (e.g. note books) can become soggy when wet.

Driving

Access/Egress

- BA results in a larger size and weight of torso when getting in and out and Officers report the need to pull/manoeuvre them selves out of a vehicle, which can result in strains.
- BA reduces the flexibility of the upper body.
- Equipment worn on UB/BA can catch/snag on seat belt or door fittings.

Sitting in car

- BA can ride up into an uncomfortable position when driving and may restrict breathing.
- There is the need for a lumbar support on the vehicle back rest, but it cannot be adjusted to make it effective when wearing BA.
- Turning in seat when wearing BA is restricted; looking over shoulder etc.

Fatigue/Discomfort following vehicle duties: Some officers reported

- feeling more tired as a result of the additional weight of UB/BA,
- increased back ache, and static posture,
- that their Cuffs can stick into their stomach.

Usability issues

From the analysis of all of the data a usability matrix was created for a range of BA/UB design requirements (including the basic requirements of safety and protection) for the main duties undertaken by officers of: Normal Patrol Tasks, Driving Tasks and High intensity, physical response tasks. For this matrix a number of primary usability issues emerged which encapsulate the full range of factors giving rise to concern including:

- The distribution of equipment between BA and UB taking account of the location of the equipment and the need to access it rapidly when required,
- General comfort and specifically thermal comfort when patrolling on foot and in vehicles,
- Officer visibility when on patrol and when outside a vehicle,
- Overall weight of equipment and its effect on restricting performance both during normal duties and high intensity activities,
- The procurement process to ensure appropriate fit for the full range of officers and also issues relating to instruction/training about and with the equipment.

These issues formed the basis for discussions about possible solutions to equipment carriage with body armour, including driving with manufacturers and other Police Forces.

Conclusions

In conclusion, Hooper's (2003) review of the issues relating to equipment carriage noted that:

> "there is no simple answer to the question - what is the best way for an officer to carry his/her equipment? There are a range of issues and the solutions are inevitably a compromise"

and additionally he noted that:

> "perhaps a key final point is that, everything else being equal, some choice over carriage method for the user is likely to provide the most acceptable system, that is, that no one solution will fit all,"

In the light of these comments and the extensive range of data gathered during this study, the following recommendations were made:

Recommendations (Stubbs et al, 2006)

- the usability matrix developed during this research should be used to:
  - compare the advantages and disadvantages of different designs or modifications of available BA/UB systems prior to their introduction and
  - aid procurement of BA/UB equipment systems and
  - assist during their roll out,
- when the current issue of BA is renewed consideration should be given to reducing the anti-stab level of protection to KR1 from KR2 thus providing a more flexible BA for all officers and a small reduction in overall weight,
- OH should continue to manage the specific needs of affected officers and be permitted to provide officers with lighter weight BA (KR1) with immediate effect, following medical referral, noting that officers so provided should fully appreciate that this will not afford the same level of protection and that this should be taken into account in any risk/threat assessments,
- officers should be given greater freedom to use dynamic risk assessments (DRA) with respect to the current requirement to wear BA and other associated equipments at all times,
- officers should be permitted by means of DRA to remove their BA etc, when driving in particular, ensuring that it is carried in their vehicle and is available for donning if required,

- investigate alternatives to current equipment carriage by means of the UB or clips onto BA, to include consideration of:
  - equipment vests
  - equipment vests with hi-viz options
  - the ergonomic UB.
- the ongoing initiatives in other Forces should be followed up and used to inform the Constabulary's future strategy with respect to improved equipment carriage for all it's officers,
- Finally, all future changes and developments to the equipment carriage system are trialled, monitored and evaluated as previously recommended (Stubbs 2005).

References

Barge R. A. 2003, Does driving in body armour alter the prevalence of back discomfort amongst police officers of the Northamptonshire Police Force? MSc Thesis: Loughborough University.

Buchanan K. and Hooper R. 2005, An ergonomics assessment of police body armour. In P. Bust and P. McCabe (eds) *Contemporary Ergonomics 2005*, (Taylor & Francis, London), 340–344.

Brown J. J. Wells G. A. Trottier A. J. Bonneau J. and Ferris B. 1998, Back pain in a large Canadian Police Force, *Spine*, **23**, 7, 821–827.

Douse K. 2006, Carriage of Police Equipment. HOSDB Publication No 10/06, pp. 22.

Edmonds J. and Lawson G. 2001, The Design of Methods for Carrying Police Equipment, In M. A Hanson (ed.) *Contemporary Ergonomics 2001*, (Taylor & Francis, London), 499–504.

Hooper R. H. 1999, Why do Police Officers leave their Body Armour in the cupboard? In M. A. Hanson, E. J. Lovesey and S.A. Robertson (eds) *Contemporary Ergonomics 1999*, (Taylor & Francis, London), 358–362.

Hooper R. H. 2003, The carriage of loads on the body in the context of the patrolling police officer: Prepared for the ACPO Uniform and Equipment Committee. pp. 6.

Lomas S. M. and Haslegrave C.M. 1998, The ergonomics implications of conventional saloon car cabins on Police drivers. In M.A. Hanson (ed.) *Contemporary Ergonomics 1998*, (Taylor & Francis, London), 471–475.

Morris W. Wilson J. and Koukoulaki T. 2004, Developing a participatory approach to the design of work equipment: assimilating lessons from workers' experience. European Trade Union Technical Bureau for Health and Safety, Brussels, pp. 174.

Soldo S. 2003, Overloaded: Inquiry into problems associated with equipment belts as issued to NSW Police, *Police News*, New South Wales, **80**, 9, 19–27.

Spencer D. 2005, Expect the unexpected. *Health and Safety Practitioner*, September, 43–44.

Stubbs D. A. 2005, Utility belt pilot study for: Hampshire Constabulary, Occupational Health, Welfare and Safety Team. pp. 77.

Stubbs D. A. David G. and Woods V, 2006, Problems associated with equipment carriage with body armour, including driving, for: Hampshire Constabulary, Occupational Health, Welfare and Safety Team. pp. 153.

TO QUEUE OR NOT TO QUEUE – WHERE ARE THE BISCUITS?

Edward Chandler & Sue Hull

The Ergonomics Society Elms Court, Elms Grove, Loughborough LE11 1RG, United Kingdom
The Royal National Institute of Blind People, Innovation and Disability Access Services, Orton Southgate, Peterborough, PE2 6XU, United Kingdom

Introduction

Conferencing is big business in today's working environment and a key way for an industry to showcase the work going on in their sector but also for a discipline or subject to increase its visibility to external sources. Websites now exist which catalogue conference events for people to quickly determine what conferences are on, in which country and on which dates. Conferences can be big money making schemes to such an extent that companies exist just to run conferences for organisations.

But what really makes a conference successful? Is it the quality of papers, or maybe the network opportunities, or even the venue? Well they probably are all very important and would rank highly on delegate feedback forms but what can really make or break a conference is the organisation behind to not only run the conference but to get the conference to organise it.

Conference inception

The first hurdle in running a successful conference starts at the conference instigation phase where the aims, themes and design of the conference are set by the conference team. Questions such as "should there be parallel sessions or just plenary session" and what are the current hot topics within in the discipline" and "should there be any special themes/symposiums running through the conference?" and finally "should we be providing accommodation or leaving that to delegates to arrange?" It is vitally important to aim the conference at the right people so initially the appropriate papers will be received but also people will be interested in attending. Therefore the marketing of the conference is a major task and needs to be fully considered at this stage.

It is crucial that several venues have been short listed and one has been identified and secured to prevent it being taken by another conference otherwise it is likely the conference will never happen in the first place.

The venue

The venue plays a crucial role in the production of a successful conference because without a venue a conference cannot be held. A conference venue should be within easy reach of major transport routes (airports, roads and train stations) and ideally located in a city that will attract people as well (for instance Slough is not going to have the attraction that Cambridge does).

Conference rooms

If the conference is to have parallel sessions, the venue should support sufficient rooms for the number of sessions. It maybe that all session rooms will be of equal size but if not, the organizing team must decide which sessions go in which rooms. Failure to get this right could mean a big auditorium being virtually empty and a small conference room packed with people standing at the side. If the conference is going to have a single strand and everyone will be in plenary all the time, the venue must have a room that will fit all the delegates comfortably.

Whilst the room(s) must be of an adequate size for the conference, they must also be equipped sufficiently to allow the audience to hear the speaker and for the audience and speakers to hear questions raised after the presentation is finished. Ideally all conference rooms should be equipped with microphones and speakers to enabled all in the room to hear what is being said by the presenter, speaker and chair.

Ideally the venue should have availability for break out rooms for people who don't want to go to sessions but want to meet up with industry colleagues to discuss business or to catch up with work. As most conferences are in the week and aimed at working people, having facilities to allow delegates to catch up on email and work is important.

The venue should have a sizeable space to accommodate the conference when all the delegates are having coffee. Generally speaking, the conference organisers will want to keep the conference delegates together to promote networking during the breaks. Having all the delegates in the same place in a big enough area will allow the organisers to bring in exhibitors which are relevant to the conference. Exhibitors normally pay to be at the conference so it is important to keep them happy by providing them with opportunities to meet delegates. Having a large open space allows for this.

In the interests of time and delegate satisfaction it is recommended that this area is close to the session rooms, conference reception desk and to venue facilities. It is likely that delegates will want any of these after a session so having them to hand cuts down on delegates walking distances whilst maximising delegate and exhibitor networking time.

Meeting the needs of delegates

In today's society it is important to meet the needs of delegates whoever they are. The conference venue should be equipped with modern and up to date equipment

such as AV equipment, access to internet etc but also make sure the venue is accessible to disabled and elderly people. This doesn't just mean a ramp at the entrance and an accessible toilet (not a disabled toilet!) but all areas of the venue need to be considered. Session rooms should have an induction loop system to enable hearing aid users to benefit fully, and ideally a place for a sign language interpreter to stand in a place which is well lit where they can see the speakers but also they can been seen easily by the audience. Signage should be clear and easy to read. Permanent signage should conform to accessibility guidelines[1] and temporary signage (e.g. that is specific to that conference) should conform to clear print guidelines[2]. Signage should also be meaning full and convey important information. Signs such as "this door is not a door" only attract negative attention and flag up that the venue has not thought through its signage properly. Failure to provide good clear usable signage will result in delegates getting lost. Ultimately the venue should commission an access audit by one of the many access consultancies in the UK. This will provide the venue team with the information they need to make their facilities more inclusive.

Venue staff

The staff at the venue are very important and are essential to keeping the conference running smooth. Timely delivery of refreshments so that the tea and coffee are hot and food is fresh. As part of the access audit the venue should consider investing in staff training such as disability awareness and disability equality training so that in the event that a disabled delegate needs assistance the staff are fully versed in being able to assist them fully. Whilst they need to be able to do their job it is likely that delegates will want to know things specific to the area such as where the nearest bank, pharmacy or shop if they are not local to the area.

Refreshment stations

One of the most important aspects for any delegate is where the coffee and food is coming from. Whilst it seems easy enough to set up a coffee or food table, the effect of how this is set up is rarely considered.

A key consideration of the refreshment stations is delegate throughput. For instance having a coffee station with one staff member on serving tea and coffee for 200 delegates is likely to cause a long queue. If the break is 15 minutes long and it takes 10 minutes to get a coffee the delegates are not going to be happy and it also means that networking opportunities are dramatically reduced.

Having a big queue can also cause problems as it takes up space. To ensure a fast throughput of delegates at a coffee station, place the biscuits, sugar and milk at the end of the run, or if possible on a separate table. Delegates will move away from the serving area to find the biscuits thereby speeding up the queue.

The food table should be designed so that there is an entrance point and an exit point for delegates. By doing this delegates avoid meeting in the middle or not knowing where to enter and exit from. If possible the food should be reachable

from either side of the table so that the food station can handle as many people as possible to minimize long queue times.

The placement of food is also very important. It has been found that the vegetarian options are taken first if put at the front of the buffet. Therefore it is recommended that the vegetarian options are placed further down the table to avoid them being eaten before the vegetarians get them! Labelling of food to inform delegates what food is what and what is in the food is helpful as this can avoid the delegate choosing something they don't like by accident.

Running the conference

Prior to the start of the conference the conference team need to set up the conference. Some conferences have a secretariat team who do that. If the conference is not going to have a team to run the conference the conference organisers need to arrange sufficient cover so that the conference is set up and run efficiently. This may be done by an external company if the conference organisers don't have the capacity to do it in-house.

The conference team

The Ergonomics Society has such a team and it draws the team from the student population studying ergonomics. The team is ultimately responsible for the running of the conference. They need to prepare signage for the session rooms and directions around the conference, set up a conference office to run the conference from, set up a registration desk, check that the session rooms are ready (water is available for the speakers and the AV equipment is working) and generally familiarize themselves with the conference layout and facilities before the conference begins. As well as their role prior to the start of the conference they are responsible for keeping the conference going. The conference team should usher delegates into session rooms at the appropriate times so that sessions start on time.

Reception desk

The team need to man the registration desk at all times during the conference as this is the first point of contact for delegates. Registering delegates for the conference is another queue management conundrum as delegates usually have to get a badge, a delegate pack, a proceedings and attendance certificate (if applicable) as this process can take time to get people through. Therefore the conference team needs to have enough people on the desk at key times to maximize delegate throughput. If space allows, set up more than one reception desk splitting by surnames alphabetically.

Session timing and chairs

Keeping the conference running to time is a huge task the bigger the conference is and the more sessions there are. One method that has been employed by

The Ergonomics Society Conference Team is to have a member of the team in each session rooms with a chair to time keep. They provide cue cards for the speaker to inform them when they have five minutes left, when they have two minutes left and when they need to finish. This provides the speaker with the information they need to wrap up their session in order to keep the conference to time. They can also be available for problem solving if and when things go wrong.

The chair's responsibility is to introduce the speakers, politely stop the presenter if it is clear they have gone over their time slot and to direct questions from the audience. The chair should insist that the person asking the question not only waits for a microphone but to make sure the questioner states their name and their affiliation. This is important as their question and the speakers answer might be affected by their affiliation. Should the discussions become side tracked from the original topic, it is the chairs responsibility to intervene and bring things back. Should their not be any questions from the audience, it is advised that chairs pose a question to the speaker as this may in turn generate more questions.

Social activities

A very important part of a successful conference is the social activities. A residential conference can become a very lonely place when a delegate has attended alone and is new to this conference. Successful events can be repeated but a combination of new and old is usually a good formula. Room preparation with comfortable seating, a bar adequately manned and the trusted barrel of real ale will guarantee a successful evening.

Summary

Ultimately if the rooms are the wrong size, or at the wrong temperature or have inappropriate lighting and the speakers cannot be heard it is likely that the delegates will be unhappy and that means. However there are additional factors such as the placement of refreshment stands, delegate facilities, staff and finally the team that run the conference to see it through to a successful conclusion. Running a conference is by no means and easy task and takes time and experience to get it right. A good conference team is constantly evolving and one which will look at past successes and failures in order to make the next conference better.

References

http://www.ergonomics.org.uk
1. RNIB sign design Guide
2. RNIB Clear print guidelines

TEACHING ERGONOMICS AT A DISTANCE: A VIRTUAL WORKPLACE

Miles Richardson & Andrew Baird

Centre for Psychological Research in Human Behaviour, University of Derby, Derby, DE22 1GB

Teaching 'applied ergonomics' away from the realities of a workplace represents a real challenge, particularly for students learning at a distance. At the University of Derby, a Virtual Workplace has been developed for e-learners on the MSc Health Ergonomics programme. This web-based resource has been developed from both a pedagogic and practical/professional perspective. In addition to facilitating the development of practical skills such as environmental surveys, it mimics the complexities of real-world situations which require a systematic approach with the appropriate use of tools and techniques. Positive initial feedback indicates that the workplace has been successful both in engaging students with materials and ensuring 'realistic' assessments.

Introduction

One of the greatest challenges when teaching ergonomics is ensuring that students grasp the complexities inherent in 'real-world' working situations. Whilst it is tempting to teach discrete subject elements, research shows us that 'classic' workplace ergonomics problems like musculoskeletal disorders are typically not associated with a simple, single cause but are rather complex and multi-factorial. It is therefore vital that students can build a 'bigger picture' so that they can apply what they have learned effectively.

Teaching this holistic systems approach is always a challenge, but it is a particular issue for distance or online 'e-learning' students who cannot take part in on-campus workshops and practical activities (Watson and Horberry, 2003). Distance learning education programmes have a long history and have recently proved a popular option for those wanting to study ergonomics. It is incumbent upon educators to ensure that the delivery of materials and nature of assessments facilitate a true grasp of 'real-world' issues. It has been proposed that technology can be used to create authentic situations and high-quality interactions if the activities are designed carefully (Tolmie and Boyle, 2000).

To this end the Virtual Workplace for Learning was developed to host interactive e-learning activities associated with the MSc Health Ergonomics programme at the University of Derby. It allows students to develop practical skills and provides a vehicle for 'applied' assessments. This paper describes the Virtual Workplace for

Figure 1. General view of the Virtual Workplace showing a video clip 'interview'.

Learning, the pedagogic benefits it provides, initial student feedback and proposed future developments.

The Virtual Workplace for learning

The Virtual Workplace (Figure 1) is web-based and allows students to view and interact with a number of 'scenarios'. It has been developed to allow future expansion. It is currently centered on a small, problematic call-centre in the fictional organisation – NonsPacifc PLC, but further parts of the organization (e.g. kitchens and workshops) will be developed to enable the generation of further scenarios aimed at teaching and assessing additional skills (e.g. posture analysis and general risk assessment). The current scenario in the Virtual Workplace allows students to investigate issues in the call-centre. Some staff and elements are 'clickable' providing further detail, some of which may or may not be truly relevant to a particular brief.

Figure 2 shows how simple environmental assessments can be made and how some elements can be studied more closely simply by clicking on them. This allows for a closer inspection of the workplace. The Virtual Workplace allows students to utilise a variety of tools, thereby creating a resource to deliver applied topics in the ergonomics curriculum in an innovative way and equipping students learning at a distance with applied skills and knowledge.

In addition, a discomfort survey can be administered, but once again further options could be added. Each employee has a set of characteristics which are used to populate questionnaires that are given to them. Results are returned on screen and can be exported for analysis.

Figure 2. Environmental assessment and close up of a workstation.

The information available from the resources and interactions is supplemented by statements from the staff, see Figure 1. Scripts were based upon real-world experience and actors were used to bring the workers featured in the Virtual Workplace to life. Careful scripting allows the subjective information from the workers to match, or mis-match, data gathered from other sources, thus illustrating the complex interrelationships present in any workplace and reinforcing the need to take a holistic, evidence-based approach.

Pedagogic benefits

It can be seen from the description of the Virtual Workplace above that a wide variety of rich data can be obtained about the workplace and staff by the student. This reveals the complexity of issues, but also how information can be redundant or relatively trivial. That is, the issue is not just one of understanding a tool, but rather the appropriate use of appropriate tools.

Research evidence suggests that it is important to promote robust and usable knowledge through engaging learners in authentic tasks and situations (Hung and Wong, 2000). The Virtual Workplace for Learning does this and offers a number of further benefits. In particular, it allows staff greater flexibility to implement imaginative teaching scenarios that allow students to explore complex issues. It adds greater interactivity, increases the variety of teaching methods and allows students to engage with applied scenarios that demonstrate task based principles of theoretical module content. It is known that students are more actively engaged with a topic when a variety of teaching is used and students have the opportunity to relate material to a real environment (Savery and Duffy, 1996). This stimulation of active involvement is one of the primary goals in e-learning (Hedberg, 2003).

Current use of the Virtual Workplace

The Virtual Workplace is currently used in two modules, Introduction to Ergonomics and Musculoskeletal Disorders and Ergonomics. In the introductory module students are asked to consider the call centre workers within the Virtual Workplace and use available resources to collect relevant information. They are then asked to critically evaluate the physical and psychological issues that may impact on the workers, to identify key ergonomics problems and propose appropriate solutions. The scenario has been developed so that some sources of evidence support certain conclusions, but there are also differences between subjective and objective information. The aim being that students understand that issues can be, and indeed tend to be, complex and multi-factorial.

In the Musculoskeletal Disorders module the Virtual Workplace is used as a fundamental part of the module assessment, which relates to the evaluation of an individual who is experiencing musculoskeletal problems. In addition to the 'interviews' describing the job and workplace in general terms, students can view clips from the operator and supervisor dealing directly with the reported symptoms. This mimics a real-world scenario with contradictions both within and between each of the accounts. The emphasis given to potential issues by the actors is not necessarily a true reflection of the scale of the problems and some issues are referenced very subtly. Perhaps most importantly, the videos allow for the introduction of 'psychosocial factors' in a much more realistic way than could be achieved with a paper-based scenario. The assessment therefore reflects both practical considerations and current research.

Future development

Funding has been secured to develop an additional room for the Virtual Workplace (a canteen kitchen) together with additional video clips. Additional 'interviews' will be used within the Stress in Organisations module and will support an introduction to qualitative analysis. The kitchen will be used for further environmental and risk assessments together with posture analysis of kitchen activities.

Initial student feedback

As can be seen from the table of results below, initial student feedback has been largely positive. Firstly, and most importantly, all students felt that the Virtual Workplace was a valuable addition to the text based materials.

Whilst the vast majority of students agreed the Virtual Workplace was authentic, this wasn't strongly expressed, although given the technical and resource limitations this is still a pleasing result. Further comments indicated that extra attention to details, such as providing more realistic environmental tools, could provide greater authenticity.

The video interviews were seen as useful by all students, as they 'helped with intonations' and were 'great for getting a feel for the personalities involved'. The

Table 1. Results of student survey.

Question	SA	A	D	SD
The tasks and situations in the Virtual Workplace were authentic.	27%	67%	7%	0%
The interactivity of the Virtual Workplace was more engaging then the text based materials.	20%	60%	13%	7%
The assessment tools (environmental measures and questionnaire surveys) were useful.	0%	93%	0%	7%
The video interviews were useful.	47%	53%	0%	0%
Coursework based on a scenario within the Virtual Workplace is better than a simple text based scenario/description of a call-centre.	47%	40%	13%	0%
The Virtual Workplace helped me gain a true grasp of the complexities of 'real-world' ergonomics issues.	40%	47%	13%	0%
The Virtual Workplace is a valuable addition to the text based materials.	47%	53%	0%	0%

Key: SA = Strongly Agree, A = Agree, D = Disagree, SD – Strongly Disagree.

comments reveal how the Virtual Workplace and interviews also added to the coursework by revealing the complexity of the situation. Again a large majority of students agreed that the Virtual Workplace helped them gain a true grasp of the complexities of 'real-world' ergonomics issues, although a recurring comment was that the text based information in the modules also facilitated this.

Whilst referring to the text based module content, the lowest level of agreement from the student survey relates to the Virtual Workplace being more engaging than the text based materials. This result affirms the value of text-based information and the quality of current module materials.

The survey included the opportunity for students to make final written comments and many did, although there were few recurring themes. Repeat comments related to provisions of more detail and a wider range of video interviews (which reflects the goals of the authors), and finally to the provision of information off-line. Whilst transcripts of the videos have been made available (albeit that students may not have this luxury in real-world situation) a paper version of the Virtual Workplace is not considered feasible. Indeed it could be argued that greater realism could be gained by limiting the time the student could spend assessing the virtual workplace – such a restriction could be built into assessments in the future.

Conclusions

The student feedback shows that the Virtual Workplace has met its aims; it provides authentic situations and complex issues in a manner that makes it a valuable addition to text based materials. The Virtual Workplace is a novel development aiming to improve the practical understanding and skill development of e-learners in ergonomics. It has been developed by considering both pedagogic and practical

issues, but it also allows for the development of teaching and assessment which reflects current research (e.g. the consideration of psychosocial factors in MSDs). The development has been well received by both staff and students and its further enhancement is planned, to allow for more scenarios developing additional skills.

References

Hedberg, J.G. 2003, Ensuring quality E-Learning: creating engaging tasks,*Educational Media International*, 40, Numbers 3–4, 175–186

Hung, D. and Wong, A. 2000, Activity theory as a framework for project work in learning environments. *Educational Technology*, 40(2), 33–37

Savery, J.R. and Duffy, T.M. 1996, Problem Based Learning: An instructional model and its constructivist framework. In BG. Wilson (ed.), *Constructivist Learning Environments: Case Studies in Instructional Desig*n. Englewood Cliffs, NJ, Educational Technology Publications, pp. 135–148

Tolmie, A. and Boyle, J. 2000, Factors influencing the success of computer mediated communication (CMC) environments in university teaching: a review and case study. Computers and Education, 34, 119–140

Watson, M. and Horberry, T. 2003, The value of an online, distance education program in human factors. Contemporary Ergonomics 2003

INFLUENCE OF VIBRATION ON WORKLOAD WHILE READING AND WRITING ON INDIAN TRAINS

R. Hancock[1], N.J. Mansfield[1], V.K. Goel[2] & R. Narayanamoorthy[2]

[1]*Environmental Ergonomics Research Centre, Loughborough University, Loughborough, UK*
[2]*IIT Roorkee, Roorkee, Uttaranchal, India*

The effect of tri-axial random vibration on reading and writing was investigated. Additionally the influence of posture and language on these two tasks was tested. 20 participants (10 males, 10 females) were assessed using an adapted version of the Word chains methodology for the reading test, and a dictation task for the writing test. Difficulty in reading or writing while under four magnitudes of vibration, adopting a posture leaning forward on the table or back against the chair, or conducting in English or Hindi was determined using NASA-TLX, a measure of subjective workload. There was a significant effect of posture on reading, and vibration on both reading and writing, but no significant effect for language. Task difficulty while undergoing whole body vibration was also matched to the theoretical framework of Hockey's compensatory control model (1997).

Introduction

Reading and writing by hand are some of the most common activities while traveling by train. The difficulty of these activities can be increased by the motion and vibration that is induced in the bogie, car body and seats. This is due to irregularities in the track, passing over turnouts, rough sections of track and cross winds. The difficulties these vibrations cause to reading and writing is potentially a problem for certain passenger demographics, for example business people or researchers who wish to work while commuting. Factors influencing the degree of difficulty experienced include the vibration magnitude (Mansfield, 2005), sitting posture (Sundström, 2006), and whether reading or writing is in first or second language.

Measures of perceived exertion have been used previously to determine the degree of difficulty on reading and writing while riding on trains (Sundstrom, 2006). For this experiment the NASA-TLX was used as a more sophisticated means of finding subjective workload, a related, but separate concept (Louhevaara and Kilbom, 2005). Differences in posture, magnitude and task language were investigated to determine their effects on the difficulty in reading or writing, as measured by the NASA-TLX.

Method

20 participants drawn from the student population of IIT Roorkee in India took part in the experiment at the Vehicle Dynamics Laboratory of the Mechanical Engineering department. The custom built shaker platform, air conditioning and speakers were used to re-create what would be experienced in a typical Indian train while the two tasks were completed.

The participants were given either a dictation task (to determine writing difficulty during the experimental conditions) or pseudo semantically related word chains (to determine reading difficulty). These two tasks were given in one of two sitting postures (leaning against the backrest with the tasks on the lap, or leaning forwards with the task materials supported by the desk), in three different vibration magnitudes, and in Hindi or English. Random vibration with a frequency range of 1–20 HZ was used. The three vibration magnitudes were 0.2 m/s^2 rms, 0.4 m/s^2 rms and 0.6 m/s^2 rms, and a stationary control. The *x* axis remained a constant 0.2 m/s^2 rms while the other two axes were varied. Reading and writing tasks were given as pre recorded ambient train noise was played through loudspeakers, and lasted one minute each. Each condition was followed by the computerised NASA-TLX.

The pseudo semantically related word chains were derived from earlier work by Miller Guron and Lundberg (2004), and created from English and Indian newspapers to increase ecological validity. This required the participant to find where the break in a chain of four to five words (lacking spaces) should be (Table 1). The writing task required the participant to write down as much as possible from English or Hindi pre recorded audio dictations in the given time allowed. Each participant was exposed to twenty four conditions: three blocks of four vibration magnitudes for the reading trial, and then the same for the writing trial (Table 2). Each block comprised the three vibration magnitudes plus the control. The order of the vibration magnitude within each block was randomised, as was the order of each block for the reading or writing trial. Which task started the experiment was alternated with participant.

Table 1. Example Hindi and English word chains.

केसालइतिहासमेंआज	infactsuchfeehysection
एकअध्यायजुड़गया	reportfinalontheissue
एकमहिलाकेरूप	asrecommendedbythe

Table 2. Experimental Order.

Task	Block 1	Block 2	Block 3
Reading	Leaning forward\Hindi	Leaning back\Hindi	Leaning forward\English
Writing	Leaning forward \Hindi	Leaning back\Hindi	Leaning forward\English

The experiment was approved by the Loughborough University Ethical Advisory committee.

Results and discussion

There was a significant increase in workload score as vibration magnitude increased (Figure 1). Analysis of variance showed there was a significant effect of vibration magnitude on reading and writing workload score ($p = <.05$). For the posture variable, a second analysis of variance showed that there was a significantly lower workload score for the reading task when sitting back in the chair as compared to the sitting forward position, with the reading material supported on the lap (Figure 2). However, this was not the case for the writing task. No significant difference in workload score was found between Hindi and English (Figure 3).

Workload score increased with vibration magnitude likely because of the increased difficulty in maintaining performance in either task. While reading, task difficulty would have increased as the vibration could also have interfered with the discrete motor control required to mark the correct place on the word chains. During writing, workload increase with vibration magnitude was likely due to disruption of the fine motor control, and the interference with the reading process, an essential component of the feedback process. There were possibly detrimental effects to the graphemic output lexicon, where words are stored in short term memory.

The influence of posture on reading was most likely due to greater support of the upper body when sitting back, and the avoidance of vibration transmitted through the table that is particularly disruptive to the two activities. This is supported by Mansfield and Maeda (2007), who found the resonant frequency that most affects reading and writing was transmitted through the body when leaning forward. The muscles would have been tenser while holding a position resting on the table, which may have increased the amount of vibration transmitted. No significant difference was found between writing on the table or on the lap likely because the vibration was

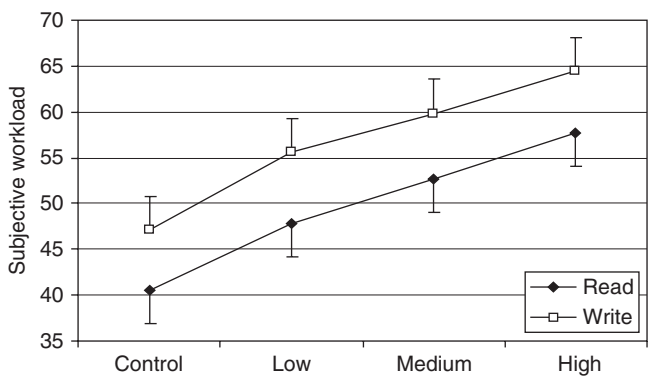


Figure 1. Effects of vibration magnitude on workload while reading and writing.

so disruptive to the writing process, which relaxed muscles could not adequately dampen.

No language difference was found. This was because the standard of English of the sample was very high, so the word chains and dictation tasks would have found to be easier than a less educated sample. This is supported by the majority (75%) of the sample at postgraduate level with a high self reported confidence in English. However, more work needs to be completed on the pseudo semantically related word chains to determine whether any semantic priming effects were occurring.

From this study a new framework for understanding the effects of random vibration on activities was tested. Hockey's compensatory control model (1997) is shown below (Figure 4). Loop A represents the automatic nature of the reading and writing process, with the switch to effortful control for loop B representing when the activity interference is occurring, and effort becomes a conscious process. However, it

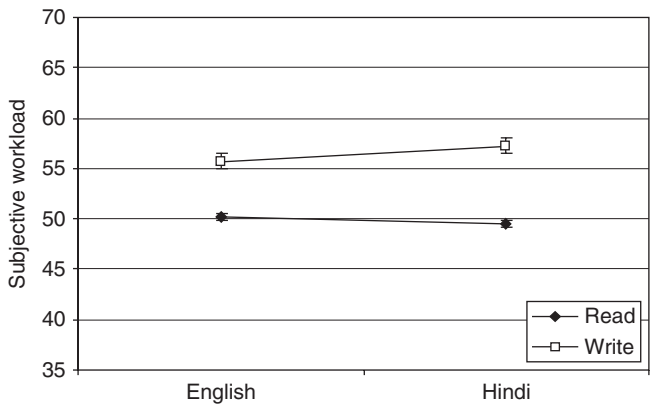


Figure 2. Effects of language on subjective workload while reading and writing.

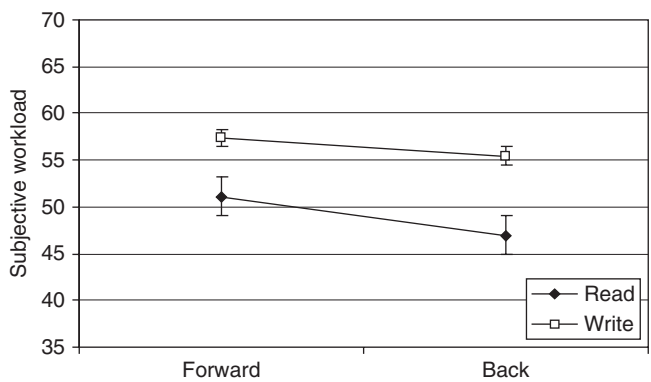


Figure 3. Effects of posture on subjective workload while reading and writing.

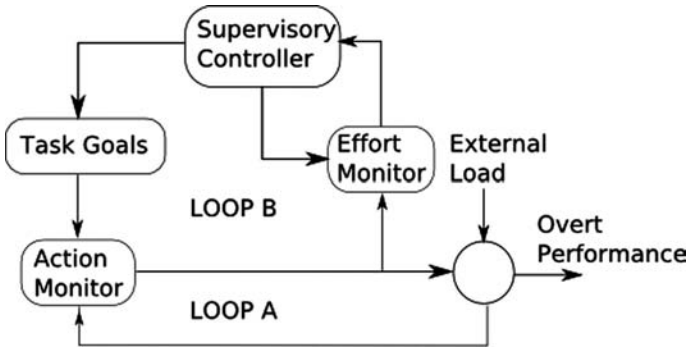


Figure 4. Hockey's Compensatory control model.

was likely that the tasks remained consciously controlled (i.e. in loop B) due to the novel nature of the tasks.

The effort and frustration subscales of the NASA-TLX increase with magnitude, showing that compensatory costs are likely being incurred to maintain performance. While the use of objective measures were considered it should be noted that these are not accurate means of inferring workload, and nonetheless are also often correlated with subjective measures. The use of the compensatory control model allows for further developments in vibration research to be built upon this framework. Further work into when switching from loop A to loop B or a change in motivational strategy occurs could also lead to the development of predictive models for random vibration, and new ISO standards.

Conclusion

Subjective workload increased with vibration magnitude for both tasks, but only while reading for the posture condition. Language difference of the task had no effect. The magnitude effects were most likely due to disruption to fine motor control for both reading and writing. The effect of posture was due to a combination of reduced stability of the torso, and the transmissibility properties of the table and backrest. No difference between workload scores for the two languages was found, most probably because of a combination of the high ability in written and spoken English, and the nature of the tasks. The compensatory control model goes some way in explaining the mechanism underlying the changes in workload scores between conditions.

Acknowledgements

This research was funded by the EU Asia-Link ASIE/2005/111000 CIRCIS (Collaboration in Research, and Development of New Curriculum in Sound and Vibration).

References

Hockey, G.R., 1997. Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, 45, 73–93.

Louhevaara and Kilbom, 2005. Dynamic Work Assessment. In J. Wilson and N. Corlett (Eds.) *Evaluation of Human Work.* 3rd Edition. (CRC Press, London).

Mansfield, N.J. (2005). *Human Response to Vibration.* (CRC Press, Boca Raton).

Mansfield, N.J. and Maeda, S. (2007). The apparent mass of the seated human exposed to single-axis and multi-axis whole-body vibration. *Journal of Biomechanics*, 40, 2543–2551.

Miller Guron, L. and Lundberg, I. 2004. Error Patterns in Word Reading Among Primary School Children: A Cross-orthographic Study. *Dyslexia,* 10, 44–60.

Sundström, J. 2006, Difficulties to read and write under lateral vibration exposure: contextual studies of train passengers' ride comfort. KTH Doctoral Thesis, Stockholm, Sweden.

COMPLEX SYSTEMS

USING ORGANISATIONAL SYSTEMS ENGINEERING TO IMPROVE THE BID PROPOSAL PROCESS

G.A.L. Kennedy, E-M. Molloy, C.E. Siemieniuch, M.A. Sinclair & R.J. Clayton

Systems Engineering Innovation Centre, Loughborough University

Loughborough University have been researching into various fields of Organisational Systems Engineering (OSE) which involve treating the enterprise as a system and modelling softer, organisational characteristics (such as role interactions, cultural values, competencies, decision making systems and enterprise strategy). This paper describes a case study undertaken looking at how OSE can be used to provide insight into problems being experienced within DefCo's (a large multinational engineering systems company) Bid Proposal phase. During the case study a number of "as is" models for process, roles and decision making were developed, gap analyses undertaken and recommendations for change made. In addition to this, a new framework of activities was introduced to provide a holistic view of the phase. The exercise provided DefCo with a clearer picture of the relationship between the process, the review activities and the roles therein.

Introduction

Organisational Systems Engineering (OSE) is about viewing any enterprise (regardless of size) as **a system**; made up of human, process and technical sub-systems which interact and interface with each-other. Typical characterising problem areas in their design and/or reconfiguration include:

- *Changing business models/transforming organisations*. How do you design an enterprise system that has the agility and resilience necessary to adapt and thrive in the business environments of the future? (Rouse, 2005)
- *Enterprise system architectures*. What are the key attributes a "resilient" enterprise system architecture needs to exhibit? Need a tool to simulate dynamic enterprise system architectures capable of re-configuration in response to rapidly changing commercial/technological contexts
- *Validation of enterprise systems*. How do you validate a proposed new enterprise system?

In 2007 Loughborough University (LU) were approached by DefCo to use a subset of OSE techniques to provide insight and recommendations for "decongesting" the Bid Proposal phase. The phase is one of 13 outlined in DefCo's project lifecycle. It is concerned with the planning and preparing of bid proposals and getting

approval. It takes place after a decision has been made whether to produce a bid and ends with the Request for Bid Approval. The phase is heavily burdened with the number of different types of reviews. The main impetus for this work was due to the following initial symptoms being noted:

- There was overload in terms of the number of reviews being conducted during the phase
- Each type of review has a different customer (commercial, technical and project management) and the reviews are not linked together leading to:
  - Duplication of work
  - Multiple/duplicated information flows
  - Unclear role boundaries and too much workload on those involved
  - Bottlenecks around reviews

Currently there are 6 different types of standard review that occur during (or within the chain of events) in the Bid Proposal phase. Each of the reviews is described in terms of their purpose in Table 1.

Due to the nature of the problem being addressed and the short timescales of the project (3 months) it was decided that the process, roles and decision making aspects of the organisational system would be investigated. A set of corporate baseline models were created that were sourced from formal material available (e.g. company documented processes) and these were validated by appropriate representatives from DefCo. 6 workshops were held with representatives who had experience of holding roles within each of the different types of review (Pink and Red being held in the same workshop). Where possible consensus was reached by those present in each workshop and any changes/differences from the baseline

Table 1. The Different Types of Review held within the Bid Proposal Phase.

Type of Review	Purpose
Pink Team Review	Review proposal (storyboards) from a **customer perspective**. Making sure the bid strategy is likely to fit with the customer
Red Team Review	Review proposal (near final) from a **customer perspective**. Ensuring the proposal sells the solution well
Design Review	Ensure that the **technical solution** being developed is mature enough to be used in the bid. Ensure that the solution suggested meets the customer's technical requirements.
Bid Status Review	Enable business leadership to review the **progress** of the bid proposal process, decisions made and coaching that is required.
Phase Review	To ensure that the bid is mature from a **project management** perspective, in terms of plans, timescales, resourcing, project risk and processes.
Technical Bid Review	Provide **engineering line management** endorsement of the solution, its costs, plans and technical risks.
Request for Bid Approval	Provide link management approval of the whole bid from a **commercial** perspective.

model were recorded. A gap analysis was performed on the different models and a revised baseline model was developed. Finally a set of issues was drawn out and a new framework for the activities suggested as an improvement.

Enterprise modelling

During the case study three types of enterprise modelling technique were utilised. Models of the process were first created (down to two levels of detail/decomposition), roles were then added onto the activities within the process, and the decision making systems (DMS), where each review is actually a decision making activity being made, were explored. It is the former two modelling techniques that will be described in detail in this paper. Further information about DMS modelling may be found in the accompanying paper (Molloy *et al*, 2008).

Process modelling

LU used the IDEF0 notation (NIST, 1993) in order to represent the process/activities that are being undertaken during the bid proposal phase. IDEF0 enables the modeller to represent each activity as a rectangular box. For each activity, there are inputs (what is being used by the activity), outputs (what are the outcomes/products of the activity), mechanisms (things that are enabling or doing the activity) and controls (things that can start or stop the activity). Figure 1 shows the notation.

Activities can be shown as sequences of activities and can be decomposed into sub activities. Activities in sequence on the same diagram are labelled in increasing numbers in the bottom right hand corner (e.g. the first rectangle has a “1”, the next a “2”, etc). For activities that have been decomposed the number moves on a decimal place (e.g. activity “1” contains sub activities “1.1”, “1.2”, “1.3”, etc)

Role modelling

LU have been developing a technique called the Role Matrix Technique for over a decade. This enables the modeller to represent the nature of involvement of a role for

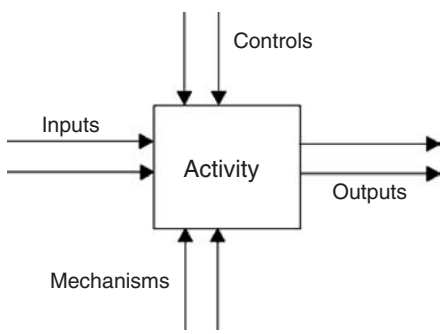


Figure 1. IDEFO Notation.

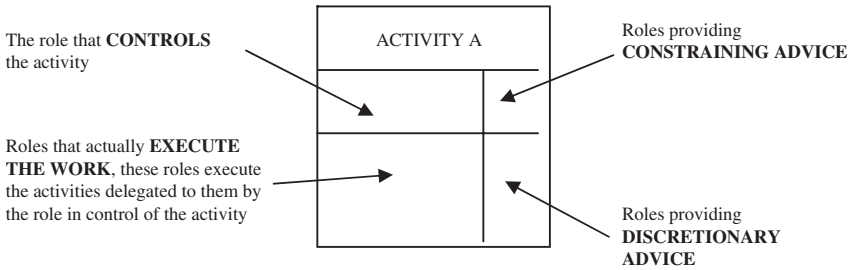


Figure 2. Role Quadrant Notation.

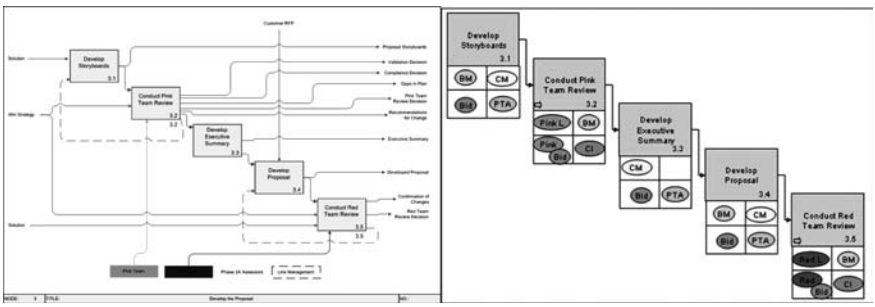


Figure 3. Corporate Baseline Models of Proposal Development.

a given activity and also how roles are interacting (Callan *et al*, 2006). For the case study Role Quadrant diagrams were used (figure 2). For each activity in IDEF0, there is a quadrant located directly below the activity. Roles are allocated to the activities depending upon if they are controlling, executing, providing constraining or discretionary advice. There is only ever ONE main role in control of any activity. There may be any number of roles executing and providing any form of advice.

Case study models

Due to the brevity of this paper it is not possible to show in detail the full sets of models that were created; instead one of the stages (the Proposal Development stage, which includes the pink and red team reviews) will be used as an example. The corporate baseline model of process and roles is provided (figure 3) to show the initial state of the organisation based on the formal documentation available. A second set of models (figure 4) shows the revised baseline version which takes into account any differences in how the "as is" is seen from the various viewpoints from the workshops. It is not intended that the exact details be analysed in this paper, but more the effect of comparing the views in a visual manner. It was evident that the same activities are performed, however in practice these tend to be done in a different sequence. Also the role modelling showed that not all envisaged roles are necessary for certain activities and the involvement can be streamlined.

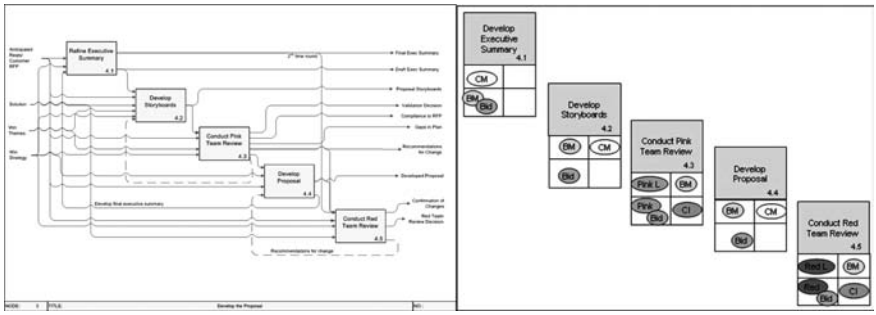


Figure 4. Revised Baseline Models of Proposal Development.

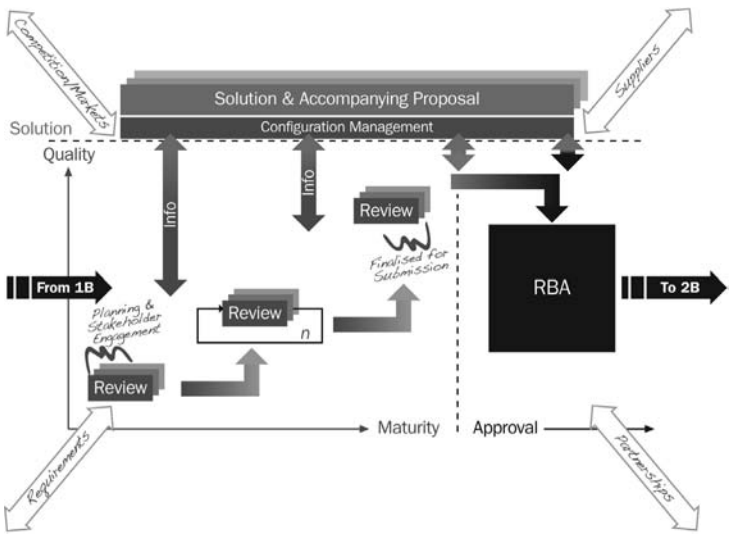


Figure 5. Proposed Integrated Activities Framework.

A new integrated activities framework

There are no major problems with the individual activities that are currently being undertaken within the phase. There was seemingly an issue of "disjointed" thinking occurring (i.e. compartmentalising into lots of separate independent reviews). In order to enable those within the phase to understand where the work they are doing fits in, it is important to ensure that there is visibility of the bigger picture. The integrated activity framework (figure 5) is intended to be an empty shell into which the activities in the revised baseline can be placed in order to build a holistic view of the phase.

Conclusions

Based on the findings of the case study, three key areas of change were recommended;

1. Enable tailoring – provision of a tailoring guide with examples for different contract types.
2. Avoid "too little, too late" – more left shift and planning will alleviate the lack of time towards the end of the phase
3. Take a holistic view – taking a bigger picture view will enable better planning and organisation in order to avoid confusion and duplication of work.

The case study demonstrates the value of using OSE through a small subset of the techniques. The case study enabled the company to gain insight into its current issues (with indicative evidence for possible root causes), identify various improvements and promote discussion/sharing between individuals involved in the process. The main conclusion fed back to DefCo was that the current set of activities they undertake are correct, but that the phase needs to be better organised and tailored in order to make it more effective. The results of the case study have been warmly received by the company and are being included within future work being undertaken by DefCo in order to make improvements to the phase activities.

References

K. Callan *et al.* 2006, *A Case Study Example of the Role Matrix Technique*, Int. J. of Project Management, **24**, Issue 6 (Elsevier).

E-M. Molloy *et al.* 2008, *A New Decision-Making Systems Framework*, Proceedings of the Ergonomics Society Annual Conference 2008, Nottingham.

NIST. 1993, "Integrated Definition for Function Modelling", FIPS publication 183.

W.B. Rouse. 2005, *A Theory of Enterprise Transformation*, Systems Engineering Journal, **8**, No. 4 (Wiley).

A NEW DECISION-MAKING SYSTEMS FRAMEWORK

Ella-Mae Molloy[1], Carys Siemieniuch[2] & Murray Sinclair[1]

[1]*Department of Human Sciences*
[2]*Department of Electronic and Electrical Engineering,*
Loughborough University

This paper introduces a Decision-Making Systems (DMS) Framework from the EPSRC-funded Grand Challenge, 'Knowledge and Information Management – through life' (KIM).

The aim of the DMS Framework is to help orient organisations to the classes of issues they experience and to assist them in developing DMS to cope with long life, complex, engineered projects and systems.

The framework has been developed by meta-analysis of accident reports, investigative studies and testing through a series of case studies, including aerospace and construction sectors. The use of the Framework is not to provide the 'right' answer, but to better inform organisations about the characteristics of their DMS and equip them to be resilient enough to deal with the situations they experience.

Introduction

The KIM Grand Challenge

The Engineering and Physical Sciences Research Council (EPSRC) in the UK has funded five Grand Challenges. All involve academic and industrial partners and aim to improve the performance of UK businesses (specifically in manufacturing and construction). 'Knowledge and Information Management – through life' (KIM) is the biggest of these Grand Challenges, being funded at approximately £5 million.

KIM involves 11 academic institutions and a number of industrial partners. The overall aim of the project, as stated on the project website (www.kimproject.org) is:

> 'The identification of approaches to information and knowledge management that may be applied to the through-life support of long-lived, complex engineered products.'

The work is split into four work packages and each of those into sub tasks. The focus of the research reported in this paper is the third sub task within the third work package – 'Managing the Knowledge System Life Cycle'.

More information may be found on the project website, www.kimproject.org.

The scope of task 3.3

The team personnel are from the Universities of Salford, Cambridge, Reading and Loughborough.

The scope of the research is to investigate the effect on decision-making and decision support of a shift from product delivery to through life service support. Support timescales may be 30 years or more, during which time the 'information and knowledge' will be stored, accessed, used and received many times over in many different situations and contexts.'

The issues which arise from consideration of such a time frame can be summarised as follows: beyond 30 years we will be designing (and servicing) products and systems whose requirements and uses are not yet known, using materials and processes not yet invented or developed, using suppliers who will be very different to now and who may have a shorter lifetime than the system itself, and all of this will be done, including maintaining system information and knowledge, with people who are not yet born and who will not be in post for more than a decade. This has been termed as the 'looking forwards problem'. There are also things to be learned 'looking backwards' i.e. how are legacy systems dealt with that were first designed and produced many years ago, e.g. the RAF Canberra photo reconnaissance aeroplane, designed in the 1950s and still flying today?

Building the framework

Our starting point for the research work is that bad things happen to good people. This research looked at two issues from this initial fact:

- Can these bad things be categorised?
- What contributions do decisions make?

Meta-analysis of accident and incident reports

Accidents are often linked to bad decisions, but in the increasingly complex systems and organisations in which we work, accidents can be the output of a series of decisions which seem good at the time they are made. Accidents often happen as a result of people simply performing their normal duties. Decision-making is affected by a number of things. There will be external environmental and commercial pressures, but there are also internal effects and pressures.

The investigation of these incidents led to the identification of a number of commonalities in the incidents, such as: complacency; ineffective process; inadequate review activities, inadequate resource allocation, poor allocation of responsibility and authority; poor communication and ineffective reporting processes, for example.

Pilot studies

There were three pilot studies. Two involved student groups completing Systems Engineering projects and Multi-User Design projects. The third involved a construction study on a cutting-edge educational building.

The pilot studies were used both for initial framework development and to develop the methodology to be used in future industry-based studies. The studies investigated how decisions were made, what decision support was available, what decision support was used and if any key decision points could be identified. A variety of techniques were used in the studies, including workshops, observations and surveys. All studies included document analysis.

Initial framework development

A DMS (Decision Making System) is defined as consisting of the following:

- Agents – software or humans and who are involved in decisions;
- Activities – the decision-making activities which enable decisions to be made;
- Infrastructure and technology – underpinnings which enable decisions to be made;
- Knowledge and information – that which is necessary for decision-making.

The knowledge and information provides support. It flows round the infrastructure, through activities to agents to allow decisions to be made. This is the function of a DMS.

The key decision-making issues identified within the accident investigations were categorised and the DMS framework created, as in figure 1. There are a number of facets of the organisation or system which could affect or be affected by the components of the DMS. These were identified by grouping the commonalities identified within the accidents. They are:

- Internal variable – for example, at what stage are you in the Engineering Life Cycle and what impact does this have?
- Environmental variables –for example legislation and health and safety rules.
- Organisational culture – issues such as: power distance, risk aversion/acceptance, regimentation and collaboration (individual work vs. collaborative work).
- Level of decision-making – strategic, tactical or operational.

Testing case studies

This consisted of three industry-based case studies. All of the case studies follow the same basic methodology; existing documentation is analysed, followed by workshops and/or interviews to elicit viewpoints across the organisation. All case studies were within an aerospace organisation.

The information gathered refers to processes, role allocation and decision-making. From this, evidence is extracted and formatted within the DMS framework.

Further framework development

Following the case studies, it is evident that there are some areas which the framework does not address, such as workload and training. Other issues are also being identified, it will be decided later whether the framework should or should not address these.

Feature of DMS	Issue
A. Agents/ Roles	Poor role/ agent definition
	Poor role/ agent allocation
	Non-availability of roles/ agents
B. Activities	Inappropriate activities
	Poor definition of activities (unclear or fuzzy boundaries)
C. Infrastructure and Technology	Inappropriate infrastructure
	Non-availability of infrastructure
D. Knowledge and Information	Inappropriate knowledge and/ or information
	Non-availability of knowledge and/ or information

Figure 1. Snapshot view of Framework.

The key aims of the DMS framework are:

1. Identify attributes of decision-making
2. Provide examples to explain what happens if the attributes are not right
3. Enable analysis/ diagnosis of DMS
4. Identify useful tools

Future work

The first exercise will develop a method for identification and analysis of patterns within the framework. Initial patterns that can be identified are concentration (where a particular cell, row or column has a high proportion of issues) and voids (where a particular cell, row or column has an absence of issues). The future work aims to identify the meaning of groups of these voids and concentrations, and to identify whether there is further insight to be gained through comparison of similar patterns

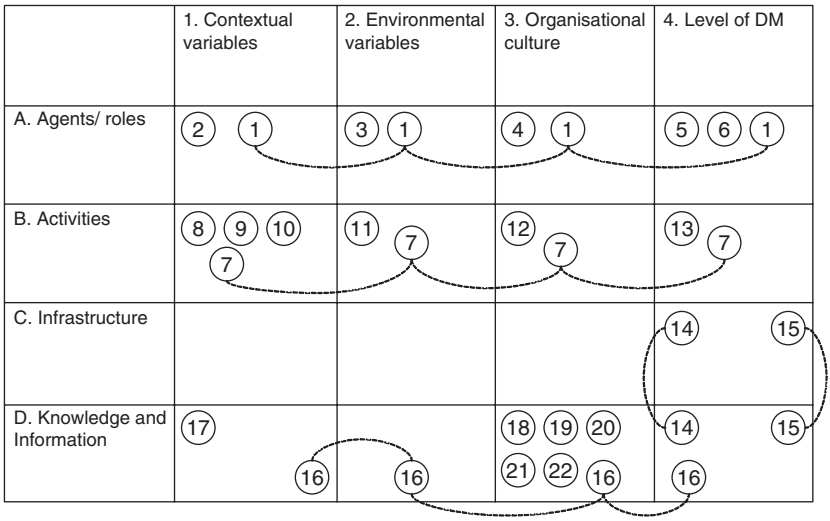


Figure 2. Representation of a case on the framework.

in multiple cases. There will also be a further investigation into the implications of time within decision-making systems. This work will be integrated into wider Enterprise Modelling work, ongoing within the research group.

The main step to be taken is the development of a demonstrable tool which will better indicate how the DMS framework may be used by organisations.

Conclusions

There are some drawbacks to the framework. It relies on successful honest identification of issues and subsequent categorisation. This is a very subjective process and can yield different results from different people and also from the same person at different times. This must be taken into account when the final framework tool is used.

It is not a tool which designs effective decision making systems, nor does it suggest an outline for the perfect decision making systems. In this situation, one size does not fit all. The aim is to provide more information to an organisation about their decision-making system so that they may make more informed alterations and adaptations to better achieve their overall aim.

Anecdotal evidence from the studies completed so far support a number of observations:

- Not all decisions are consciously made, which has serious implications for decision-making systems and any form of decision support.
- Individuals tend to view themselves as outside the organisation. This makes it easier to say that the organisation is failing rather than admitting self-insufficiencies.

- Especially when complex systems are involved, success in small scale testing does not always ensure success when the real thing is implemented on a much larger scale.
- The product-service shift demands new processes, not shoehorning new projects into old, inefficient processes, thus making the organisation blind to the risks they might encounter.

Poor decision-making is often a symptom of underlying organisational issues. This is the reason for taking a wider view of decision-making systems within the organisation.

A step change is needed in the knowledge of how to design, integrate, operate and evolve systems that are not fully understood by all stakeholders, whose behaviour may not be fully predictable and which function in an environment that cannot always be controlled.

The end goal of this work is the ability to identify key decision points and guide the appropriate configuration of DMS to enable an organisation to better deal with complex, long-lived, engineering projects.

MANAGING HUMAN FACTORS TRADE-OFFS IN ARMOURED FIGHTING VEHICLE PROJECTS

Robert Cummings

Atkins Ltd, Bristol, BS32 4SY

The management of Armoured Fighting Vehicle (AFV) requirements presents a particularly challenging task for Systems Engineering. Not all AFV stakeholder requirements are compatible; therefore a degree of requirements Trade-off is required to ensure that the system meets Cost, Time and Performance targets. Trading human related requirements often has an impact on the overall capability delivered. Therefore, the role of the Human Factors Engineer within the Systems Engineering team is to identify requirements that may conflict and manage the trade-space between them to ensure that the system delivered meets the customers' operational needs.

This paper discusses a process that is suitable for use in the identification and management of Trade-offs for AFV requirements. The paper focuses on AFVs; however, the methods discussed are equally applicable to other domains.

Introduction

A trade off occurs when one requirement is compromised in whole or in part to deliver another more important requirement. One of the primary objectives of Systems Engineering is to develop a System through the identification, management and optimisation of System Requirements. In Ministry of Defence (MoD) projects, these are recorded within the System Requirements Document (SRD) which is developed to provide the bridge between the customers User Requirements (URs) and the contractor's system design (Cummings and Ranasinghe, 2007). The development of System Requirements and the requirements optimisation activity will often require a degree of Trade-off between the different Systems Engineering segments. This will ensure that an optimised solution that meets customer Cost, Time and Performance targets is produced.

Inevitably, all systems are the product of a degree of requirements Trade-off as not all requirements are complementary. For example, in AFVs the requirement for a small, manoeuvrable, stealthy vehicle will often conflict with the requirements for armour protection and accommodation of the user population. When considering all of the elements of an AFV, (such as weapons, armour, and sensors) Human Factors requirements are sometimes compromised by the more prominent demands of other subsystems. However, if the vehicle cannot perform its primary function because the user cannot complete their tasks, then it becomes obvious that too greater Trade-off has been made.

Assessing the impact of a Trade-off

The Trade-off process can be considered in 6 key stages as follows; Identify the boundaries of the system; Identify relationships between requirements; Identify evaluation criteria; Prioritise evaluation criteria; Analyse system solutions; Optimise system solution.

Identify the boundaries of the system

The first stage of any Trade-off process has to be to identify the boundaries and constraints of the system design i.e. the absolute minimum or maximum requirements that the design must comply with (Blanchard and Fabrycky, 1990). Within Defence Projects many of these requirements can be obtained by consulting the 'User Requirements Document' (URD) developed in the early phases of acquisition by the MoD.[1]

In AFVs, the role of the vehicle is the key to identifying the areas where trades can and can't be made. This is because understanding the scope of the role will define the underlying boundaries and constraints imposed on the vehicle thereby defining the available trade-space.

The human will often present both an enabler and a constraint within the overall systems design (Hutchison et al, 2006). Therefore, it is important to understand the acceptable limits of human performance when assessing design options. For example, it may be a requirement to accommodate the 5th to 95th percentile user; it may also be a requirement to accommodate the population in a certain posture. However, a degree of trade may be taken on the posture adopted as long as the user can still safely perform their tasks. In this instance the constraint is to accommodate the user population as the user enables the system to deliver the capability.

Identify relationships between requirements

Once the boundaries of the system are understood, the Systems Engineering team must determine how different requirements are related. Identifying how different requirements affect one another is the key to understanding where Trade-offs can be made. The Human Factors Integration Defence Technology Centre (HFI DTC) suggests a variety of methods that can be used in formulating a HF argument within a Trade-off process. These include Goal Structured Notation, Quality Function Deployment and Cognitive Mapping (Hutchison et al 2006.) Each of these methods are excellent in their own right, however, the simplest and easiest to apply is probably cognitive mapping, as no specialist tools are required. Cognitive mapping enables the user to rapidly illustrate how different requirements relate to one another. They aim to simply illustrate how different design options and trades affect one another, thereby providing a quick reference for assessing the affect of trades. A simple example of how cognitive mapping may be used in the Trade-off process is given in Figure 1. In this example there is a positive relationship between crew comfort,

[1] Often the Concept and Assessment phases.

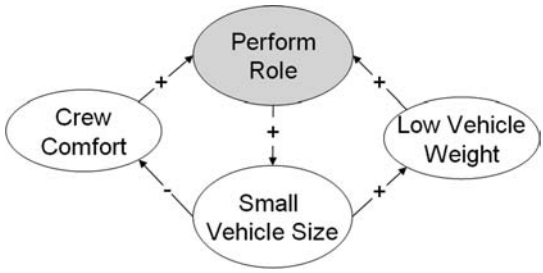


Figure 1. Cognitive mapping example.

the vehicle being small & light and the vehicle performing its role. However, it also illustrates that a small vehicle size may negatively impact the users comfort.

Understanding the relationship between requirements will go some way to understanding the impact of Trade-offs. However, considering the impact that a trade may have on an overall capability will require further analysis.

Identify evaluation criteria

Multi Criteria Decision Analysis (MCDA) is a Systems Engineering technique that facilitates the evaluation of an option that has multiple attributes. Therefore, this technique is useful when trying to understand the impact of different Trade-off decisions on overall capability. The first stage of this process will be to identify the criteria that the Trade-off decision will be judged against.

The first level criteria are Cost, Time and Performance. Performance may then be broken down into a series of lower order criteria. The second order criteria may represent the User Requirements,[2] which can then be broken down into highest priority and second order System Requirements as illustrated below in Figure 2. Not all System Requirements will be employed within the model because not all requirements have a direct influence on capability. Therefore, only the requirements that discriminate between different design options are employed within the Model. This should ensure that a degree of sensitivity is maintained as well as reducing the time taken to produce and run the Model.

When selecting assessment criteria, it is important to ensure that the hierarchy captures requirements concerning all of the Human Factors Integration (HFI) Domains (Manpower, Personnel, Training, Human Factors Engineering, System Safety, Health Hazards and Organisational & Social Factors). This will ensure that the full breadth of HFI issues is captured within the assessment process. HFI requirements will generally cut across all segments within Systems Engineering, therefore not all HFI requirements will sit within the Human Factors hierarchy. Some requirements, such as those for Training may fit within another segment, such as Supportability. The high level Training requirements may include requirements

[2] As Defined in the User Requirements Document (URD).

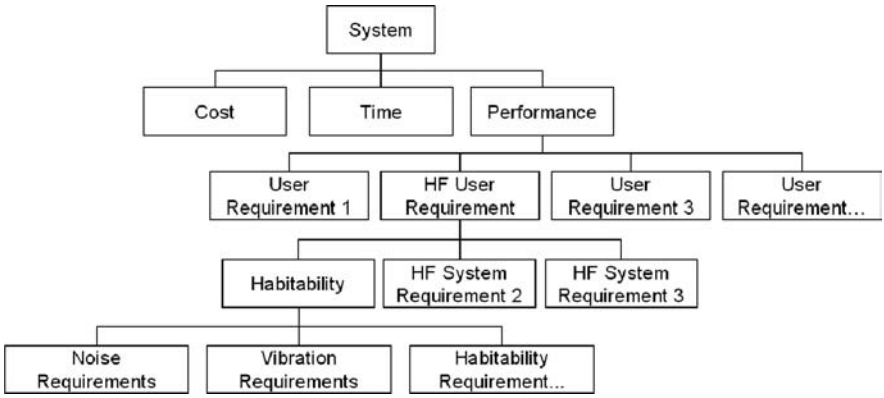


Figure 2. Hierarchy of parameters.

to integrate with existing Training solutions, or to support the adoption of a specific Training strategy.

A scoring system must then be determined for each of the Evaluation criteria identified. The scoring system (also known as Value Functions) must be identified for each of the requirements within the hierarchy. The value functions must be unique and measurable to ensure that all design options can be judged on the same criteria and to ensure consistency of scoring. A score of 0 to 4 is awarded based on how closely a candidate option comes to meeting required and desired performance. An example of how the scoring system may be implemented for a requirement to accommodate the 95th percentile user is given below:

1. Design fails to accommodate 95th percentile User (design is unacceptable)
2. Design accommodates 95th Percentile User but not within the postural limits of comfort (design falls well below the requirement)
3. Design accommodates 95th Percentile User and enables them to perform their role but may cause them to suffer minor discomfort (design falls just below the requirement)
4. Design accommodates 95th Percentile User and enables them to perform their role within a comfortable posture (design meets the requirement)
5. Design accommodates 97th Percentile User in a comfortable posture (design exceeds the requirement)

Inevitably there will be some evaluation criteria that are critically important to delivering the capability that the customer desires. However, these requirements may appear anywhere within the hierarchy. In these instances, a ‘Red card’ must be defined. Red card requirements cannot be traded if the system is to deliver the capability required. These criteria will therefore rule out a design option if the design fails to meet the requirement. Examples of requirements that may have a

Red card associated with them include the requirements to comply with Safety standards, such as the Health & Safety Executive at Work regulations.

Prioritise evaluation criteria

Once the hierarchy of requirements for the system is understood, the relative importance of each of these requirements then needs to be judged. Pairwise comparison provides a means of deciding the relative importance of each requirement within the hierarchy. Requirements are compared in pairs across each level in the hierarchy to judge which of each pair is preferred. For example at the top level Cost is judged against Time and then Performance. Subsequently each of the sub-hierarchies is judged against one another, ultimately producing a weighting factor for every branch and leaf within the hierarchy.

The Pairwise process evaluates attributes against one another, scoring relative importance against a predetermined scale.[3] When producing the hierarchies both the customer and all major stakeholders must be involved in producing the weightings, to ensure that the meaning of requirements is maintained and the importance of the requirements represented.

Analyse system solutions

Once the scoring system is in place and the weightings have been applied to the hierarchy, design options can be judged against each of the criteria and an overall system score produced. Some designs may be instantly 'Red-carded' due to failing to meet a critical requirement. However, the method will provide a means of comparison between design options, at a system and segment level. Therefore, an overall score, as well as a score for each branch in the hierarchy can be obtained. The selected alternative will usually be the one with the best overall weighted score (Haskins, 2006). However, it is advisable to perform a sanity check to ensure that the model has properly taken into account the relationships between requirements discussed above. The strength of the methodology is that it enables the user to determine where small design changes in one area may have a great affect on performance.

Optimise system solution

In some circumstances requirements will not be achievable irrespective of where trades are made. This may be due to the User requirement being too onerous or simply the technology that meets the requirement not being mature enough or affordable at this time. Therefore, a strategy needs to be put in place to ensure that the capability delivered can be grown to meet the requirement. This can be done through planned product improvements over time or by adjusting the original requirement. If a capability is to be grown the hierarchy may be used to illustrate the improvement over time.

---

[3] Often a scale of 1 to 9 is used

Conclusion

In AFV projects it is a significant challenge to deliver a solution that fully meets all user requirements within the boundaries of Cost, Time and Performance. In order to design balanced systems, a multi-disciplinary approach must be adopted; this must encompass the requirements of each segment and weight them in accordance with the customer expectations of the final systems solution. MCDA provides a quantitative approach to assessing different design options that can aid in balancing the requirements of complex systems.

Although focused on a defence industry case study, the methods presented in this paper are equally applicable in all industries. The paper highlights the need to involve all stakeholders in the Trade-off process and the importance of understanding the role that the system must perform in order to identify where trades can and can't be made.

References

Blanchard, B.S. and Fabrycky, W.J. 1990, *Systems Engineering and Analysis*, Second Edition, (Prentice-Hall, London)

Cummings, R.C. and Ranasinghe, S. 2007, Techniques for Defining Human Factors System Requirements. In P.D. Bust (ed.), *Contemporary Ergonomics 2007*, (Taylor and Francis, London), 79–84

Haskins, C. 2006, *Systems Engineering Handbook, A Guide for System Life Cycle Processes and Activities*, Version 3, INCOSE-TP-2003-002-03

Hutchison, G. Maddock-Davies, K. Mock, M. and Newman, P. 2006, *HFI DTC: Improving HFI in Design Trade-offs*. HFIDRC/WP 3.7.7/1 Version 1

DESIGN

ERGONOMICS ASPECTS OF VIDEO CONFERENCING

[1]Niels C.C.M. Moes, [1]Adrie Kooijman, [1]Ellemieke van Doorn & [2]Nuša Fain

[1]*Delft University of Technology, The Netherlands*
[2]*University of Ljubljana, Slovenia*

Team cooperation in physically distributed design teams often uses video conferencing (VC) systems. Since humans have developed specific characteristics for interpersonal communication, VC should support such characteristics accordingly. Students were educated for such cooperation in masters design courses, where international, interdisciplinary teams solve a practical design problem in cooperation with a company. A key aspect was the formation of an Academic Virtual Enterprise. While they used new communication technologies, such as VC, several ergonomics problems were detected in relationship to the progress of the design task. This paper presents the main ergonomics aspects that we found for VC.

Introduction

Globalization in product design manifests in (i) internationalisation of cooperation and (ii) application of new communication technologies, that have found a way into e.g., video conferencing (VC), electronic white board and file sharing. However, as it happens often with new technological development, the consideration of human factors is not always trivial, for instance the compatibility of effects and human expectations. We think that a graduated Industrial Design Engineer (IDE) must be prepared for such cooperation. The education for working in global context requires extended design competence (Horváth, 2006), for which purpose the concept of an Academic Virtual Enterprise (AVE) was developed (TU Delft), as well as Global Product Realisation (GPR), a design course for master students.

The basic issues in bridging physical distance for having communication as natural as possible means talking to and seeing each other. During a normal meeting, documents are shown, paper drawings are made, technical information is exchanged, etc. VC offers a simulation of the most important aspects.

During the GPR courses we observed that the product design process was significantly slowed down. Apart from multi-cultural, international and disciplinary backgrounds, we think that the communication via VC showed several ergonomics imperfections. To find the ergonomics clues for VC, we studied the recordings of the last couple of years of the GPR courses.

Lacking further research on the behaviour of people in a VC session, we assume that VC communication should resemble regular communication. Although new

technologies and software for VC, file exchange, short messaging, white board sharing, etc., have emerged during the last decades, we stress that distance communication is always different from in person communication, but some technicalities can make a lot of a difference.

Hine (2001) defined three usage scenarios for VC. *Video meetings* focus on exchanging information, using mainly audio channels, but video adds visual cues. Typically it has a small (2–5) number of participants per (less than ten) location. Communication dynamics are expected to follow that of face to face meetings with minimal interference from technology. *Presentations,* usually lectures, are often conducted by a slide show and managed by a moderator. In *Collaboration* meetings small groups communicate using not only audio-visual channels, but in addition a set of communication and collaboration tools.

This paper discusses (i) the main human factors of VC in a global product design context and the connection to technical means, (ii) experiences from educational courses on global product design and international cooperation of multi-cultural, international and multi-disciplinary (MX) teams of students.

Ergonomics aspects of VC communication

This section discusses the main ergonomics aspects of VC, and how they can be controlled by the technical characteristics of the VC system, based on our findings during seven GPR courses. First the concepts of AVE and GPR will be clarified, then our ideas about the visual, auditive and audio-visual ergonomics aspects of VC.

Since product design is usually based on multi-disciplinary cooperation, the different disciplines in a design team have a varying contribution. The disciplines use their own dictionaries, peculiarities of drawing, document types etc., so that the information transfer must happen according to a common protocol. In virtual context this is relatively difficult, and must be facilitated by appropriate technologies. Additional difficulties are the multi-cultural and multi-national aspects, where the virtual context may hide the differences of habits, language, lawgiving, hierarchical structures etc. In academic IDE education only little attention is given to the MX aspects of designing and the arising issues of communication (Chan et al., 2006).

An AVE (Horváth et al. 2003) has been invented to be a virtual cooperation between an industrial company and an international, multi-disciplinary and multi-cultural design team. The problem to be solved is forwarded by the company. The team members are taught to cooperate using advanced technical communication systems. The GPR course is an actual embodiment of an AVE in design education. It is organised as a one semester course for several international design teams. It has three main phases: (i) forming a vision of the product in relation to the company, the user, the society, etc., (ii) generating ideas for and the selection of a concept for final elaboration, (iii) the materialisation of the final concept and the assembly in a one week workshop at a location close to the company. before this week the cooperation is solely by VC techniques. The student teams represent three or four universities.

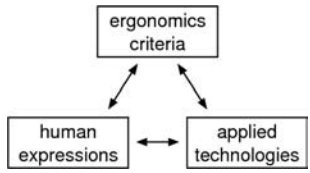


Figure 1. Ergonomics aspects of VC

The most important ergonomics criteria for optimal VC relate to enabling the transmission of human visual and auditory expressions, which must be enabled by appropriate technologies, see figure 1.

Visual aspects

In VC the visual cues for human expressions, such as facial expression, eye contact (Vertegaal et al., 2003), posture, gestures (Verhulsdonck, 2007), and clothing, are as important as in normal human communication. They can be effectively communicated if certain ergonomics aspects are taken care of. The most important aspects are (i) the smoothness of the projected movements, (ii) the size of the projected image, (iii) the contrast and the brightness, and (iv) the composition of the image, including zoom. The applied technology must enable this by (i) sufficiently high resolution of the image and (ii) have a sufficiently high refresh rate. This means that the bandwidth must be sufficiently high and a suitable codec must be applied. Usual values of the transmission bandwidth are between 56 kB/s and 2 MB/s. The picture resolution and the refresh rate are closely related. Often used are low (352×288 pix) or medium (704×576 pix) bandwidth and high definition (1280×720 pix, 2 MB/s, 30 frames/s and Hi-Fi audio) (HD). HD means practically that the thumbnails in a $1 + 4$ picture (see below) have about the same detail as the large picture in lower resolution.

In a one-to-one conversation natural communication happens if the image of the partner is about life-size and at a natural distance from the speaker (Koolstra and van Daalen, 2006). In a (large or small) group this can be achieved using a large screen.

The visual details related with the intensity distribution depend on the brightness and the contrast of the ambient light conditions, the clothing and the background. Problems may result in white faces and black clothing.

In natural communication partners look each other in the eyes. Looking in a different direction is experienced as uncomfortable and disturbing the communication. Therefore, the cameras should be placed close to the centre of the screen (otherwise it is better to look in the camera). During a lecture this effect is less pronounced. Koolstra and van Daalen (2006) did an experiment using partial reflecting mirrors enabling the effective camera position in the centre of the screen. They found a significant improvement of the eye contact, if the angle of the axes of the camera and the eye is smaller than 3°.

When more than two sites join a VC a multipoint conference unit (MCU) is used. It enables the virtual presence of the participants with the following possibilities: (i) speaker activated so that one of the sites is the main presence, (ii) continuous presence (multiple monitors) and (iii) split screen (two or four split) or a 1 + n split. The disadvantage of voice activation is the obscuring of the other sites. Multiple monitors make it impossible to place the camera at the best position. A particular problem is displaying more partners on a relative small screen such that (i) it is clear who is talking, (ii) contact with the silent partners is maintained, (iii) enough detail remains in the picture. Usually the principle of continuous presence is applied showing the speaker in a large part of the screen, keeping the others in small sub-screens (1 + n). The disadvantages include unexpected switching behaviour, the speaker in slightly lower resolution, non-speakers in low resolution. This can be improved by HD or voice activation, showing the speaker full screen to all participants. New systems offer the possibility for one site (lecturer) to have an overview of all participants and the other participants seeing only the lecturer.

Web cam and computer based programs (Skype, net-meeting, Breeze) have low video bandwidth (56 kB/s), causing low interoperability with clients using other networks. Collaboration software (Breeze) have low/medium video quality. The focus is on sharing documents together with video and audio for collaboration purposes. Standalone solutions (e.g., Polychrome, Lundberg) offer high quality video and audio, support multiple screens, camera's and microphones, easy operation, and are suited for conference rooms. The focus is on professional use: boardrooms, distance education and team meetings in a multi-x environment. Focus is on professional person to person meetings.

Auditive aspects

In communication by speech the human expressions can be understood if the technology supports the main ergonomics aspects related to intelligibility, direction of the sound, handling the background noise, intonation and loudness. The technological aspects include the frequency spectrum, volume control, reverberation of the room, the acoustical feedback, the location of the microphone(s) to the speakers and representation of the direction of the sound (mono, stereo). According to Parker and West (1973) optimal conditions exist if (ii) the frequency spectrum of the transmission media reflects human hearing characteristics, (ii) the background noise is limited, (iii) image and sound are synchronised correctly, (iv) the directional information corresponds with the visual image of the speaker. Figure 2 shows summarises the auditory aspects of a VC system. The system must allow switching on/off the microphone for local discussion, reduction of background noise and acoustic feed back.

Traditional VC systems have mono sound, which is sufficient for one-to-one conversation. In small or large groups stereo sound helps locating the talking person. Correct placement of the microphones allows that (ii) the speaking person is heard sufficiently loud and undistorted, (ii) the background noise (e.g., moving paper, tapping fingers) is minimised. (iii) the echo from either the room (reverberation) or via the VC system (Unintended Auditive Feedback, UAF) is minimised. In a

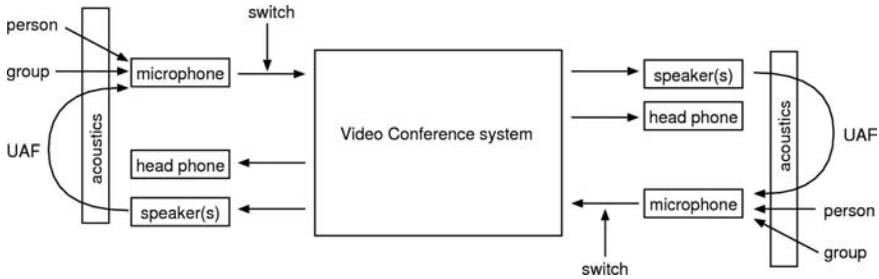


Figure 2. Auditive aspects of communication using VC system

one-to-one or small group meeting this is achieved using a modern VC system, but in a larger group it requires careful placement of the microphones(s).

Observations during GPR courses

Up to now seven GPR-courses have been conducted, which made us aware of unfulfilled ergonomics requirements. We categorised them in two main areas: purely technical aspects (e.g., vision, sound, synchronisation, document sharing, and controls) and social aspects (body language, cultural, regional and disciplinary aspects, language).

Although hearing each other is equally or even more important than seeing each other, sound is often considered of secondary importance. Most efforts are put into the visual aspect, while the quality sound is often bad (caused by background noise, echo's, sound artefacts). We have also experienced sound related problems caused by ill-adapted equipment (distortion and low level of sound). In practice remote microphones are often left switched on for interruptions or questions, but depending on the level of background noise it may be better to mute remote sites. Particularly so called "room microphones" turned out to be very sensitive for ambient noise. Most standalone VC systems come with only one room microphone that can do an excellent job if used correctly: not too far from the participants, not too close to loudspeakers or other noisy equipment. The ambient noise can be almost unnoticed locally while it is very much disturbing during a VC, for instance if the microphone is placed near a projector or a laptop.

Using a beamer projection as the main screen makes it difficult to place a camera close to the line of sight so that it is difficult to achieve eye contact, which is a disturbing factor particularly in discussions and team meetings. Using systems with partial reflecting mirrors can give a significant improvement.

Lecturers often forget that they are in a VC, loose contact with remote locations, do not look in the camera (wrong camera placement) or address only the local audience. If a slide show is presented it did often happen that the lecturer points to the local screen, where it is better to use a mouse pointer instead.

A new aspects of human interaction arose since the students first met through the VC system and worked together without ever meeting in person for most of the

course time. Development of trust, acknowledgement of each others competences and formation of the team structure was done virtually. All the communication channels that usually acquire physical presence were transmitted through VC.

Conclusions

Many ergonomics problems related to VC/GPR courses were solved relatively easily. Three examples: (ii) increasing the distance to the camera improved the eye contact, (ii) the continuous moving of the camera for changing the prime presence was solved by applying pre-sets for direction and zoom, (iii) extremely slow document sharing was solved by distributing the document to local computers and changing slides on command.

Important issues for refinement and research remain such as (ii) the direction of sound in a group VC, (ii) optimisation of the balance between image resolution and refresh rate, (iii) the minimal resolution for facial expressions, (iv) the effects of the image zoom parameter, (v) guidelines for the best light conditions (ambient, background, clothing, contrast), (iv) the correspondence between the optical axes of the camera and the eye, (vi) the size of the projected presenter during a lecture or presentation for a group to (partially) compensate for not being present in the room.

References

Chan F, Kamminga R, van de Langkruis AS, Lankhaar J, Pastoor JL, Zeedijk W and Moes CCM, 2006, *Communication in International Academic Virtual Enterprises*, Proc. DESIGN2006:1085–1092, D Marjanović editor, Dubrovnik, Croatia.

Hine N, 2001, *Issues of H.323 low cost client development*, Ukerna/VIP, U. of Dundee.

Horváth I, DuhovniK J and Xirouchakis P, 2003, *Learning the methods and the skills of global product realization in an academic virtual enterprise*, European J of Engin Education, 28–1:83–102.

Horváth I, 2006, *Design competence development in an academic virtual enterprise*, Proc. IDETC/CIE sept. 2006, ASME, Pennsylvania, USA.

Koolstra A and van Daalen T, 2006, *Verbetering van oogcontact bij videoconferenties*, Internal Report Onderzoeksleer, Delft Univ. of Techn, fac. Industrial Design Engin.

Parker JF and West VR, 1973, *Bioastronautics Data Book*, Second Edition, (national Aeronautics and Space Administration, Washington).

Verhulsdonck G, 2007, *Issues of designing gestures into on-line interactions: implications for communicating in virtual environments*, SIGDOC'07:26–33.

Vertegaal R, Weevers I, Sohn C and Heung C, 2003, *GAZE-2: Conveying eye contact in group video conferencing using eye-controlled camera direction*, Conf. Human Factors in Comp. systems, Ft. Lauderdale, Florida, USA.

AVOIDING PROJECT FAILURE: PUTTING PEOPLE FIRST

Luke Barrett & Marc McNeill PHD

ThoughtWorks Ltd., Berkshire House, London, WC1V 7AA

The consistent failure of software development projects to deliver what is expected of them is a significant headache for businesses. It has been claimed that only around a third of such projects can be regarded as successful and that one of the primary reasons for this lack of success is poor requirements management. Requirements are what drive any development process, yet without a shared understanding between those who produce the requirements (the business), those who turn them into tangible software (the developers) and those who ultimately use the software (the end-users), project failure becomes ever more likely. Using a case study this paper discusses the use of participatory techniques and lightweight models to describe and explore the problem and its potential solutions; applying highly iterative, feedback-driven and people-centred techniques from the outset to provide clarity in what is really required and so position teams for success.

Introduction

Software projects have a poor record of delivery. Whilst methodological issues have been raised (Glass 2006), the Standish Group's periodic CHAOS Report (Standish Group International 2004) suggests that only around a third of such projects can actually claim success. Defects in requirements are a major source of the defects that are later identified during testing, and problems with requirements are among the top causes of project failure (Schwaber 2006), indeed it has been estimated that 71% of projects fail due to poor requirements management (Lindquist 2005). The Standish Group point to lack of user input together with incomplete / changing requirements as the leading causes of 'challenged' projects. On the flip side they identify user involvement, executive sponsorship and clear requirements as the most important in ensuring success.

This paper discusses the role of effective participation of the right stakeholders in helping to drive project success. Focusing that participation around lightweight models that are tangible and visible, and doing 'just enough' analysis, helps ensure buy-in from the outset and increases the likelihood of successful project delivery.

Participation

Ultimately the vast majority of software solutions are about helping people (end-users) to achieve their desired goals more effectively. The challenge for the software development community is how to go about consistently delivering solutions that do just that. Addressing this challenge begins with communication between three important groups:

1. The people paying for the solution (in commercial software development often referred to as 'the business')
2. The people who are going to use the solution (the end-users)
3. The people who are going to implement the solution (the do-ers: project managers, business analysts, quality analysts and software developers)

While the importance of including end-users in the development process seems obvious (after all they are the people whose improved productivity, reduced error rates and increased morale established the business case for the solution in the first place) all too often they are missing from the equation. Despite the work of Norman and Nielsen (and many others) over the last decade to raise the profile of end user involvement; despite Standish's work pointing clearly to the importance of end-user involvement; and despite the rise of the Web (and particularly Web 2.0 with its emphasis on rich user interactions not previously associated with web applications), too few software development projects take a user-centred design approach.

The increasing popularity of 'Agile' software development approaches (e.g. Martin 2002) provide an opportunity to address this. Agile approaches are inherently people-centric being highly iterative and feedback-driven with that cycle of feedback being remarkably short in comparison to traditional methods. Whereas a traditional 'waterfall' approach to development might see a nine to twelve month gap between the business specifying a requirement and seeing that implemented, in an agile project that might be as little as a week. Such a rapid turn around of a requirement into functioning software allows teams to 'fail fast' – that is obtain feedback from end-users at the point of need, not many months after an implementation decision has been irrevocably (or at least expensively) embedded in a design.

While the omission of input from end-users is one of the more obvious issues to address, more challenging is ensuring the correct balance of representation from the business. The goal is to achieve frequent input from those who have the best grasp of the business need – this may mean input from multiple departments and certainly means input from various levels of seniority. In addition there will be members of cross-cutting functions who will also have necessary contributions to make (e.g. legal, security, operations, etc.) although the frequency of their required input is likely to be less.

The point about seeking input from multiple levels of seniority for the key owning business area (or areas) is worth restating. It is likely that on a day-to-day basis the business area representative on the team will be someone of moderate seniority, that is they are senior enough to have a good grasp of the business (and to be empowered to make decisions concerning the solution up to a point) but junior enough that this

is their day job (or a good chunk of it). Again one leading cause of project failure is lack of executive sponsorship so using the highly tangible output of the process (see below) at regular showcases (weekly/twice weekly) to engage with, and extract feedback from, senior business representatives is essential.

Lightweight models

With a multi-disciplinary team in place the challenge becomes to find a quick and clear way for the team to generate and evolve potential solutions to the problem at hand. The solutions must be captured in such a way that all three key constituencies can grasp them and provide rapid feedback. It is here that simple lightweight models of the problem domain and possible solutions to it come into their own, taking the following form:

1. **Strategic**: A simple prioritised list might be used to call out and agree key project objectives with their associated metrics while a financial model might be used to quantify the benefits case. It remains remarkable how many projects proceed with no clear quantification of the benefits they may return.
2. **Process**: The high value, high frequency processes that the solution needs to support are modelled. Typically these models explore how key user types achieve their highest value goals. Again the models are lightweight ideally at a level of abstraction that requires no more than seven to nine high-level steps. The aim here is to provide a framework for the more detailed decomposition of the solution which will occur at the next level down – that of implementation.
3. **Implementation**: The team is now distilling the output of the higher level models and capturing them as business requirements of sufficient detail that the development team can estimate the delivery effort. Importantly, at this point, the team's understanding of the written requirements should also be validated by the creation of low fidelity prototypes of key usage scenarios. On the development side the need to provide estimates should also prompt, as required, the creation of a proposed architectural model and options for key technology choices.

The process by which these artefacts are evolved is of at least equal importance to success. Experience suggests that, for a project of sensible scope, two to four weeks of effort evolving the models, distilling the requirements, prioritising them (from a business perspective) and estimating them (from an implementation perspective) will prepare a team for release planning and then development.

The initial portion of the process (the first week or so) tends to be heavily workshop focused. The team work together in short timeboxed sessions to collaboratively evolve the models and requirements. Where possible updates occur 'live' in the workshop, otherwise feedback is incorporated in dedicated consolidation time between sessions. As an example, using marker pens and index cards, it is easy to model and refine a high-level process 'live': cards can be re-ordered, torn up if necessary, replacements provided, new ones added. Likewise high-level requirements captured on index cards have the same flexibility.

As the process continues the intensity of workshops reduces (although a basic heartbeat of workshops remains in place to ensure focus) allowing more time for basic consolidation of the artefacts, 'offline' detailed analysis, workplace observation and usability testing. Key to this processes ability to address the two themes of engagement and clarity of requirements, is the way in which the multi-disciplinary team is intimately involved in the evolution of the artefacts that represent their shared understanding. The team see these artefacts taking shape in front of them based on their continuously requested feedback. And that feedback is in response, not just to written requirements, but to models (particularly the low-fidelity prototypes) that give them a really tangible feel for what's being discussed. Combine this with regular (e.g. weekly) presentations to senior stakeholders and the approach provides a powerful mechanism for engaging stakeholders across multiple areas and levels of seniority as well as for driving clarity around any proposed solution.

Process in practice

A case study for this approach follows an investment bank developing a new Client Relationship Management (CRM) tool. Four weeks were set aside for requirements definition with a focused core team and subject matter experts identified when appropriate. A dedicated project room was allocated. The project commenced with a kick-off meeting where all stakeholders were invited. The executive sponsor introduced the project, stated its importance and the need for people to make time in their diaries if required. Borrowing from Hohmann (2006), innovation games were played to help the team create a vision of the future, and what the risks to the project were. From this the project objectives were distilled and an initial risk log drawn up. A second workshop, (again with the whole team) identified high value high frequency user types and created personas (Cooper and Reimann, 2007) for each of them. Each workshop lasted ninety minutes, giving the team time to consolidate in the remainder.

The core team then walked through the 'as-is' business process, using index cards to illustrate each process step. These were stuck on the wall. On a second wall the 'to-be' process was mapped out. Using cards allowed elements to be moved around or removed and torn up if not necessary. Having them on the walls ensured visibility, and interest from stakeholders. They could see progress being made and found it easy to comment and make suggestions.

As the 'to-be' process stabilised (i.e. became less volatile following a number feedback loops), a lo-fi prototype began to take shape. Again, this was modelled on paper and key screens were again put on the walls. A technical architect was present during all workshops and was able to begin to propose a technical solution to support the process. By the end of the first week the team were able to present ('showcase') a proposed 'to-be' process supported by pen and ink drawings of how the application may look / behave. The stakeholders provided feedback and some radical changes were made – with the low fidelity of the artefacts nobody felt precious to the work done and changes were easily accommodated.

As the volatility of the design reduced, the lo-fi prototype was committed to PowerPoint and a business analyst started to document requirements as stories in the format 'as a [user], I want to[requirement], so that[value]'. Guerrilla usability testing (Nielsen, 1994) of the prototype was undertaken to refine the interaction design. Technical 'spikes', rapid and time-boxed technical investigations of potential functionality, were performed (such as a Google map – Microsoft Outlook calendar 'mash-up') to demonstrate feasibility and help with the estimation process.

By the third week the team had an extensive list of validated requirements that had high level implementation estimates. These were printed on cards with values on them (these values were arbitrary numbers for the exercise, reflecting an indicative magnitude of effort but not expressed in real days or cost). The business were then invited to prioritise the cards by 'buying' features. This was done in four rounds simulating four releases, with each release comprising of end to end functionality (rather than cherry picking high value features without the supporting low-value functionality). This process was iterative, indeed the planning continued through the development lifecycle.

At the end of the four weeks the stakeholders had a shared and common vision, prioritised and estimated requirements, a release roadmap and a lo-fi prototype that articulated how the application should look and behave. Four months later release one development was complete with the application rolled out to the business two months after that. During the development process changes and improvements were made based upon the real software rather than the lo-fi prototype. Release one delivered functionality that delivered immediate value to the business, with further features being delivered in subsequent releases. As users had been involved in the design and usability had been built into the process from the start, the tool was sufficiently intuitive to require no training.

Conclusions

The seeds of success or failure are often sown early in the life of a project. By applying highly iterative, feedback-driven and people-centred ways of working right from the get-go businesses have a better chance of addressing the lack of engagement with key stakeholders (senior management, end-users, etc.) and poorly defined requirements that, time and again, lead to project failure.

References

Cooper, A. & Reimann, R., 2007, *About Face 3: the essentials of interaction design*. Wiley Publishing.

Glass, R. L., 2006, *The Standish report: does it really describe a software crisis?* **49**, 8, 15–16.

Hohmann, L., 2006, *Innovation Games.* (Addison Wesley).

Lindquist, C., 2005, *Fixing the Software Requirements Mess*. CIO Magazine, November issue.

Martin, R. C., 2002, *Agile Software Development, Principles, Patterns, and Practices*. Prentice Hall.

Nielsen, J., 1994, *Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier*. http://www.useit.com/papers/guerrilla_hci.html

Schwaber, C., 2006, *The Root Of The Problem: Poor Requirements.* Forrester Research, Inc.

Standish Group International, *Chaos report: Chaos Chronicles.* http://www.standishgroup.com/

THE DREAM ECONOMY – DESIGN, EMOTION AND THE USER EXPERIENCE

Patrick W. Jordan

Patrick W. Jordan Ltd., Candlesby House, Butleigh, Somerset, BA6 8TH, UK

After going through various economic eras – commodity, manufacturing, service and information – we have now entered the era of the Dream Economy.

In the Dream Economy, success in the marketplace is not only about meeting people's practical needs, but also about meeting their aspirations and providing a positive emotional experience.

The key to success is an in-depth and holistic understanding of people – their needs, their hopes, their fears, their aspirations, their dreams.

This paper gives an overview of the Dream Economy and includes three case studies illustrating companies who have succeeded through a deep understanding of consumers and the delivery of an offer that meets their needs and desires at a deeper level. A framework for in-depth understanding of people is given.

The dream economy

As economies and markets have developed, they have gone through a number of distinct phases. Originally, markets were driven by trade in commodities. The Industrial Revolution then created the Manufacturing Economy, which until the 1980s, was the dominant model in the Western world.

Then came the dominance of the Service Economy which was rapidly complimented by the Information Economy. Recently these two have been dominating both the market place and employment in the world's leading economies.

Now we are moving into a new era. Consumers increasingly want goods and services that reflect their attitudes, lifestyles and aspirations. People want great experiences and an enhanced self-image, they want to express their values and convictions through their purchase choices.

The key to success in today's marketplace and the marketplace of the future is to understand consumers as people – their hopes, their fears, their values, their aspirations, their dreams. We are now in the era of the Dream Economy.

This paper looks at how to succeed in the Dream Economy and at the role of user-centred design in this success. It looks at how to gain an in-depth understanding of consumers and three case studies are given of products, services, brands and marketing campaigns that have succeeded through companies' understanding of consumers and their ability to connect with them in an emotionally effective way.

The four pleasures

The Four Pleasures is a framework that was developed by Canadian anthropologist Lionel Tiger (Tiger, 2000). Tiger looked at societies all over the world and analysed the different types of positive or 'pleasurable' experiences that people can have.

He concluded that, for all people, there four broad categories of positive experiences that we can have – he calls these the Four Pleasures. They are as follows.

Physio-Pleasure. This is to do with the body and the senses. It includes pleasures associated with touch, taste and smell, as well as feelings of sensual pleasure. It also includes pleasures associated with physical enablement, such as being able to perform physical tasks.

Psycho-Pleasure. Pleasures associated with the mind such as being able to understand things and positive emotional states. Mental challenges come into this category as do things that people find interesting.

Socio-Pleasure. This is to do with relationships, both in the concrete and abstract sense. Concrete relationships are those with specific people, such as friends, family, co-workers, neighbours and loved ones. Abstract ones are concerned with our relationship with society as a whole, such as our social status, image and memberships of social groups.

Ideo-Pleasure. These include our tastes, values and aspirations. Tastes are to do with our preferences – what colours we like best, what kinds of music and art we like for example. Values are to do with our moral belief system and our sense of right and wrong. Meanwhile, our aspirations are to do with our sense of who we want to be and the self-image of ourselves that we want to have.

To understand people deeply and holistically we need to know what is important to them with respect to all four of these dimensions. This knowledge can enable us to give people positive experiences, enhance the quality of people's lives and help them to fulfil their dreams (Jordan, 2002).

Case studies

The following are case studies showcasing three companies that have enjoyed huge success in the Dream Economy era.

Harley-Davidson

Harley-Davidson have enjoyed phenomenal success in recent years. After falling on hard times in the 1970s and 1980s the company has fought back and become one of the world's most successful automotive firms. My the middle of the first decade of this century, the company's total stock market value exceeded that of General Motors.

Harley achieved this success through understanding the appeal of their brand image and realising that although they make motorcycles, they are not really in the motorcycle business. Rather their business is about selling lifestyle.

Harley has a heritage that has led to the brand developing certain associations – freedom, rebellion, youthfulness and the American way of life. Their research has shown that people buy Harleys because of these associations and not because they

believe the motorcycles to be superior technically, or in terms of performance, when compared to other motorcycle manufacturers.

Instead of trying to make motorcycles that were faster than Kawasaki, more reliable than BMW or better built than Honda, Harley-Davidson simply concentrated on their image and heritage.

The design of their motorcycles has changed little over the years and they have not added that many new motorcycles to their range. Instead they have concentrated on emphasising the positive values that people already associate with the brand and this has been the focus of their marketing and advertising campaigns.

They have proved particularly successful with the 'baby-boomer' generation – people who were born in the decade following the Second World War and who grew up in the 1960s. The values associated with Harley closely mirror the values of the 1960s generation and now that many boomers are at the top of their professions and have a large disposable income, they can go out and buy the motorcycle that they have always dreamed of.

The question in the minds of many of these people is not, 'shall I buy a Harley or a Honda?', but rather 'shall I live the dream and buy the Harley I have always wanted, or shall I do something "sensible" such as invest in the stock market or build a conservatory on my home?'

Harley's strategy has addressed this question head-on. A recent TV commercial portrayed an older man in his 70s or 80s who regrets investing in stocks rather than having brought a Harley. The message is that if you are in your 50s or 60s you still have 20 years to enjoy riding a Harley, so don't do the 'sensible' thing, go out and live your dreams.

They also created a special club, The Harley Owners Group (HOG) which people automatically become a member of when they buy a new Harley-Davidson. HOG organises group ride-outs and offers merchandise and regalia reminiscent in style to the Hell's Angels.

By emphasising the Harley lifestyle, the company have made owning and riding their bikes into an escape from the routine of every day life and every trip a little adventure. It has proved an extremely appealing approach and made Harley-Davidson one of the most successful firms in the Dream Economy.

Starbucks

Founded in Seattle in 1971, Starbucks have established themselves as the world's leading coffee house and have become a huge global brand. Starbucks put a huge amount of effort into finding and blending great coffee and training their baristas to prepare it just right. However, it is not just the coffee that has driven their success, but the whole Starbucks experience.

The company's strategy is to try and make their stores what they call the 'third place' in people's lives – the home being the 'first place' and work the 'second place'. The idea is that Starbucks is a place you can go to relax and read the paper, hang out with friends, do some work, or even have a business meeting. And the longer that people spend in the store, the more money they are likely to spend.

Initially Starbucks core customer was highly-educated professional women based on the West Coast of the USA (where Starbucks started out). To many of these

women, the idea of the European street café or coffee-house was attractive. In Europe there was a tradition of people spending time over their coffee reading, chatting or playing games.

Compared to the American diner or fast food outlets, the atmosphere in the European cafés was relaxed and sophisticated and very appealing to Starbucks core customer base. What was not appealing to them however were European standards of service. For all the sophistication of the European café, the thought of having to wait a long time for service from a surly waiter did not appeal, especially to professional women with busy lives.

Essentially, what Starbucks offered was a sophisticated European atmosphere, but with American standards of service. It was a combination which proved irresistible and made the company into one of the worlds most well recognised and appealing brands.

As the company grew, it appeal rested on a combination of perceived status, luxury and sophistication. Even if people just got a coffee to go, there was a sense of pride in walking down the street with a Starbucks cup, the prominent logo there for all to see. Buying a coffee at Starbucks sent out a message – I am a successful, sophisticated person and I can afford to pay $2 for a cup of coffee. Indeed the high price was all part of the appeal, giving the perception of exclusivity.

The success of Starbucks has created a whole new market sector with many competitors coming into the market with similar offers and also proving successful. Examples include Seattle's Finest and Peets in the USA and Costa Coffee and Café Nero in the UK.

L'Oreal

Cosmetic company L'Oreal have proved extremely successful in recent years, experiencing double-digit growth year on year for over a decade. Remarkably, they have achieved this at a time when many of their competitors have been having difficulties.

One of the reasons that L'Oreal have been so successful is that they have an in-depth understanding of the social and lifestyle trends that affect their market. In particular, they have understood contemporary ideals of femininity and have positioned the brand perfectly to reflect these.

In post-feminist society 'tough glamour' has become the new ideal. This is about women being strong, smart and assertive but also being feminine and taking care of their appearance at the same time. This is reflected in the way that women are portrayed in the media. Buffy the Vampire Slayer and Xena the Warrior Princess have proved huge TV hits for example.

Looking at the role of women in James Bond movies – often considered a barometer of female glamour – it is noticeable that they have become far more assertive over the years, Where once they may have simply been the love interest or someone who needed to be rescued by Bond, now they are just as likely to be the ones rescuing Bond or the ones trying to kill him!

It is not only in the make-believe world of movies that we see tough-glamour. Women sports stars such as the Williams sisters also embody this trend. Perhaps

more importantly, the tough glamour ideal is something that many women recognise in their everyday lives.

More and more women are juggling busy lives and building impressive careers. While in the past their might have been the perception that success in the workplace required women to 'behave like men' – something that was often said of Margaret Thatcher for example – this perception has now disappeared. There is no contradiction between being a successful career woman and being feminine and taking care of your appearance.

L'Oreal encapsulate the spirit of tough glamour in their tagline, 'Because I'm worth it.' They are not trying to tell women how they should look, but rather empowering them with a message that it is up to them to decide how they wish to look and how they wish to express their beauty. And whatever they decide L'Oreal will provide them with the products that will enable them to achieve this.

In their TV commercials, the company uses well known women to promote their products, such as successful actors and musicians. They are not just unknown models, but rather women of recognisably high achievement.

Their website features articles about many of their senior scientists, who are all women, talking about their work, their interests and how they have built their careers. The company is also one of the biggest sponsors of the Women Into Science and Engineering (WISE) initiative which encourages girls and young women to study technical subjects at school and university. They also award substantial prizes to the world's leading female scientists.

L'Oreal have understood contemporary femininity better than any of their competitors, have built a brand to reflect this, and then, through their initiatives, built a community around the brand. Their in-depth understanding of their customers has been rewarded with huge success.

Conclusion

In the era of the Dream Economy understanding people is the key to success. The Four Pleasures provides a framework for gaining an in-depth and holistic understanding of users. As well as looking at people physically and psychologically, it also looks at their values and aspirations – their hopes, their fears and their dreams.

In order to create a design and marketing strategy that is relevant it is important to understand the behaviours, attitudes and lifestyles that are prevalent in society both now and in the future and to create products and services that connect with these. By understanding people, knowing what they want and meeting their needs, we can create products that are useful, usable and a genuine joy to own and use, both now and in the future.

References

Jordan, P.W., 2002, *How to Make Brilliant Stuff that People Love* (Wiley).
Tiger, L., 2000, *The Pursuit of Pleasure* (Transaction Publishers).

THE CULTURAL IMAGERY USED IN ICON DESIGN

Hsiu-Feng Wang & Kai-Li Wang

Department of Commercial Design, Ming Chuan University, Gwei-Shan, Taoyuan County, Taiwan

The aim of this research is to investigate the appropriateness of Taiwanese cultural imagery in icon design and to compare it with standard imagery. The method we used in this experiment was recognition testing. Seventy-four Taiwanese citizens participated in this experiment and were divided into groups according to their computer experience. A questionnaire for assessing the appropriateness of icons was used in the experiment. The results reveal that, overall, Taiwanese computer users recognised cultural icons more accurately than standard icons.

Introduction

Most software deemed appropriate by its manufacturers for a particular nation, has help files and menus which are translated into the language of that nation. However, language is only one way in which nations differ. Nakakoji (1994) reported that software designed and launched successfully in one country does not necessarily suit people in another country because differences in culture can cause misunderstanding. It is thus essential to consider other factors, such as cultural symbolism, metaphors, imagery and colour usage (Sears *et al*., 2000).

Po and Chuan (1999) compared two teams of designers who worked, independently, for Motorola on the design of pager products for the Chinese market. One design team was based in the USA and the other in Singapore. The results from the two teams were significantly different and reflected the cultural backgrounds found in each country. Another experimental study conducted by Fang and Rau (2003) examined the effects of cultural differences between Chinese and US users on the perceived usability of World Wide Web (WWW) portal sites. They found significant differences in satisfaction between each group, and in the number of steps each group used to perform the same task. The study indicated that cultural differences affect usability and task performance and that there is a need to investigate the effectiveness of icons on specific populations.

Onibere *et al*. (2001) carried out research in Botswana, a multi-cultural country which has two official languages (Setswana and English). One area of investigation was to find out whether Batswana (Botswanian citizens) users would prefer a localised interface. The survey results indicated that Batswana overwhelmingly desired such an interface. Another piece of research (Wang, 2007), conducted by the author, suggested that cultural icons do assist Taiwanese computer users, but

the majority of test icons used in the Wang's experiment were alphabetic. It is not known, however, whether concrete icons with cultural imagery were more appropriate than concrete icons with standard imagery for Taiwanese computer users. Therefore, this paper aims to investigate the appropriateness of Taiwanese cultural imagery in icon design, and compares it with standard imagery.

Method

Participants

Seventy-four Taiwanese citizens participated in this experiment. The age of the participants ranged from 21 to 35 years. The experiment took the form of a recognition test (Zwaga and Easterby, 1984) using icons on a computer screen. Participants were asked to match the icons shown to referents written on a list. The term *referent* is used in this paper to describe the name given to an icon by its program designer.

Prior to the commencement of the experiment, all participants completed a questionnaire. The questionnaire collected personal details and data relating to English ability and computer experience. After participants had finished the icon recognition test, they completed another questionnaire to assess the appropriateness of the testing icons used in the experiment.

All participants had similar English ability. Depending on their computer experience we divided participants into two groups: Group A participants familiar with computers and Group B, participants not familiar with computers

Materials

The recognition test used 14 icons shown in an environment which emulated the widely used word processing package 'Microsoft Word'. This method of icon presentation was undertaken by drawing the icons to be used in a computer graphics program and then importing them into Microsoft PowerPoint for presentation on a computer screen. 28 labels were also provided on a sheet of paper.

The icons employed in the experiment are shown in Table 1. The icons in the table are shown categorised according to whether they are *standard* icons or *cultural* icons. A standard icon is one where the design and the concept are to be found in international versions of software. A cultural icon is one that has cultural, national, or local features. In order to avoid the bias of icon design, standard icons used in the test are similar to, but not the same as, those frequently used in software packages such as Microsoft Word (see Table 1 for details).

This experiment uses the term 'label' to refer to a possible name for an icon. A label can be either a referent (the name given to an icon by its program designer) or a dummy label. In this experiment, the term 'dummy label' refers to a label which is not an icon's referent. Dummy labels were used in the experiment to increase the number of alternative answers from which a participant could pick. The labels used in this experiment were written as a list in Chinese (all referents in the list were the names of icons as used in the Chinese language versions of the software packages). The order of the labels on the list was randomised for each participant.

Table 1. Testing icons.

Referent	Standard icons	Cultural icons
1. Calendar		
2. System		
3. Library		
4. Stop		
5. Trash can		
6. Home		
7. Shadow		

In total 28 labels were shown in the experiment, of which 14 were referents and 14 were dummy labels.

Procedure

Prior to the tests, participants who had never used a computer before were shown Microsoft Word and were given a brief but succinct explanation of what the program was for and how it worked. This was done not to introduce the participant to Microsoft Word *per se* but to give an overview of how computer interface tools enable computer users to carry out tasks.

Each participant tested was individually shown a mix of cultural and standard icons; this mix consisted of both concrete and alphabetic representations. The icons were shown on a computer screen one after the other in a random order. Participants were asked to match each icon shown with a label from a list of 28 on a sheet of paper. Participants were told that there was no restriction on how often a label could be used and that they should take as much time as they needed to respond to each icon shown. After completing the experiment, each participant was individually asked to give reasons for their choice for each icon at an informal interview designed to elicit their opinions. The experimenter recorded all comments made during the process.

Results

Recognition of standard and cultural icons

Table 2 shows the number of correct labels chosen by each group with respect to standard and cultural icons. A total of 518 icon viewings were performed in the recognition test for each group (14 different icon designs were shown to all of the 37 participants in each group).

Table 2. The number of correct labels chosen for each icon.

Standard icons	Group A	Group B	Cultural icons	Group A	Group B
Calendar	34	18	Calendar	30	28
System	15	6	System	30	20
Library	27	9	Library	28	26
Stop	35	22	Stop	36	31
Trash can	34	18	Trash can	13	4
Home	33	14	Home	35	16
Shadow	15	9	Shadow	35	28
Total number of correct labels for each group	193	96	Total number of correct labels for each group	207	153
Total across Groups A and B	289		Total across Groups A and B	360	

Table 3. The number of correct labels chosen with respect to icons.

	Group A	Group B
Standard icons	193	96
Cultural icons	207	153
All icons	400	249

Combining the results of both participant groups, the total number of correct labels chosen for the standard icons shown in the icon recognition test was 289, compared with 360 for the cultural icons (289 and 360 in Table 2). In other words, Taiwanese computer users recognised cultural icons better than standard icons ($\chi^2 = 7.77$, 1df, $p < 0.05$).

When comparing the number of correct labels chosen for standard icons with cultural icons across the two groups, the pattern is different ($\chi^2 = 35.13$, 1df, $p < 0.05$; see Table 3).

When Groups A and B are looked at separately, the results show that participants familiar with computers (Group A) selected a similar number of correct labels for standard and cultural icons (193 and 207 in Table 3, $\chi^2 = 0.49$, 1df, $p > 0.05$).

This result differed from that for those participants not familiar with computers (Group B) who found cultural icons easier to recognise than standard icons (96 and 153 in Table 3, $\chi^2 = 13.05$, 1df, $p < 0.05$).

To summarise, the results seem to indicate that cultural icons assisted Group B in the task of icon recognition. The same icons, however, did not seem to further assist or hinder Group A.

Results of interview

Most participants in Group A based their decisions on their knowledge of word processing packages and/or other computer packages and, knowing what sorts of commands are found in programs could figure out what the cultural icons

represented. As Participant SL stated: ‘I use the computer everyday. Even though I am not sure what the bucket is (cultural trash can icon) for, a bucket is intuitive. A bucket can only mean something to do with getting rid of rubbish’. Participant HN stated: ‘I see this icon (standard home) when using the Internet . . . . I also see it when I design websites. Even though it is not identical to the one I used, I can recognise it’.

In contrast, the participants in Group B, who had little computer experience, almost certainly did not know what the majority of the icons shown were for and thus based their decisions on what they saw and on what tasks they thought computers might be able to perform. For example, Participant LL stated: ‘I can see a bucket (cultural trash can icon), but I don’t know why I need a bucket for using a computer’. Participant JC stated: ‘It looks like a Chinese old house (cultural home icon). It’s common to see a Chinese old house on the street, but why do I need one when I’m using a computer?’

Generally, icons that portray objects together with words were considered more appropriate than icons that portray only objects. The standard *stop* icon, which depicted a hand and the word STOP (see Table 1), was chosen as being the most appropriate icon for the referent *stop* by all participants in Group A and the majority of participants in Group B. Participant YW stated: ‘I read the English word STOP and a hand indicated the meaning. The meaning of this icon is very clear’.

Participants thought icons showing Chinese text were more appropriate than icons showing no text or English text. Many participants commented that they did not know what the standard shadow icon ‘s’ stood for, however cultural shadow icons with Chinese characters seemed to be easily understood. One participant familiar with computers (participant LL) stated that she could not say for certain that the standard shadow icon shown was the standard shadow icon, because it had not been presented next to other alphabetic icons. She said: ‘Whereas they’re next to each other (the standard text formatting icons, e.g. italic, bold and underline) you can tell which the icon is for shadow as it has a grey background around the letter S’.

Conclusion

The purpose of this experiment was to investigate whether Taiwanese computer users better understand cultural or standard imagery. The results reveal that, overall, Taiwanese computer users recognised cultural icons more accurately than standard ones. However, when participants familiar with computers (Group A) and participants not familiar with computers (Group B) are looked at separately, a difference can be seen to exist. Group A recognised standard and cultural icons with equal ease, whereas Group B recognised cultural icons more easily than standard ones. This result is in line with previous research (Wang, 2007), in which the testing icons focus on the alphabetic icons.

The results of this research can contribute to the knowledge of designers when they develop icons for Taiwanese computer users. This research is specific to the needs of Taiwanese computer users and also offers an understanding of the problems Taiwanese computer users have with respect to computer icons.

References

Fang, X. and Rau, P.L.P. 2003, Culture differences in design of portal sites. *Ergonomics*, **46** (1–3), 242–254.

Nakakoji, K. 1994, Crossing the cultural boundary. *BYTE*, **19** (6), 107–109.

Onibere, E.A., Morgan, S., Busang, E.M. and Mpoeleng, D. 2001, Human-Computer interface design issues for a multi-cultural and multi-lingual English speaking country-Botswana. *Interacting with Computers*, **13** (4), 497–512.

Po. T.M. and Chuan, T.K. 1999, Challenges in designing user interfaces for hand-held Human-Computer Interaction. In H.J. Bulinger (ed.) *Human-computer Interaction*, (Lawrence Erlbaum Associates, London), 808–812.

Sears, A., Jacko, J.A. and Dubach, E.M. 2000, International Aspects of World Wide Web Usability and the Role of High-End Graphical Enhancements. *International Journal of Human-Computer Interaction*, **12** (2), 241–261.

Wang, H.F. 2007, The appropriateness of Taiwanese cultural imagery in icon design. *Design for Safety, Health & Environment, the 14th Annual Meeting & Conference*, (Ergonomics Society of Taiwan, Taichung), CD RUM E-02.

Zwaga, H.J.G. and Easterby, R. 1984, Developing effective symbols for public information. In H.J.G. Zwaga and R. Easterby (eds.) *Information Design: the design and evaluation of signs and printed material*, (John Wiley & Sons, New York), 277–297.

ENGLISH, SLOVENIAN, SWEDISH: CAPTURING USER REQUIREMENTS ON LARGE PROJECTS

David Golightly, Alex Stedmon, Michael Pettitt, Sue Cobb, Sarah Sharples, Harshada Patel, Jacqueline Hollowood, Kate Moncrieff, Mirabelle D'Cruz & John Wilson

Human Factors Research Group, School of Mechanical, Materials and Manufacturing Engineering, University of Nottingham, Nottingham, UK

This paper explores practical issues when capturing user requirements on large projects, where user partners are often distributed over geographical and cultural boundaries (typical of large EU projects). A number of important aspects are discussed including: human factors awareness, recruiting participants, appropriate methods, stimuli (scenarios and prototypes), logistics, and language and culture. The aim is to provide an insight into the application of user-centred design in this research context and some examples of best practice for future projects.

Introduction

As human factors experts we appreciate the importance of capturing user requirements at the earliest opportunity and implementing them throughout the lifecycle of a project or design process. We also know that user involvement and participation are crucial to gaining the 'buy in', input and ongoing support of end users as well as assisting in communicating the findings back at a later stage (Wilson, 1995). Whilst we have a suite of reliable and valuable methods the pragmatics of applying them can be more than trivial, particularly on large distributed projects with partners working across geographical and cultural boundaries. Such projects are typically funded by the European Union (EU).

The Human Factors Research Group (HFRG) and Virtual Reality Applications Research Team (VIRART), at the University of Nottingham, have a wealth of past and current experience of working on EU projects. Examples include Sounds and Tangible Interfaces for Novel product design (SATIN; FP6-IST-5-034525), Computerised Automotive Technology Reconfiguration System for Mass Customisation (CATER; FP6-IST-5-035030) and Innovative Collaborative Work Environments for Design and Engineering (CoSpaces, IST-5-034245). Typically the HFRG and VIRART capture user requirements for innovative technologies and investigate fundamental human performance issues such as using virtual environments for design and training applications.

A key feature of large projects is that they usually involve a number of user partner organisations. User partners have a particular role to play in giving us access to end users, typical tasks, contexts of use and user evaluation input. User partner

involvement is fundamental in focusing activities on usable, and potentially marketable, systems, providing the EU with a return on its often substantial investment.

The reality of working with such partners is that they have their own needs and constraints. In addition, there can be practical barriers presented by time, distance and language. Human factors experts apply methods that meet the needs of the project while making best use of the resources available. While general reviews exist of applying ergonomics in industry (e.g. Haines & McAtamney, 1995) we focus here on user requirements capture for Information and Communications Technology (ICT). Our aim is to provide an insight into capturing user requirements in this context and provide examples of best practice.

Issues with user requirements gathering

Establishing the role of human factors

The first challenge when capturing user requirements is establishing the role of human factors, although this is not always a simple task. The goal, therefore, is to ensure that user partners understand why we are there and what value we can add. By having this understanding early on, partners are more receptive to our work and have a greater appreciation of the challenges we might face in getting the right participants, choosing the correct method, etc. Without this understanding, partners can be unprepared for the role we play and fall into familiar traps (e.g. selecting managers of end users rather than end users themselves to participate in interviews or discussions).

It is often useful to present our approach in terms that are relevant to the organization and draw comparisons between other user requirements gathering which the organisation may be more familiar with (e.g. using focus groups when dealing with organisations delivering consumer products). It is also critical to use terminology that is appropriate for the organisation rather than terms that may be obscure or even threatening to participants, especially when crossing a language divide.

It is tremendously useful to have a single point of contact or 'champion' within the user partner organisation who might have some understanding of human factors (this person often goes on to take the 'chaperone' role discussed below). This person can liaise with the user partner organisation while appreciating what is required for successful user requirements gathering. Sometimes partners may see user requirements as something that just happens at the start of a project rather than an ongoing process. Therefore, it is wise to get access to participants again in the future, even if this is just via email or phone in order to send results and conduct any follow up activities.

Getting the right participants

With general support onboard, the next activity is to identify the correct user group. Obtaining access to users is often the largest hurdle to overcome especially if they will be away from their work. Also, in organisations unfamiliar with human factors there can be a tendency to choose representatives of end users, rather than the end

users themselves, to save on participant time and perhaps in the mistaken belief that it will save us time conducting fewer interviews. The risk is that we only receive an abstracted or sanitised view of tasks and requirements and not one that is drawn from different perspectives. It is a common issue and it is worth being prepared for this to happen.

When dealing with smaller organisations or departments it is useful to see an organisation chart in advance. You can use this to choose the roles or personnel that you wish to talk to. If the specific people are not available, the organisation will have a much clearer idea of the role, and level, of the person that you want to talk to. This can be more difficult across larger organisations where the 'product planner' role in one organisation may be quite different from the same job title in another organisation. Ultimately, in work with large complex organisations one of the outcomes of your study may be to find out who the real end users are going to be (requiring a return visit later).

Choosing the right methods

There is a significant field of research in requirements engineering (Pew & Mavor, 2007). However, much of this assumes ideal situations (e.g. one user group or one set of developers, requirements elicitation at the beginning of the project, opportunity to define project scope, etc). In reality, and especially in the limited timescales of large projects, system development and requirements work tend to start simultaneously. The challenge is therefore to use the methods that are going to deliver the most value in this context. For example, development may be starting with a relatively impoverished view of user tasks or the real work of the target user group and in such cases a functional task analysis of the target domain would be urgently required.

Another aspect to bear in mind is to leave plenty of opportunity in the method for participants to become comfortable with the situation. A brief 15–20 minute lightweight task analysis (about their role, inputs, outputs, challenges, etc) will allow the participant to develop confidence by discussing issues with which they are familiar and will also provide valuable context for later questions. As it is difficult to predict what will happen when you get on site, it is often best to develop semi-structured rather than highly structured interviews. It may be more efficient to collect some baseline data that can be refined later, either with repeat visits or follow up activities over the phone, email, etc.

Finally, it might only be clear after initial visits that it is possible to video users or gain other permissions to record data. Again, a single point of contact within the organisation can expedite this process. Permission to record should not be assumed, particularly if there are sensitive issues (e.g. security, commercial, union, etc). If so, the method should be simplified as much as possible, so it is practical with paper-and-pencil note taking.

Scenarios and prototypes

Scenarios and prototypes (paper, mocked-up screens, videos or interactive demos) are a powerful way to introduce a proposed technology and elicit feedback and

comments for requirements (Carroll, 2000). Experience has shown that prototypes do not have to be of the highest standard as 'rough-and-ready' versions (e.g. powerpoint or paper-based) indicate the product is work in progress and may elicit more comments from participants than a system that looks like it is already complete. Also, in products where there are a number of user groups or partner organisations, it does not appear to matter that scenarios are less relevant for some groups than others. If anything, participants may focus on small, irrelevant details if they are very knowledgeable in the scenario domain than if the scenario is removed from their day-to-day work. Participants can also be swayed by very innovative, visually rich prototypes that present highly stylised visions for future products. There is a danger that this can create high expectations, risking disillusionment if the end product fails to achieve a set standard.

Logistics

With geographically distributed projects it can be difficult to travel to sites or meetings making it necessary to conduct user requirements in one visit. An advantage is that it is good to get all information at the same time and there is an opportunity to become well practised in the method. A downside is that there may be little chance to substantially re-working the procedure if something is missing or incorrect. This requires interviewers/observers to make sure that as much information as possible is prepared in advance, though this also limits how much you will be able to deviate from the proposed material on the day. Being away from your home site means this is a situation where piloting the method and fine tuning before you set off is critical.

Language and culture

Most of us at the HFRG and VIRART have English as our first language, often only with a smattering of secondary school French and German (better suited to ordering drinks than requirements elicitation) so we are both impressed and grateful to non-English speaking hosts and participants for their excellent standard of English. That said, language and culture can still present some barriers.

There will always be concepts that are difficult to express and a bi-lingual chaperone (ideally the existing point of contact) is invaluable. They can introduce the project in the participants' native tongue and explain difficult areas. However, this means that there are occasions when the chaperone and participant might break-off to discuss a point at length and then return with a short summary answer with the risk that critical information is lost in the exchange. When studying tacit knowledge researchers must also be aware that it is often difficult to verbalise concepts and trying to probe and discuss this knowledge is further impeded by a language barrier. This also applies to translated questionnaires that may lose the right emphasis or be misinterpreted altogether.

Finally, there is a risk that the chaperone takes over the direction of the session. The human factors expert must keep polite but firm control over the session. It is also important to remain respectful of what the end user thinks might be useful.

Table 1. Summary of issue areas and best practice.

Role of human factors	• Use relevant language • Give examples of HF approach • Set up point of contact
Participants	• Establish need for 'real' users • Refer to an organisational chart
Methods	• Consider faster methods • Lightweight task analysis • Semi-structured
Stimuli	• Does not need to be high-fidelity • Manage expectations
Logistics	• Compile and pilot material in advance • Prepare way for follow-up contact
Language	• Use a bi-lingual chaperone • Be cautious when eliciting tacit knowledge

For instance, it might be rude to decline the offer of guided tours of facilities even though time is ticking away for interviewing end users. However, such tours often provide background knowledge that can then be used to help develop a common understanding.

Despite these issues, there is a real value to face-to-face interactions and first and foremost in the initial activities, it is important to build a rapport and take a little time to do some homework so that you have some non-work topics to talk about, ideally with a local slant: skiing in Slovenia, the cuisine of Turin, or the price of taxis in Gothenburg.

Conclusion

While the current issues are based on large projects, they are not unique to this context. With the globalisation of products and product development, many software projects are becoming larger and more distributed, as well as relying more on the need to address user requirements. Also, issues of convincing user partners about the value of human factors is something we have all experienced much closer to home.

We have presented some of the issues that, in our experience, face user requirements capture in large projects. These are summarized in Table 1 above. The list is by no means exhaustive and future work should address practical issues such as how requirements data are analysed, implemented and subsequently evaluated.

Whilst there are many challenges with capturing user requirements, they can be anticipated and hopefully overcome. Requirements gathering may sometimes be difficult but there can be substantial reward for the effort: whether that be through better requirements and better projects, the promotion of human factors to a wider audience, or a greater understanding of the tasks, contexts and cultures that exist on large projects.

Acknowledgements

The authors would like to thank the user partners, participants, and all project partners of past and present EU projects that the HFRG and VIRART have been involved with for their time and input.

References

Carroll, J.M. (2000) *Making Use: Scenario-Based Design of Human-Computer Interactions*. Cambridge, MA: MIT Press.

Haines, H. & McAtamney, L. (1995) Undertaking an ergonomics study in industry. In, J.R. Wilson & E.N. Corlett, *Evaluation of Human Work*, 2nd Ed. London, Taylor & Francis.

Pew, R.W. & Mavor, A.S. (2007) *Human-System Integration in the System Development Process: A New Look*. Washington, D.C.: National Academies Press.

Wilson, J.R. (1995) Solution ownership in participative work redesign: the case of a crane control room. *International Journal of Industrial Ergonomics*, **15,** 329–344.

ERGONOMICS 4 SCHOOLS

DIAGNOSING FAILURES IN ORGANISATIONAL BEHAVIOUR USING INFLUENCE DIAGRAMS

Mike Tainsh

Krome Ltd, Fetcham, Surrey, KT22 9QR

This paper describes Influence Diagrams (IDs) which are often used in ergonomics, and engineering to help the diagnosis of failures. Their application to organisations appears less frequent, even though "organisational factors" may be included within the investigation of accidents. A case study is presented based on the work of a school's Board of Governors. The failure in this case is the occurrence of back-pain in school-age pupils using the information available from the charity BackCare. It is concluded that IDs can elucidate relationships between organisational variables which influence the occurrence of failures, and hence lead towards opportunities for mitigation of the risk of failure.

Introduction

The subject of organisational behaviour addresses those topics within ergonomics where one is addressing issues of occupations, jobs and tasks, including the relationships of individuals with their equipment and physical environment i.e. the socio-technical system (Ref 1).

The study of organisational behaviour, failure and its mitigation appears to be a somewhat neglected area of scientific study within ergonomics. Yet, ergonomists are frequently invited into a client's organisation (Ref 2), as consultants, because it is believed that there has been a failure, which needs to be diagnosed and fixed.

This paper addresses organisational behaviour within a school. It originates from work within the "ergonomics4schools" Special Interest Group (SIG). In this case, the failures are reports of back pain in school students. These are important failures as the school student runs the risk of discomfort and harm, while the school's Board of Governors runs the risk of failing in its "Duty of Care".

Root Cause Analysis and Influence Diagrams

Root Cause Analysis (RCA) is a technique which starts from the identification of sequences of dependencies or influences that precede an event or user action – so that one can establish the root cause rather than the immediate cause(s). RCA is used in the context of Total Quality Management (TQM) to help identify the causes

Table 1. Outcomes, and Their Characteristics Including Norms.

Outcome	The characteristics of the outcome
O is an assessment of an outcome	V is the report, measure or observation U is a norm, expectation or requirement – the norms must be considered in terms of organisational characteristics

of the failure so that they can be mitigated in the future. RCA is widely used in the engineering industries and health services (Ref 3, 4 and 5).

Within RCA, there is a variety of diagrams which can be used to indicate causal or dependent relationships: one of these is Influence Diagrams (IDs) (Ref 3). IDs are used here are they appear the most able to include organisational characteristics (Ref 5). They offer the possibility of helping to understand the dependencies between the variables which may lead to the reporting of symptoms of back pain, and support the process of diagnosis.

Using IDs when describing organisational behaviour

General Statement of Relationships

The starting point for an ID is an outcome O which is the characteristic of an object or event. O is the result of two variables V and U, where V may be different from, or equal to U. O may be judged to be an error or other sort of failure. Each outcome is considered in terms of observed and normative characteristics, and the sequence of events that generate the outcome are traced back to root causes with multiple levels of a hierarchy of influences.

V is the variable of immediate interest (report or other) and U is the normative (or expected or required) value of V. When the value between the two variables is different from that which is the norm, it may be brought to the attention of members of the organisation. Once the outcome is assessed as sufficient priority (failure or success), then it may be necessary for the members to understand it and diagnose its cause. In turn V and U depend on sequences of influences that may have the form of a hierarchy. These may be events, conditions associated with the events or boundary constraints, or others (Ref 3).

Case Study: Back-pain in school students

The problem

The example here is based on data taken from BackCare documents in a presentation to Members of Parliament (Ref 6). The problem is school students reporting back pain to medical practitioners, physiotherapists, teachers and others who have a duty of care.

This study is approaching the problem from the point of view of the school's Board of Governors. The Governors (and other specialist teams), in conjunction with local authorities, medical experts, architects, engineers and suppliers have to decide on issues of health and safety. This will include buildings, curriculum activities, and the purchase and use of furniture.

Major influences on the occurrence of back pain include:

- exercise regimes
- the design and use of furniture, especially seating
- diet including fizzy drinks
- age related conditions.

Screenshot 1 was achieved using TaskArchitect which is a commercially available tool (Ref 7). Additional information is held within the tool to support discussion.

Observed and Normative characteristics

From a medical point of view, the first topic to note is the presentation of symptoms. However this has an immediate normative aspect as one asks:

- Is the presentation rate higher or lower than one might expect?
- Is the rate acceptable?

Mitigation, and by whom

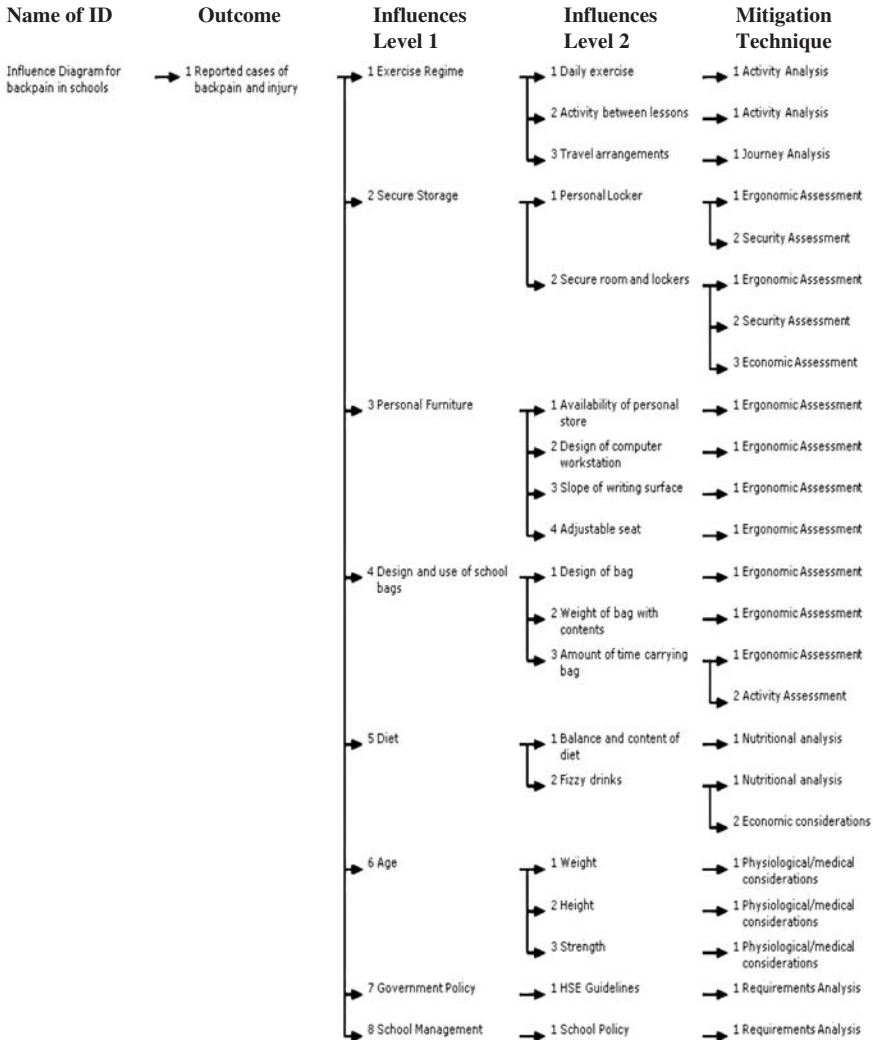
There is a need to agree that there is a problem requiring diagnosis and resolution. Possible scientific/technical mitigation techniques are shown in Screenshot 1. The Governors might seek advice on the characteristics of influences and possible causes. They will compare observations with known norms, expected values or requirements. The ergonomics practitioner may be able to raise issues of design or organisation of activities in conjunction with the local authority or purchasing organisation, or suppliers. Parents and students may be involved.

The Governors may be able to influence some of the variables immediately but equally issues associated with the provision of storage, the design and use of school bags may be more difficult. Additional storage may require funds from the local authority while parents may be unwilling to purchase new bags and school students may be unwilling to use them. Health and safety legislation and school policy may provide boundary constraints in the short term or issues to be addressed in the medium or longer term.

Conclusions

This case study provides an indication of how IDs might be used in studies of organisational ergonomics to understand influences, and diagnose sources or causes of failures (or successes). The IDs can be used to infer the relationships of influences to the outcomes that are being investigated.

Screenshot 1. This shows influences on the reporting of back pain in school students. The data comes from BackCare, 2006.
The extreme right-hand column is not usual in the presentation of IDs.



The use of IDs does appear to help elucidate some of the organisational relationships by showing:

- Sets of influences that contribute to organisational outcomes. In turn this can help identify those variables that are open to control/manipulation and those that are not;

- Sequences of dependencies, helping the understanding of sources of influences on outcomes, or root causes.

The study shows that techniques associated with quality management and RCA which are more usually associated with engineering systems can be applied to problems associated with organisational ergonomics when there is the assessment of failure.

This technique has been used in other applications, and experience so far suggests that this form of diagram and analysis is useful when considering organisational failures.

References

1. Eason, K. (1988) *Information Technology and Organisational Change*. Taylor and Francis.
2. Tainsh, M. A. (2004) *Human Factors Integration. Chap 2 in Human Factors for Engineers by Sandom C and Harvey R.* IEE, London.
3. US Dept of Energy (1992) *DOE Guideline Root Cause Analysis Guidance Document.* DOE-NE-STD-1004-92.
4. Doggett, A. M. (2004) A Statistical Comparison of Three Root Cause Analysis Tools. *Journal of Industrial Technology*. Volume 20(2) pp. 1–9.
5. Embrey, D. (2004) *Qualitative and Quantitative Evaluation of Human Error in Risk Assessment. Chap 8 in Human Factors for Engineers by Sandom C and Harvey R.* IEE, London.
6. BackCare (2006) *Pupil Back Pain.* Submission to UK Houses of Parliament.
7. Stuart, J. and Penn, R. (2004) *TaskArchitect – taking the work out of task analysis.* ACM International Conference Proceeding Series, Vol. 86.

A POST OCCUPANCY EVALUATION OF FIVE PRIMARY SCHOOLS BUILT TO COVENTRY'S MODEL BRIEF

Michelle Newman[1], Andree Woodcock & Philip Dunham[2]

[1]*School of Art and Design*
[2]*Faculty of Business, Environment and Society, Coventry University, Coventry, CV1 5FB*

Since 1998 several primary schools have been built in Coventry which follow the guidelines contained in a model brief developed by Coventry City Council. This paper describes a systematic approach to evaluating the effectiveness of these buildings through the development and implementation a Post Occupancy Evaluation (POE) toolkit designed for primary schools. The rationale for performing a POE in the educational setting in which all stakeholders can give their feedback will be presented prior to a description of the development of the toolkit. This comprised an in-depth questionnaire for all adult stakeholders and a scheme of work designed for children. An analysis of the preliminary results from the first stage of its implementation will be presented.

Introduction

This paper will explain the conceptual and practical development of a post-occupancy evaluation tool. This was developed in order to examine the suitability of recently built primary schools. This paper will first provide a rationale for conducting Post Occupancy Evaluation (POE), with particular reference to the educational setting. It will go on to critique existing methods of POE that have been used previously in schools before outlining the development of the toolkit from early pilot stages to implementation.

Coventry City Council has in recent years developed a model brief for future primary school builds. Six of these schools have been built since 1998, however there has been no systematic post occupancy evaluation (POE) performed on the schools. This is despite it being widely accepted that the environmental conditions of schools can have a significant impact on the health and safety of children and adults in school, affecting the ways in which teachers teach and children learn. It is argued that one way of improving conditions in schools is through effective evaluation and implementation of corrective action (Higgins, 2005; DCMS, 2000; Price Waterhouse Cooper, 2001; Bellomo, 2006).

A rationale for post occupancy evaluation

POE has its origins in the UK as early as the 1960s, with the publication of "The Architect and his Office" (RIBA, 1962), which called for architects to look at buildings from the users' perspective, as well as technical and design aspects. This came at the same time as an increasing academic interest in environmental psychology which aimed to create a science linking interaction with the built environment and the behaviour of occupants. Arguably, environmental psychology became disengaged from the design process throughout the 1970s and 80s, when feedback from post occupancy surveys into behavioural interactions with buildings were not fed into the design process (Cooper, 2001). However the Egan Report (1998), provided a clear rationale for using POE. It states that the construction industry needed to address the needs of the client and consumer, including ongoing research into the requirements of the end-user.

Until recently POE has not been seen as a necessary part of the design and construction of new buildings (Douglas, 1994). However, there is now an increasing body of literature on the necessity of post occupancy evaluation as part of an ongoing process of design. It is defined as "the process of systematically evaluating the degree to which occupied buildings meet user needs and organizational goals" (Lackney, 2001:2). POE provides "an appraisal of the degree to which a designed setting satisfies and supports explicit and implicit human needs and values for whom a building is designed" (Friedman, 1978:20). POE has also been described as a way to answer two questions: how is a building working, and is this as intended (Leaman, 2003). The assessment of whether a building is functioning according to the intentions of the designers is a key concept in POE (Horgen, 1996). It has been suggested that POE is a means of improving the quality and sustainability of buildings (Bordass and Leaman, 2005a) and should become a routine part of the design process(Bordass and Leaman, 2005a, Way and Bordass, 2005, Bordass and Leaman, 2005b).

POE in the educational setting

The need for effective POE in the educational setting has been recognized (Lackney, 2001, Sanoff, 2001) and should aim to assess the extent to which the building supports the educational goals of the school by measuring its physical appropriateness to its function (Hawkins, 1998). POE should "describe, interpret and explain the performance of a school building" (Sanoff, 2001:7). However building assessment and evaluations have rarely examined the relationship between the physicality of the school and the educational goals of the establishment (Lackney, 2001). The toolkit developed for this research is explicit in its aim to interrogate this relationship, establishing which elements of the school building are the most important to good design.

The role of the end-user in POE

Building assessment in the past has relied on the judgements of "experts" as to the success or otherwise of a school building. However this research begins with the

premise that those who are the most expert are the end-users. As Sanoff says "the experts are the people who know most about using it – the user." (Sanoff, 2001:8).

The questionnaire that was designed as the first stage in the POE was tailored to ask questions specific to each adult user group, which were senior management, teachers and teaching assistants, administration staff, lunchtime supervisors and kitchen staff, cleaning staff and parents. Rather than one generic questionnaire for every user group, or only asking more senior members of staff about the building, the questionnaire assesses each adult user group's needs and the extent to which these needs are met.

To assess the views of children a workbook was developed which asked questions of children about their daily routines in the school environment and assessed the appropriateness of the environment for each daily activity. It used simple language encouraging children to reflect upon their experiences and their school environment. It also used techniques such as asking children to draw maps and drawings of their school, which may be analysed in terms of content to assess what aspects are of importance. However, to date the results from this part of the toolkit have not been analysed, so the remainder of the paper will discuss the implementation and results from the questionnaire only.

The development and implementation of the questionnaire

A Likert Scale was adopted to assess the stakeholder's judgements and opinions as to the appropriateness of the school design Because of the tendency for respondents to agree with statements in a Likert scale, approximately one third of the questions were phrased in what may be termed a "negative" way. Each part of the model brief used by Coventry was worked through in order to develop a set of questions that addressed its primary intentions.

A paper based questionnaire was considered more appropriate than an electronic version because many of the stakeholders had limited access to the internet and would therefore be unlikely to complete the survey. Also teachers admitted that they would be more likely to fill in a paper copy as this could be achieved in informal settings, such as during a coffee break or at home.

Six questionnaires were devised to address specific questions appropriate to the various stakeholder groups. As well as evaluating the building the questionnaires also provided an opportunity to reflect on the process of designing the school, for example whether there was sufficient user involvement in the process or whether users felt the design was improved by their involvement. Each of these questionnaires asked questions specific to the respondents' role in the school. For example he head teachers' questionnaire relates to the management of the school, for the teachers' the extent to which the school supports the curriculum and their teaching philosophy. Other members of staff are asked about the suitability of the school design for their particular job. All questionnaires ask about physical comfort.

Following the pre-pilot stage and subsequent amendments, the questionnaire was sent to all staff and parents at five primary schools. Returns from teachers, teaching assistants and administration staff were almost 100%. Cleaning staff, kitchen staff

and lunchtime supervisors had relatively low returns. The rate of returns from parents was dependant on the school. Two of the schools had very high returns, three were low. The low returns were from schools in areas of social deprivation. The head teachers indicated that this is a problem for any letters that are sent home. Although reasons for poor parental returns are unclear two possible reasons suggested by the head teachers were poor literacy rates and not feeling part of the school community.

Results

The results are based on returns from 269 questionnaires. An in-depth analysis of each individual school was necessary in order to assess the efficacy of each specific design. However a broad analysis of all the schools was also needed in order to give an overall view of whether the model brief as used in Coventry is appropriate as a set of guidelines. It is the general set of results, providing an overview of all schools surveyed that will be presented.

Positive common features

The five schools that were surveyed were built between 1998 and 2006. Generally the more recently built the greater the satisfaction with the design, with one exception. Staff from the earliest school expressed some dissatisfaction with the layout. Another more recent school, whose architect had received awards for the school design, left the teaching staff extremely unsatisfied.

A key part of the model brief is the inclusion of sliding doors between two classrooms, so that the room may be opened up to make the space more flexible. This was generally seen as a good thing, the exception being the earliest school where they were considered a distraction.

Another key idea is to have a shared area outside classrooms which may be used to add another space to allow for flexibility of the space and teaching. This was universally seen as a positive attribute.

ICT provision was generally considered excellent.

School was considered a safe, healthy environment by teachers and parents.

Classrooms generally had enough room for all the day to day activities. This was helped by having the additional space outside, allowing for flexibility.

The school buildings generally supported the delivery of the curriculum, teaching and learning. The exception was the recent school with high levels of dissatisfaction in most aspects of the building were indicated.

All of the schools were considered aesthetically pleasing.

Negative common features

The control of temperature was a problem in most of the schools, particularly in one school where the temperature was reported as being too high even on days where the external temperature was low.

Issues with lighting was another common problem, with either natural light causing glare and reflections off white surfaces such as whiteboards, or too little natural light meaning that artificial light had to be used most of the time.

Many of the schools complained that they did not have enough large, outside play equipment.

All schools bemoaned the lack of storage space. At one school there is only one central space for storing all curriculum equipment. Staff felt this was woefully inadequate.

Generally it was felt that more small group rooms would be beneficial, where children could go if they needed extra help or for group work. Only one school had a nurture room.

Although most stakeholders, the exception being head and deputy head teachers, had little input into the school design, they generally did not see this as a major issue. Again the exception being the school with high dissatisfaction rates, where most respondents felt that the school would have benefited from more input from user groups.

Conclusions

Post Occupancy Evaluation is a way of providing necessary feedback into the design process. It is effective when it ascertains the extent to which a building fulfils the functions it was intended for, in this case supporting the delivery of the curriculum, teaching and learning. It must also assess how the building facilitates the smooth running of auxiliary services such as those provided by kitchen and cleaning staff. Other tools designed for POE are flawed in that they do not address the needs of each user group. Neither do they ask about specific areas in the school building, rather they take a generalised approach, often overly focusing on aesthetic considerations rather than functionality. This toolkit is specifically designed for the purpose of assessing these considerations.

Results indicate that although there is general satisfaction with schools built to the specifications of Coventry's model brief, there are common problems identified by the POE. These results will be used to provide evidence for a set of guidelines that will inform designers involved in the future build of primary schools. It is anticipated that the use of the toolkit will be extended to the national level and will continue to provide feedback for designers and stakeholders involved in the design process.

References

Bordass, B. and Leaman, A. 2005a, Making feedback and post-occupancy evaluation routine 1: A portfolio of feedback techniques, *Building Research & Information*, **33**, 347–352.

Bordass, B. and Leaman, A. 2005b, Making feedback and post-occupancy evaluation routine 3: Case studies of the use of techniques in the feedback portfolio, *Building Research & Information*, **33**, 361–375.

Cooper, I. 2001, Post-occupancy evaluation – where are you? *Building Research and Information*, **29**, 158–163.

Douglas, J. 1994, Developments in appraising the total performance of buildings. *Structural Survey*, **12**, 10–15.

Egan, J. 1998, Rethinking Construction: The Report of the Construction Task Force. (The Stationery Office, London)

Friedman, A., Zimring, C. and Zube, E. 1978, *Environmental Design Evaluation*. (New York, Plenum)

Hawkins, H. L. and Lilley, H.E. 1998, *Guide for school facility appraisal*, (Council for Educational Facility Planners, Scottsdale, AZ, USA)

Horgen, T. and Sheridan, S. 1996, Post-occupancy evaluation of facilities: a participatory approach to programming and design, *Facilities*, **14**, 16–25.

Lackney, J. 2001, The state of post-occupancy evaluation in the practice of educational design. *32nd Annual Meeting of the Environmental Research Association.* (Edinburgh, Scotland).

Leaman, A. 2003, Post-occupancy evaluation. In S. Roaf (Ed.) *Benchmarking Sustainable Building Performance.* (RIBA, London).

RIBA 1962, The Architect and his Office: *A Survey of Organization, Staffing, Quality of Service and Productivity*, (RIBA, London)

Sanoff, H., Pasalar, C. and Hashas, M. 2001, *School Building Assessment Methods*. (National Clearing House for Educational Facilities).

Way, M. and Bordass, B. 2005, Making feedback and post-occupancy evaluation routine 2: Soft landings: involving design and building teams in improving performance. *Building Research & Information*, **33**, 353–360.

A CHILD CENTRED APPROACH TO THE DESIGN OF E-LEARNING MATERIAL FOR THE ICT CURRICULUM

Ruba Abu Hassana & Andree Woodcock

The Design and Ergonomics Applied Research Group, Coventry School of Art and Design, Coventry University, Coventry, CV1 5FB

Due to the increasing demand for an IT (Information Technology) literate workforce and the potential for IT to enrich learning, ICT (Information, Communication and Technology) is used in schools to support learning and teaching across various subject areas and as a subject in its own right. Although IT has been used to support the delivery of other subjects in primary schools it has not been used effectively to support the delivery of ICT itself. This research considers how IT may be used more effectively in this context, in primary schools in the UK and Saudi Arabia, and whether a user centred approach to its design may result in more effective teaching and learning material.

Introduction

ICT is used in teaching and learning in primary schools to deliver material, support discussion groups, or replace paper based material. The design of educational resources to support the teaching of ICT has not received attention, even though it underpins teaching and learning across the curriculum. The main aim of this research was to explore the effectiveness of e-learning (i.e. the application of IT (computer hardware and software) as a tool in the learning process; particularly, if used by the learner rather than the teacher) in supporting the teaching of the primary school ICT curriculum in UK and Saudi Arabia.

The research involved four main stages: (1) problem recognition; (2) solution development and refinement; (3) evaluation; (4) knowledge transfer. This paper presents an overview of the research stages. However, its main focus is on the relative lack of enthusiasm shown towards the adoption of user centred design by developers of educational resources, even though this may lead to improved learning outcomes.

Research stages and outcomes

To identify issues with the current teaching of ICT a literature review, classroom observations, market evaluation, interviews with teachers and designers were conducted.

These indicated that;

1. The use of ICT is increasing. However, most research concentrates on the use of ICT in teaching and learning and not on the teaching of ICT itself (Hammond, 2004).
2. ICT as a subject is not taught effectively in UK primary schools (Ofsted, 1999). This was shown in the literature and supported by classroom observations and interviews.
3. UK schools are more developed in their use of ICT than Saudi private and public schools. Therefore, there was potential for knowledge transfer and also to influence the way in which educational material might be developed in Saudi.
4. CAL (Computer Assisted Learning) products associated with ICT curriculum could be divided into a) computer applications which enable pupils to practice their ICT skills b) products that support teachers in delivering the National Curriculum such as lesson plans, worksheets, and assessment sheets; c) video recorded tutorials that illustrate how to use a specific computer application. However, in both countries little computer based material was available to support the teaching of ICT per se.
5. E-learning was often applied inappropriately in Saudi Arabia. For example, no company producing such material evaluated its effectiveness on learning attainment; no guidelines were available for teachers on how to incorporate e-learning or move to a full e-learning position, there was little or no consultation with teachers on the design of e-learning material.
6. An investigation of current design practices in Saudi educational publishing companies showed that the ADDIE (Analysis, Design, Development, Implementation, and Evaluation) instructional design model is used in which each step feeds into the next. As part of this the designers indicated that they converted the text and graphics in the textbooks straight into electronic versions (with no adaptation), and then added extra graphics and sounds when they feel this was needed – but not from a knowledge of pedagogy or current teaching practice. No external supervision, evaluation or quality assurance took place.

In the second stage these findings were used to propose that current design and teaching processes could be made more effective by developing a set of design principles for the development of e-learning material that was more sensitive to the usage context (schools) and end user needs (teachers and children). A participatory design approach could give designers greater insight into user perspectives and preferences, allowing designers to rethink, design goals and solutions (see for example Laurel, 1990; Norman, 1988; Raskin, 2000). Puphaiboon (2005) applied such an approach to the iterative design of educational material to teach mathematics. However, his collaboration with just teachers produced designs which needed alteration when evaluated by children (for example in terms of font size and timing). Thus, it was proposed that children should be involved directly and earlier in the design cycle.

Such a method was developed and tested in the production of material to support the current ICT curriculum for primary schools. This involved five steps; firstly, a detailed design specification of the e-learning material was determined through interviews with children and teachers which included consideration of children's

ICT interactions, teaching methods, communication concepts using ICT, and the learning context. Secondly, a prototype was developed from the requirement specification which in the third stage was evaluated by children in terms of its usability and the enjoyment, and by teachers in terms of the extent to which it facilitated achievement of the learning outcomes. A full working prototype was produced taking into account this information which was then assessed in terms of it usability prior to evaluation in schools.

Results

Evaluation of the e-learning material produced using the participatory approach in three schools, with thirty children revealed that all teachers agreed that the material had educational value, was suitably designed for the intended age range, complemented the school systems, was easy to use without having a high knowledge of IT and could be used effectively to support the teaching of the ICT curriculum in year 5 and 6. They would also have liked to adapt it to other parts of the curriculum.

Evaluation with pupils did not show a quantitative improvement on learning outcomes (this was a very limited experience, and only occurred in one lesson). However 90% of the pupils liked the material (story, speech, music, navigation) and the observational study indicated that pupils were engaged with the material, understood it and competed with their peers to achieve higher scores. Their comments whilst using the material showed their enjoyment, such as “ohh, cool”, “I like it”, “that’s interesting” and mostly their questions reflect the same impression such as “can I find this game on the internet if I want to play it again?”, “How can I get the CD of the game”, “are we the first one to play it?”.

Knowledge transfer

The last stage of the research sought to address how these findings can be embedded in Saudi Arabia, to lead to the design of more effective material (and teaching methods). Six designers from the major educational software producing companies in Saudi Arabia were interviewed to determine whether they thought that a more user centred approach would be considered useful and usable by designers. The semi structured interviews started by presenting the proposed method and the material developed from its use.

Though this and previous research has shown the benefits of a more user centred or participatory approach to the development of teaching material, and the commercial success of companies (such as Ragdoll Ltd in UK), which have pioneered the use of children’s responses in their product development process, have increased interest in this area, Branton (2003) indicated that the majority of companies producing children’s technology products have failed to see the huge value that child research can add to their products. Their development decision process is mainly based on cost of goods, product margins, manufacturing issues and release dates. In addition, Kelly, et al. (2006) indicated that user centered design methods can be

seen as intensive and not suitable to a commercial environment. As such products might be developed with a small user group; they may not generalize to the wider population (and as such may have low impact).

Also, evaluation of the effectiveness of resultant products in educational environments is problematic owing to the number of uncontrollable variables (such as motivation, learners' characteristics, teacher buy in, timing and resources). So, although both teachers and the children may like the material, it is often not possible to provide designers with quantitative evidence of learning gains.

The designers interviewed believed that the proposed method could provide a deeper understanding of the psychological, organizational, social and pedagogical factors, thereby assuring that the product would be suitable for its intended purpose in the usage environment. In addition, it could provide an explicit understanding of the end user needs; thereby enhancing the usability and efficiency of the material.

However, designers thought that it would not be possible to apply such a method in their organization due to the following problems:

1. Time and cost: Child centered design approach is time consumer. Particularly in communication with stakeholders.
2. Communication difficulties: The development of the prototype does not take long, however, the testing and refinement process consume 2/3 of the design process time as communication with schools, teachers, and children was complicated. 1/3 of the design process time was consumed in managing and arranging communication with schools and not in the connection it self.
3. Children as a design partners: it was so difficult to involve the children because each child has his/her own characteristics, requirement, and when the majority of their feedback considered and inputs to further design stages, they change their mind.
4. Lack of design team members: The process requires resources, both financial and human. User-centered design teams generally benefit from including persons from different disciplines, particularly psychologists, sociologists and anthropologists whose job it is to understand user's needs and communicate them to the technical developers in the team. And most of the educational software publishing companies' human resources is limited to technical developer and maximum two designers in each company. However, they rarely have an educationalist, psychologists, sociologists or anthropologists working with the design team.
5. Competence: Companies are not sure what skills and expertise are required in a UCD project in terms e.g. social competence, technical skills or expertise on the work activities.

The designers agreed that parts of the method could be modified to suit the business world. Thus, following are some recommendations to tailor the proposed method to meet the business world requirements.

1. The communication with stakeholders should be minimized. E.g. Instead of holding three meetings to understand the design specification and children requirements, initial discussion with teachers and children should be undertaken in one single meeting, also to identify mistakes or weaknesses in the design or

usability in early stages, communication with stakeholders should take place once just to test the prototype suitability to the user preferences and needs rather than working continually with children as a design partners in developing and refining the prototype.
2. To facilitate the communication difficulty, the number of participants can also be minimized. E.g. a smaller group of participants (one school, one teacher and maximum 10 children's) might provide sufficient preferences and requirements.
3. The designing companies should employ people from different disciplines to enable the use centered design to take place efficiently.
4. The teacher participation can be limited to two times only once at the beginning in determining the needs and in the final evaluation the effectiveness of the material, as this research found that their inputs in testing and refinement were mainly to defer decisions to the children.
5. Training should be offered to the projects manager, which should highlights the following topics: (1) Ways to create a cultural context for user participation; (2) Organization and individual attitudes about users and usability as well as the role of the designer; (3) Project organization; the role of management in a project; (4) Methods, techniques and tools that suit user participation.

There are, however, inherent weaknesses in limiting user involvement in the design process – it can become tokenistic and unrepresentative. In order to ensure that this is not the case it is also proposed that steps are undertaken (at least in Saudi)

1. Schools work as research partners with the developers of new educational material.
2. A clear set of guidelines is produced for the developers of educational material which includes sections related to the human factors associated with children as learners, teachers as educators and a clear specification of the usage environment (i.e. the situation in which the material will be used).
3. A children's champion is established in each of the companies who can lead on children's issues and evaluation of the effectiveness of the material in terms of educational outcomes.
4. The Saudi Ministry of Education takes steps to ensure that the highest quality educational material and training are available in all schools.

Conclusion

In conclusion, this study shows that it is the personal experience and preference of the authors, profit and marketing that lead to design decisions, rather than the effectiveness of the materials and the satisfaction of teachers and students. The research has been contributed to knowledge with the following:

1. Indicated that e-learning can support the teaching of the ICT curriculum in UK primary schools as proofed from the evaluation of e-learning with 30 students in UK.
2. Demonstrated the impracticality of user centred design approach outside academia as confirmed by the designers' interviews in Saudi Arabia.

3. Provided a set of recommendations to designers, for developing a practical user centered design approach for designing e-learning for children.

References

Branton, A. 2003, *Developing a child-centred design process*, available online at <http://education.guardian. co.uk/elearning/story/0,10577,976205,00.html> last accessed 10/08/2007.

Hammond, M. 2004, *The peculiarities of teaching information and communication technology as a subject: a study of trainee and new ICT teachers in secondary schools*, Technology, *pedagogy and education*, **13**(1), 29–42.

Kelly, S. Mazzone, E., Horton, M., Read, J. 2006, Bluebells: *A design method for child centred product development*, Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles NordiCHI, **189**, 14–18.

Laurel, B. ed. 1990, *The Art of Human-Computer Interface Design*, (Reading MA, Addison Wesley).

Norman, D. 1988, *The Design of Everyday Things*, (New York, Doubleday).

Office for Standards in Education 1999, Annual Report of Her Majesty's Chief Inspector of Schools Section, **20** London: HMSO.

Puphaiboon, K. 2003, *A Method for the Iterative Design and Evaluation for Visual Learning Materials in Mathematics*, (PhD) Coventry University.

Raskin, J. 2000, *The Humane Interface: New Directions for Designing Interactive Systems*, (Boston, Addison-Wesley).

ERGONOMICS AND SECURITY

POSTURE, POSITION & BIOMETRICS: GUIDELINES FOR SELF-SERVICE TECHNOLOGY

James Ward, Chris Riley & Graham Johnson

Advanced Technology and Research, NCR Financial Solutions Ltd, Dundee, DD2 4SW, UK

Biometric technology provides an opportunity to improve identification security across a range of different transactions. This user-centred investigation examined the effect of position on the usability and accessibility of biometric devices. Using an approach based upon ISO 9241-11 Standard for Usability, the performance of fingerprint and palm vein technology was assessed for a self service context. Postures were also recorded and scored using the RULA posture assessment tool. The devices was tested at three heights, 1000 mm, 1100 mm and 1200 mm and three angles 0°, 15° and 45°. Device position was found to significantly affect participants' satisfaction ratings and the postures they adopted. The palm vein device out-performed the fingerprint device. This investigation shows how the physical placement of biometric devices can affect the systems' performance, and has implications for its use in the self service environment.

Introduction

Biometric authentication systems are beginning to find application in customer-facing, commercial environments. Traditional authentication methods include token-based authentication, such as keys or cards and knowledge-based authentication such as passwords or personal identification numbers (PINs) (Jain, Hong & Pankanti, 2000). Knowledge and token based methods do not provide a high level of security as tokens can be borrowed, lost or stolen, while passwords and PINs can be shared with others, written down and forgotten. Biometric identification is an alternative approach to user authentication which is not thought to suffer from the same limitations as knowledge or token based authentication. Various physical and behavioural characteristics can be used in biometric systems such as fingerprints, irises, and signature or keystroke behaviour. Biometric characteristics cannot be borrowed, forgotten or easily stolen. Biometrics are often described as the future of user authentication and are seen as guaranteeing the presence of the legitimate user in a way that other approaches cannot.

Implementations of biometrics are more likely to be successful and accepted by users, if the technology is usable. Maguire (2001) highlights the benefits of having usable products as; increased productivity, reduction in errors, reduced learning and training times, improved user acceptance and a competitive advantage. All of these

attributes would be beneficial to an implementation of biometric authentication technology.

Although the use of biometrics could improve security there are numerous issues that need to be considered before the technology can be implemented on a large scale. Previous research has shown that the physical placement of biometric devices has an effect on performance. Stanton, Theofanos, Orandi, Micheals, & Zhang (2007) found that the height of a four-digit, 'slap' fingerprint sensor affected usability. They found significant differences in transaction time, image quality and user preference as the height of the system was varied. Though height was assessed in this experiment, the effect of the angle of the sensor was not measured. Currently, there is no information about how the angle of biometric sensors affects the performance or usability of the technology.

If biometrics is to be widely used in self service technology such as self-checkout systems or automatic teller machines (ATMs) usability and accessibility become key issues. The self service environment is particularly demanding for technology as 'walk up and use' systems must accommodate a wide range of users with little or no training, supervision or external guidance. In these applications, usability and accessibility are critical; design features that have a small impact on usability may have a substantial impact on the success of the system as a whole.

Assessment approach

The aim of this research was to investigate the effect of position on the performance of biometric sensors and determine the optimum placement location for such technology. This evaluation was designed to expand on the research carried out by Stanton et al (2007) and investigate the effect of device angle on usability. In this experiment a fingerprint sensor and a palm vein sensor were tested as examples of biometric technology based on characteristics of the finger and hand respectively. Measures of usability were based on the ISO 9241 Standard for Usability, which defines usability as 'The extent to which a product can be used by specific users to achieve specified goals with effectiveness, efficiency and satisfaction in a specific context of use.' In this evaluation these terms were defined as follows:

Effectiveness – measured through image quality scores
Efficiency – measured using participants verification times
Satisfaction – measured through a subjective comfort rating given by participants

The postures adopted by participants when using the technology, were assessed using the rapid upper limb assessment (RULA) method (McAtamney and Corlett, 1993). RULA is a postural assessment tool with a focus on the positions of the trunk and upper limbs, developed to distinguish between acceptable and unacceptable postures.

The mixture of objective and subjective measurements used in this study helps to provide a more comprehensive assessment of device usability.

Figure 1. A participant using the palm vein device at three different angles.

Test positions

Positions of the devices were selected using numerous design guidelines including: *Access to ATMs: UK Design Guidelines*, the Canadian B651.1-01 *Barrier-Free Design for Automatic Banking Machines*, and the *Australian Bankers' Association Industry Standards*. A pilot test was conducted to refine the selection of heights and angles to be used in the main study. A range of angles (−15° to 75° in the sagittal plane) and heights (950 mm–1250 mm) were tested in the pilot study. The heights and angles that were used in the main evaluation were selected from the range that performed best during the pilot test.

Method

A within subject design was used with each of the 37 participants using both the fingerprint and palm vein devices. Participants ranged in age from 18–65 years old and ranged in height from 1490–1980 mm. 13 participants were female and 24 male. All participants were experienced ATM users and participated voluntarily this evaluation.

Participants enrolled using their right index finger and right palm. Enrolment involved participants presenting their palm or finger three times, so that a template could be created by the software. The technology was then tested at three heights (1000 mm, 1100 mm and 1200 mm) and three angles (0°, 15° and 45°). As seen in figure 1, the devices were mounted on camera tripods allowing height and angle to be varied. Participants verified twice at each of the 9 height/angle combinations, using each of the biometric systems. Both the order of technology use and order of positions were counter balanced across participants. Once the participants had used the sensor in each position they were asked to rate how comfortable they found the experience using a 7 point Likert scale. At the conclusion of the experiment, participants' demographic information was recorded. Postures were scored using the RULA tool from video footage taken during the evaluation.

Results

The relevant parametric and non-parametric statistical test were used to analysis the data and significant results were significant to the .05 level or greater. All

Table 1. Summary of the effect of position on performance.

	Efficiency (Time)	Effectiveness (Image Quality)	Satisfaction (Comfort Rating)	Posture Score
Fingerprint	No	No	Yes	Yes
Palm Vein	Yes	N/A*	Yes	Yes

* The software did not allow for an image quality score to be recorded.

participants were able to enrol with the palm vein device. One participant could not enrol with the fingerprint device and two participants had difficulty enrolling. The position of the sensors was found to have a statically significant effect on their usability during verification. Table 1 below summarizes the significant effects of device position on the measures taken. These results suggest that the positions used in this experiment had a greater effect on the subjective measures of satisfaction and posture than on the objective performance measures.

Comparison of sensor positions

In the interest of space not all significant effects are presented here. Overall, both the height and the angle had an effect on device usability. The results suggested that lower heights together with steeper angles should be avoided when positioning a fingerprint sensor. The 45°, 1000 mm position for the fingerprint sensor caused participants to adopt the worst arm postures and participants gave this position the lowest comfort rating. The palm vein system also performed poorly in this position with 45°, 1000 mm position rated significantly less comfortable than all of the other positions. Steeper angles at the 1200 mm height were found to be more effective and than horizontal positioning.

Stature was found to have a significant effect on the postures adopted for both technologies, with shorter participants adopting significantly worse posture for higher positions and taller participants adopting significantly worse postures for the lower height positions. Other participant characteristics did not significantly affect interaction with the devices.

Comparison of device performance

Interaction with the palm vein device took significantly less time than interaction with the fingerprint sensor. The postures that the participants adopted whilst using the fingerprint device were significantly better than those adopted when using the palm vein device however. Participants preferred the palm vein device to the fingerprint device, giving reasons based on its ease of use, comfort and the design of the housing.

Discussion

The height of participants affected the measures taken suggesting that the height of the user population also needs to be considered when implementing biometrics. Participant heights in this evaluation covered the 2nd to 99th percentiles of European adult stature (values from People Size application) so the people who took part in this evaluation are fairly representative of the wider population.

The position of biometric technology affected the usability of these systems. The height of the sensors affected usability as did the angle that the technology was presented to the user. These differences were seen only in the subjective comfort ratings and the posture analysis however. Measures of effectiveness and efficiency were not significantly affected by device position. This suggests that the subjective or 'user centric' measures of this evaluation were more sensitive than the objective measures taken. Future studies of biometric technology should focus on subjective measures of usability rather than solely use measures generated by the technology itself. Using a greater range of subjective measure in future could help to provide better understanding of the usability issues of biometric technology.

Participants preferred the palm vein device over the fingerprint device. One possible explanation of this difference is that the palm vein system seemed to be less demanding, as the fingerprint system required careful finger placement. Another reason for the difference could be due to fingerprint device requiring blind placement of the finger, as the finger covers the sensor when it is placed onto the device and there is little indication of correct or incorrect placement.

The interface that guides the user through the process of enrolment and verification is also crucially important, as this will potentially have a large effect on the ease with which the device can be used. The device packaging, the casing around the sensor, will also have an impact on comfort and user interaction with the technology. Participants' comments in the experiment reflected this. A well designed device housing is likely to lead to better placement and better image quality.

Conclusion

This study found that the position of the biometric technology significantly affects the comfort rating and the postures adopted by the users. The positions that were most usable tended to the middle height tested and the horizontal angle. The palm vein device out performed the fingerprint sensor in this evaluation, although changes could be made to the fingerprint device to improve performance. The results extend on previous research described in the introduction to this paper. Specifically, the results presented here show the importance of considering the angle of biometric technology in addition to the height at which the sensor is positioned.

Biometric authentication does provide an opportunity to improve security in self service contexts though there are numerous issues that need to be considered before an implementation of biometric technology would be successful. This investigation addressed the problem of placement locations for biometric sensors, though other issues need to be investigated to better understand the usability and accessibility

of biometrics. For instance an evaluation carried out with elderly users by Riley, McCracken & Buckner (2007) found that the vein based technology performed better than a fingerprint device. Participants who use wheelchairs should also be included in future evaluations to ensure that the most usable locations for biometric technology are also accessible. Wider contextual issues are important when it comes to biometrics and privacy issues and user acceptance also merit investigation.

Acknowledgements

We are grateful for the support and assistance provided by the Human Sciences Department at Loughborough University. We would also like to thank those involved from NCR Corporation Dundee.

References

BSI/ISO 9241-11 (1998). Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) – Part 11: Guidance on Usability. British Standards Institute.

Jain, A., Hong, L. & Pankanti, S. (2000). Biometric Identification. *Communications of the ACM*, 43, 91–98.

Maguire, M. (2001). Methods to support human-centred design. *International Journal of Human-Computer Studies*, 55, 587–634.

McAtamney, L. & Corlett, E.N. (1993). RULA: a survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics*, 24, 91–99.

Riley, C., McCracken, H. & Buckner, K. (2007). Fingers, Veins and the Grey Pound: Accessibility of Biometric Technology. Proceedings of the 25th Anniversary conference of the European Association of Cognitive Ergonomics (ECCE) 2007, London.

Stanton, B., Theofanos, M., Orandi, S., Micheals, R. & Zhang, N.-F. (2007). *Effects of Scanner Height on Fingerprint Capture* Paper presented at the Human Factors and Ergonomics Society 51st Annual Meeting, 2007, 592–596.

A HUMAN-CENTERED APPROACH FOR IMPROVING AIRPORT SECURITY SCREENING

Vassias Vassiliades[1], Paul Evans[1], Lloyd Kaufman[2]
Jer Wang Chan[1] & David Downes[1]

[1]*Nottingham Trent University, Nottingham, NG11 8NS, UK*
[2]*New York University, Long Island University, New York, USA*

Aviation security personnel can encounter problems when interpreting x-ray images of hand luggage. Our research seeks to determine whether we can increase the performance of the human operator, and thus human reliability, through the employment of our novel multiple-view x-ray imaging technique. This paper presents the results of some of our new experiments. Again we found that targets hidden in Kinetic Monoscopic (motion 2-dimensional) x-ray scans of hand luggage are more readily detected, relative to Static Monoscopic (standard 2-dimensional) scans of the same luggage. Furthermore, we now find that targets in KM images are also more likely to be identified as threats. These new results confirm and strengthen our previous findings. We feel that the remarkable improvement in performance afforded by the kinetic images warrants further future study. Overall, we expect that when deployed in the field this technique will revolutionize x-ray visual security screening.

Introduction

The experiments described in this paper are a continuation of our efforts to improve human reliability in the field of x-ray aided visual security screening. The most frequently used systems today consist of a linear x-ray sensor, which produces 2D line-scan images; color-coded for average atomic mass. The screener must rely upon his or her training and experience in order to visually sift through the cluttered x-ray images and make a decision if they contain a threatening object or not. However, this type of visual search imposes strains on the perceptual and cognitive capacities of the screeners (Harris, 2002). Furthermore, our experience in the field has led us to identify limitations in this technology, primarily due to the inherent lack of depth cues in traditional x-ray images. This complicates the screener's job, and raises the issue of human reliability in visual inspection tasks. Our research however, seeks to improve human reliability and promote the human-centered model for security screening, as opposed to automated processes.

Kinetic depth x-ray images

We know that motion parallax provides a powerful visual cue to depth (Kaufman, 1974). However, Wallach and O'Connel (1953) described a special case of motion

parallax, termed the kinetic depth effect or KDE. In their research, they discovered that they could produce a very strong depth effect by casting shadows of fully or partially rotating objects on a display. Specifically, if this was done so in a way that caused their contours or lines, to simultaneously change both their length as well as their direction. An observer could then accurately deduce object shapes from the depth information in these shadowgraphs. Kaufman (1974) suggested that since motion perspective (a form of motion parallax) entails a semi-rotation of elements along an extended terrain, it can therefore be considered a form of KDE.

Taking advantage of these findings our research team has developed a technique which produces KDE image sequences from a single x-ray source and a static configuration of multiple x-ray detector arrays (Evans and Hon, 2002; Hon and Evans, 2002). The KDEX imaging technique enables us to collect a range of views, of the item of hand-luggage under inspection, over different perspectives to produce sequential parallax. KDEX sequences may be produced in the same time period as conventional 2D x-rays, and this ability to essentially look around an object using multiple views at different angles is particularly more cost effective than using CT scans.

Human factors

Our human factors research combines our imaging technique with the bias-free procedures of signal detection theory (Macmillan and Creelman, 2005), in order to improve human performance and reliability in visual inspection tasks. We again set out to compare x-ray images incorporating depth through motion, termed "kinetic monoscopic" (KM), with the industry standard 2-dimensional, or "static monoscopic" (SM) images.

In our previous experiments we demonstrated that KM x-ray images afford a better chance for threat detection, compared to SM images. In the experiments described here we went one step further and set out to establish that KM images also lead to the better identification of threatening objects.

Methods

Stimuli

We used grayscale x-ray images with varying degrees of background clutter, and target saliency. We used images of 18 unique bags, and a control bag that appeared in both experiments, presented in 4 different orientations. Each threat-present bag image was complemented by its respective non-threat image, and all were presented 4 times during the course of the experiment resulting in a total of 320 trials. Subjects completed their experiment in 4 sessions; each consisting of 80 randomly presented images.

Equipment

Both experiments were conducted in our human factors laboratory, located in the Computing and Informatics Building at the Clifton Campus of NTU. The room

dimensions were 2.6 m × 4.9 m, with no windows and white painted walls. The room was equipped with two experimental stations, placed next to each other on a continuous bench. A 1.7 m × 2.0 m privacy screen was placed between the two experimental stations.

The experiments were entirely computer based. Stimulus presentation, and data collection was accomplished using custom software, developed by our group for this project. Subjects sat approximately 45–50 cm away from a 20" Iiyama HM 204 DT A CRT monitor (resolution 1024 × 768 pixels, refresh rate 150 Hz), with their line of sight at about the middle of the screen.

Participants

Subjects for the study were recruited from the student and staff population of NTU, and the surrounding community. We recruited 18 subjects (15 male, 3 female), with normal (or corrected) visual acuity. They ranged in age from 19 to 34 years old, and were compensated for their time with gift vouchers.

Procedure

The subjects were first explained the requirements of the project, and given informed consent. Following this, checked their visual acuity and contrast sensitivity by administering two short visual tests. Subjects' results fell within normal parameters.

The study involved two experimental groups, each combining two experiments, Static Mono (SM) and Kinetic Mono (KM), and using two sets of images, Image Set A (A) and Image Set B (B). The targets in set B images were much less salient than in set A images. Half of the subjects would complete SM(A)-KM(B), and other half KM(A)-SM(B); thus we employed a counterbalanced test-retest design where the two subject groups would serve as each others' control.

After subjects were randomly assigned to one of these groups, their training began. The focus was to familiarize them with how objects appear in luggage x-ray images and more importantly how threatening objects appear in these images. Once the subjects became familiarized with the procedure, they started regular experimental sessions, and their responses were recorded. Each session lasted between 20 to 30 minutes.

At the end of the experiment, subjects filled out a short feedback questionnaire. They were then thanked for their participation and given their compensation

Static experiment

The subjects were instructed to pay attention to each x-ray image that appeared on the screen, and when ready decide if the preceding image contained a threat or not. They indicated their response by left clicking, with the mouse, on one of six tabs below the image, giving them the option of: "YES Certain", "YES Probably", "YES Guessing", "NO Guessing", "NO Probably", and "NO Certain". If their answer was negative they could proceed directly to the next trial. However, if their answer was affirmative they were asked to place the mouse pointer over the object

Table 1. Aggregate results for image set A.

Experiment	H	F	c	s	A_z	log (β)	RT (sec)
SM Identification	0.28	0.17	0.76	0.63	0.58	0.28	12.98
KM Identification	0.62	0.08	0.54	0.46	0.82	0.92	14.52

Table 2. Aggregate results for image set B.

Experiment	H	F	c	s	A_z	log (β)	RT (sec)
SM Identification	0.18	0.07	1.20	0.62	0.59	0.70	14.33
KM Identification	0.37	0.04	1.03	0.41	0.67	1.46	12.92

they thought was a threat and left click on it. They could then proceed to the next trial.

When the new image appeared they were instructed to proceed through this, and each following trial as before. As a general rule subjects were instructed to react as quickly and as accurately as possible, and their response time during each trial was recorded.

Kinetic experiment

The procedure of this experiment was identical to the former, with the exception that all the images were presented in the Kinetic Monoscopic display. Additionally, before being able to click on an object they considered a threat, subjects had to first pause the image motion using the right mouse button. They subsequently proceeded as before.

Results

ROC analysis

As before, we pooled together the results for all subjects and using the confidence ratings data, we conducted a ROC analysis using Prism 4 (Graph Pad Software, 2005). Additionally the data were plotted in z-coordinates and linear regression was employed to obtain the slope of the lines. They resulted in a good linear fit; thus confirming the normality assumptions of the underlying distributions. However, because the slopes were non-unit, we deemed the A_z to be the appropriate performance index for our data (Macmillan and Creelman, 2005). The results are summarized in the tables below.

Discussion

In our previous experiments (Evans *et al*, 2007) we found that KM images resulted in a superior rate of detection when compared to SM images. That was indicated

by the considerable increase (20%) of the A_z value. Here we were able to not only replicate our earlier findings, but also extended the range of our research to include the actual identification of targets. For image set A, composed of images with a moderate to severe degree of difficulty, we found a 24% performance increase, in terms of A_z, There was a much smaller effect for image set B, which was composed of images with very elusive targets, which we feel would compound even professional screeners. This is hardly surprising given the fact that we used inexperienced subjects in our experiments. However, despite this sharp increase in the level of difficulty, results still showed an 8% performance improvement for the KM display, over SM.

The current results further support our hypothesis that viewing images in the KM display results in superior threat detection rates, compared to the SM display. We can now also confidently state that when taking actual target identification into account the performance difference holds up, even when confronted with very difficult targets such as those in image set B. By providing depth information to the human operators, our technique clearly helps them overcome current limiting factors such as ordinary masking, where a masking object conceals the target, and crowding, where the masking object distorts the shape of the target (Pelli *et al*, 2004).

In our current results, we again find that the image display mode does not appear to significantly influence average response times (RT). In fact, the fluctuations in RT ($\sim$1.5 sec) swung either way, and can rather be attributed to subject differences, and not the difference in viewing mode. Thus, despite the great performance advantage our technique has, over traditional displays, it does not appear to add any additional processing time to the task.

Ultimately, our research seeks to promote the human-centered model for visual security screening, as opposed to automated processes, by working to improve human reliability. To that effect, further work still needs to be done, and our group plans to expand the scope of our research to include a larger variety of threatening objects. We also plan to utilize images with false color-encoding in our future experiments.

Acknowledgements

This ongoing research is funded by a grant from the US DHS Transportation Security Laboratory's Human Factors Program. The authors would like to personally thank the program's director, Dr Josh Rubinstein, for his unwavering support throughout the project. The authors also acknowledge Professor Dick Lacey, of the UK HOSDB, for his support of this program of research and in particular for the provision of an experimental x-ray system, on a permanent loan basis.

References

Evans, J.P.O. and Hon, H.W. 2002, Dynamic stereoscopic X-ray imaging, *JNDT&E*, **35**, 337–345.

Evans, J.P.O., Vassiliades, V., Chan, J.W., Kaufman, L. and Downes, D. Human factors evaluation of kinetic depth x-ray imagery for aviation security screening, *Contemporary Ergonomics 2007*, (Taylor and Francis, London), 190–196.

Graph Pad Software 2005, Prism 4 for Windows, Version 4.03, (San Diego).

Harris, D.H. 2002, How to really improve airport security, *Ergonomics in Design*, 17–22.

Hon, H.W. and Evans, J.P.O. 2002, Multiple view line-scan imaging, *IEE Proc. Optoelectronics*, **149** (2), 45–50.

Kaufman, L. 1974, *Sight and Mind*, (Oxford University Press, New York).

Macmillan, N.A. and Creelman, C.D. 2005, *Detection Theory: A user's guide*, Second Edition, (Cambridge University Press, New York).

Pelli, D.G., Palomares, M. and Majaj, N.J. 2004, Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, **4**, 1136–1169.

Wallach, H. and O'Connell, D.N. 1953, The kinetic depth effect, *J. Experimental Psychology*, **45** (4), 205–217.

KEEPING TRACK OF A SUSPECT ON CCTV: HOW HARD CAN IT BE?!

Sarah Harris, Alex Stedmon, Sarah Sharples and John Wilson[1]

Human Factors Research Group, University of Nottingham, Nottingham, UK
[1]*also School of Safety Science, University of New South Wales*

Successful closed circuit television surveillance relies on the skill and competence of operators performing complex visual tasks. Various techniques can be employed during the design of control rooms that support operator performance by increasing the probability of them noticing an incident and reducing workload. This experiment investigated the effectiveness of one technique for facilitating operator performance during a person tracking task: manipulating the geographic layout of images on the screen. The findings failed to produce evidence of a change in performance or workload when geographic layout was manipulated. However the research provides a novel insight into what behaviours are considered suspicious within the context of a university campus and raises questions regarding the generalisability of suspicious acts to other locations.

Introduction

The design of closed circuit television (CCTV) can affect how successfully the system operates and how well the operator performs surveillance activities. It is important that monitors in a control room are displayed in a structured arrangement as this is likely to increase operator comfort and facilitate surveillance performance. However, in reality, CCTV control rooms are sometimes an ad-hoc set up of monitors which have evolved over time (as new cameras have been installed) and where the experienced observer often builds their own method of monitoring them.

Specific techniques can be employed to increase the probability of an operator seeing an incident on CCTV. For example, performance may be enhanced by grouping the displays by their site location or coverage area (Donald, 1998). Monitors can also be arranged according to the sequence of use of cameras demonstrated by the operators. If multiple images are presented on a single monitor, each image is smaller and the technical quality of the images is degraded, making visually demanding tasks such as person identification difficult (Pethick & Wood, 1989). Therefore, if multiple images are viewed on a single monitor, they should be geographically laid out in a way to facilitate performance of the task being undertaken (Donald, 1998). These techniques available for arranging CCTV monitors should encourage systematic monitoring and help the observer scan and evaluate images more effectively (Wallace & Diffley, 1998).

This paper presents a study that investigated the impact of geographic layout of CCTV images on a person tracking task. An experiment was conducted to investigate the effect of screen layout on speed and accuracy at tracking a target across 6 video screens displayed on a single monitor.

Method

Participants

24 participants were recruited, 3 male and 21 female and were paid for their time. The participants were staff and students of the University of Nottingham aged between 18 and 40 years (mean = 23 years, SD = 5.73 years). A pre-experimental test ensured that all participants were familiar with the university building shown in the video footage.

Apparatus

The CCTV set-up was simulated in a laboratory using a computer monitor showing 6 separate videos running simultaneously (Fig. 1). The videos were footage taken on the University of Nottingham with a 'target' walking around one of the buildings. The videos were set up so that the target initially appeared in screen 1 and then appeared in all of the 5 other screens at some point during the film. The footage ran for a total of 1.5 minutes. A camcorder was set up on a tripod in front of the monitor to record participant verbalisations and behaviour to enable post-experimental analysis. An electronic version of the NASA TLX questionnaire was administered to participants using a laptop.

Figure 1. Layout of video screens on computer monitor.

Design

A repeated measures within-subjects experimental design was used. Two independent variables were investigated: the geographic layout of the video and the type of tracking task. The geographic layout of the video screens on the monitor was manipulated to produce 2 conditions. In the **'logical sequence'** the video footage followed the logical footprint of the building and the movement of the target between adjacent screens across the top and then bottom row of screens. In the **'random sequence'** the layout of video screens was randomised so that the target could appear on any screen on the top or bottom row at any particular time and never appeared in adjacent screens. Two types of tracking task were used, classified as **'easy'** and **'hard'**. Tracking the target in the 'hard' condition was more difficult due to a more convoluted route taken by the target through the building than the easy route coupled with more ambiguous camera locations. The independent variables were combined to produce 4 conditions: easy logical, easy random, hard logical and hard random. Each participant completed all conditions, which were counterbalanced to reduce fatigue and learning effects. The dependent variables were tracking task performance (accuracy score and reaction time) and subjective workload (NASA-TLX).

Procedure

Experimental sessions lasted 45 minutes. In each condition, participants viewed the footage and visually tracked the target across the 6 screens, stating aloud as quickly as possible the screen the target appeared on. Participant accuracy and time taken to identify the target on each screen was recorded. When the footage ended, participants rated their workload using the NASA TLX tool and were asked questions regarding the task. Participants also searched for suspicious behaviour in the footage (there was an additional activity of a bag being dropped and abandoned by a second actor). Following completion of all conditions, participants answered questions about how they tracked the target during the trials.

Results

Mean data scores for tracking task and workload are shown in Table 1. The data was analysed using a series of within-subjects ANOVAs.

Table 1. Mean scores for tracking task and subjective workload.

	Accuracy score (%)		Reaction time (secs)		Subjective workload (/100)	
Condition	Mean	SD	Mean	SD	Mean	SD
Easy logical	90.83	14.42	2.57	1.13	54.04	16.29
Easy random	95.83	8.30	2.93	1.68	56.79	17.13
Hard logical	90.83	11.76	3.72	1.45	60.25	13.77
Hard random	89.17	20.41	3.95	1.59	61.38	14.98

Tracking task accuracy

No significant main effect was observed for task type or screen layout ($p > 0.05$) illustrating that accuracy scores in the tracking task did not alter as a function of either the route the target took through the building or the geographic layout of the screens.

Tracking task reaction time

A significant main effect was observed for task type ($F(1, 23) = 14.81$; $p < 0.05$). Raw data indicates that reaction time was slower when the target followed the hard route through the building (mean = 0.548 secs) than when the target followed the easy route (mean = 0.384 secs). No significant main effect was found for screen layout ($p > 0.05$) illustrating that reaction time did not alter as a function of screen layout.

Subjective workload (NASA TLX)

No significant main effect was observed for subjective workload ($p > 0.05$) illustrating that workload did not alter as a function of either task type or screen layout.

Unusual behaviour

Only one of the 24 participants noticed the bag drop. When questioned about suspicious behaviour, participants commonly referred to people loitering or disappearing and re-appearing, people wearing unusual clothes or being dressed strangely for the location, people moving in an unusual way such as suddenly changing direction and abnormal people for the location such as a policewoman.

Tracking technique

Table 2 presents the self-reported techniques used by participants to track the target.

Table 2. Strategies and techniques used during tracking task.

Technique used to track target	% response
Visualised building and anticipated route target would take	24
Randomly scanned video screens	18
Searched for target's appearance/ clothing/ bag	18
Eliminated screens target had previously appeared in	13
Logically scanned video screens	13
Viewed entry/ exit points and searched for people entering the scene	6
Looked for person on their own and ignored groups of people	4
Looked for person on the move/ ignored people standing still	4

Discussion

The aim of this study was to examine the effectiveness of a recommended technique for structuring CCTV images to enhance operator performance. On the basis of suggested guidelines it was predicted that the accuracy and speed at which participants could track a target across multiple screens would increase when the screens were logically displayed (Wallace & Diffley, 1998). However, performance in the tracking task was not significantly affected by screen layout and workload did not alter between conditions.

Although the lack of significant data for tracking task performance seems to contradict the guidelines this is possibly due to the nature of the tracking task (e.g. the task may have demanded more of a visual search activity as participants scanned the screens for the target rather than continuously following the target across and between screens). Further research should determine whether other types of surveillance activity, such as object recognition or person identification are more sensitive to changes in screen layout and if particular image arrangements can facilitate performance and lower workload.

The results for the unusual behaviour analysis illustrated that only one participant noticed the bag drop. This may be due to the fact that the tracking task was the primary task and so attention was focused on the target with background activity intentionally ignored or unintentionally missed. However it is also feasible that participants did not consider the bag drop to be suspicious on a university campus and so did not report it. This raises interesting questions regarding the idea of what behaviours are considered suspicious in different locations and contexts. Participants in this study provided a set of behaviours that they considered suspicious on a university campus but how generalisable are these activities to alternative locations, such as a rail station? To develop this study, further work has explored what behaviours are considered suspicious within a range of contexts and locations (Stedmon *et al*, 2008).

This study also provided an insight into how people track a target across multiple video screens. The majority of participants reported that they visualised the building and camera locations and anticipated the route the target would take. This finding highlights the importance of real-world CCTV operators having in-depth knowledge of the site locations they are viewing. Further research is required to investigate the techniques used when performing a wider range of visual tasks using CCTV.

Conclusions

The findings of this study are an important step in understanding how operators perform complex tasks using CCTV. Alternative techniques for displaying screens should also be investigated using a range of methodologies. The research also raises interesting questions regarding what behaviour is considered as suspicious in different contexts and by different observers.

Acknowledgements

Work funded through the EPSRC Rail Research UK, with the cooperation of Network Rail.

References

Donald, C. 1998, Ergonomic considerations in CCTV viewing. *Hi-Tech Security Systems Publication, May 1998*, **4(3)**. http://www.leaderware.com/hitech/ergonomic%20monitoring.htm.

Pethick, A.J. and Wood, J. 1989, Closed circuit television and user needs. In E. Megaw (ed.) *Contemporary* Ergonomics *1989*, (Taylor and Francis, London).

Stedmon, A.W., Harris, S., Carse, A., Sharples, S. and Wilson, J.R. 2008, Tracking a suspicious person using CCTV: but what do we mean by suspicious behaviour? In P. Bust (ed.) *Contemporary Ergonomics 2008*, (Taylor and Francis, London).

Wallace, E. and Diffley, C. 1998, CCTV: Making it work: CCTV control room ergonomics. *Police Scientific Development Branch Publication*, ***14/98***. http:sciencean dresearch.homeoffice.gov.uk/hosdb/publications/cctvpublications/1498CCTVControlRoomErgonomics?version=1.

TRACKING A SUSPICIOUS PERSON USING CCTV: BUT WHAT DO WE MEAN BY SUSPICIOUS BEHAVIOUR?

Alex Stedmon, Sarah Harris, Anthony Carse, Sarah Sharples[1] & John Wilson[2]

[1] *Human Factors Research Group, University of Nottingham, UK*
[2] *School of Safety Science, University of New South Wales, Australia*

A number of examples illustrate that the police can make mistakes in identifying suspicious behaviour. Criminals will try to blend in and not draw attention to their behaviour and so a fundamental question emerges with the use of CCTV: what do we mean by suspicious behaviour? A laboratory experiment and field based survey were conducted to investigate different concepts of suspicious behaviour. The findings illustrate that people can interpret innocent behaviour as potentially suspicious and that they also have consistent ideas of what behaviours are suspicious across different locations and public spaces. The work is being developed into a basic framework for profiling suspicious behaviour in different contexts.

Introduction

On 22 September 2005 an innocent man was making his way across London to see his girlfriend. Upon entering a tube station he was arrested on suspicion of terrorism and detained overnight (Mery, 2005). With tensions running high after the bombings a few weeks before, this man had been caught on closed circuit television (CCTV) and considered a suspect because: he did not look at police officers standing at the station entrance, two other men entered the station at the same time as he did, he was wearing a jacket that was considered too warm for the season, he was carrying a bulky rucksack and kept it with him at all times, he looked at people coming onto the platform, and he played with his mobile phone. More recently the situation with Jean Charles Menizes also illustrates that tragic mistakes can be made.

Whilst this example was context specific to recent terrorist activities, it illustrates that people make assumptions about what might constitute suspicious behaviour in a given situation. In most cases a terrorist, or someone about to commit a less extreme crime such as pick-pocketing, will try to blend in and not draw attention to their behaviour. This presents a problem in distinguishing between everyday behaviour that could be misinterpreted as suspicious and behaviour that is actually suspicious and perhaps a pre-cursor to a criminal act. As a result, a fundamental question with the use of CCTV is: what do we mean by suspicious behaviour?

Previous research has illustrated that both expert and novice CCTV users are able to identify suspicious behaviour leading up to a criminal act and that CCTV users

could mistake innocent activities for suspicious behaviour (Troscianko, et al, 2004). It is important, therefore, to understand what people might regard as suspicious behaviour in different contexts as certain behaviours might be common across situations or context specific. By developing a better understanding of these factors, it may be possible to profile a suspicious person based on key traits that can be used as metrics or a training tool by CCTV users. In order to investigate what constitutes 'suspicious behaviour' two independent activities were undertaken:

- laboratory experiment to classify suspicious behaviour
- field based survey of suspicious behaviour

Method

Participants

In the laboratory experiment 24 participants (12 male and 12 female) were recruited from the student population at the University of Nottingham. In the field based survey 150 participants were recruited from members of the public.

Apparatus

In the laboratory experiment video footage was presented on a computer monitor representing six CCTV screens running simultaneously. The video was the same footage used by Harris et al (2008) captured on the University of Nottingham. Participants' comments were recorded by the experimenter and a stop watch was used to log specific events. In the field based survey a simple questionnaire was developed as a basis for structured interviews.

Design

A within-subjects design was employed for the laboratory experiment with two CCTV scenarios counterbalanced between-subjects. Each scenario ran for 90 seconds and portrayed normal student behaviour on the campus as well as the suspicious behaviour of the bag drop from the earlier study. Data was recorded from participant comments on any suspicious behaviour (including the bag drop) that they observed. The field based survey was conducted at three separate locations:

- Nottingham Train Station
- Grosvenor Shopping Centre (Northampton)
- University of Nottingham campus

Procedure

In the laboratory experiment, participants viewed each of the two CCTV scenarios in turn and commented on anything they perceived as suspicious. They were not given any guidance on what behaviours might be regarded as suspicious and they were not informed prior to the event of the bag drop. Participants were paid for their time. In the field based survey 50 volunteers from each location were interviewed

about their perceptions of CCTV, security, and suspicious behaviour. The interviews were designed to take a few minutes and volunteers were not paid for their time.

Results

In the laboratory experiment participants identified a number of behaviours which they felt were suspicious. These are presented in Table 1 below for both scenarios. Only the top responses for each scenario are illustrated. In total there were 15 behaviours listed for Scenario 1 and 17 for Scenario 2. Responses further down the original tables were only mentioned by one participant at a time.

The top suspicious behaviours observed in each scenario were the bag drops although in each case less than 50% of participants noticed this behaviour. Loitering, looking around and running were also considered to be suspicious behaviours.

Results for the field based survey are presented for each question below.

Question 1 – when you see signs asking you to keep a look out for unattended bags or suspicious behaviour do you generally think about looking out for suspicious

Table 1. Suspicious behaviour from video footage (n = 24).

Scenario 1	Suspicious behaviour	Responses	Reasons given
•	Bag drop on stairs	9	– Hood up
•	Girl running towards screen through doorway	9	– Running
•	Man waiting holding a red bag	5	– Loitering, red bag
•	Man in suit walking upstairs looking around	4	– Wearing a suit – Looking around
•	Policewoman in uniform walking away from the camera	4	
•	Two people standing outside the entrance	3	– Loitering – Not talking
•	Man with moustache wearing a brown jacket walking upstairs and looking around a lot	3	– Moustache – Jacket – Looking around
Scenario 2	**Suspicious behaviour**	**Responses**	**Reasons given**
•	Bag drop next to bin	10	
•	Man wearing long coat, standing around	8	– Loitering – Big jacket – Older person
•	Woman looks through window before entering a room	6	– Loitering – Looked through window – Looked confused
•	Man in long coat returns	5	– Loitering – Wearing a suit
•	Unattended bag on the floor	4	
•	Man wearing green trousers waiting	4	– Loitering – Green trousers

Table 2. Suspicious behaviour from different locations (top 5 ranked answers).

University	Shopping Centre	Train station
1 Loitering individuals	**1** Loitering individuals	**1** Loitering individuals
2 Older men	**2** Loitering groups	**2** Looking around a lot
3 = Staring, looking at me/someone else	**3** Scruffy clothing	**3** Running, pushing
3 = Running, pushing	**4** Running, pushing	**4** Big jacket
5 = Looking around a lot	**5 =** Looking around a lot	**5** Backpack, large bag
5 = Stands out, looks out of place	**5 =** Stands out, looks out of place	
5 = Scruffy clothing	**5 =** Avoiding eye contact	
5 = Hanging around bikes, cars, labs	**5 =** Backpack	
	5 = Hat on indoors	

people or items? 37% of people said they would generally look out for suspicious behaviours. People were more likely to be vigilant at a train station (46%) than in a shopping centre (40%) or on a university campus (24%). In all cases the majority of people would not look out for suspicious behaviour.

Question 2 – Does it make you feel safe knowing that CCTV is being used in public spaces? 73% of people surveyed felt safer knowing CCTV was being used in public spaces. People felt safer at a train station (76%), than on a university campus (74%) or in a shopping centre (68%). People felt safer with CCTV in use than without it.

Question 3 – what kind of things would make you think someone was being suspicious here today? The top five ranked answers for each location are presented in Table 2 above.

Loitering is considered the most suspicious behaviour in all three locations. More generally, people who are running, pushing, and/or looking around a lot are also likely to be regarded with suspicion.

Discussion

From the results of the laboratory experiment a number of activities were regarded as suspicious (such as loitering, running and looking around) even though the people in both scenarios were people innocently going about their daily activities. In each scenario, participants noticed the bag drop as a key activity presumably because they were primed to look for anything suspicious. In another study using the same scenarios, when participants were not instructed to look for anything suspicious only one participant noticed the bag drop (Harris, et al 2008).

The findings from the field based survey supported the laboratory study: loitering, running, and looking around appeared to be key behaviours that are

considered suspicious. The findings from the laboratory experiment were consistent with the university campus survey location which suggests that the survey results reflect the visual cues that are used to interpret suspicious behaviour. Some elements of behaviour are regarded as suspicious across all three locations, however, there were also behaviours which seemed to be location specific.

On the university campus older men were regarded as suspicious because they looked too old to be students. The university campus is biased towards a particular age range (18–24 years) and so people who do not 'fit' a particular stereotype might be regarded as suspicious (Stangor, 2000). In the shopping centre a unique behaviour that was considered suspicious was not carrying shopping bags. This behaviour in itself is very interesting as it implies that observers might infer a certain purpose in different locations. Avoiding or paying attention to security staff was a unique suspicious behaviour to the train station. This location had the most obvious security measures in place and so anyone actively trying to avoid contact or monitor security staff would arouse suspicion. This is linked to general looking around behaviour which was regarded as suspicious across all three locations, however the response was given more times at the train station than the university campus or shopping centre. Other behaviours specific to the train station were wearing a large jacket or carrying a large back/rucksack. These responses might have been due to new stereotypes emerging in relation to the context of the 'London bombings'.

From the general questions on CCTV attitudes and security, people feel reassured that CCTV is being used in public spaces. Perhaps this reliance on CCTV also helps to explain why people do not actively look for suspicious activities when they might think that someone else is doing the job for them.

Conclusion

The findings of this study illustrate that people can interpret innocent behaviour as potentially suspicious and that they also have consistent ideas of what behaviours are suspicious across different locations and public spaces. The work from the laboratory experiment and field based survey is being developed into a basic framework for profiling suspicious behaviour in different contexts.

Acknowledgements

Work funded through the EPSRC Rail Research UK, with the cooperation of Network Rail.

References

Harris, S., Stedmon, A.W., Sharples, S.C. & Wilson, J.R. (2008) Keeping track of a suspect on CCTV: how hard can it be? In, P. Bust (ed.) *Contemproary Ergonomics 2008*. Taylor & Francis: London.

Mery, D. (2005) Suspicious behaviour on the tube. *The Guardian* (22nd September, 2005).

Stangor, C. & Shaller, M. (2000) Stereotypes as individual and collective representations. In, C. Stangor (ed) *Stereotypes and Prejudice*. Taylor & Francis.

Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D. & Wilson, A. (2004) What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception* 33. pp.87–101.

EXPLORING ESSENTIAL SKILLS OF CCTV OPERATORS: THE ROLE OF SENSITIVITY TO NON-VERBAL CUES

Anastassia Blechko, Iain Darker & Alastair Gale

Applied Vision Research Centre, Loughborough University, Loughborough, LE11 3UZ

The aim of the present research is to investigate the importance of individual differences in the recognition of emotional state from non-verbal, visual cues in relation to the work of CCTV operators. An experiment was conducted to determine whether the detection of a gun carrier can be predicted on the basis of an observer's ability to recognize emotion conveyed by non-verbal, visual cues. There was a relationship between the emotional state reported by an individual whilst carrying a firearm and the extent to which an observer was ability to determine whether or not this individual was carrying a gun. It was also found that observers with a high ability to recognize emotion from facial expressions were able to spot a gun carrier more accurately than observers with a low ability in this respect. This small-scale pilot study requires replication with a larger number of participants and real CCTV images.

Introduction

The selection of good Closed Circuit Television (CCTV) operators is essential for effective CCTV system functioning (Donald, 1999). Through dialogue with operators and their managers in several control rooms in the UK it has been indicated that good performance amongst CCTV operators depends not only on proper training but also on the 'innate abilities' of candidates. However, there is little empirical research to support this assertion.

Previous research involving those convicted under the UK firearms act indicates that carrying a firearm evokes an emotional response in an offender (Hales, Lewis, and Silverstone, 2006). It is possible that the ability of operators to perceive this emotional response will predict their performance in spotting armed criminals. This ability might be based on body language analysis and particularly on an understanding of emotional state derived from non-verbal, visual cues.

The ability to recognize accurately emotional expressions transmitted by others through their non-verbal behaviour (e.g. from facial expressions and gait) has been widely studied in experimental psychology (e.g. Ekman, and Friesen, 1969; Montepare, Goldstein, and Clausen, 1987; Atkinson, Dittrich, Gemmell, Young, 2004). However, little applied research has been done on this topic with respect to the work of CCTV operators. The aim of the present research is to investigate the importance of individual differences in sensitivity to non-verbal cues and the ability

to recognize emotions from these cues for a better understanding of the performance of CCTV operators. An experiment has been conducted to determine whether the detection of a gun carrier can be predicted on the basis of an observer's ability to recognise emotions from non-verbal, visual cues. Specifically, it is hypothesised that those who are better able to recognise emotional state on the basis of facial and bodily cues will be better able to detect concealed and unconcealed firearms in CCTV footage.

Experiment

The experiment was designed to examine the relationship between an observer's sensitivity to non-verbal, visual cues relating to emotional state and their ability to detect a gun carrier. Further, the possible relationship between the self-reported emotional state scores of gun-carriers, and the ability of observers to recognize that gun carrier was investigated.

Method

Firstly, idealised, 'staged' CCTV video footage of twelve males, each carrying either a firearm or a matched innocuous object, was generated. The firearm and matched innocuous objects were either carried in plain view or they were carried concealed on the person (Figure 1). Whilst the footage was filmed, the potentially differing influences on affect of carrying a firearm and carrying an innocuous object were assessed. Seventy-two clips of people carrying firearms were and 216 clips of people carrying an innocuous object were generated (half unconcealed and half concealed). For each type of object (firearm or innocuous) half the clips featured an unconcealed object and half featured a concealed object. See Darker, Gale, and Blechko (forthcoming) for a complete description of this aspect of the experiment.

Subsequently, eight staff members from Loughborough University (6 female and 2 male) volunteered to participate in the study relating to an observer's ability to detect a person who is carrying a firearm. The experiment took place over two sessions. In the first, participants were administered the Face and Body PONS test, which is a shortened version of the full PONS test (Rosenthal, Hall, DiMatteo, Rogers, and Archer, 1979). The Face and Body PONS test measures the ability to decode non-verbal cues expressed by the face and body. The test consists of video fragments of a young woman acting in different naturalistic, emotional situations. In the present experiment, all audio content was excluded. After viewing each video fragment the participants (i.e. 'observers') were required to make a choice between one of two descriptions relating to the emotive content of the clip. Participants were asked to select the description which best described the situation acted out in the fragment. Then each participant viewed the idealized footage of people carrying either a firearm or an innocuous object. In the second part of the experiment participants viewed each clip of idealised footage of people carrying either a firearm or an innocuous object and, after each clip, rated whether or not they thought the person in the clip was carrying a firearm. This is referred to as the Gun Carrier Detection (GCD) task.

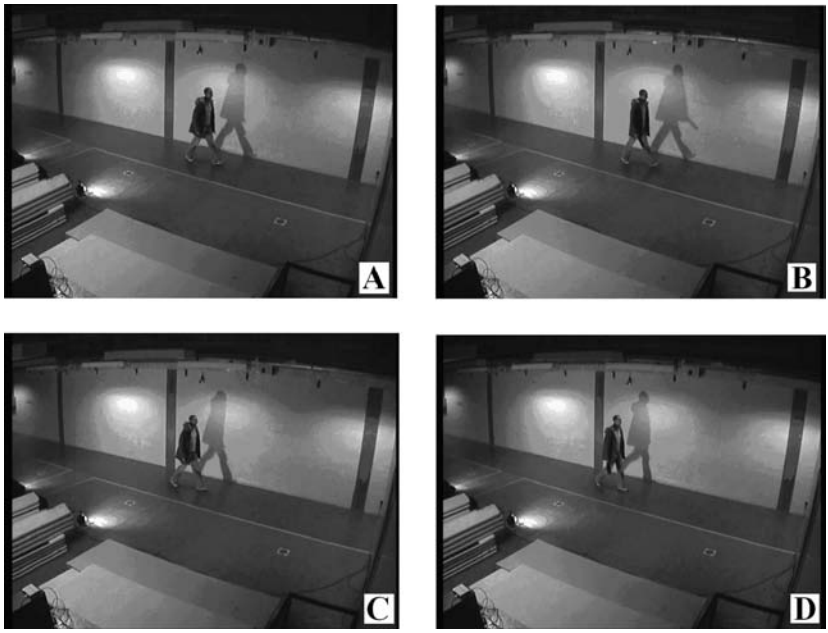


Figure 1. Examples of still images of the idealised footage (A – walking with a concealed firearm; B – walking with an unconcealed firearm; C – walking with a concealed two-litre bottle; D – walking with an unconcealed two-litre bottle).

Results

Performance on the PONS test was analysed against two measures: Face PONS test score; Body PONS test score. These scores were used to split the participants into two groups for each of the measures, based on a median split. Thus, for the Face PONS scores a high sensitivity (n = 4) and a low sensitivity (n = 4) group were formed. Similarly, for the Body PONS scores a high sensitivity (n = 4) and a low sensitivity (n = 4) group were formed.

Performance on the GCD task was measured in terms of the percentage of true positives (hits: the percentage of times a gun is detected when it is present), false positives (false alarms: the percentage of times a gun was reported, when in fact no gun was present), and the sensitivity parameter *da* derived from a Signal Detection Theory (SDT) based analysis. On this basis, sensitivity (*da*) represents the ability of the observer's sensory process to discriminate a target (the gun) from noise(the innocuous object). (For a review of SDT see Harvey, 2001).

In order to analyze the influence of concealment and sensitivity to facial and bodily non-verbal cues on the performance of observers two, 2 × 2 mixed ANOVAs were conducted: concealment (concealed, unconcealed) (within subjects) × PONS face group (high sensitivity, low sensitivity) (between subjects); concealment (concealed, unconcealed) (within subjects) × PONS body group (high sensitivity, low sensitivity) (between subjects).

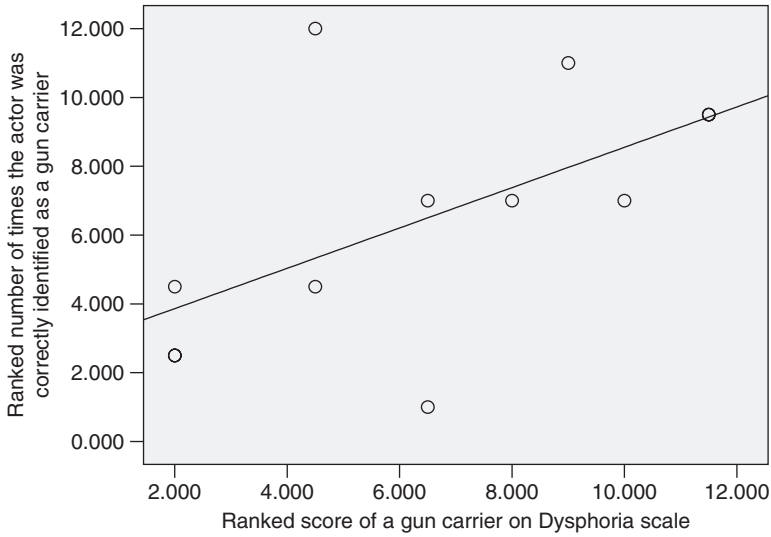


Figure 2. Correlation between the ranked score of a gun carrier on the Dysphoria scale and the ranked number of times the participants was correctly identified as a gun carrier in the Gun Carrier Detection task.

The analysis showed a significant main effect for concealment on all dependent variables in both analyses. Participants performed significantly better in the unconcealed condition than in the concealed condition. There was a significant interaction effect found for the Face PONS score and Concealment on *da* scores ($F_{(1,6)} = 7.391$, $p = 0.035$). The group of observers with a high score on the Face PONS test had higher recognition sensitivity than the observers with a low score on the Face PONS test, but this effect was found only in the unconcealed condition. There was no significant effect of Body PONS test on any of the dependent variables, or other main or interaction effects.

A Spearman rank correlation test was conducted to explore the relationship between the ability of observers to recognize correctly a gun carrier and the emotional state scores reported by gun carriers whilst carrying a gun (see Darker *et al.*, forthcoming). The results of the test (Figure 2) showed that there was a positive correlation between the ranked Dysphoria (MAACL-R, Lubin and Zuckerman, 1999) scores of gun carriers and the ranked number of times the actor was correctly identified as a gun carrier (rho = 0.586, n = 12, $p = 0.045$, two-tailed).

The current experiment revealed that observers' performance was associated with more true positives and fewer false positives when the object (firearm or innocuous) was in view, compared to when it was concealed. In contrast to those in the low sensitivity group for the Face PONS score, observers in the high sensitivity group for the Face PONS score performed significantly better in the detection of a gun carrier than other observers when the gun or innocuous object was not hidden. However, there were no significant differences between these groups when the

objects were concealed. The Body PONS score did not influence gun detection performance.

General discussion

Prior research indicates that firearms induce an emotional response in the bearer and that an observer can detect this affect through non-verbal, visual cues conveyed by the face and body of the person expressing the affect. Therefore, it was hypothesised that those who are better able to recognise emotional state on the basis of facial and bodily non-verbal, visual cues will be better able to detect concealed and unconcealed firearms in CCTV footage.

The results of the current study showed that those who are more sensitive to emotional cues conveyed by the face are better able to detect an unconcealed firearm, whilst it appears that differing sensitivities to bodily cues did not influence gun detection performance. Thus, it might be argued that non-verbal, emotional cues conveyed by the face are most useful in detecting the emotion evoked in a gun carrier by the act of carrying a gun. However, the ability to detect a concealed gun was not influenced by sensitivity to either facial or bodily non-verbal emotional cues. Thus, it is possible that the ability to detect a gun amongst those with high sensitivity to facial, non-verbal emotional cues is not subserved by an ability to decode such cues. However, performance on the unconcealed was at base level, consequently the true significance of this result remains obscured. There is further evidence that the ability to detect a gun carrier is related to the ability to decode visual, emotional cues. Those gun carriers who reported more negative affect on carrying a gun were most likely to be successfully recognised as carrying a gun.

Thus, the present study provides an initial indication that visual, emotional cues conveyed by the face of a person who is carrying a gun can aid in the detection of the gun by a third party observer such as a CCTV operator. Future work will build on this small-scale pilot study to quantify the gun detection abilities of CCTV operators and lay people in relation to their abilities to decode non-verbal, visual emotional cues conveyed by the face and body. It is intended that this work will help inform policy and practice with regard to the recruitment of CCTV operators and the detection of firearms via CCTV.

Acknowledgments

This work is a part of the Multi Environment Deployable Universal Software Application (MEDUSA) project which is funded by EPSRC grant number EP/D078830/1.

References

Atkinson, A.P., Dittrich, W.H., Gemmell, A.J. and Young, A.W. (2004). Emotion perception from dynamic and static body expression in point-light and full-light displays. *Perception, 33*, pp. 717–746.

Darker, I.T., Gale, A.G. and Blechko, A. (forthcoming). *An affective component to the weapons effect.*

Donald, C.H.M. (1999). *Assessing the Human Vigilance Capacity of Control Room Operators.* In: People in Control: An International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres, Conference Publication No. 463, IEE, pp. 7–11.

Ekman, P. and Friesen, W.V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32*, pp. 88–106.

Hales, G., Lewis, C. and Silverstone, D. (2006). *Gun crime: the market in and use of illegal firearms*. Home Office Research Study 298.

Harvey, L.O. Jr. (2001).'*RSCORE 5.3.2*''. Department of Psychology, University of Colorado, Boulder, Colorado.

Lubin, B. and Zuckerman, M. (1999). Manual for the MAACL-R: Multiple affect adjective check list–revised. Educational and Industrial Testing Service, California.

Montepare, J.M., Goldstein, S.B. and Clausen, A. (1987). The identification of emotions from gait information. *Journal of Nonverbal Behaviour, 11(1)*, pp. 33–42.

Rosenthal, R., Hall, J.A., DiMatteo, M.R., Rogers, P.L. and Archer, D. (1979). *Sensitivity to nonverbal communication: The PONS test.* Baltimore: Johns Hopkins University Press.

HAZARDOUS INDUSTRIES

DO I REALLY NEED A LIMO? FIT FOR PURPOSE ERGONOMICS/HUMAN FACTORS WORKSHOP

Alan Jackson & John Wilkinson

Health & Safety Executive, Redgrave Court, Bootle, Merseyside L20 7HS

The introduction of the Control of Major Accident Hazards (COMAH) regulations in 1999 raised the profile of human factors within the major hazards sector in the UK. Few, if any, of the top tier COMAH sites in the UK employ full time, qualified ergonomists/human factors practitioners. Therefore, it is common for sites to buy in their expertise from consultants. However, there have been occasions where consultants have failed to deliver services which allow sites to come into compliance with their legal duties. The relationships between site operators, consultants and the competent authority charged with enforcing the regulations needs to be better understood if the needs of all three parties are to be met. This workshop aims to explore these relationships and compile data from the discussions for future papers.

Background

The introduction of the Control of Major Accident Hazards (COMAH) regulations in 1999 raised the profile of human factors within the major hazards sector in the UK. One of the requirements of the COMAH regulations is for sites with more than the prescribed inventory of hazardous substances (known as top tier sites) to complete a COMAH safety report which sets out what their particular hazards are and how they maintain control of the risks posed by those hazards. The enforcement of the COMAH regulations is carried out by the UK competent authority which comprises the Health and Safety Executive (HSE) and the Environment Agency (EA). The competent authority has legal duties to report back to sites their comments on the COMAH safety report and follow up with site inspections. The Health and Safety Executive has published guidance on how to complete COMAH safety reports (HSG190) and an online safety report assessment manual (SRAM) (available via the HSE web site human factors pages). The SRAM contains a section on human factors and sets out the criteria used by the competent authority in the assessment of human factors.

Few, if any, of the top tier COMAH sites in the UK employ full time, qualified ergonomists/human factors practitioners. Therefore, it is common for sites to buy in their expertise from consultants. The production of a COMAH safety report is no small task so it is common for some or all of the sections of the report to

be written by consultants. Where consultants work closely with sites this can be an effective way of completing the report at a reasonable cost in a reasonable time frame. However, experience has shown that the quality of input from consultants can vary considerably. In some cases, even with the 'help' of consultants, the COMAH safety report fails to meet all the criteria. We are not suggesting that this is all the fault of the consultants. The sites need to understand what they will need to buy in and work in partnership with the consultants to ensure they get the right product. Ultimately, the legal duty falls to the site operator.

It is not just work on COMAH safety reports that is contracted out. Much of the human factors work that sites are required to carry out results from either planned inspections or incident investigations by the competent authority. In some cases these interventions result in enforcement action being taken. The competent authority has the power to serve improvement notices which place a legal requirement on the site to carry out specific work by a set date to achieve compliance with specified legislation and/or regulations. Some of the human factors issues addressed in improvement notices have included human factors in accident investigation, human factors risk assessment and competence assurance (Jackson et al, 2007). Faced with legal deadlines, many sites rightly seek expert input from human factors consultants. Again, experience has shown that without a good understanding of what the site requires and what the consultant is providing, the work carried out may not fully meet the requirements of the improvement notice. Failure to comply with an improvement notice by the specified date is in itself an offence under the Health and Safety at Work etc Act 1974.

In summary, companies operating within the major hazards sector in the UK have a real need to employ external human factors consultants if they are to fulfil their legal duties, particularly those covered by the COMAH regulations. The experience of specialist human factors inspectors in the HSE Hazardous Installations Directorate (HID) Onshore Human and Organisational Factors team is that, on occasions, the work carried out by consultants fails to deliver what is required for the sites to fully meet their legal requirements. This has raised some questions within the HID Onshore team: Do the consultants understand what the real needs of major hazard companies are? Do the major hazard sites understand what it is they need? What advice should the competent authority be providing to help companies and consultants understand what is required to ensure sites are able to meet their legal duties? These experiences and questions have prompted the workshop to be presented at the 2008 Ergonomics Society Annual Conference.

Workshop outline

The workshop will be facilitated by experienced human factors health and safety inspectors from the UK competent authority (HSE). Issues arising from human factors support provided to COMAH sites will be illustrated through case studies. This should open up points for discussion amongst workshop attendees. The basic framework for these discussions will be to consider how best to satisfy the needs of

the customers, the consultants and the competent authority (the law). The workshop aims are as follows:

- To discuss the type of support provided to the oil, gas and chemical industry by ergonomics/human factors consultants in relation to COMAH regulations.
- To discuss the expectations of the competent authority charged with the inspection of human factors issues relating to compliance with COMAH legal duties.
- To discuss the potential legal implications of support provided to the oil, gas and chemical industry by ergonomics/human factors consultants in relation to companies legal duties under the COMAH regulations.
- To explore ways of satisfying the needs of customers, consultants and the competent authority.

Brief outline of case studies

Case study 1: Top tier COMAH site safety report

As outlined above, top tier COMAH sites are required to produce a safety report stating what their hazards are and how they control the risks. This includes how they assure the performance of humans engaged in safety critical activities which may initiate or propagate a major accident. When assessing the safety report the human factors specialist inspectors in the HSE are generally looking for the report to demonstrate a structured, systematic approach to the management of human performance. Like many, the site in this case study employed the services of human factors consultancies. The consultants had carried out several studies/projects to provide data for inclusion in the safety report. Some assistance had also been given in producing the report itself. The report contained some good examples of a structured approach to human factors issues such as fatigue management and the determination of adequate staffing levels for safety critical operations. There was also evidence of some pilot work on safety critical procedures. In these examples there were demonstrations of well structured approaches. However, the report failed to demonstrate how these techniques were embedded in the overall safety management systems and that they were systematically applied across the various units on the site. Later inspections at the site concluded that what the consultants had provided was akin to hiring out a chauffer driven limousine when what the site really needed was a more sensible mode of transport and a driving instructor. The question is, was the human factors service provided fit for purpose? As the safety report failed to fully meet the human factors criteria the answer is no. The site was left with insufficient knowledge and confidence to take the work forward and demonstrate a structured, systematic approach to human factors.

Case study 2: Human factors risk assessment

As part of the requirements to comply with an improvement notice served on a top tier COMAH site, task analysis and human error analysis was required on several

safety critical tasks with a major hazard potential. The site had no human factors expertise in house so they brought in human factors consultants to help them. The onsite team, consisting of process engineers and health and safety managers, worked with the consultants on the initial task analysis and human error analysis. The consultants subsequently left, leaving the site to complete the work, having shown them the way to do it. When the HSE human factors inspector assessed the final results it was very obvious at which point the consultants had left. The consultants had used a relatively complex hierarchical task analysis technique and human error assessment technique. What they had not done was provide the site with a practical vehicle by which they could complete the task required of them. Another example of the chauffeured limousine approach. Following further advice from the HSE inspector the site adapted the techniques to suite their purpose and they now have a structured, systematic approach to human factors risk assessment which is part of the overall safety management system.

Case study 3: Human factors risk assessment DIY

Following assessment of their COMAH safety report, it was clear that this site had not addressed human factors in their major hazard control measures. During the feedback on the report assessment at the site they admitted to a lack of knowledge in human factors but they demonstrated a willingness to learn. It was decided that to help the site comply with their legal duties under the COMAH regulations, HSE specialist human factors inspectors would give the site a short tutorial based around the human factors guidance published by HSE, starting with HSG48 – Reducing error and influencing behaviour, and the material published on the HSE human factors web pages. The site committed to a sensible timescale to complete the work thus avoiding the need for enforcement action (improvement notice). The site went on to use the HSE material as a distance learning package and they eventually produced a structured approach to human factors risk assessment which they were able to systematically apply to all their safety critical activities. The results were achieved without any input from external human factors consultants. The results were so impressive that the HSE has encouraged the site to share their techniques with other sites in similar industry sectors. No limousine in this case, more of a home made kit car project!

Points for discussion

There are three main parties involved in fulfilling the human factors aspects of the legal duties set out in the COMAH regulations; the site operators, human factors consultants and the UK competent authority (HSE & EA). Each of these parties has their own basic needs. Both the site operators and the consultants have obvious business needs. If their business is not successful, they will cease to be. The site operators and the competent authority need to fulfil their legal duties under COMAH. If the site operator fails in their legal duties they will face enforcement action by the competent authority. Even worse, if the site operator fails to control the risks posed by their major hazards, they may suffer the kind of incident which would

destroy their business and have very serious consequences for the surrounding communities and environment. It is to the benefit of everyone for these three parties to foster an effective working relationship. One point for discussion is what form should these relationships take?

It would be of benefit to sites to become intelligent customers for human factors. The question is how much in house intelligence do they need and when is it appropriate to bring in specialist consultants? It is unthinkable that a major hazard site such as an oil refinery could operate without qualified process engineers on their staff. However, almost all operate without any qualified ergonomists/human factors specialists on staff.

What level of expertise should human factors consultants be supplying to major hazard sites? How much responsibility should the consultant take on board for helping to fulfil the sites legal duties? Consultancies need to maintain business at these sites in order to survive. It is important for consultants who can supply an appropriate level of service to be available to industry and this importance is clearly recognised by the competent authority. However, consultants need to identify the appropriate vehicle for the job in hand. There may be occasions where deep topic specialism is required e.g. quantified human error assessments in making ALARP demonstrations (limousine?). There is also a need for consultants to help bring sites up to speed with more routine human factors applications, the driving instructor approach rather than the chauffer.

Post workshop

The intention is to stimulate discussion amongst the three main parties at the workshop and record the comments and ideas. It is hoped that the data gathered through this exercise can be collated and analysed to be published in a future paper as a follow up to this workshop.

References

HSE, 2005, L111: A guide to the Control of Major Accident Hazards Regulations 1999 as amended, HSE Books 2005.

HSE, 1999, HSG190: Preparing safety reports, HSE Books 1999.

HSE, 1999, Reducing error and influencing behaviour (HSG48), HSE Books 1999, ISBN 0 7176 2452 8.

HSE, 2002, Developing and maintaining staff competence (Railway Safety Principles and Guidance Part 3 Section A), ISBN 0 1 7176 1732 7, HSE Books 2002.

HSE, 2003, Competence assessment for the hazardous industries. RR086/2003. http://www.hse.gov.uk/research/rrhtm/rr086.htm.

HSE Human Factors web site, http://www.hse.gov.uk/humanfactors/comah/index.htm.

Jackson, et al, 2007, Competence Assurance in the Major Hazards Sector: A Regulators Viewpoint, Contemporary Ergonomics 2007, Taylor & Francis.

CURRENT TRENDS IN HUMAN RELIABILITY ASSESSMENT

Barry Kirwan[1] & Huw Gibson[2]

[1] *Eurocontrol Experimental Centre, Bois des Bordes BP15, Bretigny/Orge, F-91222 CEDEX, France*
[2] *School of Electronic, Electrical and Computer Engineering, The University of Birmingham, Edgbaston, Birmingham, West Midlands*

Human Reliability Assessment (HRA) is more than four decades old, and this paper seeks to review current trends in its use in applied settings. It focuses on nuclear power, rail and air traffic industries, with a more cursory look at space exploration, aviation and medical risk. Three trends are discussed: a shift towards data-informed approaches; a preference for so-called 1st Generation approaches, whilst 2nd Generation techniques remain under development; and a debate over the relative utility of quantitative and qualitative approaches in safety cases. The conclusion is that whilst debate continues over HRA's validity and 'simplification' of human error contexts, people in industry still find it helps in the assessment and management of human-related aspects of safety.

HRA in nuclear power

Nuclear power is often seen as the birthplace of HRA, although its true origins were in the defense domain in the US in the early 1960s. However, the first 'real' HRA technique was undoubtedly THERP (Technique for Human Error Rate Prediction (THERP: see Kirwan [2007] for a review of all the HRA techniques listed in this section), which was fast-tracked for release following the Three Mile Island (TMI) nuclear power accident in 1979. Nuclear Power has remained the ultimate 'testing ground' for emerging HRA techniques, and has seen a plethora of techniques and tools over nearly three decades since TMI. Several other techniques arose in the 1980s, which might be considered the 'boom years' of HRA, such as SLIM (Success Likelihood Index Method) and HEART (Human Error Assessment and Reduction Technique), for example, which (together with THERP) are still in use today. These '1st Generation' HRA techniques generally predict human error probabilities based on an initial nominal error probability modified by a set of performance shaping factors, seen as the dominant Human Factors affecting human error in industrial contexts.

In 1990 a prominent HRA expert (Dougherty, 1990) suggested, based on the experiences of incidents/accidents in the nuclear power area, that most HRA approaches did not pay enough attention to context, i.e. to the detailed scenarios people found themselves in. Essentially, the argument was that considering human

reliability in an abstracted format of fault and event trees (which still dominate safety case approaches) was insufficient to capture the local situational factors that would actually dictate human behaviour, and lead to success or failure, to recovery or disaster.

Therefore work on a set of so-called 'Second Generation HRA' techniques began in the early – mid 90's. The most notable of these were 'A Technique for Human Error Analysis (ATHEANA) and the Cognitive Reliability Error Analysis Method (CREAM). Actually used in various nuclear safety assessments were also MERMOS and CAHR (Connectionist Approach to Human Reliability). ATHEANA is notable because it has had more investment than almost any other HRA technique. The qualitative part of the method is used for identification of safety critical human interventions in several instances, but its use has so far been marginal due to residual problems over quantification. CREAM is a more straightforward technique that has had mixed reviews, although currently it is in use in the nuclear power domain in the Czech Republic.

Recently the HEART technique has been re-vamped using human error data collected over a ten year period in the CORE-DATA database, to develop a new nuclear-specific HRA technique called NARA (Nuclear Action Reliability Assessment). NARA has been successfully peer reviewed by the nuclear regulator and industry, and independent HRA experts in a formal peer review process. It has been seen as a highly transparent use of human performance data to derive a HRA approach. It is currently awaiting regulatory approval for usage in the UK nuclear power arena.

In the US a two-pronged approach is being pursued – a 1st generation approach called SPAR-H (Standardised Plant Analysis Risk HRA Method) which has some similarities with NARA, and is in very active use in the US, and the ATHEANA approach, which is not yet activated by the regulatory authorities – partly because of concerns over consistency of usage/results (the approach is quite complex) and over-simplicity in its quantification approach.

Currently methods in use are, from 1st generation, HEART, SLIM, APJ and THERP and from 2nd generation ATHEANA, MERMOS, CREAM, and CAHR. An interesting development is a set of studies ongoing in the Halden Reactor Project in Norway, firstly to try and underpin HRA approaches using simulation data (whether 1st or 2nd Generation), and secondly a large-scale 'Benchmark' study contrasting a set of different techniques against a set of nuclear power scenarios.

Whilst there is continued interest in 2nd Generation approaches in the area of HRA development, in practice in nuclear power, '1st Generation' techniques are the ones that are mainly being applied in real risk assessment and safety assurance work.

Railways

Industry-specific regulatory requirements are frequently seen as a key motivator for use of HRA within an industry, in particular the use of human error quantification. Human reliability assessment is not generally mandated as a regulatory requirement for railway operations. However, both quantitative and qualitative human reliability

assessments are used as part of railway human factors and safety studies (e.g. see Bell et al, 2005).

The development of human error probability data is one indicator of the extent to which quantified risk assessments incorporate human reliability assessment. Examples of data collection efforts for railways include:

- The collection of driver and signaller communication error probabilities (Gibson et al, 2006).
- Operator errors in Yugoslav Railway control centres have been quantified using the SLIM technique (Grozdanovic, 2005).
- Wreathall et al (2005) have quantified errors as part of an assessment of a train monitoring and control system.
- Gibson et al (2005) have collected human error probability data on train driver fault diagnosis using train cab simulators.
- Absolute probability judgement sessions have been used to quantify a range of train driver errors and violations (Rail Safety and Standards Board, 2004).

In common with other industries it is not always possible to collect HEP data as the basis for studies and human error quantification techniques are used. The most commonly used quantification technique is HEART. For example, Reinach et al (2006) have applied the HEART technique as part of a comparative risk assessment of existing yard switching operations and remote control locomotive operations in the United States.

However, there are indications that the HEART technique requires adaptation to the railway context and, for example, Reinach et al (op cit) deemed it necessary to support their HEART assessments with expert opinion (Absolute Probability Judgement) sessions. A rail-specific human reliability assessment technique is under development in the UK and sponsored by the Rail Safety and Standards Board (Rail Safety and Standards Board, 2004). An indication of further developments in the context of HEART was the presentation at a railway conference of data relating to the quantification of age as an error producing condition to be used within a HEART type approach (Bell et al, 2005).

Human error quantification is therefore being applied in the railway context, but it should be noted that non-probabilistic approaches are also being developed. For example, 'Analysis of Consequences of Human Unreliability' (ACIH) has been developed by Vanderhaegen (2001) and considers the wide context of railway reliability. Also, looking at second generation type HRA approaches, Cacciabue (2005), has applied the 'Human Error Risk Management for Engineering Systems' (HERMES) approach in the railway domain and this technique seeks to consider wider management factors which impact on systems reliability.

In summary, HRA forms a key part of railway Human Factors. That HRA is a live issue in the railways can be seen not only in the use/adoption of existing HRA methods for railway assessments but also in the collection of railway-specific quantification data, and adaptation/development of railway-specific HRA methods.

Air traffic management

Air Traffic Management (ATM) in Europe has to demonstrate that it is acceptably safe, against an overall target level of safety (1.55 accidents per 10^8 aircraft flight hours). This means that the overall risk management framework has a quantified output. Since human error is a key contributor to both risk and safety, work is ongoing (Kirwan, op cit) to ensure that human reliability is a key component of the ATM 'risk equation'. This means that key human error risks, whether induced by human error, or representing air traffic controllers failing to control system or environment-induced risks, can be identified, evaluated, prioritised and mitigated (i.e. managed). The approach of Eurocontrol, a European organisation concerned with the safety of European air navigation, is to develop a quantified HRA approach, which is currently being tested in several safety case applications. The CARA (Controller Action Reliability Assessment) approach is a first generation approach following in the footsteps of HEART and NARA, and is based on human error data from the Human Factors literature and real-time simulation studies. It is intended that a second generation HRA approach will be developed by the 2010 timeframe (based on CAHR), to be used for more detailed studies.

A cursory look at space exploration, aviation and medical risk

These three industries are relevant to HRA because human error plays a significant role in their incidents/accidents, though in space exploration this has largely to date been at the 'blunt end' (management) rather than at the sharp end (operators, i.e. astronauts) – yet none of them has a definitive HRA approach. All of them however manage human-related risks.

NASA is currently considering development of a HRA method for space missions to Mars. Towards that end, a review was carried out (Mosleh et al, 2006) of a dozen existing mature HRA techniques. The following were selected as candidates: NARA, HEART, SPAR-H, CREAM and ATHEANA. What is perhaps of interest to note here is that NARA, HEART and SPAR-H are 1st Generation techniques.

Aviation has a wealth of literature on cockpit errors and Human Factors, but does not boast a HRA culture. Instead, there is on the one hand a strong research culture on Human Factors in incident and accident analyses, and a corresponding focus on Human Factors by the airframe manufacturers (e.g. Boeing and Airbus). Airframes (new aircraft) are certified by competent authorities, and so Human Factors aspects must be demonstrated to be adequate. This risk management paradigm is therefore different from a quantified risk assessment paradigm, which is dominated by a numerical target level of safety (as in nuclear power and ATM, for example).

Medical risk is a burgeoning topic, because of the high rate of deaths caused by medical error. A scan of the Human Factors studies published in this area certainly reveals a focus on human error, but suggests a dual qualitative approach of trying to understand why such errors are occurring (including studies of safety culture), and of developing approaches that can identify errors, their causes and appropriate mitigations. This may also reflect the way risk is managed in medicine: there is a

more localised approach. This means that individual medical units (e.g. hospitals) may identify task or situational factors that can contribute to risk, and then turn immediately to rectification, based on more straightforward prioritisation rather than detailed quantified risk assessment. HRA might therefore only be needed in more cross-cutting assessments of widespread procedures (e.g. laparoscopy), or if an equivalent medical regulatory authority wanted to have a 'top-down' assessment and prioritisation of human error risks. Another reason quantified HRA may be less necessary in medical is that in other industries, there is a strong interplay between human and other system elements (e.g. hardware and software) – much of the time HRA is focusing on how the human copes with a system failure. In medical risk, human error often leads to the fatal event on its own. There is therefore less need to consider non-human elements in the 'risk equation'.

Disclaimer

This paper expresses the opinions of the authors and not necessarily those of parent or affiliate organisations.

References

Bell, J.L., Williams, J.C., Daniels, C. and Warren, N.D. 2005, Improving the predictive power of an existing human reliability assessment method. In *Proceedings of the Second European Conference on Rail Human Factors*, London, 21–23 November 2005. No page numbers.

Cacciabue, P.C. 2005, Human error risk management methodology for safety audit of a large railway organisation, *Applied Ergonomics*, 36, 709–718.

Dougherty E.M., 1990, Human Reliability Analysis – Where Shouldst Thou Turn? Reliability Engineering and System Safety, **29**, 283–299.

Gibson, W. H., Halliday, M.W., Sutton, L., Shelton, J. and Bond, M.A. 2005, Train driving simulator experiment to investigate driver fault diagnosis. In *Proceedings of the Second European Conference on Rail Human Factors*, London, 21–23 November 2005. No page numbers.

Gibson, W.H., Megaw, E.D., Young, M. and Lowe, E. A. 2006, Taxonomy of human communication errors and application to railway track maintenance, *Cognition, Technology and Work*, 8, 57–66.

Grozdanovic, M. 2005, Usage of human reliability quantification methods, *International Journal of Occupational Safety and Ergonomics*, 11(2), 153–159.

Kirwan, B. 2007, Technical Basis for a Human Reliability Assessment Capability for Air Traffic Safety Management, EUROCONTROL Experimental Centre, EEC Note No. 03/07, April. http://www.eurocontrol.int/eec/public/standard_page/eec_notes_2007.html.

Mosleh, A. et al. 2006, (Draft) Evaluation of current HRA methods for NASA use. NASA HRA Methods Workshop, Kennedy Space Centre, Florida, Jan 16–18.

Rail Safety and Standards Board 2004, *Rail-Specific HRA Tool for Driving Tasks T270 Deliverable 4 – APJ Workshops Summary Report* (RSSB, London).

Reinach, S., Fadden, S., Gamst, F., Acton, S. and Bartlett, S. 2006, *A Comparative Risk Assessment of Remote Control Locomotive Operations versus Conventional Yard Switching Operations* (U.S. Department of Transportation, Washington D.C.).

Vanderhaegen, F. 2001, A non-probabilistic prospective and retrospective human reliability analysis method – application to railway system, *Reliability Engineering and System Safety* 71, 1–13.

Wreathall, J., Bley, D., Roth, E., Multer, J. and Raslear, T. 2004, Using an integrated process of data and modeling in HRA, *Reliability Engineering and System Safety* 83(2), 221–228.

NUCLEAR ACTION RELIABILITY ASSESSMENT (NARA), FURTHER DEVELOPMENT OF A DATA-BASED HRA TOOL

W. Huw Gibson[1], Barry Kirwan[2], Richard Kennedy[3], Jim Edmunds[3] & Ian Umbers[4]

[1]*University of Birmingham, B15 2TT, UK*
[2]*Eurocontrol, BP 15, F-91222 Bretigny/Orge, France*
[3]*Corporate Risk Associates, Leatherhead, Surrey, UK*
[4]*British Energy, Barnwood, Gloucester, UK*

Nuclear Action Reliability Assessment (NARA) is a tool for quantifying human reliability. NARA uses generic task types and associated human error probabilities as one part of the quantification process. A key feature of NARA is that the human error probabilities are underpinned by human performance data. This paper describes the methodology and results for development of the NARA generic task type human error probabilities.

Introduction

In the UK, the principal tool used to quantify the reliabilities of human interactions in Nuclear Power Plant (NPP) Probabilistic Safety Assessments (PSAs) has been the Human Error Assessment and Reduction Technique (HEART: Williams, 1992). Although HEART has undergone some development since its inception, in principle, it has remained the same technique, based on the same original data. Meanwhile, additional sources of human error probability data have become available. In particular, effort has been expended in establishing the human error probability database COREDATA (Computerised Operator Reliability and Error Database: Gibson et al, 2006). Additionally, an internal industry review of the application of HEART in actual PSAs revealed some limitations of the technique, in particular that HEART did not always 'fit' very well with the NPP tasks being assessed. It was therefore felt, by users of the technique in the UK nuclear industry, that a new tool could be developed. This new tool is known as NARA (Nuclear Action Reliability Assessment).

NARA development has been funded by British Energy. NARA is based around the elements of the HEART approach but tailored specifically to the nuclear context and to the data available to underpin the technique. The HEART approach is summarised in Kirwan (1994) and an initial overview of NARA is presented in Kirwan et al (2004). The key elements of both HEART and NARA are:

- **Generic Task Types (GTTs).** A task requiring quantification is assigned a nominal human unreliability by classifying it into one of a number of generic task

types. NARA GTTs have been developed based on a review of NPP PSAs and human performance models. This paper focuses on the development of the NARA GTTs and the derivation of their human error probabilities (HEPs).

- **Error Producing Conditions (EPCs)**. For each task being assessed, the factors negatively influencing human performance, compared to the 'ideal' conditions associated with the GTTs, are identified based on a set of EPCs. A revised set of EPCs has been developed for NARA, again based on NPP PSA requirements and reviews of other performance shaping factor taxonomies. In addition, estimates have been provided for the EPC 'maximum affects' i.e. the maximum predicted increase in the nominal GTT HEP which could result from the adverse performance shaping factors. Further efforts are currently ongoing to underpin the maximum affect values with data from the experimental literature available through the Ergonomics Information Analysis Centre.
- **Assessed proportion of affect (APOA)**. For each identified EPC, the analyst makes a judgement on the strength of the affect that the EPC has on successful task performance. This is known as the assessed proportion of affect in both HEART and NARA and has a numerical value between 0.1 and 1, where 0.1 represents a very weak effect and 1 represents a full affect. NARA provides guidance for determining the appropriate APOA.
- **Calculation Method**. Both HEART and NARA use a simple calculation method to combine the GTT HEPs with the EPC/APOA values. Examples of this calculation are presented in Kirwan (1994).

A long standing criticism of most human error quantification techniques is the absence of an explicit link between the HEPs used and the available HEP data which have been collected in actual studies of human performance. Without this linkage, and despite validation studies (e.g. Kirwan et al, 1997), there has been some lack of confidence within the human factors community that the derived human reliability estimates reflect actual human performance reliabilities which can be expected when tasks are carried out in the real world. This paper focuses on how NARA addresses these concerns by underpinning the GTT HEPs used with reported human performance data.

Quantification of generic tasks types using data

For NARA, the GTT HEP values are explicitly underpinned by HEP data which have been collected through accepted methods in applied settings. These underpinning data were selected from COREDATA, a database of error probability data and associated background information (Gibson et al, 2006). Some additional data sources, not contained within COREDATA, were available to certain of the NARA team members, and these were also used for the NARA development. COREDATA contains data collected in the following industries: Nuclear, Offshore, Manufacturing, Railway, Chemical and Aviation. The data have been collected using a number of approaches, which are all based on looking at the real-world context in which errors occur. The approaches include: direct observation/recording; simulator observation/recording; incident data; and expert judgement. It should be

Table 1. Data Quality Scoring Criteria.

Dimension	Rating	Description
Data Type	3	Field data; incident records; simulator data; training records.
	2	Formal expert judgement data (using formal expert judgement procedures and statistical aggregation processed); experimental data.
	1	Synthetic data.
Collection Process/ Method	3	Formal and robust data collection processes.
	2	Good incident recording system; semi-controlled study (informal experiments); transparent process of making data generic; convergence of data (synthetic).
	1	Possible under-reporting of incident rate; informal study.
Statistical Integrity	3	Large sample, large denominator and relatively large numerator.
	2	Medium – large denominator, but small numerator.
	1	Very low denominator; non-statistical aggregation process; engineered judgement of actual data, e.g. for synthetic techniques such as CREAM or JHEDI.
Relevance	3	NPP HRA task-relevant.
	2	Universal behaviour less affected by context, e.g. some skill-based tasks such as reading, communicating, etc.; high task relevance but different context, i.e. not NPP or process control.
	1	Non-representative sample population, e.g. students, in task affected significantly by training/experience; task not relevant to NPP context.

noted that for expert judgement, these are judgements made by task experts (e.g. using nuclear power plant operators to estimate error likelihoods) using structured techniques such as absolute probability judgement or paired comparisons. COREDATA does not simply contain these probabilities, but a wide range of supporting information, which allows a human error probability to be understood in both its practical and methodological context. The COREDATA database therefore contains text fields and taxonomies which aim to capture: error types; task context; data collection methodology; and numerical data underpinning the human error probability

The available data from COREDATA and additional sources have been used to develop the GTT HEPs using a process of screening and averaging. Generally, two types of criteria can be applied when screening data – 'satisficing' and 'relative'. Satisficing criteria ask whether the data are good enough for initial consideration. They are binary in nature, i.e. the data are deemed to be either suitable or not suitable. This represents a 'coarse screening' process. Relative criteria ask how good the data are for the purpose, representing a 'fine screening' process.

For NARA, satisficing criteria were applied as a first pass to exclude data which were clearly not applicable, primarily because they did not match the description of the GTT closely enough. This first pass naturally led to the accepted datapoints being associated with specific GTTs. Ideally, a number of datapoints would be associated with each GTT which could then be taken forward to the relative screening step.

Table 2. NARA GTTs.

ID	GTT Description	HEP	95% t-confidence
Task Execution			
A1	Carry out a simple single manual action with feedback. Skill-based and therefore not necessarily with procedures.	0.005	0.002–0.01
A2	Start or reconfigure a system from the Main Control Room following procedures, with feedback. The procedures may require some diagnosis of alarms/indications before the need for the action is recognised.	0.001	N/A (only 3 data points)
A3	Start or reconfigure a system from a local control panel following procedures, with feedback.	0.002	0.0007–0.006
A4	Judgement needed for appropriate procedure to be followed, based on interpretation of a situation which is covered by training at appropriate intervals.	0.006	N/A (only 3 data points)
A5	Completely familiar, well designed highly practised, routine task performed to highest possible standards by highly motivated, highly trained and experienced person, totally aware of implications of failure, with time to correct potential errors.	0.0001	0.000004–0.002
Ensuring correct plant status and availability of plant resources			
B1	Routine check of plant status.	0.02	0.003–0.2
B2	Restore a single train of a system to correct operational status after a test, following procedures.	0.004	0.0008–0.02
B3	Set system status as part of routine operations using strict administratively controlled procedures, e.g. top up tank to correct level.	0.0007	N/A (only 3 data points)
B4	Calibrate plant equipment using procedures, e.g. adjust set-point.	0.003	0.0003–0.03
B5	Carry out analysis.	0.03	N/A (only 1 data point)
Alarm/Indication Response			
C1	Simple response to a range of alarms/indications providing clear indication of situation (simple diagnosis required). Response might be direct execution of simple actions or initiating other actions separately assessed.	0.0004	N/A (only 1 data point)
C2	Identification of situation requiring interpretation of complex pattern of alarms/indications.	0.2	0.15 – 0.33
Communication			
D1	Verbal communication of safety-critical data.	0.006	0.002–0.009

Relative fine screening was accomplished by reviewing the data using a quality scoring method. The scoring method was based on a set of explicit criteria for judging data quality so that numerical quality ratings could be consistently assigned. This aimed to provide an explicit link between specific datapoint judgements and the subsequent use of the datapoints to underpin the NARA GTT HEPs. The key

elements of the quality scoring were based on the dimensions presented in Table 1, where higher values represent preferred quality attributes.

The quality ratings were used to further screen-out datapoints which did not pass a defined numerical quality criteria. The data review process also included a review of any EPCs associated with each of the data points. This review aimed to ensure that the data used to underpin a GTT were free of EPC elements (other than those inherent in the GTT description). The number of available data points associated with a GTT determined how the final GTT HEP was derived. In the majority of cases, an averaging process using the geometric mean was applied. For the remainder, a degree of judgement was used. In addition, the data underpinning a GTT have been explored from a statistical perspective where appropriate, in particular identifying the 95 percent confidence interval where possible.

The key element of this process, even when a single datapoint has been relied on to provide a GTT HEP, is that the relationship between the GTT HEP and the datapoint(s) are explicit. In addition, the rationale for selecting datapoints and the means of assigning quality scores to them is open to scrutiny. This means that the critical link between data and the GTT human error probability has been made explicit. Although this approach could be criticised as relying on subjective judgements, the impact of this has been reduced by: (a) the review being scrutinised by the NARA project team, who are all HRA/PSA practitioners having experience in the nuclear context, and (b) making the judgements and quality criteria explicit such that they are open to wider scrutiny.

Another benefit of this approach is that in the future, as more data become available, these can be added to the NARA dataset and used to develop increased confidence in the GTT values. Table 2 presents the full set of NARA GTTs. It should be noted that the values are for information only as the project is currently undergoing final review. The NARA technique is fully described in the Technical Basis report which summarises the GTT quantification process in detail, Kirwan et al (2007).

Conclusions

This paper has outlined the methodology used in developing the GTT human error probabilities used for the NARA human error quantification technique and the results have been summarised. The most important feature of this development has been the use of data from the COREDATA database of human error probabilities to develop explicit links between the underlying performance data and the resultant GTT human error probabilities.

Other key developments relating to the NARA methodology (not all discussed in this paper) have been: an international peer review process, supported by the Nuclear Installations Inspectorate; specification of maximum affect values for EPCs using performance literature; development of guidance related to Errors of Commission, applying assessed proportions of affect and taking account of operator dependencies; and production of a user manual.

References

Gibson, W.H., Hickling, B. and Kirwan, B. 2006, *Feasibility Study Into The Collection Of Human Error Probability Data*, EUROCONTROL: Bretigny-Sur-Orge, France.

Kirwan, B., Kennedy, R., Taylor-Adams, S. et al. 1997, 'The Validation of Three Human Reliability Quantification Techniques – THERP, HEART and JHEDI: Part II – Results of Validation Exercise *Applied Ergonomic*s (28)1 pp. 17–25.

Kirwan, B., 1994, *A Practical Guide to Human Reliability Assessment* Taylor and Francis: London.

Kirwan, B., Gibson, H., Kennedy, R., Edmunds, J., Cooksley, G. and Umbers, I. 2004, Nuclear Action Reliability Assessment (NARA): A data-based HRA tool. In Spitzer, C., Schmocker, U. and Dang, V.N. (Eds.) *Probabilistic Safety Assessment and Management 2004*, London, Springer, pp. 1206–1211.

Kirwan, B., Gibson, H., Edmunds, J. and Kennedy, R. 2007, A Method for Human Error Quantification based on an Update of HEART. Report CRA-BEGL-POW-J032 R1, Issue 5, April 2007.

Williams, J. C. 1992, *A User Manual for the HEART Human Reliability Assessment Method*, Technica Report C2547, March 1992.

RECURRING FACTORS IN REPORTABLE EVENTS OF NUCLEAR POWER PLANTS

Mandy Richter & Tom-Oliver Solisch

EnergieSystemeNord GmbH, Germany

On behalf of the nuclear regulatory authorities we assess reportable events in German nuclear power plants, in which human and organisational aspects played a major role for the event genesis. As a result of our assessment we have identified certain event-relevant factors (initiating, causal or contributory) which occurred repeatedly in different events and settings respectively. This raises the question regarding common event-relevant factors occurrence patterns and the effectiveness of the licensee's previously realised countermeasures. This paper presents our assessment method as well as the results of a meta-analysis we have conducted to investigate the apparent recurrence of similar event-relevant factors in different events/settings.

Introduction

Events occurring in German nuclear facilities have to be reported by the licensee of the facility to the nuclear regulatory authorities according to the current version of the federal reporting criteria. The Event Report comprises information about the causes for the event, the course of the event, the impacts as well as the presently implemented and furthermore planned countermeasures. The nuclear supervisory authority assesses the report submitted by the operator of the nuclear facility. In individual cases the regulatory authority assigns independent experts with this assessment. If necessary, the supervisory authority orders remedial actions which have to be taken by the licensee.

It must be pointed out that the assessment of reportable events is subject to certain basic conditions, e.g. the licensee is responsible for the event analysis and the derivation and realisation of countermeasures. It is essential that an independent assessment of the event does not undermine this principle. Therefore this external evaluation applies to the adequacy of the licensee's event analysis and includes the verification that the licensee's analysis identified the appropriate deficiencies associated with the event and that the licensee has initiated reasonable and comprehensible corrective actions.

EnergieSystemeNord GmbH (ESN) works as an independent expert on behalf of the nuclear regulatory authorities, primarily for Human Factor Events (Sträter, 2000, distinguishes between Technical Events, Human Factor Relevant Events and Human Factor Events). As it is demanded in German nuclear power plants that all relevant operational and safety related procedures and plant regulations are described in an

instruction manual, the focal points in our assessment of reportable events are the suitability of and the compliance with these regulations and procedures.

Assessment of reportable events

The assessment of reportable events by ESN pursues a holistic approach, which means that all aspects of the socio-technological system are taken into account. As mentioned earlier, the evaluation is carried out on basis of information, which is provided by the licensee. A multidisciplinary team sifts, utilises and assesses this information in the following steps:

1. Reconstruction of the chain of events (actual course)
2. Identification of all event-related actions/processes and deduction of the nominal course on basis of given regulations and procedures for this actions/processes and expert knowledge (nominal course includes human, organizational and technical sequences).
3. Comparison of the nominal and actual course with the target to identify all deviations
4. Analysis of the identified deviations: Following the "Swiss Cheese" model of human error by James Reason (1990) we examine in this step, which factors (latent failures) have led to these deviations (active failures). The focus is here not on individual operator performance (unsafe acts), but rather on the performance of the organizational system which gives the context to the individual's action (NEA, 2006). This step also includes an assessment of the suitability of the nominal course. The identified factors are then assessed in respect of their relevance for the event (initiating, causal and contributing factors).
5. Assessment of the licensee's previously implemented and furthermore planned countermeasures: On basis of the causal factors identified by the licensee, the results of the nominal-actual-course comparison (step 3) and the additional identified event-relevant factors where required (step 4) we assess in this step the adequateness of the licensee's proposed remedial actions. Where required we deduce and formulate additional remedial actions (in form of recommendations).

With the described method we have assessed so far nearly 50 reportable events over the last 10 years. Beyond the causal factors examined by the licensee in our evaluations we have in large part identified more and/or other causal and contributing factors for the event under observation, primarily in the range "suitability of given work orders" and "compliance with given regulations and procedures". For these additional factors we have recommended the realisation of supplementary remedial actions (e.g. implementation and improvement of trainings, adaptation and concretion of specific regulations and procedures).

Furthermore in recent evaluations we have found event-relevant factors which recur in similar as well as in different events and settings. This raises the question regarding common event-relevant factors occurrence patterns and the effectiveness of the licensee's realised countermeasures. Therefore we have decided to conduct further investigations in form of a meta-analysis.

Meta-analysis: methods and first results

Method

For the meta-analysis we selected reported events (INES classification 0) of the last ten years in one nuclear power plant, where the event was not explainable by only a technical cause (HF-events), and where the event assessment was performed by ESN (21 events selected). The data bases for the evaluation were the Event Reports of the nuclear power plant operator and our expert reports on these events.

Following our event assessment procedure the meta-analysis encompasses three focal points:

1. The identified deviations
2. The identified event-relevant factors
3. The previous countermeasures

To establish comparability between those 21 different events we choose a process-oriented procedure, which means we do not focus on the specific event but on the concerned processes within the event. As in this nuclear power plant a quality management system is implemented, every process should encompass the process phases "plan", "do", "check" and "act" (PDCA-Cycle, Deming, 1986), so we have taken this PDCA-Cycle as a framework for our analysis.

For achieving a standardised data basis for this meta-analysis we have firstly carried out a subsequent classification of the identified event-relevant factors based on the taxonomy HFACS (Shappell and Douglas, 2000), an IAEA-guideline (IAEA, 2001) and an inspection manual by the NRC (2000) while consolidating the individual assessment results. According to the given basic conditions for an external assessor (limited information accessibility), this taxonomy distinguishes between three different levels of event-relevant factors: unsafe acts (refers to all aspects of individual human performance), organisational failures and interorganisational failures. The latter level was included as it has become clear that there is an increased tendency to involve contractors and external personnel, which places great demands on the adequate design of interfaces involved (OECD, 2006).

Results

So far we have realised the first two steps of the meta-analysis in a large part, the third step is being carried out at present. As a result of the **first step** (meta-analysis of the identified deviations) 99 deviations were found over the 21 events under observation. The processes to which the most event-relevant factors could be assigned were "planning and realisation of modifications" (37 identified deviations), "planning and realisation of repair" (16), "planning and realisation of recurrent tests" (12) and "planning and realisation of maintenance" (9). Over these processes the aspects with the most frequently observed deviations were "planning and realisation of isolations" and "planning und realisation of acceptance and operational tests", whereas the former principally occurred as a causal factor and the latter primarily as a contributing factor for the genesis of the event.

In the analysis of the causal and contributing factors for these deviations (second step of the meta-analysis) we have summarised 87 event-relevant factors, which occurred 305 times within the 21 events. The most frequently identified event-relevant factors on the level "unsafe acts" were

- Given procedures, documents, drawings, or other references not (fully) applied,
- Execution of faulty/incorrect/erroneous orders,
- Independent checking applied ineffectively,
- Lack of questioning attitude,
- Maintenance performed incorrectly.
- On the level "organisational aspects" the most frequently observed factors were
- Procedure/document technically incorrect or incomplete,
- Non-conformance not identified/recognized,
- Training ineffective.

Interorganisational aspects, e.g. "inadequate control of contractors", have been of little importance within the 21 events involved.

With regard to the event-relevance of each identified factor we have found that 21% of factors can be described as "initiating" (primarily the factor "execution of faulty/incorrect/erroneous orders"), 29% can be classified as "causal" (the factor with the highest emphasis was "modification planned inadequately") and 50% operated as contributing factors (primarily the factors "independent checking applied ineffectively" and "non-conformance not identified/recognized").

As a result of the process-oriented consideration it can be stated that 32% of all identified event-relevant factors occurred within the process "planning and realisation of modifications", 17% within the process "planning and realisation of repair", 14% within the processes "planning and realisation of recurrent tests" and "planning and realisation of maintenance" and 7% within the process "planning and realisation of preventive maintenance". This allocation is similar to the one we have found within the analysis of the identified deviations (step 1).

Furthermore can be declared that 39% of the 305 mentions of event-relevant factors occurred during the planning phase of these activities (primarily concerned were the factors "independent checking applied ineffectively", "original design inadequate" and "lack of questioning attitude"). 47% of all factors could be assigned to the phase "do", the most relevant factors in this connection were "training ineffective" and "procedure/document technically incorrect or incomplete". The factor "non-conformance not identified/recognized" occurred mainly in the phase "check" to which altogether 10% of the identified factors can be assigned. No event-relevant factors were identified within the phase "act".

In the second step of this meta-analysis we have also investigated the connections between the identified event-relevant factors. Significant correlations between certain factors were found, e.g. the factor "no approved procedures/documents available" shows a significant correlation to "violated trainings rules" ($r = 0.81$), "unsafe working practices applied" ($r = 0.40$) and "training ineffective" ($r = 0.51$).

According to our assessment focus (compliance with and suitability of event-related regulations and procedures) we have also conducted a further investigation concerning the factor "given procedures, documents, drawings, or other references

not (fully) applied", which is the most frequent identified factor within this meta-analysis. In this context a graduation is possible and useful regarding the degree of non-compliance (compliance, formal compliance, inconsequent compliance, deficient compliance, non-compliance). For instance we have found that in 10% of all cases the given event-related regulations and procedures were complied with formally, but this adherence turned out to be ineffective, as errors in planned isolations or functional tests were not identified within the realised quality assurance (factor "independent checking applied ineffectively"). Furthermore we have found that in 55% of all cases event-related regulations and procedures were entirely not applied (non-compliance). The regulations and procedures concerned within these cases of non-compliance were primarily inspected and approved (by the regulatory authority) maintenance rules (32%), followed by regulations which lay within the individual responsibility of the licensee (21%).

With regard to the suitability of these event-related regulations and procedures it can be stated that in 58% of all cases involved the regulations and procedures could be classified as appropriate. It should also be noted that in cases where procedures were assessed as "inappropriate" (7%) or "erroneous" (5%) the procedures in question were the work orders which resulted from the planning phase (e.g. planned isolations or functional tests). These erroneous orders were in large part executed and led to the reported event.

Conclusions and next steps

Based upon the present results of the meta-analysis (which have been described exemplary in this paper) we have so far drawn several conclusions. First of all it can be said that there are different processes and aspects which have been more relevant for the genesis of events than others (e.g. planning and realisation of isolations and functional test during modifications and repair). In this context a pattern of event-relevant factors for instance refers to erroneous planning of isolations, the non-identification of this error during the independent checking and the execution of this erroneous orders which finally resulted in the event. In these cases the non- or rather inconsequent compliance with given regulations and procedures has been proven as a comprehensive event-relevant factor. The meta-analysis has also shown that most of the recurring event-relevant factors are not restricted to certain processes; instead they have played a role in the event genesis as a comprehensive factor (generic aspect). But also relevant is the finding that some event-relevant factors appeared only in certain process phases.

These results suggest that event specific remedial actions (aiming to avoid the recurrence of the same event) should be supplemented by generic/comprehensive improvement measures. This is an important point as it has major impacts on the assessment of measures deduced after a specific event. It should also be further examined whether different trainings/measures should be implemented for different personnel groups (e.g. planner, operator) which focus on the specific requirements of the different process phases.

These aspects will be taken into account while performing the third step of this meta-analysis (analysis of the licensee's remedial actions).

References

Deming, W. Edwards (1986). *Out of the Crisis*. MIT Center for Advanced Engineering Study.

IAEA (2001). *Working Material. Guidelines for Describing Human Factors in the IRS (Human actions and related causal factors and root causes)*. IAEA.

NEA (2006). *State-of-the-art report on systematic approaches to safety management*. NEA/CSNI/R(2006)1.

NRC (2000). *Inspectional manual, Inspection procedure 71841, Human Performances*. NRC.

OECD (2006). *Nuclear Power Plant Operating Experiences from the IAEA/NEA Incident Reporting System 2002–2005*. NEA No. 6150.

Reason, James (1990). *Human Error*. Cambridge University Press.

Shappell, Scott.A. and Wiegmann, Douglas.A. (2000). *The Human Factors Analysis and Classification System–HFACS*, Final Report. US Department of Transportation.

Sträter, Oliver (2000). *Evaluation of Human Reliability on the Basis of Operational Experience*. GRS-170. Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) mbH.

HEALTH AND WELL BEING OF CONSTRUCTION WORKERS

RECOMMENDATIONS FOR ERGONOMIC DESIGN OF WORKPLACES AND MACHINERY IN CONSTRUCTION

Alicia Piedrabuena, Alberto Ferreras, Purificación Castelló, Carlos García, Alfonso Oltra, José Gálvez, Lourdes Tortosa & Álvaro Page

Instituto de Biomecánica de Valencia, Universidad Politécnica de Valencia. Edificio 9C. Camino de Vera s/n, 46022 Valencia, España

This paper presents a study carried out by the *Instituto de Biomecánica de Valencia* (IBV) and the *Fundación Laboral de la Construcción* (FLC), with grants from the Spanish *Fundación para la Prevención de Riesgos Laborales*, about ergonomic conditions of representative workplaces, machines, vehicles and hand tools in the construction sector. The results of this study are applicable to a wide number of companies and workplaces in the construction field and have been put together in two documents edited by the FLC.

Introduction

Working of safety and health problems is one of the most worrying issues in the construction sector. This sector presents the highest levels of work-related accidents each year among all sectors of the economy in Spain. According to official data (INSHT, 2004), in the year 2005 there were 250.376 accidents at work with a subsequent sick or disability leave in the Spanish construction sector.

In Spain, overexertions constitute the first cause of accidents with sick leave in the construction sector (more than 25% of the total accidents). The main ergonomic problems in this sector are basically associated to the following factors: manual material handling, execution of repetitive tasks, adoption of awkward postures and inadequate use of machines and tools.

The increasing importance of ergonomic risks, as evidenced by the rate of accidents at work, in a sector where safety problems have always been the highest concern, motivated the *Instituto de Biomecánica de Valencia* (IBV) and the *Fundación Laboral de la Construcción* (FLC) to open a new line of work with the basic goal of obtaining recommendations that permit the ergonomic improvement of equipments and working conditions in the construction sector by means of actions concerning information, education and public awareness.

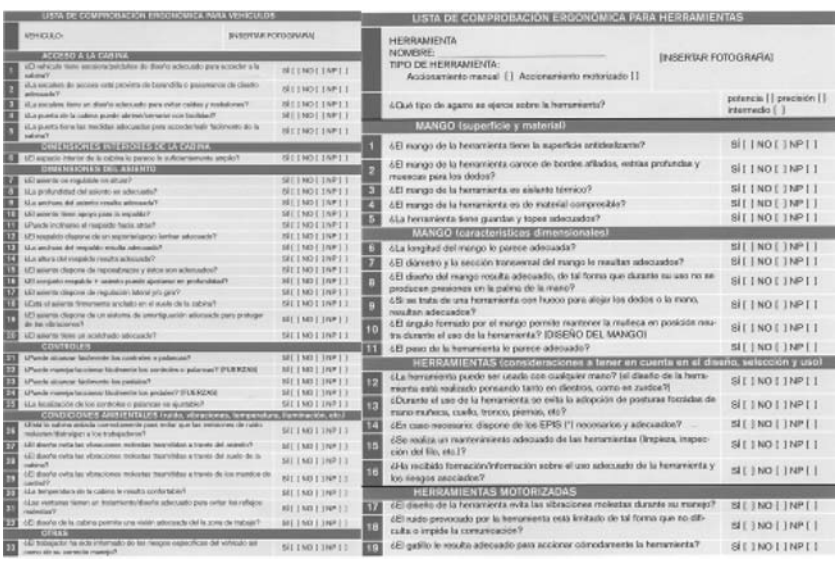
Methods

In order to obtain this goal a study was development that consisted of the following phases:

1. **Identification and selection of the study sample** focused on vehicles, machines, tools and workplaces that are most representative of the sector. For this purpose, a discussion group comprising specialists in ergonomics, technicians in the occupational risks, manufacturers and distributors of machinery, workers' representatives, labor unions, etc was established.
2. **Review of literature related** with ergonomic studies of machines, tools and jobs performed in the construction sector (standards, manuals, scientific publications, manufacturer catalogues, etc.).
3. **Generation of verification criteria**. Starting from the information gathered in the previous phases and from the ergonomic characteristics of each machine tool selected in phase 1, an initial proposal of Ergonomic Checklists was completed.

   For the completion of the checklists, it has been taking into account general ergonomic requirements adapted to the construction sector (NIOSH, 2004), as well as specific requirements for the design of vehicles, machines and tools coming from standards and other bibliographic sources, and from the experience of the technicians and other professionals that participated in the project.

   Each of the development checklists is composed of:
   - A section for identification of the vehicle, machine or tool.
   - A list of items or questions about ergonomic aspects arranged in thematic subsections (Figure 1).


LISTA DE COMPROBACIÓN ERGONÓMICA PARA VEHÍCULOS

VEHÍCULO: [INSERTAR FOTOGRAFÍA]

ACCESO A LA CABINA

1	¿El vehículo tiene escaleras/peldaños de diseño adecuado para acceder a la cabina?	SÍ [ ] NO [ ] NP [ ]
2	¿La escalera de acceso está provista de barandilla o pasamanos de diseño adecuado?	SÍ [ ] NO [ ] NP [ ]
3	¿La escalera tiene un diseño adecuado para evitar caídas y resbalones?	SÍ [ ] NO [ ] NP [ ]
4	¿La puerta de la cabina puede abrirse/cerrarse con facilidad?	SÍ [ ] NO [ ] NP [ ]
5	¿La puerta tiene las medidas adecuadas para acceder/salir fácilmente de la cabina?	SÍ [ ] NO [ ] NP [ ]

DIMENSIONES INTERIORES DE LA CABINA

6	¿El espacio interior de la cabina le parece lo suficientemente amplio?	SÍ [ ] NO [ ] NP [ ]

DIMENSIONES DEL ASIENTO

7	¿El asiento es regulable en altura?	SÍ [ ] NO [ ] NP [ ]
8	¿La profundidad del asiento es adecuada?	SÍ [ ] NO [ ] NP [ ]
9	¿La anchura del asiento resulta adecuada?	SÍ [ ] NO [ ] NP [ ]
10	¿El asiento tiene apoyo para la espalda?	SÍ [ ] NO [ ] NP [ ]
11	¿Puede inclinarse el respaldo hacia atrás?	SÍ [ ] NO [ ] NP [ ]
12	¿El respaldo dispone de un mecanismo para bloquear adecuado?	SÍ [ ] NO [ ] NP [ ]
13	¿La anchura del respaldo resulta adecuada?	SÍ [ ] NO [ ] NP [ ]
14	¿La altura del respaldo resulta adecuada?	SÍ [ ] NO [ ] NP [ ]
15	¿El asiento dispone de reposabrazos y éstos son adecuados?	SÍ [ ] NO [ ] NP [ ]
16	¿El conjunto respaldo + asiento puede ajustarse en profundidad?	SÍ [ ] NO [ ] NP [ ]
17	¿El asiento dispone de regulación lateral y/o giro?	SÍ [ ] NO [ ] NP [ ]
18	¿Está el asiento firmemente anclado en el suelo de la cabina?	SÍ [ ] NO [ ] NP [ ]
19	¿El asiento dispone de un sistema de amortiguación adecuado para proteger de las vibraciones?	SÍ [ ] NO [ ] NP [ ]
20	¿El asiento tiene un acolchado adecuado?	SÍ [ ] NO [ ] NP [ ]

CONTROLES

21	¿Puede alcanzar fácilmente los controles o palancas?	SÍ [ ] NO [ ] NP [ ]
22	¿Puede manipularlos/accionarlos fácilmente los controles o palancas? (FUERZAS)	SÍ [ ] NO [ ] NP [ ]
23	¿Puede alcanzar fácilmente los pedales?	SÍ [ ] NO [ ] NP [ ]
24	¿Puede manipularlos/accionarlos fácilmente los pedales? (FUERZAS)	SÍ [ ] NO [ ] NP [ ]
25	¿La localización de los controles o palancas es ajustable?	SÍ [ ] NO [ ] NP [ ]

CONDICIONES AMBIENTALES (ruido, vibraciones, temperatura, iluminación, etc.)

26	¿Está la cabina aislada correctamente para evitar que las vibraciones de cuerpo rígido lleguen a los trabajadores?	SÍ [ ] NO [ ] NP [ ]
27	¿El diseño evita las vibraciones molestas transmitidas a través del asiento?	SÍ [ ] NO [ ] NP [ ]
28	¿El diseño evita las vibraciones molestas transmitidas a través del suelo de la cabina?	SÍ [ ] NO [ ] NP [ ]
29	¿El diseño evita las vibraciones molestas transmitidas a través de los mandos de control?	SÍ [ ] NO [ ] NP [ ]
30	¿La temperatura de la cabina le resulta confortable?	SÍ [ ] NO [ ] NP [ ]
31	¿Las ventanas tienen un tratamiento/diseño adecuado para evitar los reflejos molestos?	SÍ [ ] NO [ ] NP [ ]
32	¿El diseño de la cabina permite una visión adecuada del/ la zona de trabajo?	SÍ [ ] NO [ ] NP [ ]

OTRAS

33	¿El trabajador ha sido informado de los riesgos específicos del vehículo en cuanto de su correcto manejo?	SÍ [ ] NO [ ] NP [ ]

LISTA DE COMPROBACIÓN ERGONÓMICA PARA HERRAMIENTAS

HERRAMIENTA
NOMBRE: _____ [INSERTAR FOTOGRAFÍA]
TIPO DE HERRAMIENTA:
Accionamiento manual [ ] Accionamiento motorizado [ ]

¿Qué tipo de agarre se ejerce sobre la herramienta? potencia [ ] precisión [ ] intermedio [ ]

MANGO (superficie y material)

1	¿El mango de la herramienta tiene la superficie antideslizante?	SÍ [ ] NO [ ] NP [ ]
2	¿El mango de la herramienta carece de bordes afilados, estrías profundas y muescas para los dedos?	SÍ [ ] NO [ ] NP [ ]
3	¿El mango de la herramienta es aislante térmico?	SÍ [ ] NO [ ] NP [ ]
4	¿El mango de la herramienta es de material compressible?	SÍ [ ] NO [ ] NP [ ]
5	¿La herramienta tiene guardas y topes adecuados?	SÍ [ ] NO [ ] NP [ ]

MANGO (características dimensionales)

6	¿La longitud del mango le parece adecuada?	SÍ [ ] NO [ ] NP [ ]
7	¿El diámetro y la sección transversal del mango le resultan adecuados?	SÍ [ ] NO [ ] NP [ ]
8	¿El diseño del mango resulta adecuado, de tal forma que durante su uso no se producen presiones en la palma de la mano?	SÍ [ ] NO [ ] NP [ ]
9	¿Si se trata de una herramienta con hueco para alojar los dedos o la mano, resultan adecuados?	SÍ [ ] NO [ ] NP [ ]
10	¿El ángulo formado por el mango permite mantener la muñeca en posición neutra durante el uso de la herramienta? (DISEÑO DEL MANGO)	SÍ [ ] NO [ ] NP [ ]
11	¿El peso de la herramienta le parece adecuado?	SÍ [ ] NO [ ] NP [ ]

HERRAMIENTAS (consideraciones a tener en cuenta en el diseño, selección y uso)

12	¿La herramienta puede ser usada con cualquier mano? (el diseño de la herramienta está realizado pensando tanto en diestros, como en zurdos?)	SÍ [ ] NO [ ] NP [ ]
13	¿Durante el uso de la herramienta se evita la adopción de posturas forzadas de muñeca, cuello, tronco, piernas, etc?	SÍ [ ] NO [ ] NP [ ]
14	¿En caso necesario, dispone de los EPIS (*) necesarios y adecuados?	SÍ [ ] NO [ ] NP [ ]
15	¿Se realiza un mantenimiento adecuado de las herramientas (limpieza, inspección del filo, etc.)?	SÍ [ ] NO [ ] NP [ ]
16	¿Ha recibido formación/información sobre el uso adecuado de la herramienta y los riesgos asociados?	SÍ [ ] NO [ ] NP [ ]

HERRAMIENTAS MOTORIZADAS

17	¿El diseño de la herramienta evita las vibraciones molestas durante su manejo?	SÍ [ ] NO [ ] NP [ ]
18	¿El ruido provocado por la herramienta está limitado de tal forma que no dificulta o impide la comunicación?	SÍ [ ] NO [ ] NP [ ]
19	¿El gatillo le resulta adecuado para accionar cómodamente la herramienta?	SÍ [ ] NO [ ] NP [ ]

Figure 1. An example of a cheklists.

- A guide of design recommendations and criteria, which contains specific measures to perform the evaluation of each item included in the checklists, as well as recommendations, clarifications, etc (Figure 2). This mini-guide aims to be a support document to make it easier for the technician to evaluate the different items in field study.

The Ergonomic Checklists were thought as a diagnostic tool, so that the different agents involved in the construction sector can determine the condition of the vehicles, machines and tools from the ergonomic point of view.

4. **Field study and information analysis**. Experts in ergonomics, biomechanics and anthropometry visited construction sites at different stages of the construction process and analyzed a representative number of the selected vehicles, machines, tools and workplaces.

For the aspects concerning vehicles, machines and tools, the field study allowed to evaluate and redefine the checklists, as well as the development ergonomic criteria and recommendations.

Finally, for the purpose of providing examples of use, a summary sheet for each of the vehicles, machines and tools that were the object of the study was developed based on the collected information (Figure 3). In each of the developed example sheets, the following information was collected:

- Description of the vehicle, machine or tool, main uses in the sector and posture adopted by the worker when he or she is operating it.


DIMENSIONES INTERIORES DE LA CABINA

6 CABINA, DISEÑO

Recomendaciones para el **diseño interior de la cabina**:

La altura de la cabina (a) no debe ser menor de 1510 mm. Valores recomendados 2000 mm.

El espacio para las piernas (b) medido desde el SRP (*) no debe ser menor 600 mm. Valores recomendados 1150 mm.

El espacio para las rodillas (c) medido desde el SRP no debe ser menor de 730 mm. Valores recomendados 880 mm.

La distancia hasta la parte posterior (tomada desde la altura de la cabeza del operador) no debe ser menor de 440 mm. Valores recomendados 630 mm (d).

La distancia hasta el cristal o parte anterior (tomada desde la altura de la cabeza del operador) no debe ser menor de 550 mm. Valores recomendados: 700 mm (e).

La anchura de la cabina medida a la altura de los reposabrazos (f) no debe ser menor de 920 mm. Valores recomendados 1300 mm.

Figura 11. Dimensiones interiores de la cabina y Punto de referencia del asiento

Para determinar las medidas de los vehículos es necesario considerar un punto de referencia tal y como establece la norma UNE EN ISO 6863, donde se define el Punto de Referencia del Asiento (SIP). En la práctica, este punto es difícil de determinar, por lo que se toma como punto de referencia la intersección entre el asiento y el respaldo (SRP).

El SIP está localizado 97 mm por encima y 130 mm hacia delante del SRP.

SRP: Intersección entre el respaldo y el asiento

SIP: Punto de referencia del asiento

DIMENSIONES DEL ASIENTO

1: rango regulación profundidad asiento + respaldo
2: profundidad del asiento
3: longitud del reposabrazos
4: Altura del reposabrazos
5: Anchura del asiento
6: Anchura del respaldo parte superior
7: Anchura del respaldo parte inferior
8: Distancia entre reposabrazos
9: Anchura del reposabrazos
10: Rango de altura

Figura 12. Dimensiones del asiento

MANGO (superficie y material)

1 SUPERFICIE

La superficie del mango de la herramienta no debe ser tan lisa o pulida que sea deslizante, ni tan rugosa que pueda ser abrasiva para el trabajador.

Figura 36 Mango, superficie

2 TERMINACIÓN DEL MANGO

Los extremos de los mangos deben estar redondeados para evitar que provoquen presiones localizadas en la palma. Es necesario evitar surcos profundos que puedan causar presión localizada.

Los mangos con alojamientos para los dedos no son nunca recomendables, sólo se adaptan a cierto número de usuarios, pudiendo causar presiones localizadas intensas en los dedos de otros usuarios.

En herramientas donde su agarre empuja o tornillo puede ser conveniente que la superficie sea moleada o con surcos en dirección contraria al movimiento.

Figura 37. Mangos, terminaciones

3 CONFORT DE USO

Material aislante del calor:

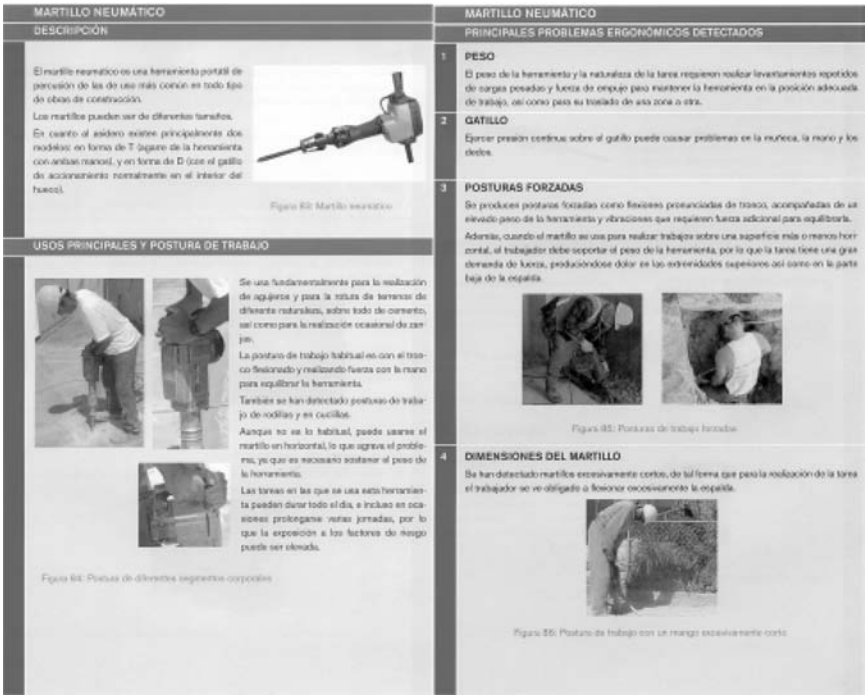
Temperatura mínima confortable de 36°.

Máximo en cortos periodos de tiempo de 43°.

Se refiere a que no transmite rápidamente el calor o frío cuando se está trabajando en un ambiente caluroso o frío.

La evaluación de este aspecto en campo se realizará subjetivamente, preguntando al trabajador.

Figure 2. An example of a guide of recommendations and criteria.


MARTILLO NEUMÁTICO

DESCRIPCIÓN

El martillo neumático es una herramienta portátil de percusión de las de uso más común en todo tipo de obras de construcción.

Los martillos pueden ser de diferentes tamaños.

En cuanto al asidero existen principalmente dos modelos: en forma de T (figura de la herramienta con ambas manos), y en forma de D (con el gatillo de accionamiento normalmente en el interior del hueco).

Figura 63: Martillo neumático

USOS PRINCIPALES Y POSTURA DE TRABAJO

Se usa fundamentalmente para la realización de agujeros y para la rotura de terrenos de diferente naturaleza, sobre todo de cemento, así como para la realización ocasional de zanjas.

La postura de trabajo habitual es con el tronco flexionado y realizando fuerza con la mano para equilibrar la herramienta.

También se han detectado posturas de trabajo de rodillas y en cuclillas.

Aunque no es lo habitual, puede usarse el martillo en horizontal, lo que agrava el problema, ya que es necesario sostener el peso de la herramienta.

Las tareas en las que se usa esta herramienta pueden durar todo el día, e incluso en ocasiones prolongarse varias jornadas, por lo que la exposición a los factores de riesgo puede ser elevada.

Figura 64: Posturas de diferentes segmentos corporales

MARTILLO NEUMÁTICO

PRINCIPALES PROBLEMAS ERGONÓMICOS DETECTADOS

1 PESO

El peso de la herramienta y la naturaleza de la tarea requieren realizar levantamientos repetidos de cargas pesadas y fuerza de empuje para mantener la herramienta en la posición adecuada de trabajo, así como para su traslado de una zona a otra.

2 GATILLO

Ejercer presión continua sobre el gatillo puede causar problemas en la muñeca, la mano y los dedos.

3 POSTURAS FORZADAS

Se producen posturas forzadas como flexiones pronunciadas de tronco, acompañadas de un elevado peso de la herramienta y vibraciones que requieren fuerza adicional para equilibrarla.

Además, cuando el martillo se usa para realizar trabajos sobre una superficie más o menos horizontal, el trabajador debe soportar el peso de la herramienta, por lo que la tarea tiene una gran demanda de fuerza, produciéndose dolor en las extremidades superiores así como en la parte baja de la espalda.

Figura 65: Posturas de trabajo forzadas

4 DIMENSIONES DEL MARTILLO

Se han detectado martillos excesivamente cortos, de tal forma que para la realización de la tarea el trabajador se ve obligado a flexionar excesivamente la espalda.

Figura 66: Postura de trabajo con un mango excesivamente corto

Figure 3. An example of development sheets.

- Ergonomic analysis, and detected main problems.
- Recommendations or proposals for ergonomic improvement to eliminate or reduce the incidence of detected problems.

With regard to the workplaces, the field study allowed to obtain a map of ergonomic risks associated to each of the jobs. For the collection of the information, it was necessary to perform recordings of each of the activities and collect data about the weights of the most representative materials, tools employed, etc. Then, in the IBV's ergonomics lab, an analysis was performed of both the recordings and the collected data, which permitted the identification of the most important ergonomic risks associated to each of the studied jobs.

5. **Elaboration of information, education and advising materials**. During this phase, information and education materials were put together to promote the dissemination of the results among the agents of the construction sector involved in projects.
6. **Assessment of the materials**. Once the information materials had been completed, the effectiveness and efficiency of them were checked by subjecting them to the assessment by a group of experts, who analyzed the contents and the means of dissemination of the materials that were more adequate and effective, as well as the parallel activities to be performed.

This multidisciplinary group of experts was composed of specialists in ergonomics, technicians in the occupational risks, manufacturers and distributors of machinery, workers' representatives, experts in the elaboration of education and information materials, etc.

7. **Dissemination**. The goal of this phase is to make known the results of the performed actions and reach the final users.

Results

The fundamental result of the realization of these studies is the completion of two publications:

Guide for the ergonomic verification of machine tools in the construction sector.(FLC and IBV, 2007a) This guide aims to serve as an instrument for the improvement of the ergonomic conditions at work associated to the use of machines and tools in the construction sector.

Manual of ergonomics in the construction sector FLC and IBV (2007b). The manual aims to serve as a tool for the identification and resolution of ergonomic risks, as information and consultation material and as a basis for the development of an education plan specific for workers in the construction sector.

Conclusions

Due to the importance that ergonomic risks have in the construction sector, it is essential to create lines of work for the generation of tools and programs that facilitate the evaluation and control of this type of risks in the sector.

The developed study is aimed to:

- Reduce ergonomic risks related to the use of machines and tools in the Construction sector.
- Facilitate the technical assistance in the field of ergonomic risks prevention for all the agents involved in this sector.
- Provide objective criteria for the selection and purchase of tools and machines.
- Allow the manufacturers to integrate the ergonomic principles into the design projects.
- Promote the culture of prevention among employers and workers.
- Offer a material the can be used for specific training of the workers in this sector.

References

INSHT (2004). *V Encuesta Nacional de Condiciones de Trabajo. Instituto Nacional de Seguridad e Higiene en el Trabajo*. Madrid.

NIOSH (2004). *Easy Ergonomics: A Guide to Selecting Non-Powered hand Tools*, DHHS (NIOSH) Publication No 2004 – 164.

FLC and IBV (2007a). *Guía para la verificación ergonómica de máquinas-herramientas empleadas en el sector de la construcción.* Fundación Laboral de la Construcción (FLC), Instituto de Biomecánica de Valencia (IBV). Madrid (Tornapunta Ediciones).

FLC and IBV (2007b). Manual de Ergonomía en la Cosntrucción. Fundación Laboral de la Construcción (FLC), Instituto de Biomecánica de Valencia (IBV). Madrid (Tornapunta Ediciones).

EVALUATING TECHNOLOGY AND RESEARCH KNOWLEDGE GAPS FOR PREVENTION OF FALLS FROM SCAFFOLDS

Hongwei Hsiao

National Institute for Occupational Safety and Health, Morgantown, West Virginia, 26505, USA

Falls from scaffolds are one of the leading causes of work-related fall fatalities and injuries. More than 9,500 workers are injured and 80 killed annually in the United States in scaffold mishaps. A literature search and review was performed to develop information on the current state of knowledge of fall protection technologies and existing knowledge gaps concerning safety during scaffold work. The search and review identified several key research opportunities to fill current safety-knowledge gaps: (1) understand causes of scaffold tipping and structure failure through fatality investigations, (2) study the life-cycle cost and ease of use of fabricated decking against those of solid sawn-lumber planks, (3) develop a set of best practices for scaffold safety, (4) identify safe designs for accessing or transiting onto or off a scaffold, and (5) evaluate whether certain scaffolding systems or designs are safer to erect and safer to use than others.

Introduction

Falls from scaffolds are one of the leading causes of work-related fall fatalities and injuries (Personick, 1996). Over the last two decades, the number of scaffold-related fall incidents has remained high (Webster, 2000). In 2006, approximately 14.2 percent (809 of 5,703) of fatal occupational injuries were falls; of those, 10.9 percent (88) involved scaffolds or staging (Bureau of Labor Statistics [BLS], 2007).

There are three basic types of scaffolds: supported scaffolds, suspended scaffolds, and other scaffolds (mainly aerial lifts). The goal of the current study was to conduct a review of peer-reviewed and trade literature for information on fall prevention methods and current state of knowledge related to work tasks performed on supported scaffolds. Supported scaffolds are defined as platforms supported by outrigger beams, poles, frames, or similar rigid support. They have historically accounted for the greatest number (3/4) of fatal and nonfatal injuries associated with scaffold use (BLS, 1983).

The overall findings of this review provided information on: (1) key fall-injury risk factors related to supported scaffold use, including OSHA's performance-based criteria to protect employees from scaffold-related hazards, (2) technologies for supported scaffold safety, (3) the state of knowledge about fall safety, and

(4) remaining unanswered questions pertaining to supported scaffold safety and existing knowledge gaps.

Method

Literature from 1950 and subsequent years related to the general topic of scaffold safety were located by searching the following computerized data bases: safety-related journals available under ScienceDirect®, American Society of Civil Engineers Online Journals, Medline, the U.S. Bureau of Labor Statistics (BLS), the Electronic Library for Construction Safety and Health, and the Occupational Safety and Health Administration (OSHA) rulemaking docket on OSHA's 1990s revisions to 29 CFR 1926 Subpart L. Expert opinion was also solicited from members of three trade associations and knowledgeable persons at two major U.S. manufacturers of scaffold equipment. Papers identified in the initial search were screened using three criteria: (1) Were new data or a synthesis of a selected group of studies presented? (2) Did the studies concern falls? and (3) Did the study results deal specifically with scaffold safety? Scaffold safety technologies published on the Internet and in trade association journals were included in the review to accommodate the fact that limited information on the subject was available in peer-reviewed journals. Fifteen scientific manuscripts and 9 scaffold-safety guidelines met the minimum inclusion criteria although their quality varied somewhat. Sixteen of these have either information on etiology of falls from scaffolds or engineering control suggestions and are reported in this presentation.

Results

Current scaffold safety regulations

In 1996, the U.S. OSHA revised standards for scaffolds (Title 29 Code of Federal Regulations Part 1926 Subpart L) in use since 1971 in the construction industry. It set performance-based criteria to protect employees from scaffold-related hazards. The key elements of the standards for fall prevention (i.e., general rules, scaffold configuration, and use of scaffold planks) are described below (Yassin and Martonik, 2004).

The general rules require employers to protect each employee on a scaffold more than 3.1 m above a lower level from falling to that lower level. Scaffolds and scaffold components must not be loaded in excess of their rated capacities. A guardrail or personal fall arrest system on all sides except the side where the work is being done must protect employees performing extended overhand work from supported scaffolds. The scaffold-configuration rules pertain to the design, manufacturing, and assembly of scaffolds. Each scaffold and scaffold component must support without failure its own weight and at least four times the maximum intended load applied or transmitted to it. Supported scaffolds with a height-to-base width ratio of more than 4:1 must be restrained by guying, tying, bracing, or an equivalent means. The height of the top rail for scaffolds must be between 0.97 m and 1.2 m. Mid-rails

must be installed approximately halfway between the top rail and the platform surface. The use of scaffold planks pertains to the selection, use, and assembly of scaffolds planks. Scaffold planking must be able to support, without failure, its own weight and at least four times the intended load. Each platform must be planked and decked as fully as possible with the space between the platform and uprights not to exceed 2.5 cm. The space must not exceed 24.1 cm when side brackets or irregular-shaped structures result in a wider opening between the platform and the uprights. The platform must not deflect more than 1/60 of the span when loaded. Finally, each scaffold platform and walkway must be at least 46 cm wide.

Risk factors associated with falls from scaffolds

The most common causes of falls associated with the use of supported scaffolds can be organized into five major categories: (1) scaffold tipping or structure failure (OSHA, 1979; Whitaker et al, 2003), (2) planks breaking, slipping, and gapping (Chaffin and Stobbe, 1979; BLS, 1983), (3) unguarded scaffold (Chaffin and Stobbe, 1979; Shepherd, 2000), (4) difficult access or transition onto or off a scaffold (Shepherd, 2000), and (5) problems with erection and dismantling of scaffold (OSHA, 1979; Shepherd, 2000).

Scaffold tipping or structure failure is attributed to various factors, such as inadequate anchoring into the walls, improper scaffold assembly, improperly secured bracing, loading beyond designed capacity, and failure of scaffold components under stresses (Chaffin and Stobbe, 1979; OSHA, 1979; Whitaker et al, 2003). Plank breaking, slipping, and gapping may result from a heavy load, physical damage of a plank, misinformation about the type of plank or its rating, inadequate overhang over the bearers, unsecured planks (no cleat), sideways movement of planks, and missing planks (OSHA, 2007; A-1 Inc., 2007; BLS, 1983). Conditions of unguarded scaffolds include missing guardrail and inadequate cross bracing (Chaffin and Stobbe, 1979; Shepherd, 2000). A variety of factors may influence safe access to the working level of the scaffold: the width of a run of integral ladders in scaffold end frames, the distance between two runs, and the difficulty transiting onto or off a scaffold platform to integral ladders due to the required plank overhang (Chaffin and Stobbe, 1979). Problems with erection and dismantling of scaffold may result from environment conditions, the weight of scaffold units, and the availability of handholds (Hsiao et al, 2008).

Current measures and technologies to control falls from scaffold

OSHA requires an anchorage to control scaffold tipping or failure when the scaffold reaches four times its width at the base (NAHB, 2004). Halperin and McCann (2004) reported that properly erected scaffolds were correlated with supervision by a competent person. Modern modular scaffold systems that are simpler to erect were reported to be safer because they reduced the possibilities for errors in construction (DOE, 1999). OSHA (2007) has developed a guide for selecting lumber for scaffold planks in construction with a goal to reduce plank-breakage incidents. Use of metal catwalks or platforms is also a solution to the problem. Plank locks are available in the market to prevent scaffold planks of solid sawn lumber from

slipping (Safway, 1998). Fabricated steel planks in 15- and 23-cm widths are also commercially available, thereby making it possible to arrange the planking to reflect the constraints of work situations involving irregular surface contact. Because of this fact, the maximum room between planks and the uprights of the scaffold frame can be lessened to below 7.5 cm instead of the allowed 24 cm, to provide better slippage control (Excel, 2001).

Essentially all scaffold manufacturers currently provide guardrail components that can be used to ensure compliance with current OSHA requirements; some frame-scaffold systems use a modular guardrail and mid-rail system that does not rely on cross-bracing to provide fall protection. As to the control of difficult access or transition onto or off a scaffold, engineered scaffold planks which hook onto the horizontal scaffold member, and do not extend beyond the end frame (as compared to the required 15-cm hangover to the supporting member for sawn lumbers) provide increased ease of access to the working surface from a ladder built into the end frame. Finally, lightweight scaffold components are available to ease the problem of falls during scaffold erection and dismantling (Hsiao et al, 2008). Fabricated scaffold planks are also available that are lighter than wood planks and thus decrease the risk of imbalance-related falls during the erection and dismantling phase.

Knowledge gaps remain to be answered

Several research hypotheses to address current knowledge gaps concerning scaffold safety are apparent from this review. A new study based on the findings of OSHA fatality investigations may help improve our understanding of the extent to which scaffolds tip when they are not anchored or are improperly anchored. Industry literature generally suggests that fabricated planking or decking appears to offer certain advantages over sawn lumber planks; it may be useful to conduct a series of experiments to validate these putative safety differences and the life-cycle costs. In addition, if one considers exposure to scaffold falls as a function of time, then the initial access phase during which the worker accesses or steps onto the scaffold may be the most dangerous part of scaffold work. The challenges in negotiating overhanging planks or guardrail components bear serious consideration. Evaluations are also needed on whether certain scaffolding designs are safer to erect and safer to use, given current advances in scaffold technology and work practices. Moreover, what opportunities are there for scaffold erectors to use personal fall arrest systems? How can technology be used to reduce fall exposures for this group? The development of a comprehensive set of "best scaffold safety practices" is suggested. Finally, the most recent detailed BLS survey on scaffold use and injuries dates to 1978 (BLS, 1983), and is now over two decades old; this survey may not be fully representative of the conditions under which workers perform tasks on scaffolding today. An enhanced national survey on scaffold use and injuries is long overdue.

Summary and disclaimer

The existing literature and industrial sources of subject expertise demonstrate the potential for occupational fatal-injury data and expert opinion sources to help

provide a data- and practice-based understanding of the etiology of falls from scaffolds, but this potential has not been fully realized. The result is that a set of best practices and standardized operating procedures and tools based on scientifically validated research have not been established for industrial applications involving scaffolds. Several apparent targeted research opportunities exist in surveillance and technology assessment, and additional research topics are suggested by the review. This area should be more fully explored in meetings of standards committees, industry-wide symposia, and targeted research efforts.

The author gratefully acknowledges the technical support of Mr. Paul Keane, Dr. Alfred Amendola, Dr. Jim McGlothlin, and Advanced Technologies and Laboratories International, Inc. The findings and conclusions are those of the author and do not necessarily represent the views of NIOSH. Mention of any products does not constitute the endorsement of NIOSH.

References

A-1 Plank & Scaffold Manufacturing, Inc., (A-1 Inc.), 2007, *A-1 Scaffold Safety Guidelines*. Paramount, CA, http://www.a1scaffold.com/safety.htm.

Chaffin, D.B. and Stobbe, T.J., 1979, *Ergonomic Considerations Related to Selected Fall Prevention Aspects of Scaffolds and Ladders as Presented in OSHA Standard 29 CFR 1910 Subpart D,* (University of Michigan: Ann Arbor, MI).

Excel Modular Scaffold & Leasing Corporation, 2001, *Technical Manual – Savings Through Safety, Quality, and Technology*, EMSLC-TSM-1001.

Halperin, K. and McCann, M., 2004, An evaluation of scaffold safety at construction sites, *Journal of Safety Research*, **35**, 141–150.

Hsiao, H., Hause, M., Powers, J., Kau, T., Hendricks, C. and Simeonov, P., 2008, Effect of scaffold-end-frame carrying strategies on worker stepping response, postural stability, and perceived task difficulty, *Human Factors*, **50** (1), [in press].

National Association of Home Builders (NAHB), 2004, *Scaffold Safety Handbook*, (BuilderBooks.com: Washington, DC).

OSHA, 1979, *Occupational Fatalities Related to Scaffolds as Found in Reports of OSHA Fatality/Catastrophe Investigations*, (Washington, DC).

OSHA, 2007, *Scaffold Specifications – 1926 Subpart L App A*, http://www.osha.gov/ pls/oshaweb/ owadisp.show_document? p_id=10917&p_table=STANDARDS.

Personick, M.A., 1996, New data highlight gravity of construction falls, *Issues in Labor Statistics*, (U.S. Government Printing Office: Washington, DC).

Safway Steel Products Inc., 2007, *Safway systems scaffold technical manual*, http://www. safway.com/PDF/ORN203.pdf.

Shepherd, G., 2000, Construction fatalities – the toll of gravitational energy, *ISF Symposium* (Orlando, FL).

U.S. Bureau of Labor Statistics (BLS), 1983, Survey of scaffold accidents resulting in injuries, 1978, *Work Injury Report*, (Washington, DC).

U.S. Department of Energy (DOE), 1999, Excel automatic locking scaffold, *DOE/EM-0482, Innovative Technology Summary Report*, (U.S. DOE: Idaho Falls, ID).

Webster, T., 2000, Workplace falls, *Compensation and Working Conditions*, (U.S. Government Printing Office: Washington, DC), 17–27.

Whitaker, S.M., Graves R.J., James, M. and McCann, P., 2003, Safety with access scaffolds: development of a prototype decision aid based on accident analysis, *Journal of Safety Research*, **34**, 249–261.

Yassin, A.S. and Martonik, J.F., 2004, The effectiveness of the revised scaffold safety standard in the construction industry, *Safety Science*, **42**, 921–931.

THE IMPORTANCE OF POWER TOOL CONSUMABLES ON VIBRATION EXPOSURE

Luca Notini[1], Neil J Mansfield[1] & David J Edwards[2]

Department of Human Sciences[1]/Department of Civil & Building Engineering[2], Loughborough University, Loughborough, LE11 3TU

This paper presents data on the effect of power tool appendage selection on construction worker hand-arm (HAV) exposure. It presents an analysis of HAV data acquired from cut-off saws, angle grinders, and hammer drills utilised with different models of tool appendage. The tasks conducted during tool utilisation were typical of those conducted in the construction industry. In the case of the cut-off saws the tool/appendage combinations were used to cut concrete, in the case of the grinders stainless steel and concrete, and in the case of the hammer drills to drill through concrete. Significant changes occurred in the work one can complete before exceeding statutory HAV exposure limits when the tool appendage model changed.

Introduction

Power tool vibration transmitted from the machine to the upper limbs is a primary source of hand-arm vibration (HAV) exposure within the construction industry. The main deleterious health effects linked to excessive HAV exposure are vibration white-finger, carpal tunnel syndrome and sensory impairment of the fingers. It is estimated that within Great Britain about 5 million men and women are exposed to HAV in the workplace every week, construction contains the largest HAV exposed population (Palmer, 2000).

The European Union introduced the European Physical Agents (Vibration) Directive (PA(V)D) in order to protect employees exposed to vibration. The directive requires minimisation of risk and limits HAV exposure workers may experience during the course of a working day. It defines a metric for daily vibration exposure dose ($A(8)$) that is based on the vibration emitted by the tool and duration of exposure. If $A(8)$ exceeds the Exposure Action Value (EAV $= 2.5\,\mathrm{m/s^2}$) during the course of the working day then measures must be taken to reduce vibration exposure within the environment. The $A(8)$ must not exceed the Exposure Limit Value (ELV $= 5\,\mathrm{m/s^2}$). The number of work units one can complete before reaching the EAV/ELV ($WU_{\mathrm{EA/LV}}$) depends on both the vibration emission magnitude and exposure time.

Manufacturers' declared emission values have historically under-estimated vibration exposure during field use (Mansfield, 2006). The freely available online OPERC HAVTEC register (www.operc.com) provides machine tool field-use HAV emission values for many machine tool types in order to assist workplace HAV

risk assessment. The database presents emission values for tools used on different materials and with multiple tool appendages (i.e. consumables such as drill bits or grinding discs).

The aim of this paper is to present data demonstrating the importance of power tool appendage model on the number of work units one can complete before reaching the ELV (WU_{ELV}) for cut-off saws, angle grinders, and hammer drills.

Methodology

HAV data was acquired from five 300 mm petrol powered cut-off saws, two 115 mm angle grinders, five 230 mm angle grinders, and four hammer drills. Each tool was used to execute work units with several models of appendage; approximately 1900 work units were executed with 72 tool/appendage combinations. In the case of the cut-off saws the work unit consisted of cutting concrete with a diamond disc (600 mm × 50 mm slot, water based dust suppression unless stated otherwise), in the case of the 115 mm angle grinders and two of the 230 mm angle grinders cutting stainless steel with abrasive discs (115 mm: 60 mm × 10 mm, 230 mm: 150 mm × 5 mm), cutting concrete with diamond discs in the case of three of the 230 mm angle grinders (600 mm × 40 mm slot), and in the case of the hammer drills drilling 100 mm into concrete. Data was acquired from each tool/appendage combination with three experienced operators; each operator executed at least five work units with the combination, and the sum of the five individual times to work unit completion was at least 60 seconds.

Acquisition of the HAV data was conducted in accordance with ISO 5349. Vibration acceleration data was acquired from each tool via two tri-axial accelerometers (SEN021F); each accelerometer was mounted within the proximity of an area of hand/tool contact whilst avoiding locking the tool trigger. The HAV emission value for each work unit executed consisted of the most severe W_{h} weighted r.s.s. value. The hardware utilised (IHVM 100) for acquiring the acceleration data and applying W_{h} conformed with ISO 8041. The time required to execute the work unit was measured.

In the following, where reference is made to significant difference, this refers to statistical significance as determined by the application of a one-way ANOVA design and t-test based post hoc testing (Bonferroni corrected, $p < 0.05$).

Results

In the following, the percentage difference between two mean values of WU_{ELV} stated is equivalent to the relative difference between the two mean values of work units to EAV.

Cut-off saws

In the case of all five saws, significant differences between blades in terms of mean WU_{ELV} were observed (Figure 1). The percentage difference between the blades that

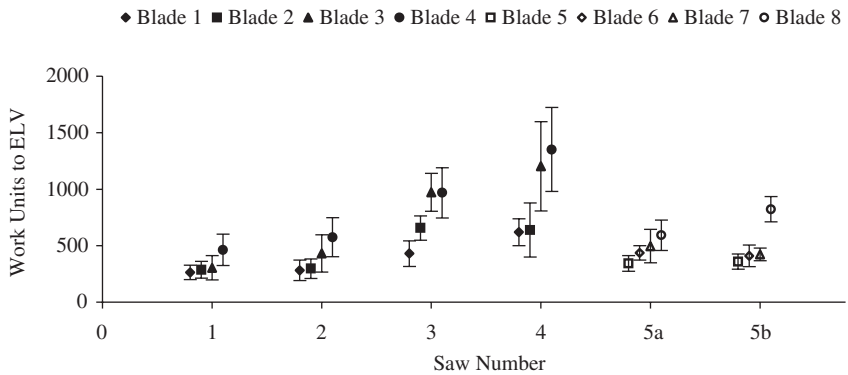


Figure 1. Plot of mean and sd of work units to ELV for the four models of blade used with machines 1–4 and the four blade models used with machine 5 (a: no dust suppression; b: with dust suppression).

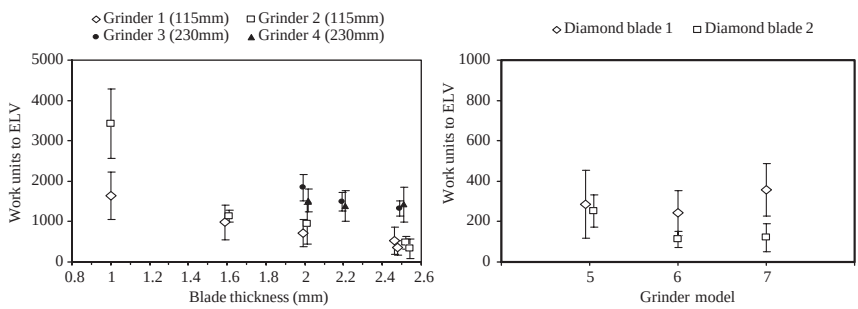


Figure 2. Mean and sd of work units to ELV for (left) the five blades used with grinders 1 and 2 (115 mm) and the three blades used with grinders 3 and 4 (230 mm) to cut stainless steel, and (right) for the two blades used with grinders 5, 6, and 7 (230 mm) to cut concrete.

produced the highest and lowest values of mean WU_{ELV} was 76%, 104%, 145%, 118%, 72% (dry cutting), 129% respectively.

The highest values of mean WU_{EALV} tended to occur with saw 4 as did the highest levels of variability in WU_{ELV}. Saw 5 was used with and without water dust suppression. The difference between wet and dry cutting in terms of mean WU_{ELV} was significant in the case of disc 8 only.

Angle grinders

In the case of the 115 mm angle grinders (Figure 2) used to cut stainless steel the percentage difference between the highest and lowest values of mean WU_{ELV} (respective to ascending blade thickness) was 109%, 11.5%, 16%, 32%, and 7%; in the case of the 230 mm grinders 7%, 8%, and 2%. A statistically significant

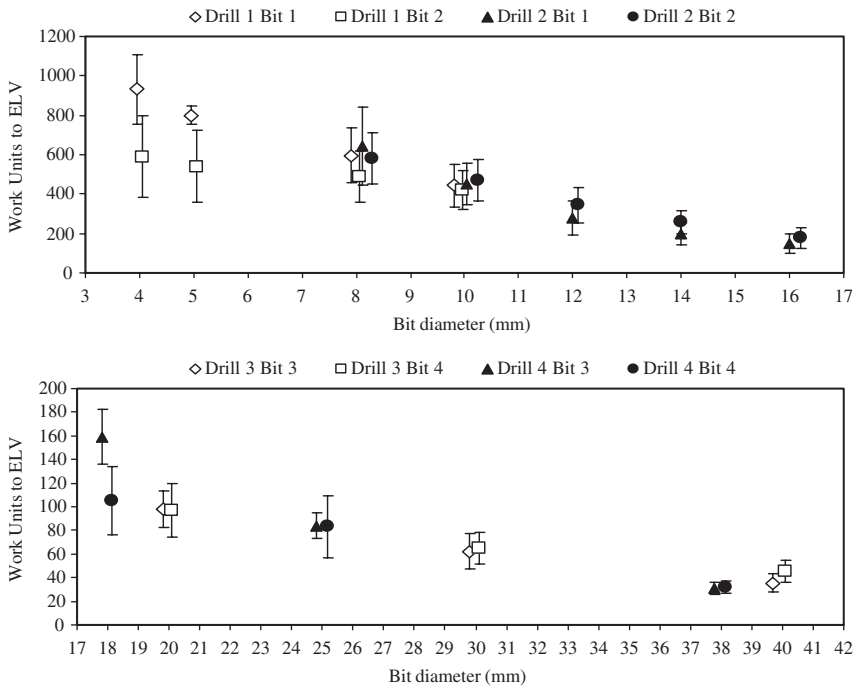


Figure 3. Plot of mean and sd of work units to ELV versus bit diameter for the four hammer drills.

difference between mean work units WU_{ELV} was observed only for the thinnest blade.

A strong negative correlation (Spearman, $p < 0.05$) exists between blade thickness and observed mean WU_{ELV} for the 115 mm grinder/blade combinations ($\rho^2 = 0.92$) and 230 mm grinder/blade combinations ($\rho^2 = 0.71$) used to cut stainless steel.

In the case of the three 230 mm angle grinders used on concrete (models 5, 6, and 7), the percentage difference between lowest and highest mean WU_{ELV} measured for the two models of diamond blades used with grinders was 13%, 119%, and 196% respectively. Significant differences between mean WU_{ELV} were observed for grinder 6 and grinder 7.

Hammer drills

Figure 3 present mean and sd WU_{ELV} for the hammer drills. The largest significant difference between highest and lowest mean WU_{ELV} occurred in the case of the 4 mm bits used with hammer drill 1 (48%), in the case of hammer drill 2 the largest difference between types occurred with the 14 mm bits (32%), hammer drill 3 the 40 mm bits (29%), and hammer drill 4 the 18 mm bits (52%). In the case of drill 1

the average difference between mean WU_{ELV} for the two bit types was 33%, in the case of drill 2 18%, drill 3 12%, and drill 4 16%.

Discussion

For each tool type, substantial changes in the allowable number of work units before reaching statutorily defined HAV exposure limits were observed when the appendage model changed. As such, it is deemed that blade model selection should be taken into account when assessing risk from HAV exposure for construction workers.

Additional analysis and previous studies (e.g. Rimell, 2006) indicate that both vibration magnitude and time required for work unit execution could be affected when change in appendage model occurs; hence both constitute contributory elements to the impact of appendage selection on HAV exposure risk.

In the case of the cut-off saws, differences between discs were observed in terms of vibration emission, with an average difference of 19% between highest and lowest mean vibration emission. In the case of the 115 mm angle grinders cutting stainless steel the average of the percentage differences observed when a change of appendage model occurred was 28%, and in the case of the 230 mm angle grinders 4%. For the 230 mm angle grinders used to cut concrete, the average change was 27%. In the case of hammer drill 1 the average percentage difference in vibration emission observed when bit type was changed was 13%, 18% in the case of drill 2, and drills 3 and 4 6%; vibration emission tended to increase with bit diameter.

In the case of all the appendages used, the material processed during execution of the work unit was appropriate for the tool. The diamond discs, as well as differing in terms of dimensions e.g. thickness, can be distinguished in terms of diamond grain and diamond/segment bonding process. The abrasive angle grinder discs used to cut stainless steel, as well as differing in terms of thickness, can be distinguished in terms of material composition of the abrasive grit, grit size, disc hardness, and bonding type. In the case of the drill bits flute design, cutting tip design, and total length constitute variables that may affect HAV exposure during execution of drilling tasks on a construction site.

Conclusions

This paper presents a summary of hand-arm vibration data acquired from cut-off saws, angle grinders, and hammer drills whilst used to conduct a task typical of those executed within the construction industry using various models of appendage.

In the case of the cut-off saws and angle grinders substantial differences between mean work units to ELV/EAV were observed when a change of appendage model occurred. For some appendages more than double the work units to ELV/EAV was measured on average than measured with another appendage.

For the hammer drills, whilst work units to ELV/EAV tended to decrease as bit diameter increased, significant differences between work units to ELV/EAV for the different drill bit types were only observed in a limited number of instances.

References

Palmer K.T., Griffin M.J., Brendall H., Pannett B. & Coggon D. 2000, Prevalence and pattern of occupational exposure to hand-transmitted vibration in Great Britain: findings from a national survey, Occupational and Environmental Medicine, 57, 218–228.

Mansfield N.J. 2006, Filling the gap between manufacturers' declared values and hand-arm vibration emission under real working conditions, Proceedings of the Institute of Acoustics, 28(4).

International Organisation for Standardisation 2001, Mechanical Vibration – Measurement and evaluation of human exposure to hand-transmitted vibration – Part 1: General requirements, ISO 5349-1.

International Organisation for Standardisation 2002, Mechanical Vibration – Measurement and evaluation of human exposure to hand-transmitted vibration – Part 2: Practical guidance for measurements in the workplace, ISO 5349-2.

International Organisation for Standardisation 2005, Human response to vibration – Measuring Instrumentation, ISO 8041.

Hand-Arm Vibration Test Centre (HAVTEC) www.operc.com (July 2007).

Rimell A.N., Mansfield N.J. & Edwards D.J. 2006, The Influence of Drill Bit on the Measured Vibration Exposure of Hammer Drills, 41st Group Meeting on Human Responses to Vibration, Farnborough.

IMPLEMENTING A HEALTH MANAGEMENT TOOLKIT IN THE CONSTRUCTION INDUSTRY

Phil Bust, Alistair G.F. Gibb & Charlotte Brace

Department of Civil and Building Engineering, Loughborough University, Loughborough. LE11 3TU

Occupational ill health affects large numbers of construction operatives in the UK each year. Ill health can have an effect on people's ability to work and on their general health and well-being. Unfortunately, in the more serious cases of occupational ill health, permanent disability and a slow death can result. Despite the dire consequences of occupational health problems, there is little management of ill health within the industry sector, possibly due to the delay between exposure to hazardous materials and activities and the onset of health problems.

With this background Loughborough University were commissioned by the Civil Engineering Contractors Association (CECA) to develop the health management toolkit (Brace and Gibb 2005).This paper reports on investigations into the implementation of the toolkit, including in depth interviews, carried out to understand how the toolkit has been used by the CECA members in part or in whole, how this affects any existing occupational health provisions that they have in place and what alternatives to the toolkit have been adopted. The responses from the CECA members are to be used in the further development of the toolkit in an iterative process to improve its design.

Introduction

Ill-health continues to disable and kill large numbers of construction workers and the delay in the effects becoming obvious is one of the main reasons why the subject should be taken seriously (ECI 1999). In 1999 the Health and Safety Executive's document Securing Health Together was released in which the UK Government set a target to reduce incidence of work related ill health by 20% by the year 2010. The stated aim in one of its five programmes of work was that a priority area was to increase the number of firms whose culture encouraged the management of health issues as an integral part of their day to day business. These events led in part to the decision by CECA to produce a health management toolkit that could be used by any contracting company and be freely available to all those who could benefit in UK construction.

The toolkit was launched at the end of 2005, consisted of 5 main sections, as listed in table 1. and is available free on the CECA website (www.ceca.co.uk).

Table 1. Respondents' Use of the toolkit.

Section of the toolkit	1	2	3	4	5	6	7	8	9	10	11
Ill health report form – collecting information on health problems	O	O, R	U	O	X	O	O	O	O	O	R
Health questionnaire information about workers overall health	O	O, R	O	U	U	O	M	O	O	U	R
Guidance on how to register with a General Practioner	O	O, R	O	O	X	O	X	O	O	X	R
Prevention of 5 key health issues - information & toolbox talks	O	O, R	O	O	U	O, P	M	O	O	O, P	R
Advice on health screening	O	O, R	O	O	X	O	U	O	O	O	R

O = own system, U = using, P = using with own system, M = using in modified form, x = nothing in place R = used for reference or guidence.

Methods

Loughborough University were asked to carry out a study to see how the toolkit had been adopted by a sample of CECA members who volunteered to take part. The sample had a geographical spread across the UK with the majority of workforces being around 500 in size although they were as small as 250 and large as 6000.

A series of questions were used to explore the toolkit's use and it was decided to use these to guide one to one interviews with the health and safety managers of the volunteer organisations. The questionnaire opened with some general questions on which parts of the toolkit were being used, how this fitted in with the use of any other health management procedures, reaction from workers and how the organisations were dealing with data protection issues. There then followed a section of questions for each of the elements (health report form, health questionnaire, GP registration, key health issues and health screening) before completing the questions by asking for information about the type of work carried out and size of workforce.

Findings

Use of toolkit

Eleven of the respondents used some part of the toolkit but none of the respondents used the toolkit in its entirety.

All but two of the respondents had an occupational health system in place prior to the launch of the toolkit. One had nothing and one had some elements in place. External occupational health services were being or about to be used by 8 of the respondents. Three of the companies used Norwich Union Health for their occupational health provision.

Personal health data was held by the external occupational health provider if used. Where occupational health was handled internally there was no consistent method of handling. Some used the human resources department, some the health

and safety department and on occasion data was held by a single individual. There was only one report of any negative reaction from the workers to occupational health intervention.

Ill health reporting

One respondent only used the ill health report form. They felt that stress should be included in the list of problems and were concerned about the confidentiality of the supervisor using the form.

Generally it was felt that workers were able to relate to the health problems on the list. One respondent felt that workers seemed have become more aware of work related health problems in recent years. One found the workers struggled with the term "musculoskeletal disorder" but understood "back pain". Another said that their workers were able to relate to the health problems because of one to one interventions with the occupational health nurse.

One respondent thought that high blood pressure and diabetes were not relevant to work while another said that high blood pressure was being recorded prior to the launch of the toolkit and a lot of workers were affected.

It was pointed out that the list of health problems documented on the CECA toolkit health report form is the same as the French model and that in France workers get regular health checks and the State pick up the bill.

Health questionnaire

Section 2 of the toolkit included a questionnaire for completion by construction workers, designed to record any existing health problems. Of the four companies that have used the questionnaire; one had a good completion rate and said that they administered the questionnaire on a one to one basis; one said it was rare for workers not to complete the questionnaire but they tended not to answer questions on hand numbness and a third one said that the workers did not always complete the questionnaire even when someone was sitting with them. One of the organisations said that, at the moment, workers who did not fill in the questionnaire were left alone but in future they would be sent to the occupational health nurse for an assessment.

Within this group one respondent said that the questionnaire highlighted breathing problems while another said that back problems and hand arm vibration cases were highlighted.

Of the companies that used their own system, one said that the workers accepted their system because they started it on their first day of work; one said that half the workers were positive about their questionnaire and half were negative (and it was not age related); and one said that because it was carried out at the recruitment stage that they just got on and did it; one had been using their own questionnaire for about 9 months prior to the launch of the toolkit and said that half the workers answered no to all questions.

One company using their own system said that their questionnaire highlighted a lot of high blood pressure problems while another found that deterioration of

eyesight was common. One of the cohorts was shocked at the size of the Norwich Union questionnaire and now use a tool box talk to prepare the workers for it.

GP registration

Section 3 of the toolkit included information on how to register with a GP. In eight of the companies they had either asked if the workers were registered with a doctor or advised them to register. One interviewee said that they had a regional workforce and didn't bother with GP registration because there was not a lot of working away from home and another said that they would contact local GPs if they had a number of men working on a project in a remote location.

One interviewee said that out of 10% referrals, 20% of these didn't turn up and another said that when their workers were referred to a doctor "the odd one disappeared but generally there wasn't a problem". One respondent said that the married workers tended to be registered and the single ones not. Another commented that the Norwich Union system did not push for workers to be registered.

Key health problems

This is the largest section of the toolkit and most used by the respondents. Of the key issues that were prevalent in different organisations vibration came out top (4 highest ranking responses) followed by noise and muscular (2 each) and then dermatitis and respiratory (1 each). When asked if there were any other issues more relevant to their work, two respondents said stress and two other respondents said problems with eyes.

Each of the interviewees were asked to place the key issues in order of relevance to their organisation. Figure 1 below shows the results with 5 signifying most relevant and 1 least relevant.

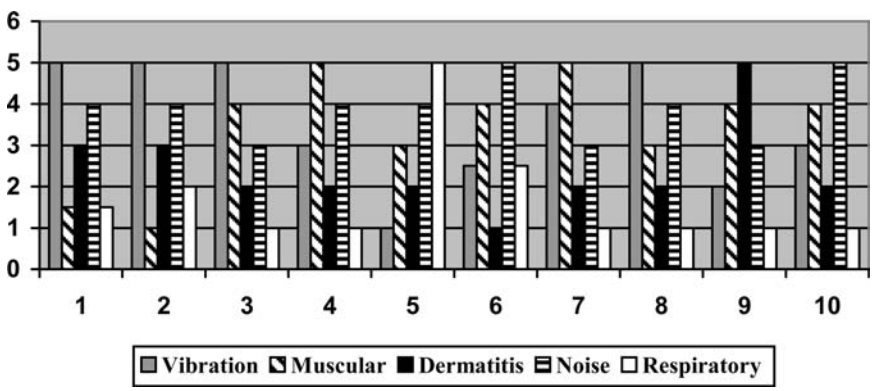


Figure 1. Relevance of key issues to respondent's organisations.

Health screening

Most of the respondents had had experience with commissioning occupational health practitioners with varied results:

- "One of them gave us no confidence they could provide a national service"
- "We looked at using NHS+ but they have limited resources"
- "It is difficult to find occupational health providers who you feel are not trying to fleece you"
- "Most providers talked about cost and not the medical issues and many couldn't provide all the services we needed"
- "The first provider didn't understand construction and this one isn't flexible with the changing work environment"
- "A lot of occupational health providers use shock treatment – if you don't do this you will go to jail"
- "There are charitable organisations that provide the service but you have to go to them"

Suggested changes to the toolkit

Two of the interviewees said that they didn't think that the GP registration section was appropriate for them. Otherwise there were no comments on the remaining sections. One person said that the case studies ought to be updated with ones more relevant to construction and another said that it would be useful to include a page to clarify the difference between health and occupational health and some statements on the benefits of health.

There were two comments on the size of the toolkit. Both suggesting it should be smaller and that some flow charts might improve it to show who does what and when. One interviewee asked if there could be information included to advise members on how to get construction equipment tested for vibration. Whilst another requested the inclusion of an approved list of occupational health providers.

Insurance

As several of the members referred to insurance during the interviews and had different opinions on the insurance companies views on the adoption of occupational health schemes (one thought that adopting a scheme would uncover health problems and premiums would rise; the other thought that adopting a scheme would lead to lower premiums) the Association of British Insurers (ABI) was contacted in order to clarify the situation.

The ABI reported that there are always a number of factors that affect levels of premiums some of which are beyond insurers' and companies' controls. However, they were of the opinion that introducing an occupational health scheme would help to bring premiums down. If the companies introduce occupational health schemes and preventative processes this can reduce incidents of absence leading to claims.

The ABI would prefer insurance brokers to tell companies of the benefits of occupational health provision. They also said that they are currently working with

the UK Department of Work and Pensions on tax incentives to improve occupational health and rehabilitation.

Discussion

In their guidance document 'Leading health and safety at work' (HSE 2007) the Institute of Directors and the Health and Safety Commission state that delivery depends on an effective management system to ensure, so far as is reasonably practicable, the health and safety of employees, customers and members of the public. In order to establish if a system is effective it's implementation needs to be investigated. That is why this study is important.

Interviewing health and safety managers from various size organisations who have day to day involvement with the occupational health systems was also valuable and the one to one interviews proved to be more beneficial because the participants often volunteered useful information about what was not included in the questionnaire. They also contributed examples of their own occupational health forms and in some cases details of general results.

Conclusions

The toolkit has proved to be successful in improving the occupational health provision in the majority of the organisations in some part. It has provided a starting point in instances when companies had no previous provision in place. It has also been useful in supplementing existing provision especially with regard to the toolbox talks and it has been used as a lever to broker deals with occupational health providers. From the feedback obtained there is now an opportunity to update the toolkit so that it can continue to be available as a useful tool for construction companies to use as a whole or to supplement their own systems.

References

Brace, C.L. & Gibb, A.G.F. 2005, *A health management process for the construction industry*, in proceedings of CIB W99 Working commission conference, South Africa 2005. ISBN 0-620-33919-5.

Gibb, A.G.F., Gyi, D.E. & Thompson, T. 1999, *The ECI guide to managing health in construction*, (Thomas Telford, London).

HSE. 2007, *Leading health and safety at work* Guidance document (HSE Books, Sudbury, UK).

SURVEY OF HEAVY EQUIPMENT OPERATORS WHOLE-BODY VIBRATION PROFILES

Gerry Newell[1], Neil J. Mansfield[2] & Luca Notini[2]

[1]*Defence Science and Technology Agency, Farnborough, UK*
[2]*Environmental Ergonomics Research Centre, Loughborough University, Loughborough, UK*

The Physical Agents Vibration Directive was implemented in 2005 it stipulates measures for controlling the health risks associated with vibration exposure. The increased awareness of this occupational hazard has resulted in widespread research trying to address the problem. This study focused on characterising features of whole-body vibration exposure among operators of heavy machines throughout a range of industry sectors including; construction, quarries and coal mining. Measurements were carried out under real operating conditions to investigate the nature of occupational exposure to whole-body vibration. The findings have identified the machines which are most likely to cause potential harm to operators from excessive amounts of vibration exposure, of particular concern were the machines fitted with tracks. They were found to pose a significant risk combined with the poor postures adopted by the operators.

Introduction

Whole-body vibration (WBV) exposure is a common problem in particular working environments, such as construction and agriculture. Excessive and prolonged exposure has often been reported to cause back problems, especially for employees whose majority of their working time is spent in vehicles (e.g. Pope *et al*., 1998). Since the introduction of the Physical Agents Vibration Directive (PA(V)D, 2005) every UK workplace that exposes their employees to vibration is required to carry out WBV specific risk assessments. The regulations stipulate measures for controlling the health risks associated with vibration exposure including consideration of "*the design and layout of workplaces and work stations*" amongst other factors. It is therefore important to consider the combination of occupational risks during evaluations of vehicle operators to ensure a holistic approach is adopted.

The daily exposure action and limit value in the Directive have been standardised to an eight-hour period. Both the limit and action values pertain to the highest vibration of the three orthogonal axes; fore-and-aft (x), lateral (y) or vertical (z). The principal method used in the UK and across Europe is the A(8). This method produces a cumulative exposure using a root-mean-square (r.m.s.) acceleration

value adjusted to represent an 8-hour working day. The second method uses the Vibration Dose Value (VDV) it is the time integral of the acceleration and can be applied to a number of vibration forms. The equation for VDV outlines the use of the fourth root to calculate the vibration value. The equations for r.m.s. and VDV are similar apart from the power values and the measurement time division for the r.m.s. The VDV responds more readily to the shocks in a signal compared with r.m.s. and it maintains this influence as time passes. It has been suggested to present a more reliable measure of the risk exposed to operators (Sandover, 1997). VDV is now used in a small number of institutes, yet seldom used in industry and it has no legal framework. The exposure values specified by the PA(V)D for the two different methods are as follows:

- Daily exposure limit value (ELV): 1.15 m/s^2 A(8) or 21 $m/s^{1.75}$ VDV
- Daily exposure action value (EAV): 0.5 m/s^2 A(8) or 9.1 $m/s^{1.75}$ VDV

The Directive has been incorporated into the UK's health and safety agenda with the publication of the 'Control of Vibration at Work' regulations (Health and Safety Commission, July 2005). Brereton and Nelson (2003) outlined the criteria for assessing whether an exposure is to be considered moderate or high:

- Moderate exposure can be regarded as exposures likely to exceed the EAV on some days or exposure below the EAV but containing occasional high magnitude shocks where VDV is usually less than 17 $m/s^{1.75}$
- High exposure can be regarded as exposures likely to exceed the EAV and a VDV of 17 $m/s^{1.75}$ and likely to exceed or approach the limit value if not adequately managed.

The aim of the field study was to quantify whole-body vibration exposures among operators of heavy earthmoving machines throughout a range of industry sectors. Some of the biggest industries; coal mining, quarries, and construction were targeted to obtain data on the types of machines for which very little was previously available. Measurements were carried out under real operating conditions to investigate the nature of occupational exposure to whole-body vibration. The criteria specified previously were used to determine if the exposures were moderate or high.

Methods

Measurements of whole-body vibration were made on 43 earthmoving machines. Repeat measurements were taken twice, three times or four times on 7 different types of machines, this made 61 whole-body vibration measurements in total. Machine groups included wheel loaders, tracked loaders, skid steer loader, bulldozers, motor graders, articulated dump trucks, rigid dump trucks, excavators, material handler, compacter, rollers and challenger (tracked tractor). Machines of the same type and model were evaluated at the different sites to allow for comparison across the variety of working environments. In total 10 different sites were visited to collect data. The study was designed so that minimal interference was caused to the operators who were required to perform their daily work tasks. In order to achieve this equipment

set-up was completed as fully as possible before approaching the machine and operator. The equipment was set-up in the machine during break times or periods of inactivity. Information about the machine and operator was also collected during these times.

Measurement durations varied depending on the operation of the machine. The average measurement duration was $131 \pm 67(\sigma)$ minutes (range 22–326 minutes). However, in common with many types of earth moving machines, the work usually required some waiting time where the machine was stationary (e.g. waiting for another operator to suitably position a lorry; queuing at a site bottleneck). The vibration measurements were conducted according to ISO 2631-1 as required by the Physical Agents (Vibration) Directive. A tri-axial accelerometer was fitted to the seat pan in a SAE flexible disc beneath the ischial tuberosities. The accelerometer measured vibration in 3 translational axes; the fore-and-aft (x-axis), lateral (y-axis), and vertical (z-axis). Statistical measurement parameters identified in ISO2631-1 (1997) were used for the evaluation of health effects and whole-body vibration exposure. The r.m.s was used on the frequency weighted acceleration data. Vibration Dose Value's were also calculated from the frequency weighted acceleration.

Observations of video data and high level task analysis allowed for identification of any distinct tasks encountered during the operating cycle, e.g. loading versus hauling. This enabled comparison of such tasks to help identify tasks that subjected operators to the greatest amounts of vibration exposure. It also enabled information to be recorded about the typical postures adopted by the machine operators.

Results & Discussion

The whole-body vibration guidance recommends a holistic approach to back pain incorporating factors such as manual handling and posture. This translates into the holistic health monitoring for those who are exposed above the action value, which will identify cases of back pain by self-reporting on symptoms and lead to action on any possible causative factors. They have taken the holistic approach to dealing with whole-body vibration because suitable methods still do not exist for detecting the onset of back pain and even if they did the back pain could not be linked to whole-body vibration exposure by itself. The Health and Safety Commission suggest that if a person is exposed to two or more factors together then their risk of getting back pain will increase, some combinations of factors include;

- WBV exposure for long periods without being able to change position
- Driving over rough terrain while looking over your shoulder to check on the operation of the attached equipment
- Being exposed to high levels of WBV and then doing work involving manual handling of heavy loads

The operators of the machines measured for this study were found to mainly be in operating environments where the top two factors applied. Their primary job was to operate the machine and to ensure any obstacles were avoided; they did not perform any additional tasks involving manual handling. However they were

Figure 1. From left to right; Challenger tractor, Tracked loader, and Bulldozer.

exposed to variety of different sitting postures, particularly the tracked dozer and loader operators who were found to regularly adopt twisted postures.

The worst individual machine measured during this work was the challenger tractor (Figure 1); in addition to being propelled by crawlers the machine was also carrying a "hex" compactor style attachment. Discussion with the manufacturers highlighted the issue of non-compliance with the use of the attachment. However, even without the attachment the machine is still likely to expose the operator to high magnitudes of vibration. Especially considering the machines fitted with crawlers and performing tracking tasks were found to produce the greatest WBV emissions (bulldozers and tracked loaders). The challenger is not typically used in construction it is more likely to be observed in agricultural environments. This machine sets a prime example for how inappropriate selection of machine and attachment for a specific task can expose the operator to unnecessary high vibration magnitudes.

Out of the remaining machines over 70% were found to exceed the exposure action value of the Directive. The bulldozers, tracked loaders (Figure 1), articulated trucks and nearly all the wheel loaders were found to exceed the EAV in half a day. A number of the bulldozers and tracked loaders also had the potential to exceed the ELV in a working day, considering many of the operators were found to work for longer than 10 hours a day.

The Control of Vibration at Work does not require employers to make vibration measurements where the value is not likely to exceed the ELV within an 8-hour day. They are allowed to use existing data (if available) on the types of machines they are using to create a WBV health risk assessment. If the machines have the potential to exceed the ELV then measurements may be needed in order to minimise the exposure. Figure 2 outlines a model of the WBV exposures from the machines measured for this study. It highlights the machines of most concern and therefore the ones that should be targeted for measurement where there is likely to be cases exceeding the ELV. Therefore anyone using the table can identify the machine of interest and if it required further measurement to determine the true WBV exposure. For example, dozers fall within the high exposure for the fore-and-aft or vertical vibration combined with twisted postures and therefore are high priority for risk management. The skid steer loader has been placed in the moderate category because although the vibration in the vertical direction was just below the ELV the exposure duration is limited to a short period each day.

The first method of reducing exposure to vibration is to try and reduce vibration at source. This can often be a difficult and costly task. In certain circumstances it

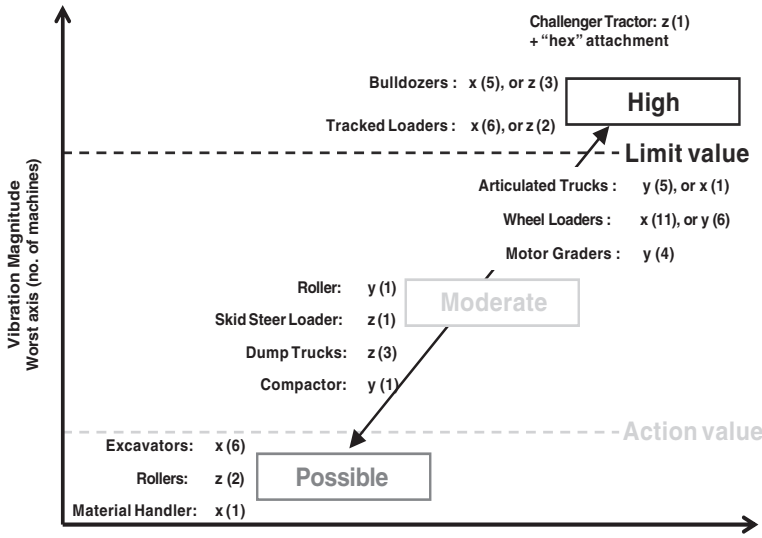


Figure 2. Characterization of earthmoving machine operator health risk exposures from; construction, mining and quarrying environments.

could save the company money at the start followed by unknown cost savings over a longer term for potential increased productivity, less wear-and-tear on machines, reduced sick leave, no compensation claims and prevention of re-training due to loss of operators. The current study provides a prime example of this with the company that were using the Challenger with a "hex" attachment. The sole task of this machine on the particular work site was to flatten the ground in preparation for building construction to commence. The vibration magnitudes were extremely high in this machine and there was only one operator that could use it for 8 hours a day. They would exceed the limit value in 2½ hours, so under the Control of Vibration at Work Act they would be required to stop the operation of the machine and attachment. The most cost effective solution would be to replace the machine with a roller, as this type of machine performs the same tasks as the Challenger and the rollers in this study were found to expose operators to low vibration. The roller appears to be the most viable solution to provide protection for their employee while also adhering to the regulations. It is often the case that large construction companies will have a number of the machines on loan this would make it even easier to implement the changes.

Terrain and driving style were found to affect the vibration exposure of operators particularly in the loaders and dozers. One of the most practical methods of reducing the vibration exposures is to educate the operators on correct driving speeds and appropriate usage of the machine. Operator driving style was also found to influence the measurement of vibration in construction, mining and quarrying machines (Scarlett and Stayner, 2005). It was found that vibration magnitudes vary according to how hard/enthusiastically the bucket of a loader was driven into a stock pile. It was

also established that three of the machines with the lowest vibration measured were owner operated. For this reason the operator's behaviour appeared to be influenced by the cost of maintenance and repairs of their machine. This phenomenon was not present for the machines that were hired or owned by "the company" nor was it a factor for those measured in this study.

Overall the most practical method of reducing the vibration exposure is through training to ensure the machine operators adjust their driving technique depending on the task and environment and to ensure they do not adopt an aggressive driving style. It is the most cost effective method of reducing exposures as it does not require replacement of equipment. Re-educating the operators regarding appropriate driving techniques could help to minimise their exposures for both vibration and postures this is particularly relevant for the bulldozers.

Acknowledgements

The authors wish to acknowledge the support provided by Caterpillar Inc., Peoria, Illinois, USA. The views expressed here are not necessarily those of the sponsor.

References

Brereton, P. and Nelson, C. 2003, HSE's consultative document on future regulations and guidance – whole-body vibration. *38th United Kingdom Conference on Human Response to Vibration*, held at Institute of Naval Medicine, Alverstoke, Gosport, England, 17–19 September.

European Council Directive. 2002, The minimum health and safety requirements regarding the exposure of workers to risks arising from physical agents (Vibration), (EN) 177, 13–19.

HMSO. 2005, Control of Vibration at Work Regulations 2005, Statutory Instrument 2005 No.1093, ISBN 0110727673.

Pope, M.H., Wilder, D.G. and Magnusson, M. 1998, Possible mechanisms of low back pain due to whole-body vibration. *Journal of Sound and Vibration*, 215(4), 687–697.

Sandover, J. 1997, High acceleration events in industrial exposure to whole-body vibration. *HSE Report, 134*.

Scarlett, A. J. and Stayner, R.M. 2005, Whole-body vibration on construction, mining and quarrying machines: Evaluation of emission and estimated exposure levels. *HSE Report 400*.

MAINTENANCE WORKERS AND ASBESTOS: UNDERSTANDING INFLUENCES ON WORKER BEHAVIOUR

Claire Tyers & Siobhán O'Regan

The Institute for Employment Studies

This paper reports on a selection of the findings from a qualitative study on the attitudes, knowledge and behaviour of maintenance workers to working with asbestos. Although supply, import and use of asbestos has been bannned for some years, maintenance workers remain at risk of new exposure and therefore asbestos related diseases (ADRs). The study was commissioned by the Health and Safety Executive, UK and completed in 2006.

Introduction and background

Asbestos is the single largest cause of work related deaths from ill-health. The long latency period between exposure and ill-health means the death rate has not yet peaked but the current annual toll is estimated at 4,000 deaths. It was used extensively as a building material between 1950 and 1980, and was used widely until as late as 1999. Asbestos is dangerous when the fibres are released into the air, which could happen during application, removal or maintenance. The occupational groups now at risk of new exposure, and therefore asbestos related diseases (ADRs), are maintenance workers. Research has indicated that this group of workers underestimate their exposure to asbestos (Burdett and Bard, 2003). The HSE has therefore targeted maintenance work in a bid to reduce the number of new cases. The duty to manage asbestos contained in current legislation[1], requires dutyholders (eg the building owner) to identify and record the location and condition of asbestos, assess the risks, and take steps to reduce exposure. The regulations mainly pertain to non-domestic premises; the HSE also ran a campaign in the autumn of 2006 to capture work not covered by the dutyholder regulations. IES was commissioned by the HSE to conduct a research project to examine maintenance workers' awareness, attitudes and behaviour towards asbestos risks including an examination of the barriers which discourage such workers from taking appropriate action.

[1] Control of Asbestos Regulations 2002, and more recently the Control of Asbestos Regulations 2006, which amalgamated all the existing asbestos legislation into one set of regulations.

Method and participants

The research involved 60 in-depth interviews with maintenance workers. Interviewees were recruited using a variety of methods, including through the contacts of trade unions and colleges, press advertisements, and liaising with large facilities management companies. The most successful method was the use of an opt-in survey of sole traders using a commercial database, and more than half of the achieved sample came through this route. Sole traders are a particularly interesting group as they are able to offer a dual perspective from their experiences of working on jobs where they self-manage and also where they have worked as virtual employees (ie as sub-contractors for other firms). All participants were given a £20 gift voucher and asbestos safety information at the end of the interview.

The obtained sample cannot claim to be representative of the maintenance worker population, however, the research did include an approximately even distribution of electricians, plumbers/heating engineers, carpenters/joiners, painters/decorators and workers in other areas of maintenance. Around half of the interviewees were aged over 50, and the majority were speakers of English as a first language.

The majority of interviews were conducted face-to-face and took place outside the work environment such as at the worker's home, or a local café or public house. The interviews were semi-structured and included a brief review of their work and training history and an exploration of their attitudes towards working with asbestos. Interviewees were then asked to rate their own knowledge about the various aspects of working safely with asbestos. Finally, a behavioural event interview technique was employed to explore their experience of working with asbestos. Transcriptions of interviews were analysed with the aid of Computer Assisted Qualitative Data Analysis Software.

Research messages

Understanding of asbestos risks

A number of the interviewees had taken part in formal asbestos training, although this was often part of a broader health and safety course rather than a specific training episode. Training messages for many, however, had not been updated since their apprenticeship or early in their career. Therefore, much of the training discussed predated 1992 when the asbestos regulations came into force. In addition, interviewees, who at the time of the interview were working as sole traders, were often relying on training which had taken place in a very different context, such as when working for a large employer. Most respondents supplemented their knowledge of asbestos obtained through training in an informal manner, through their work experience, and/or the knowledge and experiences of colleagues and family. Almost all those interviewed had picked up on the general message that asbestos was a dangerous material, with clear messages not to touch it, and to not work with it. However, their knowledge was much more limited when it came to identification or safe working procedures. In the face of high anxiety and a lack of specific knowledge about how to deal with asbestos, individuals are likely to listen to, or give credence to, sources

of knowledge that reduce their anxiety, even where they know these sources to be less reliable.

The generic awareness about the dangers of asbestos exposure were not always supported by more specific, action-oriented knowledge of how to determine when exposure may occur. Many respondents lacked confidence about their ability to recognise and identify asbestos during their work, or lacked awareness of the range of materials that potentially contain asbestos. Workers tended to focus on the different types of asbestos, relating more to colour descriptions (ie brown, blue, white), and relative dangers posed by different materials, rather than any detailed understanding of the full range of asbestos containing materials (although there were a number of exceptions).

Additionally, few felt confident about the specific procedures to follow once asbestos had potentially been identified (eg what to do , who to notify, how to test etc). Older workers were generally more confident about their ability to identify asbestos, but their methods of identification could involve disrupting asbestos and therefore potentially causing harm to themselves or colleagues. Examples of professional testing of potential asbestos containing materials were fairly rare. These options were generally only available to workers when working on larger sites, commercial or public buildings. So some workers were aware, theoretically, that they needed to report their suspicions and get some testing done (because that is what they had learned from training on a large site) but were lost in practical terms when they found themselves working on their own in the private domestic sector.

Older workers are often strong role models for younger colleagues in a more general sense, shaping social norms on work sites. This can act as a positive influence where older workers reinforce messages about the importance of safe behaviour, and older workers can be powerful champions for asbestos awareness and safety in their workplaces. However, in some cases, older workers felt that it was too late to protect themselves from the dangers of asbestos exposure and therefore engaged in behaviours which could put themselves and others at risk. There were also examples given where older colleagues had discredited safety messages about asbestos, playing down the risk. In addition, given that training on asbestos tends to be a fairly infrequent event, knowledge of what constitutes safe practice and the information to support this has changed over time, and over the course of some older workers' careers. Their ability to provide appropriate information and support for less experienced colleagues can therefore be limited, even given good intentions.

Barriers to safe working

To even begin to take appropriate action workers need a basic awareness that asbestos exposure is a risk to them and that they may come across asbestos containing materials (ACMs) in the course of their work. There does appear to be some tolerance for what is perceived as 'low level' exposure, and some confusion about how exposure levels are linked to the onset of asbestos related diseases (ARDs). There was a widespread belief that the potential levels of exposure in the maintenance trades is now very low. This was driven by the knowledge that asbestos is not placed in new builds, but also, in some cases, by a misperception that it was no

longer present in the majority of older buildings due to widespread removal. The asbestos message was also diluted by concerns about other risks. Safety risks, such as falls or other accidents, for example, were felt to be of more immediate concern, whilst 'new' materials (eg MDF or fibreglass) were often felt to pose more of a health risk to today's worker than asbestos.

As a result, many workers don't feel they need to know more, particularly about some aspects of working with asbestos that they don't perceive as relevant to them (eg decontamination, disposal or risk reduction). This appears driven by the general perception that in their working lives, the tasks they complete and the job they do, asbestos isn't really a risk. Therefore, whilst asbestos is acknowledged as a dangerous material in a general sense, individuals tend to be reluctant to accept, or be dismissive of, the actual risks to their own health. This can manifest itself in inaccurate estimations of the risks facing those conducting maintenance work. It is therefore entirely rational for workers who do underestimate the risks in this way to dismiss the need to learn anything other than the very basics about working with asbestos.

There are a range of other factors which contribute to how they actively view or manage these risks. In order to take the decision to behave safely around asbestos, workers not only need to know about the risks, but also take these risks seriously enough to cause them to change their behaviour. Reasons given for **not** adopting safe practices included financial pressures and constraints, concerns about speed and timescales, work cultures and peer pressures. Individuals were able to rationalise unsafe behaviour through their calculation of the risks of the job versus the financial benefits. Often, people were prepared to 'take a chance' if the job was felt to warrant it through a sufficient financial pay off. This was further exacerbated by the fact that a fairly common view was that there is somehow a random element to whether ARDs are contracted after exposure, often described as 'a lottery'. Therefore, individuals are able to focus much more on concrete, commercial concerns rather than less visible and less clear issues around their own health. Prospect theory details the asymmetry of human choices, that how we assess risks of losses differently from how we assess risks of gains (Kahneman and Tversky, 1979).

The general safety culture of a site (eg rigorous risk assessments, use of personal protective equipment) plays a role in how seriously workers take all risks, and certainly how they assess the risks posed by asbestos. The attitude of employers towards the safety of their staff has a very powerful effect. Where they are seen to take the risks seriously, and express their willingness to stop the job if workers' health is at risk, this is extremely effective in preventing risk taking.

There is also likely to be a complex interplay between these factors. For example, on sites where corners are cut, there is likely to be a combination of negative peer pressure, a strong drive for cost reductions and tight deadlines. In such an environment the extent to which individuals feel able or sufficiently 'in control' to take appropriate action in relation to any risk could be limited. However, given the lack of understanding about the risks posed by asbestos and how to deal with them, it is likely that individuals feel even more powerless in relation to asbestos risks than many others. This could be a particular issue for those working on a temporary or casual basis, or the self-employed working as sub-contractors.

Experiences and work behaviour

Detailed knowledge of how to actually reduce risks when working with asbestos were uncommon, largely because people didn't identify their job as involving this type of contact. A few interviewees did have detailed knowledge on this issue, however, and were able to identify a range of actions that could be taken. Decontamination was seen as an emergency procedure only, rather than something that affected their day to day working lives, and was also seen as difficult to undertake properly in practice. Disposal often involved using the council, and in a small number of cases individuals discussed examples of illegal disposal that they were aware of. Options to report or test asbestos were not common except by those working for larger employers or as subcontractors on a big site. Walking away from the job appeared to be related to an individual's perception of the strength of their market position and choice of other work.

Individuals were able to discuss a range of experience of working with asbestos. Some of these related to situations where asbestos exposure was likely, either through unwitting or deliberate dismissal of the risks. More common were instances where, although some procedures had been followed, these were in some way incomplete or incorrect, and therefore may have led to exposure. Despite this, the vast majority of more recent experiences were viewed by individuals as relatively safe, and their own behaviour appropriate. The most commonly discussed safeguards were: to pause or stop work; avoid unnecessary fibre release; use personal protective equipment or damp down materials. Overall, workers' experiences were characterised by a lack of detailed procedural knowledge of their own, and many were reliant on others to inform them about correct procedures. There were a number of examples where individuals themselves had tried to follow correct procedures, but where this required them to break with the overriding culture on a site, or stand up to colleagues. It is likely that individuals wanted to share examples where they felt they had, themselves, behaved appropriately, so it is unclear how widespread unsafe behaviours are more widely.

Influences on knowledge and awareness

Workers were positive about the effect that large scale, national, marketing campaigns could have upon their knowledge and awareness of the issue. When showed materials produced by the HSE, individuals reacted well to them, but most could not recall having seen them before. For information to effectively reach all those working within this sector, it will be necessary for a variety of formats to be adopted and information developed to fit the varied needs of a heterogeneous community.

It could be useful to move away from the basic message that 'asbestos kills', which has successfully reached this population of workers, to a more detailed message about how many and how. This clarity is what causes people to take the risks seriously. Previous marketing and awareness raising campaigns have been immensely successful in communicating basic messages about how risky it can be to work with asbestos in an unsafe way. The priority for the future should be to continue to focus on providing workers with more detailed knowledge on how to protect themselves.

Conclusions

There are a range of issues which affect how likely an individual is to behave safely around asbestos. This research suggests that they can be broken down into four main categories:

- Technical issues, relating to the complexity of messages about asbestos, its effects and how to deal with it effectively.
- Psychological issues, concerning an individual's attitudes towards risk, health and the specific risks posed by asbestos.
- Cultural factors such as pressures from their employers, clients, co-workers etc., which are largely driven by economic as well as social pressures.
- Control factors, namely the extent to which individuals feel that they are able to control their work environment. These are linked to the nature of the employment contract an individual has, and their labour market capital.

Attitudes towards behaving safely are likely to be affected by an individual's internal perception of whether the potential negative impact of exposure is outweighed by other factors. In essence, whether the economic or social costs are outweighed by the health benefits. However, even intending to behave safely is not enough if an individual lacks sufficient knowledge to support this intention.

References

Burdett, G. and Bard, D. 2003, Pilot study on the exposure of maintenance workers (industrial plumbers) to asbestos, HSL MF/2003/15

Hope, C. 1999, The impact of procurement and contracting practices on Health and Safety: a literature review, HSL report RAS/99/02, in White, J. 2003, *Occupational Health in the Supply Chain: A literature review*, HSL/2003/06

Kahneman, D. and Tversky, A. 1979, Prospect Theory: An analysis of decision making under risk, *Econometrica*, **47**, 2

This paper is based on research reported in full as:

O'Regan, S., Tyers, C., Hill, D., Gordon-Dseagu, V. and Rick, J. 2007, Taking Risks with Asbestos: What influences the behaviour of maintenance workers, HSE Research Report 558

EMPLOYING FOREIGN WORKERS IN SWEDEN CONSTRUCTION SECTOR

Radhlinah Kunju Ahmad & Bengt Hansson

Department of Construction Management, Faculty of Engineering, Lund University, Lund, Sweden

The enlargement of the European Union provides opportunities for members to be mobile in seeking various forms of permanent and non-permanent employment, particularly temporary work and contract labour. These changing conditions require quick reactions, overcoming soft issues such as language barriers and coping with culture clashes and political issues that could create problems for the worker. While the influx of foreign workers may bring much needed skills and meet the labour demands of the sector, the existence of an irregular workforce can have an impact on labour market conditions, local economy and health and safety measures. Currently, the use of foreign workers and contractors on construction sites has been identified as one of the major issues confronting clients, employers and unions. This paper will present the benefits and disadvantages of employing foreign workers and contractors in the Swedish construction sector and why the need to examine their working environment.

The changing scene

After a slump lasting almost throughout the 1990s, the Swedish construction sector has grown by double-digit rates since 2004 (Sweden, 2006). Over half of all construction firms now point to a lack of qualified labour as the main barrier to expansion. Such pronounced labour shortages have not been seen since the late 1980s (Sweden, 2006). This setback will correspondingly lead to higher construction costs for delivering projects as well as leading to a shortage of resources. Consequently, the wind of change is turning towards foreign contractors and foreign workers to fill this gap (Lawén, 2005).

Manpower supply bottlenecks can be overcome by tapping the labour pool of other countries. The enlargement of the European Union (EU), which means free movement of workers inside the EU, aims to solve this critical situation (Sandberg et al., 2004). The movement results in a flood of cheap labour from the east to the west. It is estimated there are around 3,000 foreign workers in the construction sector with concentrations of carpenters, joiners and concreters (Holmberg, 2004). While some workers in Western Europe are concerned about cheap labour coming from the east, with the right to employment, others see it as a positive development process within the sector (Boeri and Brücker, 2006). The once Swedish-dominated

working environment is now being complemented with workers from different cultural background and with different competencies.

This paper examines the threats and opportunities of employing foreign contractors and workers and discusses the options for handling this sensitive matter. A research project is proposed for which objectives and methods of enquiry are outlined. For the purpose of this paper and the research, foreign workers are categorised as those who work in another country without the initial intention of settling there and without the benefits of citizenship in that (host) country. The Swedish Immigration Board awards work permits to the following: workers who are highly specialised, where skills are not available through the domestic workforce; workers who are recruited to supplement the workforce for a limited term or to provide skills on a contractual basis; and workers who are recruited directly by an employer, who may need to certify that it cannot find domestic workers.

Opportunities and threats

The presence of foreign workers results in both opportunities and threats to the construction sector which are discussed below in the context of findings from other studies:

Opportunities

- Flexible adjustment – The presence of young foreign workers can contribute towards filling the pensions' shortfall. Simultaneously, this will give time for the younger Swedish generations to acquire the right competencies and knowledge (Tilly, 2005). These numbers can be controlled by adjusting the quota of incoming foreigners.
- Enriching knowledge – Sweden's construction sector is one of the best in the EU in regard to health and safety and working conditions. By engaging in a Swedish construction environment, foreign workers not only fulfil their own dreams but also acquire and learn from best practice (Tilly, 2005). They may take this new knowledge and experience to their homeland and put it into practice.
- Lower cost of labour – Usually, foreign labour is not bound by the host country tax system if temporarily employed and working for less than a year. These workers contribute to their homeland or their security contributions are borne by the foreign company (Werner, 1996). Werner states that the host country usually pays a wage that is roughly equivalent to the regular wage in the location in order to avoid the displacement of domestic labour.

Threats

- Increase in hiring cost – When hiring new workers, employers must conduct induction cost or training. Most companies will try to keep this cost as low as possible especially if the workers are to be engaged for a short time (Werner, 1996). This cost can only be kept low if the foreign workers are specialised in given tasks.

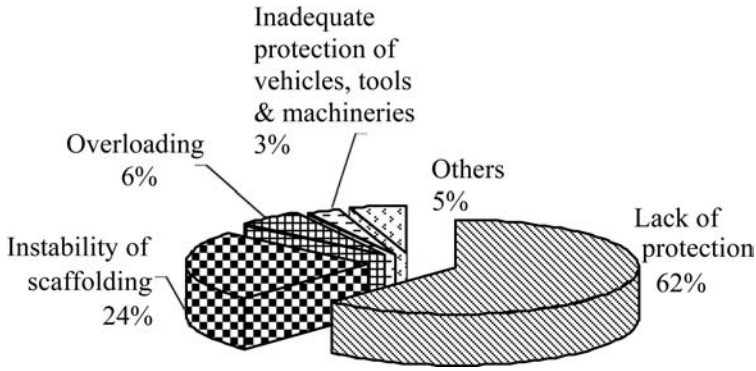


Figure 1. Common offences of health and safety committed by foreign workers on construction sites (Christensen and Fransson, 2007).

- Labour market compatibility – Werner argues that even though the wage difference between the host country and the country of origin is considerable, there is a tendency for employers to pay foreign workers wages lower than the local wages. This situation is readily accepted by most foreign workers without protest as long as the wages are higher than back home. It is claimed that a worker from Slovakia earns 11.5% of a Swedish worker; a worker from Poland earns 18% and a worker from Lithuania earns just 8.4% (Lawén, 2005). The drawback is that this could result in wage dumping and unfair competition between companies that employ such labour and those that do not (Werner, 1996).
- Working conditions – Abuse of foreign workers can lead to safety and health risks and it is an impediment to the sustained development of the national construction sector, particularly since the workers affected are not often afforded the protection of social security systems, such as sickness and sick-leave benefits (Agapiou, 2005). These workers may not possess the right qualifications, lack formal training, are without proper protection and suffer from poor induction programmes. A consequence of the lack of safety awareness is, in particular, a worrying matter, which together with different working standards and high mobility in the labour market, further raises the risk of injury and death (Döös et al., 1994).
- Safety risks – In many countries there is a higher accident frequency for foreign labours than for national labourers (Agapiou, 2005). Such studies have proven that occupational accidents among foreign workers can be as high as 5.7 per 100 000 workers compared to 4.3 per 100 000 workers for locals (Loh & Richardson, 2004). The fact that these groups are sometimes called cheap labour is also the reason for these workers being preferred for work in hazardous conditions. These groups also tend to take jobs that are temporary, require less skill and are largely unattractive to local labour, at the same time being paid lower than local workers (Rowlands, 2005). A study performed by Byggnadsarbetaren on foreign workers (Christensen and Fransson, 2007) demonstrated that the most common type of offence on construction sites is lack of protection when working at height.

Table 1. Working cultural differences between Polish and Swedish construction workers (Gustafsson and Hansson 2005).

Polish working culture	Swedish working culture
○ Focus on result	○ Focus on working time
○ High man-hours	○ Low man-hours
○ Fear of making mistakes	○ High priority for quality
○ No working order	○ Proper working order
○ Low safety priority	○ High priority for safety
○ High responsibility	○ Low responsibility

- Language barriers – Language diversity is also a barrier to communication on site which often results in unsafe behaviour. A study performed by Trajkobski and Loosemore (2006) claimed that nearly half of the respondents admitted to having misunderstood work-based instructions as a result of their poor command of the native language, while two-thirds acknowledge that they had made a mistake at some point as a result of this handicap. In Sweden, this barrier is even greater as most foreign contractors and workers face difficulty in understanding both the Swedish and the English language. Most face difficulty especially in interpreting the rules and regulations when undertaking a project in Sweden, where most information and instructions are in Swedish (Christensen, 2007). From the survey reported by Christensen and Fransson (2007), only 21% of 3,600 sites surveyed had translated the working environment plan into the language of the foreign workers.
- Job and skill mismatches – There is evidence of mismatch-unemployment for welders, technicians, electricians, occupations in the food industry and construction, carpenters and painters (Doudeijns 2005). Very often the high skill migrants occupy positions below their educational level. Additionally, a study performed by Dainty et al. (2005) demonstrated that construction companies are facing problems of poor quality of skills especially among new entrants. Acute shortages of skills were apparent across the key trades including bricklaying, carpentry and joinery, electrical installation, plumbing, pipe fitting and roofing.
- Cultural differences – Different countries have a different working culture which can influence the way of working as shown in Table 1.

The way forward

The above discussions reveal that in most situations, employment of foreign contractors and workers can lead to abuse and misunderstanding. It is important, therefore, to probe the current working conditions on construction sites where foreign contractors or foreign workers are employed. This action could even extend to foreign clients. Detailed diagnosis of the current situation will help in preparing an action plan for improvement. Examining their working environment will help

to strategically plan a better working environment and benefit both employers and employees.

Objectives of the research study

In order to undertake this research approach, the following objectives are employed to aid the process: appraise the current trend of employing foreign workers, contractors and clients in the Swedish construction sector; study the terms of appointment for employing foreign workers, contractors and, possibly, foreign clients; investigate the working environment of foreign workers and contractors; identify the problems and barriers to employing foreign workers and contractors; and suggest an action plan for improvement.

Methods

The intended method is that of a questionnaire-based survey with follow-up interviews on construction sites in Sweden. It is anticipated that interpreters will be needed to a certain extent to ensure that both questions and answers are understood.

Conclusions

Lack of knowledge about the current situation of foreign construction workers in Sweden, the scale of health and safety problems and their implications make it difficult to demonstrate accurately the extent of the problem. Employing foreign workers is possible, but the key issues regarding their working and living conditions must be considered before ploughing ahead. An urgent need is to study the current situation and to determine how this can be changed for the better. All who are involved – from the client to site workers – must understand their roles in ensuring equal opportunities and in providing decent working and living conditions. They must also be aware of the potential for clashes of culture and language barriers in the pursuit of a safer and healthier working environment.

References

Agapiou A. 2005, The impact of labour migration on the construction industry in the enlarged European Union, *The Equal Opportunities International*, **24**(7/8), 11–28

Andersson K. 2007, Brist på arbetskraft kan ge dyrare byggen, *Näringsliv* 3 april

Boeri T. and Brücker H. 2006, Why are European so tough on migrants, *Economic Policy*, **20**(44), 629–703

Christensen N. 2007, Språkproblem skapar faror för bygget, *Byggnadsarbeten*, 5 April, 14

Christensen N. and Fransson M. 2007, Utländska arbetare har sämre arbetsmiljö, *Byggnadsarbeten*, 5 April, 12

Dainty A.R.J., Ison S.G. and Briscoe G.H. 2005, The construction labour market skills crisis: the perspective of small-medium-sized firms, *Construction Management and Economics*, **23**, 387–398

Doudejns M. 2002, Labour shortages and the need for immigrants: A review of recent studies. *Metropolis Conference Workshop 48*. September, 103–127

Döös M., Laflamme L. and Backstrom T. 1994, Immigrants and occupational accidents: A comparative study of the frequency and types of accidents encountered by foreign and Swedish citizens at an engineering plant in Sweden, *Safety Science*, **18**, 15–32

Gustafsson S. and Hansson K. 2005, Utländsk arbetskraft i svensk byggsektor, En studie av två NCC-projekt ur ett kommunikations- och Kulturperspektiv, *Institutionen för bygg- och miljöteknik Byggnadsekonomi, C*halmers Tekniska Högskola

Holmberg A. 2004, *Byggchefen* 4/2004, 9

Lawèn P. 2005. Det ska vara lätt att hyra en lett! *Elektrikern*, June, 38–39

Loh K. and Richardson S. 2004, Foreign born workers: trends in fatal occupational injuries 1996–2001, *Monthly Labour Review*, **127**(6), 45–53

Robert R.D. 2004, Supervising immigrants workers. *Supervision*, **65**(5), 14–16

Rowlands C. 2005, Disposable society? *Work Management*, **50**(10), 3

Sanberg Bo., Odgaard G. Lubanski N. 2004, *Free mobility and EU's enlargement – migration of the construction workers after May 2004. (report)*, CLRdenmark

Sweden statistic, 2006, www.sweden.se (2006-10-27)

Tilly H. 2005, *Byggchefen* 1/2005, 18

Trajkobski and Loosemore. 2006, Safety implications of low-English proficiency among migrant construction sites operatives, *Intl Journal of Project Management*, **24**, 446–452

Werner H. 1996, *Temporary migration of foreign workers* 1996, Institut für Arbeitsmarkt- und Berufsforschung, Strasbourg (CDMG (96) 18E)

AGEING WORKFORCE IN CONSTRUCTION – EQUIPMENT USE AND THE PREVENTION OF EARLY RETIREMENT

Joanna Leaviss, Alistair G.F. Gibb & Phil D. Bust

Department of Civil and Building Engineering, University of Loughborough

The construction industry has a heavy physical workload, which can accelerate the ageing process. In addition natural ageing causes a decline in physical fitness. Incidence of injury and ill-health in the construction industry is high, and can result in early retirement. This has consequences for the industry itself and society as a whole. This study aimed to explore how the use of equipment and design of work process for older construction workers could prevent injuries and ill-health, and how it could assist ageing workers as their physical fitness naturally declines. The study used semi-structured interviews and small focus groups with equipment designers, equipment rental firms, older workers, site managers, and construction health and safety managers.

Participants revealed the kinds of equipment currently available to ease the physical burden of construction tasks for older workers. Participants also reported barriers to the use of this equipment. These issues related to individual attitudes, financial implications, organisational structure, and training. In addition, changes in work processes were suggested. Solutions are needed to overcome these barriers to uptake of safe work practice including better provision and design of equipment. Interventions to encourage more frequent use are also required. Design of equipment and interventions designed to promote safe practice should be inclusive. Goals should include the protection of workers before they suffer injuries and ill-health related to a career in construction in order to prevent early retirement from the industry.

Introduction

The world's population is ageing, and as a result of this, the working population is ageing. It is therefore becoming increasingly important to understand the specific needs of older workers in order that the working environment may be adapted accordingly. In addition to the value in retaining the skills and experience of older workers, the recent Employment Equality (Age) Regulations (October 2006) make it unlawful to discriminate against employees or job seekers on the grounds of age. This presents organisations with the responsibility of ensuring health and safety policy is 'age friendly'.

Declines in physiological fitness through the natural ageing process differentially affect workers in a range of jobs and industries (see McMahan and Chikamoto 2006 for review of the following physiological declines). Visual impairments are more common with increasing age – these include a reduction in visual range, loss of contrast sensitivity, increased susceptibility to glare, and problems discriminating between certain colours. Decreases in the ability of an older person to perceive depth can result in trips and falls, and compromise safety. Hearing loss becomes more likely with increasing age, which, in addition to causing safety concerns (e.g. hearing alarms, understanding verbal instructions) may also result in problems maintaining balance. Reaction times slow with increasing age, and this is more apparent the more complex the task. However, older workers can compensate for this decline as a result of their prior experience. Musculoskeletal changes occur with ageing. These affect joint dexterity (reducing manoeuvrability, posture and grip), endurance and muscular strength (Muscular strength peaks by the age of 30, declining to 75% of this peak by age 65 (Vitasalo et al. 1985). This is particularly pertinent in occupations with a heavy physical workload.

Age-targeted ergonomic intervention can reduce the occupational challenges faced by older workers for individual areas of physiological decline, for example increased area lighting, glare reduction, increased font size (visual), reduction in background noise and use of ear defenders (hearing), use of power tools and manual handling aides (musculoskeletal). For reviews of age-related ergonomic interventions see Moyers and Dale (2004), and Roper et al. (2007). Interventions that are industry and task specific may promote even greater success. The current study focuses on interventions relating to older workers in the construction industry.

The construction industry presents specific challenges to the older worker. It is well-documented that construction workers are at greater risk of a range of work-related health disorders than many other industries and the population as a whole (see Kines et al. 2007, and Arndt et al. 2007 for reviews). The tough physical workloads associated with construction work have been shown to take their toll on the workforce. Guberan and Usel (1998) showed the construction industry to have the lowest rate of survival in work without permanent incapacity at aged 65, at only 57% compared with an average (across a range of industries) of 75%. Brenner and Ahern 2000 showed the most common disabilities leading to early retirement of construction workers in Ireland on health grounds to be cardiovascular disease (31%), hypertension (16%) and musculoskeletal disorders (30%). These studies suggest that the construction industry is not a hospitable environment for the ageing worker. Although studies have identified ergonomic solutions to some of the tasks involved in construction (e.g. McMahon and Philips 1999, Vedder and Carey 2005) few studies have focused on the specific needs of the older construction worker in terms of equipment and work process.

The current study aimed to investigate how equipment and work process could improve the working environment for the older construction worker. It used qualitative methodology to explore the perceptions of older workers, health and

safety managers and equipment providers and designers in relation to age-related ergonomic intervention in the construction industry.

Methods

Participants: 55 participants were drawn from all areas of the construction industry. These included older workers (average age 56), health and safety managers, equipment and materials designers and manufacturers, equipment providers, and industry training bodies and trade unions.

Data collection and analysis: Data was collected via semi-structured interviews and small focus groups. Core questions related to the types of equipment and materials available to assist older construction workers in their everyday tasks, barriers to use of this equipment, and work processes specifically relevant to the older worker in construction. Data was examined and a thematic analysis was conducted to identify emerging issues.

Results

Participants reported items of equipment they felt were of importance to the older construction worker. They also reported specific work processes that they felt would assist the day to day construction tasks for older workers. These are listed below.

Equipment/materials:

Aim: Reduce manual handling	**Aim: Reduce repetitive motion**
Kerb-lifting equipment	Pre-mix mortar
Reduction in weights/lifting	Power hand tools
Scissor-lifts	Block splitters
Conveyor belts	Long-handle shovels
Vacuum lifters for glass	Pre-fabricated units
Forklifts	
Mini-diggers	
Electrical hoists	
Aim: Postural improvement (stretching/stooping etc.)	**Aim: Reduce noise/vibration**
MEWPS	Plant operatives equipment (e.g. noise reduction)
Ladder assists	PPE
Access equipment (e.g. stair access on scaffolding)	Low-vibration tools
Aim: Compensate for loss of balance	**Aim: Comfort and Hygiene**
Fall arrest equipment	Facilities e.g. toilets and catering

Work processes:

Reorganise the way the work is carried out – skills versus heavy work	Flexible working patterns
Loading out gangs	Self-selection Use work rotation system to avoid repetitive exposure

Barriers – Participants reported barriers to the use of the equipment/work processes.

Attitudes

Some management level participants reported a perception that older workers have a more relaxed attitude to health and safety than younger workers. This is perceived to impact on their decisions to use equipment or follow safe practice.

Craftsmanship

Older workers often reported a dislike of power tools, even though these reduce the need to for repetitive motion. They linked this to a more traditional training as apprentices, being taught to work materials with their hands, often feeling that this gave them more control over the materials to obtain a better quality finish than if they used power tools.

Provision

Participants reported that appropriate tools and equipment are not always provided. Provision of these items depends on several factors – unclear lines of responsibility, budgetary restrictions, individual site/company policy.

Training

Even if a piece of equipment was available to use on site, it was reported by some participants that on some occasions there was inadequate training on the use of that item.

Bonus schemes

The organisation of the pay structure places an emphasis on speed of work. This can often deter workers from using a piece of equipment or following a process that they perceive will slow up the work and reduce their income.

Discussion

This was a limited exploratory study that aimed to identify some of the specific needs of older construction workers and identify the kinds of equipment and work

processes that may improve their working environment. The study also identified some of the perceived barriers to the use of this equipment or the uptake of safe working practice. The first of these relates to attitudes towards health and safety.

Dester and Blockley (1995) maintain that in construction, 'unsafe behaviour' is the most significant factor in the cause of site accidents. Unsafe behaviour, for example failure to use appropriate equipment, may also result in work-related ill-health, which in turn could influence early retirement. Researchers have identified 'sub climates' for safety within an organisation, and it is suggested that age may be one such subgroup (e.g. McDonald and Ryan 1992, Glendon and Litherland 2001). If this is the case, older and younger workers may differ in their attitudes toward health and safety. There is a widely held belief, supported in this study, that older workers have a less positive attitude toward health and safety than younger workers. Many authors suggest a link between employee attitudes to health and safety and safe behaviour (e.g. Schroder 1970). However, many empirical studies do not support the belief that older workers attitude to health and safety is more negative than younger workers. Oi-ling et al. (2003) found older workers to have a more positive attitude to safety, and Cooper and Philips (2004) found older workers to be more responsive to safety training than younger workers. Further empirical work is needed to look specifically at these questions in relation to behavioural outcomes such as construction workers' use of equipment or following safe practice.

Other barriers included the use of bonus schemes to enhance productivity. Such schemes have been shown to reduce the uptake of safe behaviour, and cause workers to carry out unsafe processes (see Sawacha et al. 1999). This barrier does not appear to be specific to older workers, however. Indeed, many of the interventions – i.e. pieces of equipment/materials/work processes reported to assist the older construction worker may not necessarily be age specific. However, they may still help to reduce early retirement form the industry by preventing the onset of chronic ill-health and injury in younger workers. The research that has shown particular interventions to have a more positive effect on older rather than younger workers (Cooper and Philips 2004) raises the issue of whether older and younger workers need age-targeted interventions and is worthy of further study.

In conclusion, participants in this study shared their perceptions about the issues surrounding the use of labour-saving equipment, materials and work processes that they felt were relevant to older construction workers. Although some were age specific, for example the reorganisation of workload to utilise the skills and experience of an older worker e.g. as a safety scaffolder rather than do the heaviest work, many suggestions were appropriate for workers of all ages. For example, employers should do their utmost to ensure the appropriate equipment is available on site and that all workers are trained in its use. By protecting the health of both older and younger workers, early retirement from the industry may be reduced.

References

Arndt, V., Rothenbacher, U., Daniel, U., Zscshenderlein, B., Schuberth, S., Brenner, H. 2005, *Construction work and risk of occupational disability: a*

ten year follow-up of 14,474 male workers. Occupational and Environmental Medicine, 62, 559–566.

Brenner, H., Ahern, W. 2000. *Sickness absence and early retirement on health grounds in the construction industry in Ireland*. Occupational and Environmental Medicine, 57, 615–620.

Cooper, M.D., Phillips, R.A. 2004, *Exploratory analysis of the safety climate and safety behaviour relationship*, Journal of Safety Research, 35, 497–512.

Dester, I., Blockley, D. 1995, *Safety behaviour and culture in construction*, Engineering, Construction and Architectural Management, 1, 17–26.

Glendon, A.I., Litherland, D.K. 2001, *Safety climate factors, group differences and safety behaviour in road construction*, Safety Science, 39, 157–188.

Guberan, E., Usel, M. 1998, *Permanent work incapacity, mortality and survival without work incapacity among occupations and social classes: a cohort study of ageing men in Geneva*. International Journal of Epidemiology, 27, 1026–1032.

Kines, P., Spangenberg, S., Dyreborg, J. 2007, *Prioritizing occupational injury prevention in the construction industry: Injury severity or absence?* Journal of Safety Research, 38, 53–58.

McDonald, N., Ryan, F. 1992, Constraints on the development of safety culture: a preliminary analysis. The Irish Journal of Psychology, 13, 273–281.

McMahan, S., Chikamoto, Y. 2006, *Minimising musculoskeletal discomfort in the workplace: an age based approach*. Californian Journal of Health promotion, 4 (3) 092–102.

Moyers, P.A., Dale, L. 2004, *Challenges and opportunities for the older worker*. Work, 22 (2) 71–78.

Oi-ling, S., Phillips, D.R., Leung, T-W. 2003, *Age differences in safety attitudes and safety performance in Hong Kong construction workers*, Journal of Safety Research, 34, 199–205.

Ropar, K.O., Yeh, D.C. 2007, *Ergonomic solutions for an ageing workforce*, Journal of Facilities Management, 5 (3), 172–178.

Sawacha, E., Naoum, S., Fong, D. 1999, *Factors affecting safety performance on construction sites*, International Journal of Project Management, 17 (5) 309–315.

Schroder, H.M. 1970, *Safety performance measurement*, Journal of Safety Research, 2, 188–195.

Vedder, J., Carey, E., 2005, *A multi-level systems approach for the development of tools, equipment and work processes for the construction industry*, Applied Ergonomics, 36, 471–480.

Vitasalo, J., Era, P., Leskinen, A., Heikennen, E. 1985, Muscular strength profiles and anthropometry in random samples of men aged 31–35, 51–55, and 71–75. Ergonomics 11 (3), 283–290.

COST EFFECTIVENESS OF SAFETY INITIATIVES IN THE HONG KONG CONSTRUCTION INDUSTRY

Steve Rowlinson & Brenda Yip

Dept of Real Estate & Construction, The University of Hong Kong

The construction industry in Hong Kong has made a concerted effort over the past ten years to improve its performance in terms of safety and the initiatives introduced during this period, over 20 different initiatives, appear to have been successful. However, it would be useful to assess which initiatives have been more successful and so focus attention on enhancing these as the accident rate appears to have plateaued and a "kick start" is needed to enable the continuation of the downward trend of the past ten years. Thus, this paper reports on research aimed at identifying the most effective safety initiatives and also investigates the factors within organisations which aid or hinder these initiatives.

Introduction

This paper reviews the current situation with regard to the improvements in construction site safety performance in Hong Kong that have taken place over the past ten years and the authors identify those safety initiatives which have been most effective in performance improvement (see, for example, Rowlinson, 2003 & 2004). The approach to this research project is briefly described below.

Problem

Implementation of cost effective safety initiatives in Hong Kong at a time when construction workload has been in a steady decline:

- throughout a range of projects and organizations;
- with a focus on reducing serious accident rates.

Context

An underlying theme running throughout the study has been the need for empowerment of the industry to manage itself effectively. In order to do this many industry participants have recognized the need for capability development within individual firms and development of a sustainable industry. As far as safety is concerned, initiatives have been, in the main, very successful on larger, public projects but a plateau in the improvement trend has now been reached. In order to further reduce

accident rates a different approach needs to be adopted and the industry as a whole needs to examine the skills required industry wide for further successful accident rate reduction (Lingard & Rowlinson, 2005).

Rationale

The rationale behind the recommendations in this research is:

- successful accident prevention requires a change of mind set – a culture change;
- clients and consultants must collaborate with the contracting side to drive change;
- a fit is required between safety initiatives and capability, leading to improvement;
- a maturity gap has been identified between high performing companies and the rest.

Objectives

The objectives of this study are to advise on effective strategies for all sectors and sizes within the industry. In order to achieve this the following need to be addressed:

- an audit is required of the current situation and problems identified and targeted;
- capabilities and costs need to be identified and change focused on effectiveness;
- change needs to be implemented at all levels and in all sectors of the real estate and construction industry.

Current status

The following sections report on the current status of the study.

Cost effectiveness study

This study has given ambiguous results. For example, respondents identified the green card system, an accreditation mechanism for assuring a minimum level of safety knowledge amongst workers, as the least effective safety initiative. However, the green card system was also seen as being relatively the cheapest initiative. Hence, when the study team calculated the cost effectiveness of various initiatives the green card scheme came out as the most cost effective. Obviously, this cannot be an acceptable result. Hence, the study team are looking at alternative ways of addressing the issue of cost effectiveness in order to produce a sensible answer to this question. If cost effectiveness cannot be readily defined in a sensible manner then one of the objectives of the study cannot be achieved. Indeed, many of the respondents have expressed the view that cost effectiveness is not an appropriate measure of safety initiative performance.

Suffice it to say that the most effective safety measures identified in the study were as follows (from the highest to the lowest rank):

- Safety Management System
- Pay for Safety Scheme

- Performance Assessment Scoring System
- Contractor Performance Index System
- F&IU Ordinance and Regulations and Occupational Safety and Health Ordinance and Regulations
- Automatic Suspension from Tendering System
- Safety Supervision Plan System
- Green Card System

Benchmarking study

The results of a benchmarking study of contractors were completed in early October, 2007. The results gave rise to a number of questions from those contractors which had participated in the study. The most striking result was that many respondents, even high performing respondents, failed to perform adequately on the dimension of "project specific objectives". Indeed, there was only one respondent organisation which could be classified as a good performer and over 40% of respondents were classed as poor performers. This contrasts poorly with a Japanese study undertaken previously whereby over 40% of respondents were classed as good. As this study dealt with the role of the safety management system within the business this is an area which needs to be addressed by the industry. Indeed, feedback indicates that this is a structural problem inherent in the construction process which needs to be addressed by the industry as a whole, particularly the developers who drive the process (see, for example, Mohamed, 2003). The research team recommends that a similar study be conducted through Real Estate Developers Association in order to identify commonalities and mismatches between the supply and demand side of the industry.

Results from comprehensive survey of industry participants

The map which appears in Figure 1 (at the end of this paper) indicates the range of issues which surfaced during the detailed study of perceptions of safety initiatives and safety management in Hong Kong. Each of these will be dealt with briefly here but two underlying issues need to be highlighted in order to put the issues into context.

Maturity is an important issue

It became apparent during the study that a major issue for the industry was the maturity of organisations. If we were to take a five level maturity system then one might argue that only the top ten contractors could be classed as level five maturity. Indeed, much of the industry is seen to be of a maturity level one or two whereby their main focus in terms of safety management is compliance as opposed to the continuous improvement philosophy of the top contractors. Such a view could also be applied to the private developers, in that perhaps less than ten of these can be seen to be adopting a mature attitude to safety management on their projects. Hence, in terms of future strategy this leads us to the situation whereby a three pronged

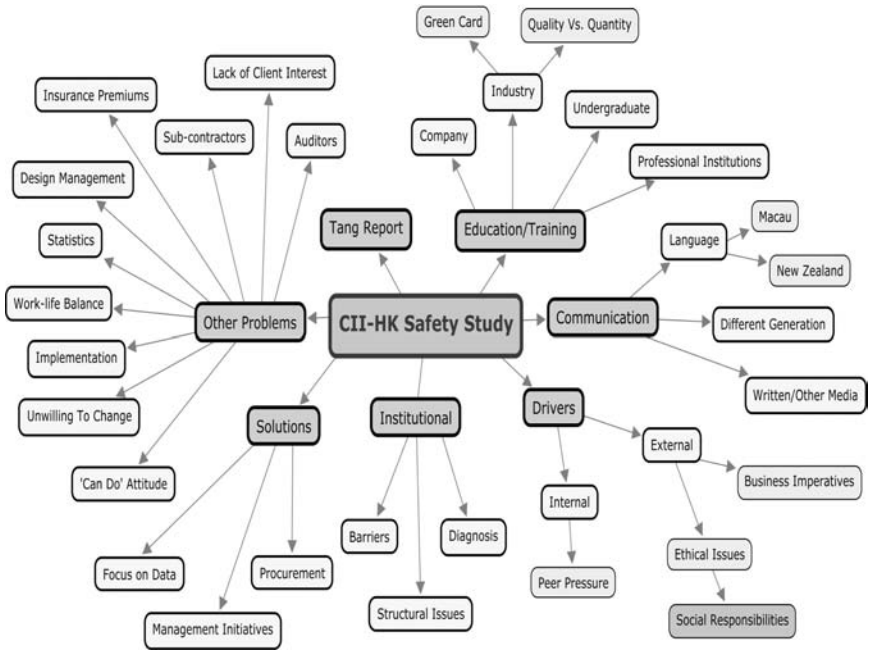


Figure 1. Map of Hong Kong safety issues.

Table 1. Maturity of organisations.

Level	Nature	Commitment	Focus
Level 1 – Initial	Immature	Compliance	Compliance
Level 2 – Repeatable [Managed]	Maturing	Continuance	Norm
Level 3 – Defined		Normative	
Level 4 – Quantitatively Managed	Mature	Affective	Continuous Improvement
Level 5 – Optimizing			

(Source: adapted from the Capability Maturity Model: http://valuebasedmanagement.net/methods_cmm.html).

approach to improving accident rates is necessary. Briefly, this means that three different foci of attention should be provided for those contractors, and developers, falling into the three categories of immature, maturing and mature.

Focus should be management not legislation

There was an underlying agreement amongst interviewees that Hong Kong had enacted sufficient legislation to deal with safety management and that the industry now needed to focus much more effectively on management of existing systems rather than the enactment of more legislation. Indeed, this fits in with the safety effectiveness study results which indicated that the enactment of safety management legislation was in fact the catalyst for better management of safety processes within

the industry. The focus now needs to be on implementing more effectively the safety management systems which already exist and educating the majority of the industry that a compliance approach, based on conforming with legislation, is not acceptable as an effective approach to safety management.

Key findings in brief

The current green card system is not effective

This finding was reported from a number of sources and the problem is twofold. Firstly, the syllabus of the green card accreditation needs to be reviewed so that it may be more carefully structured and address skill and safety issues at a more detailed level. Secondly, there is an over provision of underperforming providers of accreditation courses and this issue needs to be addressed by the industry as a whole.

Communication is a problem in safety improvement

It was reported by a number of respondents that communicating safety management ideas, procedures and instructions is often a difficult task. A major issue here is the overwhelming of middle management with documentation and systems whereas at the worker level it is the problem of maintaining a consistently focused message to the workers. It was believed that language wasn't so much a barrier as the means of communication. There was also a belief that the ageing of the workforce and ingrained ideas and attitudes were difficult to change because of this demography.

Drivers for improvement are opposing:

Business issues
Ethical issues
Social responsibility

The forces driving change in the industry are opposing. The overriding principle of business survival is in conflict with the ethical issues of occupational health and safety. However, change is taking place and the move towards corporate social responsibility is an important change in the attitudes and maturity of the participants in the Hong Kong industry. Indeed, one mechanism for promoting a more mature attitude might be to adopt SA 8000. SA stands for "Social Accountability" and SA8000 is developed as a reference standard in the context of human and social rights. This standard can be applied to all organisations, regardless of their sectors of activity.

Institutional barriers exist to a concerted effort

The construction industry is a cacophony of competing voices with conflicting interests. A major barrier to a concerted effort on safety initiatives is the lack of one voice with which the industry can speak. There are different problems at different levels within the industry and an industry-wide body which can take an

independent and considered view on all of the competing issues is a necessity. The newly formed Construction Industry Council has responsibility for addressing this issue. However, government, through Works Bureau and Housing Authority, can also help to address this issue but there is a need to bring in diverse interests such as the insurance industry and Labour Department if the effort is to be effective.

Education and training at tertiary and professional levels are inadequate

The coverage of occupational health and safety issues at tertiary institutions can be improved quite considerably. A whole series of professions are educated in Hong Kong's tertiary institutions, such as civil engineers, builders, facility managers, building services engineers, structural engineers, architects, etc. and there is no common syllabus in terms of occupational health and safety and there are no courses designed to integrate the design and management of construction projects excepting MIDIM at Hong Kong University. This is an issue which should be addressed by both the universities and the professions. It is now a cause for failure in ICE (UK) professional assessments if the candidate does not show adequate knowledge of occupational safety and health. Such an approach should be mandated for all of the professions within Hong Kong. For further details and examples, refer to "A Construction Safety Competency Framework: Improving OH&S performance by creating and maintaining a safety culture", CRC Construction Innovation, Australia.

Potential issues which can be used as drivers for improvements

The sections below give a brief indication of the issues involved in driving improvements in safety management systems in Hong Kong.

Corporate social responsibility

This is now a very important issue for listed companies and organizations such as China Light and Power provide excellent examples of how CSR can be used to promote the occupational health and safety message to employees and their families. Indeed, CLP might be considered as an exemplar developer.

Developer attitude

Following on from the previous point, it is important to recognize that developers as well as contractors exhibit different levels of maturity. Such a problem needs to be addressed at an institutional level in that the organizations representing developers and government departments interacting with them need to lay down basic principles and procedures they are expected to follow.

Client insurance

The insurance industry plays an important role for the real estate and construction industries. However, the Insurance Industry Ordinance does not allow for the active and comprehensive sharing of information on construction industry performance.

Hence, an experience rating modification system is difficult to implement in Hong Kong at the present time. This is an institutional barrier to progress which would allow better performing contractors to experience lower premiums and so higher competitiveness. Indeed, one mechanism for addressing this might be to put the insurance in the hands of the clients and so focus clients' attention on occupational health and safety management.

Design management

Design management is an important issue as far as occupational health and safety performance is concerned. Up to 60% of accidents on site can be attributed to design decisions. Hence, there must be an initiative developed in this area if the more mature contractors are to be facilitated in improving their OHS performance. The CDM regulations in the UK have been shown to be ineffective and so it is recommended that an alternative approach be considered.

Work life balance

Work life balance has been shown to be an important moderator of management effectiveness and the presence of stress in the workplace. Anecdotal evidence from Macau has indicated that an improved work life balance leads to more effective working and safer construction. This is an area which needs to be further investigated and is an area that Hong Kong University propose as a new research project. Lingard et al. (2006) and Yip and Rowlinson (2006) have indicated the positive effects of an improvement in work-life balance.

"Can do" attitude

One of Hong Kong construction industry's distinctive competences is its "can do" attitude. The ability to construct high-rise buildings on four day floor cycles cannot be matched in many places worldwide. However, this "can do" attitude comes at the cost of flooding sites with plant and equipment and a focus on long working hours in arduous conditions leading to stress all round. This is an issue which needs to be reviewed and the industry needs to be educated to take a more mature attitude to this problem. The ability to say no to unreasonable client demands for speed needs to be developed and the Housing Authority's initiatives in the 1990s in this area are an excellent, successful example.

Use of data

The industry as a whole and the Labour Department in particular collect a massive amount of data on construction site accidents and their effects. A program should be put in place to make better use of this data in order to inform contractors and developers of trends in accident causation. This work should be let competitively to an organisation outside of the Labour Department in order to ensure an unbiased opinion. Examples of issues which have been identified but not so far addressed are

the occurrence of a high rate of accidents in the summer months and the existence of two peaks in accident occurrence at different times of the day.

Procurement systems

Many of the recommendations indicated in this paper are predicated on a more collaborative and cooperative approach to procurement. Indeed, without a relationship management approach to design and construction it is impossible to further reduce the current accident rates. Hence, a move towards more innovative and collaborative procurement systems is essential in this respect (Walker & Rowlinson, 2007).

Relationship management

The process of planning, design, construction and facility management is highly complex, recursive and interdependent. However, we still maintain rigid structures for managing these processes in a separate, sequential manner. Such an approach is no longer acceptable and there must be a move to a new paradigm of relationship management across the whole development process. This will, of course, incur transaction costs but the benefits will far and away exceed these costs.

Honest auditing – focus on improvement

Independent auditors are placed in an ambiguous position in that they strive to provide honest audits and yet are under pressure to ensure that their auditees actually achieve a passing grade. Hence, there is a tendency to underplay faults in audited safety management systems in order to address this ambiguity. However, the mature contractors and developers expect to be given feedback from audits which will allow continuous improvement within their organizations. Thus, there needs to be a careful review of the existing system to protect the integrity of the auditors and provide best value for the auditees.

References

Lingard, H. Yip B., Rowlinson S. and Kvan T. (2006), The experience of burnout among future construction professionals: a cross-national study. *Construction Management and Economics*, **24**, 11, 10–12, November.

Lingard, H.C. and Rowlinson, S. (2005), Occupational Health and Safety in Construction Project Management, Taylor & Francis Books Ltd., United Kingdom.

Mohamed, S. (2003), Scorecard approach to benchmarking organizational safety culture in construction. *J. of Construction Engineering and Management*, American Society of Civil Engineers, **129**, 1, 80–88.

Rowlinson, S. (2004), (ed.), *Construction Safety Management Systems*, Taylor & Francis Books Ltd., United Kingdom.

Rowlinson, S. (2003), *Hong Kong Construction – Safety Management and the Law*, Sweet & Maxwell, Hong Kong.

Rowlinson, S., Mohamed, S. and Lam, S.W. (2003), Hong Kong Construction Foremen's Safety Responsibilities: A Case study of Management Oversight, *Engineering, Construction and Architectural Management*, Blackwell Science: Oxford, United Kingdom, **10**, 1.

Walker, D.H.T. and Rowlinson, S., Eds. (2007), *Procurement Systems – A Cross Industry Project Management Perspective*. Series Procurement Systems – A Project Management Perspective. London, Taylor & Francis.

Yip, B. and Rowlinson, S. (2006), Coping Strategies among Construction Professionals: Cognitive and Behavioural Efforts to manage Job Stressors. *Journal for Education in the Built Environment*, **1**, 2, 70–79.

HCI

HUMAN-COMPUTER INTERACTION: THE ELEPHANT IN THE CORNER

Hugh David

R+D Hastings

Although ergonomists have produced much valuable and valid data on the interaction between humans and computers, there has been surprisingly little attention paid to the 'Elephant in the Corner' – the PC (or Apple) without which few people could work effectively today. This paper summarises the development of the classic interface and discusses some of the choices that were made on the way, and how they were made. It looks at the way the personal computer has changed in response to perceived needs (not necessarily those of the user), the transition from a numerical processing device to a text handling device to an image-handling device, and potentially a 'process handling device'. It discusses the separation of 'hardware' from 'software', the development of Information Technology as an elementary life skill, the transition from the 'stand-alone' computer to the Internet-linked and Web-integrated system and the possible consequences of the contemporary trend to an increased role for 'intuitive' designers, as opposed to 'scientific' experimenters in the development of future systems. Finally, it offers a few reproaches, warnings and suggestions for ergonomists from the past for the future.

Introduction

The PC (Personal Computer) as we know it, originated as a relatively small, relatively cheap machine designed to do what large computers did, using the first integrated circuits, which were enormously faster and cheaper than the separate solid-state transistors that had replaced the slow, delicate and expensive thermionic valves.

Origins

The first machines relied on the existing ideas of how a computer should work, using a 'Teletype' approach for interaction with the machine. The Commodore, the TRS 80 and the Sinclair ZX were primarily character-oriented devices, using monochrome CRTs to display a few lines of about forty characters each. The purchaser was expected to write programs, usually in BASIC (Beginners' All-Purpose Symbolic Instructional Code), although some programs were available, usually on magnetic tape. Although these machines appealed to a relatively small proportion

of the population, they caused some concern about the intensity of their appeal (Shotton (1989) is a study of 'computer obsessives'). Although computer obsession can be identified right back to the time of Colossus, it attracted little notice while it was confined to universities and research establishments. The IBM-PC was designed by a small group in a 'skunk-works' at a remote IBM plant. Because they were a small underfunded group they bought most of their components off the shelf rather than designing their own.. (Initially, the IBM-PC had neither a mouse nor a desktop display.). The IBM designers did not develop their own operating system, but licensed one from Bill Gates, who bought it from someone else. PC-DOS and MS-DOS were line based command systems, requiring the user to memorise the range of commands available, and almost as unforgiving as the dreaded mainframe IBM JCL (Job Control Language). Significantly, they did not restrict the use of their architecture, enabling PC-compatibles galore to flood on to the market. Hardware developers were able to devise specialised input and output devices, from overhead projectors to scanners to knitting machines. Equally significantly, the free availability of operating system specifications offered free rein to developers of all sorts, producing speech recognition, text scanning, editing, spreadsheets and a host of specialised packages.

The mouse, screen and pointer interface

An alternative approach developed into the screen-based systems we all know and may or may not love. The XEROX Palo Alto Research Centre produced a networked office system (Alto) in which documents were stored, retrieved, edited and transferred between workstations electronically. Although Alto introduced the mouse, the rectangular window, the scrollbar, the pushbutton, the 'drop-down menu', the Ethernet and the laser printer, it was not intended as a stand-alone personal computer. It was a commercial failure. It was complex, slow and very expensive. XEROX was primarily a copy machine company, and, although it took up the laser printer with some enthusiasm, was not interested in the computing business.

Apple

Steve Jobs created a company with some of the PARC staff, which produced a stand-alone computer called Lisa. It also failed because it was slow and expensive. Jobs set up a 'skunk-works' inside Apple which came up with the Macintosh. The process by which the Macintosh was defined was not formally recorded at the time, and reminiscences of the designers (Moggridge 2007) show that there was little systematic user requirement analysis, little formal analysis and little testing.

According to Larry Tesler (2003) Bill Atkinson developed the pull-down menu system in one night in 1984. Douglas Englebart (2003), who had invented the mouse in 1964 was inspired by a two-axis map-measuring device. Only later was it realised that the mouse was an inverted trackball, long used in radar installations.

The question of how many buttons should be associated with the mouse was originally determined by the difficulty in training people to use an unfamiliar device. Inside Apple there was a heated discussion about how many buttons there should

be on the mouse. Eventually, the one button mouse was adopted mainly because the marketing people felt that it would be easier to learn.

Englebart also developed a five-key chord keyboard for control that used a language called NLS (oN Line System), which took at least a week to learn the minimum, and several months to understand thoroughly.

Larry Tesler (2003) at Apple was developing text-editing software, supposing that the computer could show a screen of text, and that the users could use a mouse to point to places on the screen, and click buttons to communicate what they wanted to do. For example, one button selected a cursor location, a second a word and a third a sentence. Tim Mott came in one day and suggested using a single click to select to select a place, a double click to select a word and a triple click to select a sentence. Tesler closed his eyes and tried to envisage it. 'Yes, it feels just right. Double click to select a word'. He could not envisage a triple click, so the world was spared that. A few secretaries were consulted, but no one even thought of considering elderly or handicapped users, or of conducting any formal trials. In addition to being fast and moderately expensive, the Macintosh was designed by designers.

Apple computers were, and are, relatively expensive and beautifully designed. Apple decided to maintain control of its systems, and nearly bankrupted itself trying to prevent Microsoft from using the types of interface that Apple had developed. Apple has survived by developing superbly designed hardware and software in the personal computing and other digital electronic fields, and maintaining complete control of their products.

Microsoft

Meanwhile, Microsoft was developing Windows. Versions 1 and 2 were not successful, but Windows 3 (becoming 3.1) took off and became a standard. Microsoft decided to develop a mouse and carried out an extensive series of design exercises, user consultations and trials, modifying the design in the light of their experiences. Microsoft decided on a two-button mouse which is now the industry standard. (SUN Microsystems and the UNIX community chose to use three buttons.)

Moore's Law (that the power of computers doubles every eighteen months) meant that it became increasingly easy to provide a high-level visual interface on low-level (cheap) hardware. The command-line based systems were made obsolete by page-based systems. The availability of a computer display meant that users could be offered a choice between different actions, rather than having to remember codes or computer languages.

As computers became faster, memory became cheaper and video displays became bigger, the size and quality of the image that could be presented to the user became correspondingly more detailed. Most software became ever more complicated, adding features which most users never used. The sheer size of a software package was taken as evidence of its value, although much of the size was made up of digitised images of pens and pencils or other junk. (Typically, 'bloatware' is sold on a CD-ROM in a virtually empty package the size of a dictionary.)

Contemporary interface design

Generally, the windowing process is now the accepted method of displaying information, and designers have made considerable efforts to standardise and smooth the interaction between the user and the system. Tidwell (2006) is a comprehensive and insightful tutorial on the design of interfaces, mostly for conventional flat-page designs. Designing for the disabled is mentioned in a few places, but not treated in any detail.

Cooper *et al* (2007) is a rather deeper and wider ranging review of interaction design. It is a wide-ranging work, summarised in 80 aphorisms scattered throughout 610 pages. These range from the epigrammatic – 'Do, don't ask' to the gnomic – 'Sovereign applications should exploit rich input'. Considerable effort is devoted to the identification and definition of the potential users of the interface, both for their individual differences and their changing relation with the system as they become more familiar with it.

These authors are designers, drawing inspiration from the work of J.J. Gibson (1979). Gibson, who had studied driver and pilot behaviour from 1938 onwards, developed theories of the way in which the physical characteristics of the environment influenced the behaviour of the operator. He introduced the term 'affordance' for the characteristics of a device that suggest how it should be operated. Unfortunately, he also referred to this field as 'ecology', leading to some confusion and suspicion among his potential audience. Donald Norman (2002) provides a more accessible explanation of the general concept, which has much in common with fundamental ergonomic principles.

Most PC operating systems use an implicit metaphor of filing cabinets and of files containing files containing documents of multiple pages. This was helpful when office staff were being persuaded to move to computer systems. The metaphor is now being reversed, since most people encounter files on computers before they encounter them in offices. The discovery that a filing cabinet has no 'search' facility can come as a severe shock to the modern office junior. Contemporary designers are looking at alternative metaphors, including dynamic graphics and 'Information Dashboards' (Few, 2006) which provide real-time on-line images representing the processes taking place inside a company, an office a factory, a transport system or a network of distributed facilities. Spence (2001) provides a broader outlook with historical perspective and unorthodox visual structures divorced from the traditional page and file metaphor.

The internet

Although the Alto, the ancestor of all modern personal computers, was designed to be a network, early Personal Computers were self-contained devices. The advent of the Internet transformed the public view of computers. The Internet itself developed in an intentionally unplanned way, from the ARPAnet, a US Department of Defense project, which linked mainframe computers using multiple packet-switching pathways, specifically to keep the system functioning in the event of a nuclear war. The original Internet operated essentially as a telegraph system, exchanging text messages. Rapid technical development, associated with the introduction of digital

technology into telephony, led to the introduction of the World Wide Web, providing access to ever-increasing volumes of information, not only in type format, but also in images of increasing detail. Digital photography and increasingly sophisticated digital data compression algorithms made it easy to send and receive pictures via the computer. More important socially, it transformed the world of computer gaming, from being the province of social isolates to being a source of social status. Gordon (2007) noticed in 1987 that 'every junior high school was starting to have kids who were opinion leaders about games . . . and that conferred as much status as being the quarterback on the football team'. On a darker note, increased communication possibilities have provided opportunities for abusers of confidence of both children and adults. It can, however, be argued that such abusers existed long before the Internet, and that technical solutions can be provided to these technical problems, as the authorities become more sophisticated and technically adept themselves.

As the Internet has become more widespread, huge bodies of accessible information and technical expertise are becoming more easily available. 'E-Commerce' continues to grow steadily, eliminating much physical travel, and generating unease among the pundits. The BBC and other broadcasters routinely make their output available through the Internet, so that their discussion groups receive contributions from South America and Siberia. Internet telephony and visual transmission provides direct information from trouble spots, wherever a wireless link is available – which is practically everywhere. Computer software is downloaded and updated, often without specific instruction from the user. A recent development, producing more routine disquiet from the pundits, is the social group, such as Facebook or Second Life, which can become more real than real life in some instances.

Generally, computer packages and services are now available in complete form, requiring only the minimum skill to enter data or commands. IT (Information Technology) courses concentrate on how to use the packages provided by some remote producers, who may have little or no knowledge of who is using their product.

'Programming' as a user activity has virtually disappeared. 'Source code' in an accessible form is rarely available for commercial products, so that the correct functioning of, say, a statistical analysis package must normally be taken on trust. For the most part this is an acceptable situation, but the possibility of rogue programmers, disgruntled employees or simple mistakes cannot be entirely ruled out.

The future

Moggridge (2007) includes interviews with a number of contemporary designers who are developing future systems and devices. Most seem to be motivated by a desire to create objects to provide entertainment rather than serve simple commercial purposes. Some of them are concerned with devising different ways of interacting with a computer to provide emotional responses. There is an undercurrent of emotional rather than rational response.

There appears to be a need for simpler ways to instruct computers and to interact with computer-controlled devices. Some sort of cross between BASIC and Meccano would greatly simplify the lives of these designers and others.

The existing mouse-pointer-display loop may be capable of improvement. Small children find difficulty in moving a mouse without looking at it, and there may be a case for a direct touch interface. The average human finger is a blunt instrument, but it may be possible to develop some method of touching a screen that gives sufficient precision without inducing the equivalent of Repetitive Strain Injury or 'Mouse Arm'.

A more prosaic development will be a standardisation of the multifarious remote control boxes that litter our domestic environment. It will involve largely organisational effort, similar to that involved in allocating unique bar codes or mobile phone numbers, and a common language for the identification of devices controlled and actions to be taken. The existing X10 protocol for the remote control of lights and domestic devices provides a precursor.

Generally, we may expect to see steadily more integrated systems, fewer separate identification systems and an ever-increasing richness of potential interactions mostly ignored by most people.

The role of the ergonomist

Where do ergonomists come into this? Historically, some ergonomists have had some influence on the development and validation of systems, but most development has been, and remains, in the hands of designers, who have very different priorities to ergonomists. Design has been, since the 1930's subject to, and sometimes seen as a branch of, sales (Krippendorf, 2006). Krippendorf declares that a semantic turn is in progress, away from the design of goods for sale and towards the design of devices that fit human beings. Although he dismisses ergonomics as reducing human beings to mere functional components in systems, and unreliable ones at that, his new approach to design has much in common with cognitive ergonomics. Krippendorf, in common with many other designers, is not interested in 'scientific' assessments. Designers appear to rely on 'the market' to evaluate their products. Ergonomists traditionally lament that they are called in to rectify mistakes that could have been avoided if they were involved from the outset, and that, like the quality control inspector, they are seen as causing the faults they identify.

Conclusion

The Ergonomics Society and individual ergonomists should, and should have already, approached the design profession to show how we can combine our strengths rather than sniping at each other's weaknesses.

References

Cooper, A., Reimann, R. and Cronin, D., 2007, *About Face 3*, (Wiley)
Englebart, D., 2007, in Moggridge *Designing Interactions*, 24–37
Few, S., 2006, *Information Dashboard Design*, (O'Reilly)

Gibson, J.J., 1979, *The Ecological Approach to Visual Perception*, (Houghton Mifflin)
Gordon, B., 2007 in Moggridge *Designing Interactions*, 322–336
Krippendorf, K., 2006, *The Semantic Turn – A New Foundation for Design*, (Taylor and Francis)
Moggridge, B., 2007, *Designing Interactions*, (MIT Press)
Norman, D.A., 2002, *The Design of Everyday Things*, (Basic Books)
Shotton, M. A., 1989, *Computer Addiction?* (Taylor and Francis)
Tesler, L., 2003 in Moggridge *Designing Interactions*,56–72
Tidwell, J., 2006, *Designing Interfaces*, (O'Reilly)

THE CULTURAL ISSUES IN ICON DESIGN

Hsiu-Feng Wang

Department of Commercial Design, Ming Chuan University, Gwei-Shan, Taoyuan County, Taiwan

The aim of this research is to collect the opinions of Taiwanese computer users with regard to computer icons and their appropriateness. The method we used in this research was brainstorming technique. We asked participants to draw what they consider would be appropriate icons for Taiwanese computer users. After drawing their icons, participants are then asked to discuss the reasons for their drawing. This paper offers confirmation that the participant believes that the drawn icon is appropriate and gathers information as to what aspects make it appropriate.

Introduction

Interaction between cultures has occurred throughout history but never on the scale that we see today. However, computers have brought challenges in addition to opportunities. Problems due to cultural differences have not limited themselves to the e-mails sent between computer users of different cultures but have also been occurring when a user of one culture uses software developed by persons of another (Aykin, 2001). As one would expect, these problems have led to an increase in research by industry into culture and cultural issues.

A successful technique used in computer icon design is to take advantage of a user's existing knowledge and to use metaphors. Thus it follows that as life experiences, common everyday objects and meanings attached to objects often differ from country to country, and interface designers should use metaphors and objects from the target computer users' environment.

Taiwan is one of the most intensive Internet using countries in the world (Nielsen, 2000); the number of regular users stood at 10 million at 2007, which is 44% of the total Taiwanese population (www.find.org.tw). Previous research (Chen and Chen, 2000) suggests that some Taiwanese computer users find certain cultural images easier to recognise than their standard icon counterparts and that they benefit from the use of Chinese text in icon designs. Therefore, in order to develop more appropriate icons for Taiwanese computer users, this research collects the opinions of Taiwanese computer users with regard to icon appropriateness.

Most researches in the field of icon design related to icon design are using empirical research to evaluate the appropriateness of icon design (Wang, 2007). However, interface designers do not know the needs before they design the icons for their target users. Horton (1994) stated that: 'even obviously wrong ideas can

provide a path to better ideas' (Icon book). The purpose of brainstorming is to elicit as many ideas as possible. Participants have been encouraged to offer new ideas as well as to add to earlier ideas. This will form a clearer picture of what attributes Taiwanese computer users find appropriate. This information will form the framework for a set of requirements to act as a starting point in the development of appropriate icons. In order to achieve this, information is collected through a drawing exercise and followed by a semi-structured interview.

In this paper, a standard icon is described as an icon that is found in international versions of software such as Microsoft Office. A cultural icon is an icon that has cultural, national, or local features.

Method

Participants

This research involved 10 Taiwanese participants aged 25 to 35. All the participants came from an art/design background and were chosen because they were accustomed to communicating ideas visually. Participants were told that they were contributing to research on the development of icon design and that it was the idea or meaning expressed in the drawing that was important, not the quality of the drawing.

All participants have similar English ability. According to their computer experience we divided participants into two groups: Group A, participants familiar with computers; and Group B, participants not familiar with computers.

Procedure

Prior to the task, participants in Group B were given an explanation of what computer icons are and how they are used to obtain desired results. A collection of icons, some cultural and some standard, were shown to aid understanding. The collection of icons shown did not relate to any of the referents that were to be used in this experiment.

The task involved participants individually drawing what they believed were the most appropriate icons for given referents. Referents were issued to each participant one at a time in a random order. The referents were: home, e-mail, stop, notepad, calculator, find, italic, bold, underline, type/text/font, font colour, text direction.

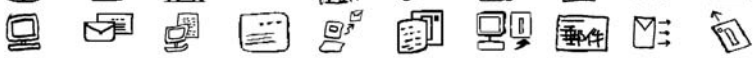
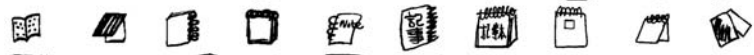
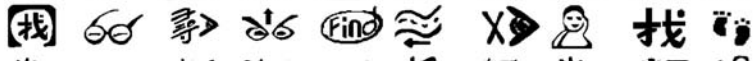
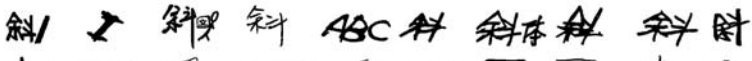
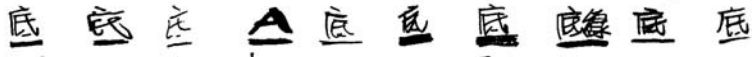
The instruction given when a referent was issued to a participant is given below using the *stop* referent as an example:

Instruction: Please draw what you consider to be the most appropriate icon for *stop*.

The *home* referent was issued using both its Chinese and English names as the referent *home* is also known as 首頁 in Taiwan (which in English translates as 'first page'). Each participant was told to take as much time as they needed to complete their drawings.

After the participants had drawn the icons, the experimenter asked them, in a semi-structured interview, the reasons behind their drawing.

Table 1. Participants' drawings corresponding to the given referents.

	Group A Familiar with computers					**GroupB** Not familiar with computers				
Referents	LL	JL	YW	JH	EC	CW	CH	HC	MC	CY
Home										
E-mail										
Stop										
Notepad										
Calculator										
Find										
Bold										
Italic										
Type/Text/ Font										
Underline										
Font colour										
Text direction										

Results and discussion

Table 1 shows each participant's drawings with respect to the given referents. Analyses of the drawings follow the table and summarize the manuscript from the semi-structured interview.

Use of text

All participants in Group B used Chinese characters, often with other image elements, in their icon designs for the referents: *italic, bold, underline, type/text/font, font colour* and *text direction.* In addition, at least one participant in Group B used Chinese characters, often with other image elements, in the icon designs for the referents: *home, e-mail, stop, notepad* and *find*. No participant in Group B used either English letters or English words in any of their icon designs. The exclusive use of Chinese text by participants in Group B is interesting as all alphabetic *standard* icons for the given referents used in the task show English letters.

In Group A, participants LL and YW exclusively used Chinese characters when they used text in their icon designs. The remainder of Group A when using text in their icon designs sometimes used Chinese characters and sometimes English text; no icon drawn by Group A was designed showing both Chinese characters and English text together. All the participants who used English text did so more than

Figure 1. (a) Icon drawn by participant JH for the *underline* referent (b) The standard icon.

once. Participant EC, the participant who used English text the most often in the icon designs, used it in half of her icon designs.

Interestingly, when English letters were used by Group A, the letters were not always the same as those found in the standard icon for the same referent. This occurred even though it seems that some participants were trying to replicate the standard icon designs for certain given referents. For example, participant JH who is familiar with computers drew an icon that showed a capital letter 'A' with a thick line underneath it for the referent *underline* (see Figure 1a). In effect, the icon drawn was that of the standard icon for *font colour* (see Figure 1b). These confusions occurred because the participants could not remember the exact representations used by the standard icons they were trying to draw. In the case of participant JH, she seemed to remember that the standard *underline* icon showed an underlined English letter but not the exact letter itself. The subject of memorisation is discussed by Norman (1999) who states that: 'it is a general property of memory that we store only partial descriptions of the things to be remembered, descriptions that are sufficiently precise to work at the time something is learned….'. Participant JH, although aware that the standard *underline* icon showed a letter, did not remember what it was, probably because she did not use the letter to distinguish the icon from the standard *bold* icon and the standard *italic* icon (all three icons show different letters and generally appear alongside each other in Microsoft Word). Instead, it is likely that she identified the standard *underline* icon using the underline and hence only remembered this and not what the letter was.

A number of referents were repeatedly given the same text elements by participants. Looking at the combined results of Group A and Group B, seven participants used the Chinese character '斜' (which in English translates as 'slant' or 'italic') to represent the referent *italic*. Eight participants used the Chinese character '粗' (which in English translates as 'thick' or 'bold') to represent the referent *bold*. Eight participants used the Chinese character '底' (which in English translates as 'under' or 'underline') to represent the referent *underline*, and seven participants used the Chinese character '字' (which in English translates as 'word' or 'Chinese character') to represent the referent *text direction.*

Icons that showed Chinese text were deemed more appropriate than icons that showed English text. Apart from one participant in Group A (participant JL), every participant chose cultural icons above standard icons for all the referents that were represented as alphabetic icons. In other words, regardless of their familiarity with using computers, participants thought icons showing Chinese text were more appropriate than icons showing English text. Many participants commented that they did not know what the letters used in the alphabetic *standard* icons stood for; however alphabetic *cultural* icons seemed to be easily understood. One participant familiar with computers (participant LL) stated that she could not say for certain that the

standard *bold* icon shown was the standard *bold* icon because it had not been presented next to other alphabetic icons. She said: 'When they're next to each other (the standard text formatting icons, e.g. italic, bold and underline) you can tell which is the icon for *bold* as it has a thicker letter.'

Translation of referent

One factor that seems to have had an impact on the design of the icon for the *home* referent is that two names are used for it in Taiwan. The results show that the majority of participants in Group A decided to base their icon designs on the referent's English name *home*, whereas the majority of Group B based their icon designs on the referent's Chinese name 首頁 (which means *first page* in English).

Style of imagery

Participants in Group B used traditional Chinese imagery more often in their icon designs than participants in Group A. For example, three participants of Group B (participants CW, CH and CY) showed a traditional Chinese envelope, a thin width envelope addressed in portrait orientation, for the referent *e-mail*. Although three participants of Group A (participants JL, JH and EC) also used envelopes in their icon designs for the referent *e-mail*, they chose not to use traditional Chinese envelopes. Another example of participants in Group B using traditional Chinese imagery more than participants in Group A can be seen with the *calculator* referent. The majority of participants in Group B decided to use a traditional Chinese abacus to illustrate the referent whereas all but one of the participants in Group A decided to use an electronic calculator. One reason Group A might have used less traditional Chinese imagery than Group B could be that they have become familiar with the Western image elements employed by most software manufacturers. It is possible that this familiarity led to participants sometimes choosing the same, or similar, image elements as those used in the icons they use on a regular basis, without questioning whether the image elements were the most appropriate.

Choice of images used

Some of the referents that expressed abstract concepts resulted in a large number of different image elements being used to represent the referent. For example, the referent *find* was illustrated by, amongst other things: a tunnel and the Chinese character 找 meaning 'find' (participant LL), a pair of glasses (participant JL), a pair of eyes (participant JH), a person looking for something (participant HC), the English word 'find' (participant EC), and a picture of some footprints (participant CY). Conversely the referent *notepad* was represented by a notepad by all participants. It should be noted, however, that in some cases participants supplemented their image of a notepad with the words: 'note' or 'notepad' written in English or Chinese.

The information gained in this research will help in the formation of a set of requirements to act as a starting point in the development of appropriate icons for Taiwanese computer users.

References

Aykin, N. 2001, Legal issues reconsidered: do the country boundaries matter anymore? In M.J. Smith and G. Salvendy (eds.) *Systems, social and internationalisation design aspects of HCI* (Lawrence Erlbaum Associates, London), 548–552

Chen, C. H. and Chen, C. H. 2000, Research on Chinese computer users' mental models in software interface design. In S.A. Robertson (ed.) *Contemporary Ergonomics 2000*, (Taylor and Francis, London), 182–186

Horton, W. 1994, *The icon book*, First Edition, (John Wiley & Sons, New York)

Nielsen, J. 2000, *Designing web usability: the practice of simplicity*, First Edition, (New Riders Publishing, Indiana)

Norman, D. A. 1999, *The design of everyday things*, Second Edition, (The MIT Press, London)

Wang, H. F. 2007, The appropriateness of Taiwanese cultural imagery in icon design. *Design for Safety, Health & Environment, the 14th Annual Meeting & Conference*, (Ergonomics Society of Taiwan, Taichung), CD RUM E-02

CHARLES BABBAGE AND THE "FACILITY OF USING"

R.B. Stammers

School of Psychology, University of Leicester, Leicester, LE1 7RH

Charles Babbage is a key figure in the history of computing. One of the driving forces for his creation of a computing machine was to eliminate human error in the production of mathematical tables. His early attempts at ameliorating the situation concentrated on improving printed tables and in this context that he describes in 1826, "… the facility of using …". He also devised a notation system for representing the complexities of machines. So was Babbage the first ergonomist? The paper explores his concern with making technology usable through matching design to human characteristics.

Introduction

Any writer on the history of ergonomics has the problem of defining a starting-point. One approach would be to begin with the earliest human tools as discovered by archaeologists. It might be easier to begin with the modern coining of the word *ergonomics*, contemporaneous with the formation of the Ergonomics Research Society in 1949. Even this approach must make some reference to what came before, and in this case it is usually a mention of the wartime research by human scientists. However, even taking this stricter definition of what is meant by ergonomics, it is possible to find isolated examples in earlier times of systematic multi-disciplinary approaches to studying the human use of technology.

One way round this problem is to describe the pre-1940s work as "proto-ergonomics" (Monod, 2000). Such an approach is convenient because it allows the isolation of examples of good ergonomics practice from pre-history onwards, without impinging on the definition of the specific field of multidisciplinary research and practice. When writing on the history of ergonomics it is tempting to ignore this point and make the case for, "was *X* the first ergonomist?". Whist this is headline-catching, it is of course a contradiction in terms, given that an *ergonomist* was not seen as a possibility until post-1949.

One aspect of human endeavour, the development of effective technology using design principles has been little studied. Industrial archaeology would provide a rich source of material that would demonstrate whether something along the lines of a natural evolution of good design could be isolated. Doing this, would involve exploring the resourcefulness of artisans, engineers and scientists who were ever ready to produce the "better mousetrap". This is not to denigrate the role of small improvements in design. Rather it highlights that need for a discipline of ergonomics only emerged when the pressure to exploit technology did not allow for a slow process of trial and error to be used to produce the best form of an artifact. The

pressures of war meant that technology exploitation had to be driven by a range of disciplines, including the human sciences. So examples of good design can be separated out from the avocation of principles or guidelines for the design process. This is a less well-documented activity, but its basis in the history of engineering can probably be established. Some examples from the work of the Nineteenth Century polymath, Charles Babbage, serve as a case of someone who stood back from the phenomenon that he was studying and tried to design artifacts and tools to take into account human limitations and capabilities. Two examples will be taken from Babbage's work, one concerning the design of mathematical tables and another on the derivation of a notation system for describing the functions of machines.

Designing mathematical tables

Babbage is probably best known for his designs for a *Difference Engine* and an *Analytical Engine*. The purpose of these machines was to automate the process of designing mathematical and other tables. As such they have been taken as the earliest examples of programmable computers (Swade, 2000). The source of Babbage's interest in this problem was the high frequency of errors in tables available at this time. Such tables were not only used by mathematicians but were also at the heart of practical activities such as maritime navigation and the use of actuarial data in the insurance industry. Babbage noted that the sources of error in such tables were manifold, arising not only from calculation but also in printing. In the latter, errors can arise in the transcription of numbers, the setting up of type, in proofreading and in the subsequent printing processes. At the root of the design of his engines was his ideal of automating table production, including direct printing, thus by-passing the whole error-prone typesetting and production processes.

An earlier stage of his work had looked at the more traditional processes of table production and in doing so he explored ways of improving the printing and layout of tables. It is in the context of this work that a concern with improving the design of information presentation for human users can be found. In the preface to some tables he produced there is the telling phrase:

> In presenting to the public a new table of Logarithms, two things particularly demand the attention of the editor – their correctness, and the facility with which they can be used by computers. (Babbage, 1827, p. v)

In this context, "computers" are individuals who are doing the computing of numbers. In order to improve the accuracy of tables he examined a wide range of existing tables and used a number of other means to improve the accuracy of his new tables (Campbell-Kelly, 1988). On the question of improving what we would now call the *usability* of the tables, he says the following:

> With regard to those arrangements, on which the facility of using a table depends, I shall offer a few remarks, although at the hazard of seeming to

> bestow more attention than may be necessary on a subject which may appear of trifling importance. (Babbage, 1827, p. vii)

In this one sentence two key points of importance to modern ergonomics can be seen. Firstly, as indicated above, it is very tempting to equate Babbage's "facility of using" with modern term *usability*. The second point puts an emphasis on the importance of "arrangements" of the printed material. All too often, ergonomists find that they are accused of tinkering with trivial matters.

Babbage, in collaboration with his friend Lieutenant-Colonel Colby collected examples of good and bad practice from existing tables, "with the view of discovering on what typographical circumstances perspicuity depended." (Babbage, 1827, p. vii). From this activity he assembled a set of rules, some of which he felt "commanded immediate assent", others were deemed to be of "more doubtful propriety". Unfortunately, no indication is given about which rules he felt were of most value, it might be assumed that there is some order of priority in the first nine items. It should not be assumed that Babbage derived these aspects of typography for himself, rather his contribution lies in collecting them together as a set of guidelines that would enhance good design in general.

In addition, perusal of the rules indicates a clear recognition of the role of perceptual and cognitive capabilities in the use of printed materials. A rule is commended because it helps in, "directing the eye". In Rule 4, "a single dark line is much more conspicuous than two fainter lines adjacent to each other". It is advocated that items should be enhanced so as to, "strike the eye readily". The very modern-sounding phrase, "the eye has not to travel so far to acquire the information sought in the table", is offered in support of another typographical feature. In addition to these observations on improving legibility of tables, Babbage also makes a number of practical points about printing. In doing this he draws attention to the need for a trade-off between the cost of printing and the usability of the tables.

A feature of Babbage's work that, with hindsight, might be seen as controversial his was exploration of the value of coloured paper and coloured type. Bearing in mind the lighting conditions prevalent at his time, he thought it important to explore any possible advantages that might arise from variation in hue of paper and type. To this end he assembled together a range of coloured papers (140 examples) and coloured inks (10 were used) and had examples of tables produced using a variety of combinations. He then experimented with this material in order to determine the most beneficial combinations. It can only be assumed that such experimentation was fairly informal. However, it is interesting that this systematic variation of stimulus material and a recording of responses predate the emergence of experimental psychology. Psychology utilised the above features as the defining nature of an appropriate experimental approach for the new discipline. Some details of his work can be found in a review of an example of the tables by in 1832 by Scottish scientists (Anon, 1832). The conclusion of the reviewers is it that, whilst some combinations of colours may be "agreeable" to observers, "they may not be the *most useful*". Their conclusion was, "that the *blackest ink on the whitest paper* ought to be adopted" (Anon, 1832, p. 149, original italics). This early practical criticism is also supported

by a more modern view that, "of course Babbage's ideas on coloured paper . . . have proved plain wrong." (Campbell-Kelly, 1988, p. 168). Whilst these proposals on the use of colour might seem a little quirky, the important point is that Babbage saw the need for an exploration of the problem space of printing, to determine whether different combinations would yield particular results under particular lighting conditions. He also did not want to take what might seem to be an obvious route, rather he suggested experimentation in order to establish optimal combinations.

A notation system for complex artifacts

The second example concerns a scheme devised by Babbage in order to capture the complexity of the functioning of engineering artifacts. Basic engineering drawings were available in his time, but as he pointed out:

> In the description of machinery by means of drawings, it is only possible to represent an engine in one particular state of its action. (Babbage, 1826, p. 250)

Several drawings and annotations would be needed in order to capture all the complexity. He therefore set out to design a notation system that would allow for the representation of a machine in all its complexity. In diagnosing the need for such a system, he made reference to the human characteristic of not being able to envisage the complexity of a device in conscious processing, saying:

> The difficulty of retaining in the mind all the cotemporaneous *[sic]* and successive movements of a complicated machine, and the still greater difficulty of properly timing movements which had already been provided for, induced me to seek for some method by which I might at a glance of the eye select any particular part, and find at any given time its state of motion or rest, its relation to the motions of any other part of the machine, and if necessary trace back the sources of its movement through all its successive stages to the original moving power. I soon felt that the forms of ordinary language were far too diffuse to admit of any expectation of removing the difficulty, and being convinced from experience of the vast power which analysis derives from the great condensation of meaning in the language it employs, I was not long in deciding that the most favourable path to pursue was to have recourse to the language of signs. (Babbage, 1826, pp. 250–251)

He sets some criteria for such a scheme of notation, it should be:

> simple and expressive,
> easily understood at the commencement,
> capable of being readily retained in the memory from the proper adaptation of the signs to the circumstances they were intended to represent. (Babbage, 1826, p. 251)

Space limitations preclude a detailed description of Babbage's scheme. Its main features were the supplementation of engineering drawings with a tabular representation of the various components, their linkages, movements and other features. Symbols and lines were used to link components and to indicate such things as direction and nature of movement. More important from the current perspective is Babbage's recognition of the need for such a scheme to supplement the normal mode of working with drawings. In advocating the use of such a system he makes the following notable points:

> The advantages which appear to result from the employment of this *mechanical notation* are to render the description of machinery considerably shorter than it can be when expressed in words.
>
> The signs, if they have been properly chosen, and if they should be generally adopted, will form as it were an *[sic]* universal language
>
> … to those who become skilful in their use, they will supply the means of writing down at sight even the most complicated machine, and of understanding the order and succession of the movements of any engine of which they possess the drawings and the mechanical notation. (Babbage, 1826, pp. 260–261)

It can be seen again that in deriving his scheme Babbage makes reference to the ways in which it can be used to good effect to supplement human cognition. This early example of a system for encoding the complexity of dynamic systems has a strong resemblance to many methods of capturing human task information derived by work-study practitioners and ergonomists in the Twentieth Century.

Discussion

In these examples can be found a concern with topics that continue to have relevance to contemporary ergonomics. The fact that in both cases there has been no subsequent development should not detract from the quality of the arguments made by their progenitor. Indeed, Babbage's engines were never successfully built and used in his time, yet he is still revered for them. Examination of such examples demonstrates how great thinkers in the past have dwelt upon issues that are central to contemporary ergonomics. The problem with such examples is that they remain isolated, there being a lack of continuity from the thinking of the past to contemporary developments. This is not to denigrate past achievements, but to recognise that a discipline will only emerge and develop if there is such continuity. To emphasise the point, there is no continuity of development between Babbage's engines and contemporary computers, even though his work is taken as the first example of a general purpose computing device.

What such examples do enable us to understand is that the discipline of ergonomics emerged successfully at a particular time and place because it met a particular need. It has survived because it has continued to show its relevance to the effective exploitation of technology in a context of human well-being.

References

Anon. (1832) Specimen of logarithmic tables printed with different coloured inks on variously coloured papers. *Edinburgh Journal of Science*, **6**, 144–150.

Babbage, C. (1826) On a method of expressing by signs the action of machinery. *Philosophical Transactions of the Royal Society of London*, **116**, 250–265.

Babbage, C. (1827) *Table of Logarithms of the Natural Numbers from 1 to 10800*. (London: J. Mawman).

Campbell-Kelly, M. (1988) Charles Babbage's Table of Logarithms (1827). *Annals of the History of Computing*, 10, 159–169.

Monod, H. (2000) Proto-ergonomics. In I. Kuorinka (ed.) *History of the International Ergonomics Association: The First Quarter of a Century.* (Gréalou: IEA).

Swade, D. (2000) *The Cogwheel Brain*. (London: Little, Brown).

INCLUSIVE DESIGN

REASONABLE ADJUSTMENTS: COLLECTING USER EXPERIENCES FOR A 'SOLUTIONS DATABASE'

Nadine Geddes & Victoria Haines

Ergonomics and Safety Research Institute, Loughborough University, Loughborough UK

The Disability Discrimination Act (1995) states that employers must make 'reasonable adjustments' to the workplace for employees with disabilities. To help provide advice on this subject a project was undertaken to build a database of work related 'solutions' that would supply information to those that support people with disabilities about their potential work related needs. This paper presents one of a series of studies designed to determine the information requirements that would assist people with disabilities to join or rejoin the workforce.

Introduction

As employers become more aware of the need to make reasonable adjustments for employees with disabilities, more initiatives are being introduced to encourage this process. Many papers have been published declaring the need for improved employment outcomes for disabled people but few have made any practical suggestions on how to improve them (Thompson, 2005; Department for Work and Pensions, 2006; Strategy Unit, 2005).

The overall purpose of this project was to develop a better system of information delivery on the work related needs of people with disabilities, referred to in this paper as a 'solutions database'. To accomplish this, several studies were undertaken and this paper presents the first study which assessed the current situation for people with disabilities, in relation to employment. The data collected from the interviews is being used to contribute to several aspects of the solutions database. However due to the space limitations of this paper only the data that relate to the content of the solutions database will be presented.

Methodology

This study explored the experiences and level of knowledge on disability and employment related topics through interviews of working age people with disabilities. A total of 33 participants took part in the study, all were of working age with some currently working ($n = 23$) and some not currently working ($n = 10$). Analysed variables included gender, age, education level, work experience, type and onset of disability.

Results

Each of the interviews was broken down into sections, an outline of these sections along with a short summary of the results is outlined below:

History of interventions

The majority of each interview entailed the participant giving an account of the professionals with whom they had come into contact from the onset of their disability. The purpose of this was to a) identify who are the stakeholders in the journey from diagnosis to working and b) to determine what information they supplied to the participants. Table 1 outlines the groups of stake holders along with a summary of the type of information they passed on to participants.

Table 1. Stakeholder groups and advice.

Stakeholders	Guidance/information
Medical professionals – general practitioners – hospital doctors – company doctors – consultant physicians – surgeons	– physical restrictions – information on medical procedure – recommended time off work – information on support organisations – information on social services – advice on condition and exercise – advice on diet and medication
Allied medical professionals – physiotherapists – occupational therapists – nurses – rehabilitation consultants – psychologists – acupuncturists	– exercise advice – told to retrain for desk job – suggested seeing a DEA – told had to give up licence – information on condition – recommended looking for jobs with low stress – gave advice on changes to lifestyle – recommended a light weight wheelchair – gave number to get some equipment – information on access to work – information on ways to cope and strategies – DFS claims
Employer	– gave information on DDA – arranged for assessments and adjustments
Employment Intermediaries – Job Centre Plus advisors – disability employment advisors – employment programmes (NDDP, Y2W, etc.) – Department for Work and Pensions – Access to Work	– gave info on benefits – recommended retraining – information on courses – info on right to stop working – told about access to work – educated on disability issues

Attitude toward the employment process

Three questions were asked to determine what the participants felt about the following: (a) what factors were most instrumental in their getting or keeping employment? (b) what information did they find hard to obtain? and (c) what recommendations they would make to improve the employment situation for people with disabilities? Table 2 summarises the responses that are pertinent to the content of the database:

Knowledge of disability provisions

Participants were asked about their knowledge or understanding of important disability related terms concerning Government legislation and support, the terms were: Disability Discrimination Act, Disability Employment Advisor, Access to Work, local support organisations, and reasonable adjustments. Each answer was rated according to the participant having knowledge of the term that was; good, some or none. Overall the results showed that the employed group had at least 200% more 'good' knowledge in all categories except for knowledge of Disability Employment Advisors, where both groups were approximately equal.

Use of adjustments

Participants were asked if they have made, or would need to make, any changes to their job or workplace in order for them to work. There was a great deal of

Table 2. Attitudes toward the employment process.

Q: Who or what would you say has been the most instrumental in your getting into work and why?	Support from employment intermediaries Encouragement and assistance from their employer Because their disability didn't affect their ability to do their job Knowing their rights
Q: Is there anything that you can think of that you wish you had known earlier or information that could have helped you but you were not made aware of it when you needed it?	More medical information on condition and/or treatments for own knowledge More information on Access to Work Information from Job Centre Plus on; courses, jobs, tax credits, housing grants, DEA services Information on support groups Information on Health and Safety Regulations (i.e. right to refuse and audits) Information on mobility schemes Information on DDA and related legislation Information on counselling services available For others to be more aware of condition and what to do in an emergency
Q: If you could make any recommendations on how to help people with disabilities concerning employment, what would they be??	A fuller explanation of access to work Better awareness training for able bodied people, so they can understand the needs of people with disabilities Better guidance and support for people who are job seeking More advice from medical professionals Better accessibility

Table 3. Impairment and adjustments.

Impairment	Adjustments/Aids
Mobility – Total body	
Rheumatoid arthritis in all joints, restricted movement	– equipment for computer: free standing arms, Posturite desk – surrounding (all within arms reach), ergo seat (body forming), speech operated computer, is able to control own hours
Neurological/Muscular Conditions	
Hyper-mobility syndrome. This limits ability to walk and stand, high levels of pain when mobile	– In the conference room employer put a roll-up foam mattress to rest back. – Access to work got an especially supportive chair, a writing slope, a grabber to get things off the floor and they suggested a lift but they haven't got it.
Vision	
Visual impairment – Aniridia (no iris in eyes) sporadic condition, has no central vision and has some cataracts. Strong light can affect vision	– When started work – company did an assessment, recommended a 17″ flat screen monitor, to enlarge text and decrease glare on the screen, the phones are controlled by a touch screen monitor too, uses a laptop for taking notes. – Is fine using public transport and has a disabled person's rail card. Shift work can be difficult because of bus times. – Got a magnifier and monocular (to see distances i.e. bus numbers and signs) – can get Access to Work to pay for taxis when doing home visits
Learning Disability	
Dyslexia	– Access to Work did an assessment, recommended a computer to do work, tried to use speech to text software but it didn't work because of office environment. – Does find typing reports easier to do than writing reports – recommended to get someone to do the filing – uses text to speech software to help with reading emails this works well – was given a hand held Dictaphone to use for home visits, but can't use it because was only given the manual to learn how to use it but can't read the booklet to learn how to use it

information acquired from this question, therefore it is only possible to present, in Table 3, a few of the responses to serve as an illustration of the type of data collected.

Analysis

All of the collected data was used to build a picture of what type of information the end user (i.e. person with a disability) would find useful when trying to find suitable employment. As can be seen from the results this information varies widely from universal information that all users should know such as details on benefits, rights and legislation, funding for adjustments etc, to detailed information that would only apply to a small number of people with a specific disability e.g. use of a helper dog for people with mobility impairments. For the solutions database to contain all

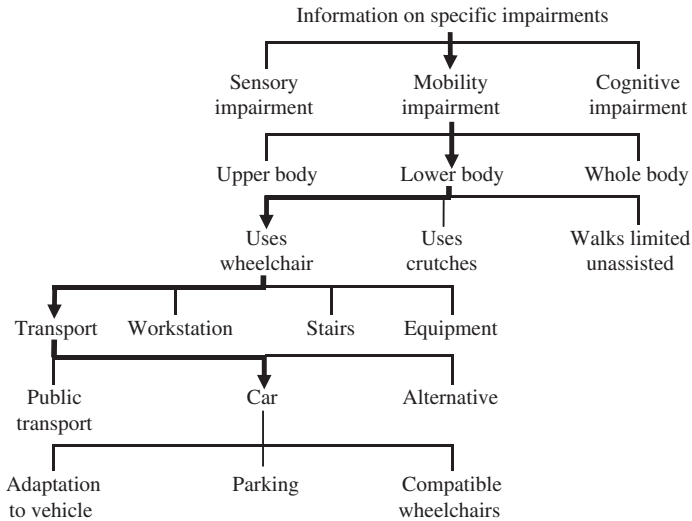


Figure 1. Functionality of 'solutions database'.

of this information but still be user friendly, a taxonomy was created. This system of classification organises the information from general to specific, with the first section separating the most basic types of information, including:

Rights and Laws
Government programmes
Support organisations
General information on making adjustments
Information on specific impairments

Each of these sections would have sub-categories and so on. To illustrate this, the diagram in Figure 1 shows an example of a person who uses a wheelchair looking for information on using a car as part of their work.

Conclusions

As the purpose of the wider project is to design a database of information on work related solutions for people with disabilities, it is necessary to learn from disabled people themselves what information they have found useful or difficult to obtain and incorporate this into the database. This study illustrated that the information needs of a person with a disability relating to work can vary widely from person to person, and in order for a 'solutions database' to be effective the content must reflect this by including many levels of information, going from general to more specific. This will allow the user to identify appropriate solutions at every stage of the person's journey from onset of disability to full integration at work.

References

Thompson, M. 2005, *Able to Work – Report of the National Employment Panel's Employers' Working Group on Disability*, (www.nationalemploymentpanel.gov.uk/)

Department for Work and Pensions, 2006, *A New Deal for Welfare – Empowering People to Work*, (TSO, www.tsoshop.co.uk/bookstore)

Strategy Unit (DWP, DH, DfES, and ODPM), 2005, *Improving the Chances of Disabled People*, (www.cabinetoffice.gov.uk/strategy/)

A FRAMEWORK FOR ESTIMATING EXCLUSION FROM WORK TASKS DURING INCLUSIVE DESIGN

David Nery, Pat Langdon & John Clarkson

Engineering Design Centre, Department of Engineering, University of Cambridge, Cambridge, CB2 1PZ, UK

In the context of work tasks, this paper presents a new framework for functionally linking a comprehensive movement capability data set for the UK population, with the process of task design. The alarming escalation in the rate of the ageing population with their associated changes in capability has resulted in people being excluded from the safe and effective use of many products. The approach presented will allow ageing user movement capabilities to be defined and applied to the design process in a practical and accurate way, thus allowing the level of exclusion to be defined very early and throughout the design process.

Introduction

Inclusive design is a user-centered design approach that focuses on designing for the widest population possible within the design process (Keates & Clarkson, 2004). For this purpose designers require tools that enable the iterative comparison of the evolving design with the capability ranges in the population (Persad, et al, 2006). This paper presents a new framework for functionally linking a comprehensive capability data set for the UK population, with the process of task design. It is well known that the proportion of older people in the world is accelerating at an alarming rate with significant increases being predicted for many nations, including Britain, with the global increase in the proportion of people over 65 years set to triple by 2051 (United Nations Population Division, 2001).

With ageing of the population comes changes in the cognitive and physical capabilities of people. For example, there are reductions in a range of capabilities, including upper limb flexibility (Doriot & Wang, 2006) and muscular strength (Voorbij & Steenbekkers, 2001). Age related changes in functional capabilities can often lead to negative outcomes such as poor health and increased difficulty performing everyday tasks with increasing age (Vercruyssen, 1997).

This global trend of an increase in the proportion of older people and the associated changes in their functional capabilities contrasts sharply with the broad proliferation of product task demands that are beyond the functional capabilities of this ageing population. This mismatch between user capability and product task demands results in people being excluded from using part or all of the task elements of products that have task demands that are in excess of their capabilities.

The analysis of human characteristics for the purposes of user centered design is a central theme within "inclusive design" research and practice.

Using this "inclusive design" approach requires a detailed understanding of the functional capabilities of the users relative to the particular product demands so the level of exclusion, that is, those who are not accommodated within the product demands, can be defined. This is not a simple issue within the context of the ageing population. We cannot just design a product for the "older generation" (Haigh, 1993). This is because the nature of people-product interactions is multi-faceted and the nature of the ageing process is an insidious one with the rate and type of ageing changes being highly variable between people (Haigh, 1993).

Improvements in the reliability and validity of human capability data sets are vital to accurately define the level of exclusion for a particular set of product task requirements. The Disability Follow-Up (DFS) Survey (Grundy, et al. 1999) showed that an estimated 8,582,200 adults (20% of the adult population) had a disability according to the criteria as defined in the survey. It was also noted in the survey that 48% of the disabled population were aged 65 or older (Grundy, et al. 1999). The strengths of this survey, as noted by Keates and Clarkson (2004), is that the large sample is representative of a broad social and geographical area of Great Britain and the survey reviewed multiple capabilities which have resulted in a detailed review of the relationships between ageing and capabilities. The limiting aspects of this survey, from a design perspective, is the survey focused on a range of simple everyday tasks such as picking up a mug of coffee or reading a newspaper which may not necessarily relate to activities which are relevant to the process of product design (Persad, et al. 2006).

The approach described in this paper is focused on utilizing the strengths of the robust DFS data set, but recasting the measurement scales that underpin the survey data so they can be more validly applied to the process movement analysis within a work context. This paper describes the framework by which the current DFS human capability scales can be reconfigured to create the building blocks of a movement analysis framework that can allow the DFS capability data set to be more accurately applied to the process of product design.

The data presented in this paper is taken from the task observations of 136 people performing three different types of tasks in a mail sorting distribution centre. The results of the movement analysis are matched against the DFS data sets to determine the level of exclusion that is associated with a particular type of task demand.

Methodology

This paper is concerned with the examination of the "locomotion", "reach and stretch" and "dexterity" scales which represent the motion-based scales on the DFS disability data set. The proposed taxonomy involves rating 14 degrees of freedom of human movement in terms of the exertion and range of movement for each rating question on each of the three scales that are reviewed in this paper. Characterising each rating question in terms of capability requirements for different degrees of

freedom of movement provides the building blocks of a language to functionally link the DFS data set with product task requirements.

The implications of achieving this are that the DFS data set can be functionally linked, via the development of a task analysis language, to product task requirements. This will allow the level of exclusion, using the DFS data set, to be more accurately defined for product task requirements.

The locomotion scale consists of 13 levels of rating, the reach and stretch has 10 levels and the dexterity scale has 10 levels. Some of the limitations of the existing scales relate to the ambiguous language that is used with the rating questions and the lack of apparent quantifiable progression in severity as one moves through the levels of a particular scale. The taxonomy used in this study provides a consistent approach to measuring the range of movement and exertion for each of the 14 degrees of movement of the body in each of the levels on the DFS motion scales. This approach provides a more objective and quantifiable basis for comparing the various levels on the individual DFS motion scales. The four steps described below provide a summary of how the previously mentioned taxonomy was developed for this study.

Step 1: Define DFS scales using a capability assessment matrix

This stage is concerned with using a 4×4 matrix which has the two dimensions of "range of movement" and "exertion" and is used to define capability for each degree of freedom of movement. Each of the 2 dimensions on the matrix is expressed on a four level scale ranging from low through to medium, high and extreme. Each of the 4 categories are incrementally defined as 25% of the variance in "range of movement" or "exertion" for each of the 14 degrees of freedom of movement that were measured.

Step 2: Redefine DFS motion scales using capability matrix

The application of the matrix to an individual degree of freedom of movement provides a measure of the "capability" for that degree of freedom of movement. This methodology was applied to each level on the three DFS motion scales. The results yielded a language of "capability" measures for 14 individual degrees of freedom of movement on each level of the DFS motion scales.

Step 3: Apply capability matrix to task demands

The same matrix-based approach that was described in stage 2 is now applied to each of the subjects performing each of the three tasks in this study. The results of this analysis is a taxonomy of "capability" for each degree of freedom of movement.

Step 4: Define population exclusion

Steps 2 and 3 involved defining the level of "capability" for the DFS motion scales and the test task demands respectively. Since both the DFS user capabilities and

the test task demands are now described in the same matrix-based language they can be compared in a meaningful way. That is, the level of best "fit" for the task demands with a particular level on the DFS scale will determine the "match" of the task demands to a particular level on the DFS scale and hence determine the level of exclusion for that particular task.

Results

Figure 1 illustrates the results from the assessment of the 14 degrees of freedom of movement for each of the three tasks that were assessed. The results clearly identify that there are more capability demands on particular degrees of freedom of movement in each task. For example, task 1 (lifting a carton from the ground) requires a high level of capability demand on the lumbar spine and shoulders, followed by the hips and knees. So these degrees of freedom would become the limiting factors in people becoming excluded from this particular task.

Climbing out of the truck clearly requires more capability on the knees and ankles of people which becomes the limiting factor in this task. This is often due to people climbing out forwards from the truck cabin and jumping out facing forwards which creates more strain on the knees and ankles when they land on the ground. In contrast, sorting cartons on the conveyor is essentially a waist height task with limited weight bearing demands for sorting the parcels. This is reflected in the capability demands illustrated in figure 1 which show some moderate level of capability demand for the upper limbs and lumbar spine. This is due to the forward reaching component of the task but there are very low levels of capability demand for the lower limbs.

The next stage in the process is to match the capability "profile" for the degrees of freedom of each task with the closest "match" on the DFS motion scales. For example, the best match for task 1 is level 11 on the DFS reach and stretch motion scale.

Discussion

The results of this study indicate that even within the domain of motion, the performance of simple tasks can be limited by a range of task demands that impact on different degrees of freedom of movement. One of the strengths of the proposed exclusion tool is the broad range of degrees of freedom of movement which are analysed when defining human capability.

The task analysis framework summarised in this paper provides a mechanism to effectively link the robust DFS data set with product design. This connectivity will allow the level of population exclusion to be defined for a particular set of product task demands. The relative ease with which the tool can be used will allow the level of population exclusion to be defined very early in the design process, thus providing designers with vital information regarding modifications to their proposed design to increase the range of users who can safely and effectively user their products. Analysing the DFS data using this capability-based exclusion approach provides

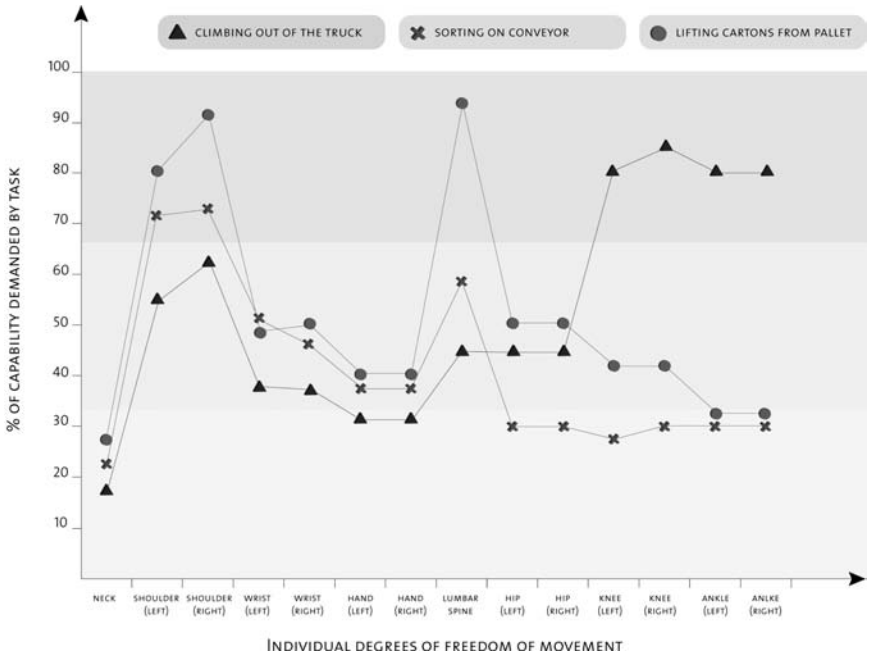


Figure 1. Measures of "capability task demand" for each of the 14 degrees of freedom of movement in each of the 3 tasks that were assessed.

a finer level of granularity for defining exclusion that will produce meaningful comparisons for why people with reduced capability become excluded from tasks. This information can then be factored into designs to produce products which are more inclusive.

Future work in refining this framework will include linking the model to a new human capability data set which has improved data collection reliability. There is also scope within this exclusion framework to redefine the measure of "exclusion" itself. That is, there will still be measures of capability exclusion as described in this paper. But, a more granulated analysis of exclusion will include measures of "task quality" and measures of the "user cost" to people when they achieve a particular level of task performance. These additional dimensions provide an opportunity to examine issues like acquired skill (e.g. errors and learning) and repeated exposure to products (e.g. fatigue and overuse injury risk). These additional aspects will add vital information to the analysis of human-product interactions and provide more predictive power to measure exclusion.

References

Keates, S. & Clarkson, J.P. (2004) Countering design exclusion An introduction to design inclusion, Springer-Verlag: London.

Doriot, N. & Wang, X. (2006) Effects of age and gender on maximum voluntary range of motion of the upper body joints, Ergonomics, Vol. 49, No. 3, 269–281.

Grundy, E., Ahlburg, D., Ali, M., Breeze, E. & Slogget, A. (1999) Disability in Great Britain: Results from the 1996/97 Disability Follow-up to the Family Resources Survey. Charlesworth Group, Huddersfield, UK.

Haigh, R. (1993) The ageing process: a challenge for design, Applied Ergonomics, 24 (1), 9–14.

Langdon, P., Keates, S. Clarkson, J.P. (2003) Developing cognitive capability scales for inclusive product design. In Proceedings of ICED, Stockholm, Sweden, August 2003.

Persad, U., Langdon, P.M. & Clarkson, J.P. (2006) 'Inclusive design evaluation and the capability-demand relationship'. In Clarkson, P.J., Langdon, P.M. & Robinson, P. (2006). Designing accessible technology. Springer-Verlag.

Voorbij, A.I.M. & Steenbekkers, L.P.A. (2001) The composition of a graph on the decline of total body strength with age based on pulling, twisting, and gripping force. Applied Ergonomics, 32, 287–292.

Vercruyssen, M. (1997) 'Movement control and speed of behavior'. In Fisk A.D. & Rogers W.A. (1997) Handbook of human factors and the older adult. Academic Press: London.

CATEGORISING DESIGN METHODS: HOW DESIGNERS VIEW THE ROLES OF USER METHODS IN DESIGN

Joy Goodman-Deane[1], Patrick Langdon[1], Susannah Clarke[2] & P. John Clarkson[1]

[1]*Engineering Design Centre, Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK*
[2]*Civil and Environmental Engineering Department, Imperial College, London*

Methods for involving and understanding users are of key importance for effective and inclusive design. However, many of these methods have had limited uptake, due in part to a poor fit with the ways in which designers think and work. We therefore conducted a card-sorting study with twenty-one product and communications designers, examining their views of design methods in general, and user methods in particular. The study was analysed using hierarchical agglomerative cluster analysis. In this paper, we present some of the initial results, describing the main categories of design methods used and discussing the designers' views of the roles of user methods in design.

Introduction

In order for products to be designed effectively and inclusively, designers need to keep in mind the end-users' needs and desires. However, this can be difficult to do, especially when seeking to include users with different abilities and situations from the designer's own. Many methods have therefore been developed to help with this. Some involve users directly in the design process, while others help designers to consider, understand and empathise with users more remotely.

However, many of these methods have had a mixed and limited uptake in design practice (c.f. Cardoso *et al*, 2005). Our previous research (e.g. Goodman *et al*, 2006) identified various possible reasons for this, such as a lack of resources and uncertainty about the methods' usefulness. In particular, it seemed that one of the key reasons was that there was often a poor fit between user methods and the ways in which designers think and work (e.g. Cardello, 2005).

We therefore conducted a card-sorting study to understand more about how designers view and think about design methods, particularly examining how user methods of various types fit into this picture. This paper describes the study and some of its initial results, particularly examining how the designers categorised the design methods and where they placed user methods within this framework. This study was based on a previous pilot study with a small number of designers, some results from which were published in (Goodman *et al*, 2007).

Related work

Various studies have examined the role of user methods in design practice, describing what methods are or are not used and for what purposes. For example, Hasdoğan (1996) described the prevalence of informal approaches, with designers often "employing themselves as model users". Porter *et al* (2005) further explained that some methods are employed "to gain a more holistic view of the user" or "to give a more intimate picture of [the] target consumer". These findings describe the roles that user methods currently play in design, and some of the reasons for those roles, but they do not really examine how the designers perceive these user methods.

Some researchers have addressed this by asking designers to explain the reasons for their use (or non-use) of user methods. For example, Crilly and Clarkson (2006) interviewed designers and consumer researchers, identifying, among other factors, limited project resources and the difficulties in applying the findings of user research. Others have conducted surveys, asking designers to rate a variety of user methods (e.g. Gulliksen *et al*, 2004).

These direct responses from designers are very valuable in understanding their perceptions of design and of user methods. However, they do not always examine the designers' underlying assumptions about design methods. We therefore sought to augment this previous work with a study that aimed to uncover more about designers' underlying perceptions of the relationships between different design methods, and of the roles played by user methods in particular.

Other researchers have examined the relationships between design methods, categorising them according to various criteria. These categorisations were often based on the authors' experience and knowledge of the methods and the design process (e.g. Aldersey-Williams *et al*, 1999; Hanington, 2003), or on studies of commonly-used methods and their characteristics (e.g. Bylund *et al*, 2003). These provide helpful ways of viewing design methods, but do not examine how designers themselves view these relationships.

Method

A large number of design methods and techniques were identified through a literature review covering fields such as product design, HCI and ergonomics. From this, 57 methods were chosen, ensuring a representative range of method types, with a focus on methods for understanding and involving users. Each method was described on a card, as shown in Figure 1, and the resultant card set was given to a group of designers.

Twenty-one product and communications designers took part. They organised the cards into groups using any criteria they liked and any number of groups and sub-groups. They then labelled these groups, as shown in Figure 1. By allowing the designers flexibility in grouping and labelling the methods, we aimed to uncover their own perceptions of design methods and the considerations that they feel are important in categorising them.

Figure 1. A completed card-sort.

Resulting method groupings

Different participants used different groups, making it difficult to determine similarities and differences by hand. We therefore employed cluster analysis, an exploratory statistical technique, traditionally used to group and classify objects in a dataset. Hierarchical agglomerative cluster analysis was conducted on the probabilities of methods occurring in the same groups as each other. We used the Euclidean distance metric and Ward's and complete linkage grouping rules. This analysis identified six main clusters of methods, some of which had stable sub-clusters, shown in Figure 2. These clusters were then named by the researchers, based on an examination of the methods within each cluster.

The locations of user methods

As explained above, user methods are design methods that help designers to consider, understand and empathise with users. They were located in the above clustering in clusters A (Examining the market), B (Analysis), C (User research without direct user contact) and F (Direct user contact). However, different types of user method were placed in different locations within these clusters.

User involvement methods are one type of user method, distinguished by designer contact with potential users of the product or service. Examples include user diaries, interviews, usability tests and observation, among others. These methods were all located in clusters C and F, and all except one were in sub-clusters C2, C3, F1 and F2. Moreover, these sub-clusters contained *only* user involvement methods, with the exception of "Videos of user needs", which is a user information method.

In contrast, **user information methods** provide designers with second-hand data about users; the designers do not gather this data nor contact users themselves. Examples include anthropometric data, personas, simulation exercises and videos of users. These methods were located in clusters A, B, C (in C1 and C4) and F (in F2). As well as user methods, Clusters A, B and C4 also contained several other methods.

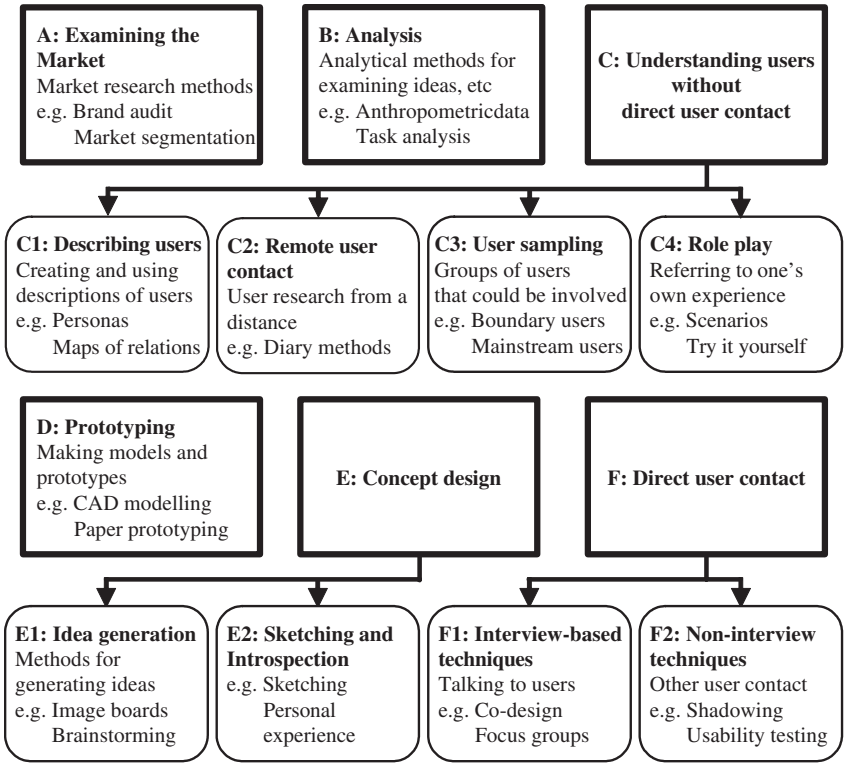


Figure 2. Clusters of methods resulting from the study.

Discussion

The results indicate that designers do have clear ideas about design methods and do think of them as belonging to distinct categories. They are also fairly consistent in this categorisation, with different designers identifying similar groups of methods. Thus, although designers do differ, it is reasonable to discuss the views of this group as a whole towards design methods.

As seen in Figure 2, the categorisation of methods was based primarily on the methods' objectives (e.g. for analysing a situation or understanding users), but was also influenced by the types of technique involved (e.g., introspection and interview-based techniques).

The roles of user methods

When designers want to conduct a core design activity (e.g. generating or analysing ideas), it seems likely that they will reach for methods categorised under those objectives. Users are very important throughout the design process and user methods can provide much valuable input to many of these design activities. Yet, user methods are rarely categorised under these objectives.

This is particularly the case for user involvement methods. Their placement indicates a clear division between them and other design methods, implying that designers think of them separately. They classify these methods under "understanding users" (whether through direct or indirect contact), suggesting that they see this as these methods' primary role, rather than more design-oriented objectives. In contrast, user information methods are more integrated with other methods. Although some were grouped under "understanding users", others were placed in clusters that were primarily for other purposes and that included other kinds of methods.

However, neither user involvement nor user information methods were located in clusters D (Prototyping) and E (Concept Design). While understanding users may have little to contribute to making models (cluster D) or to introspection (cluster E2), many user methods (e.g. user forums, scenarios and shadowing) aim to help generate ideas (cluster E1).

If user methods, particularly user involvement methods, are to be used more often, then this indicates that they need to be more integrated into designers' thinking and into the design process. In particular, designers need to see the methods' value, not just for understanding users, but for enhancing core design activities, like generating ideas and analysing designs. It is therefore important to emphasize this value when describing and disseminating user methods, and to consider these aspects when developing new methods.

User involvement methods can also learn from user information ones, which are more integrated. Further research is needed to determine why this is the case, but a likely reason is that the techniques they use are more similar to other design methods. For example, using anthropometric data requires a similar process to using other design data. Another possible reason is that user information methods require less resource and effort. We have conducted a parallel study to examine this issue and are currently analysing the results.

Conclusions and further work

This study has shown that, although individual designers differ, designers as a group have clear and consistent views about design methods. In particular, they tend to view user involvement methods as separate from other methods. User information methods are more integrated but their scope is still limited. User methods need to be integrated more closely into designers' thinking and into the design process. In particular, designers need to see the methods' value for enhancing core design activities, like generating ideas and analysing designs.

We plan to conduct further analysis of this data to understand more about designers' response to user methods and how they view different kinds of user methods. In particular, initial analysis indicates some differences between product and communications designers, and we are conducting further analysis to identify the differences in these groups' views of user methods. The results of this study are also being used to construct a framework to help designers to think about user methods and to choose ones that are appropriate for their needs.

References

Aldersey-Williams, H., Bound, J. and Coleman, R. (eds.) 1999, The Methods Lab: User Research for Design. Design for Ageing Network. Available at: http://www.education.edean.org/pdf/Tool039.pdf (Accessed: 30 Nov 2007)

Bylund, N., Grante, C, and López-Mesa, B. 2003, Increasing the usability in industry of methods from design research. In *ICED 2003*, Stockholm, (The Design Society)

Cardello, A.V. 2005, Terminology, reliability, validity, and subjectivity in the search for the "voice of the consumer", *Food Quality and Preference*, **16(3)**, 203–205

Cardoso, C., Clarkson P.J., and Keates, S. 2005, Can Users Be Excluded from an Inclusive Design Process? In *HCII 2005*, Las Vegas

Crilly, N. and Clarkson, P.J. 2006, The influence of consumer research on product aesthetics. In *DESIGN 2006*, Dubronik, (The Design Society), 689–696

Goodman, J., Langdon, P.M. and Clarkson, P.J. 2006, Providing Strategic User Information for Designers: Methods and Initial Findings. In Clarkson et al (eds.), *Designing Accessible Technology*, (Springer), 41–51

Goodman, J., Clarke, S., Langdon, P. and Clarkson, P.J. 2007, Designers' perceptions of methods of involving and understanding users. In *Universal Access in HCI (Part 1), HCII 2007*, (Springer), Lecture Notes in Computer Science 4554, 127–136

Gulliksen, J., Boivie, I., Persson, J., Hektor, A. and Herulf, L. 2004, Making a Difference – a Survey of the Usability Profession in Sweden. In *NordiCHI'04*, (ACM Press), 207–215

Hanington, B. 2003, Methods in the Making: A Perspective on the State of Human Research in Design. *Design Issues*, **19(4)**, 9–18

Hasdoğan, G. 1996, The Role of User Models in Product Design for Assessment of User Needs, Design Studies **17(1)**, 19–33

Porter, C. S., Chhibber, S., Porter, J.M. and Healey, L. 2005, RealPeople; encouraging inclusive design through empathy. In *Include 2005*, (Helen Hamlyn Centre, London)

OPERATIONALISING ANALYTICAL INCLUSIVE DESIGN EVALUATION

Patrick Langdon, Umesh Persad & John Clarkson

Cambridge Engineering Design Centre, University of Cambridge, UK

Functional capabilities in Inclusive Design may be characterised by thresholds or human performance envelopes for users with variable or reduced functional capacity. In order to validate the effectiveness of this framework, an experimental approach is reported that compared human performance and measures of perceived difficulty with actual measurements of demand for a range of products and users and with varying capability ranges. This paper explores some of the operational assumptions, for the validation study, their rationale, some quantitative results and the physical and human constraints of the test situation that had to be considered.

Introduction

Inclusive design is a design philosophy that aims to consider the needs and capabilities of users in the design process. This includes addressing the needs of older and disabled populations. The goal of inclusive design is to design products that are accessible and usable to the maximum number of users without being stigmatising or resorting to special aids and adaptations.

In previous work [Persad, 2007], we have examined the possibility of assisting the conceptual design process by providing a framework for analytical assessment that is based on a synthesis of interaction design, human factors and ergonomics and psychology. To do this it was necessary to arrive at a set of requirements for describing in quantitative terms the interaction between the facets of product demand and the statistics of human capability. This was developed in order to encompass the scope of inclusive design that includes the wider capability ranges prevalent in the population, particularly amongst the aging and disabled. The central aim is to evaluate the match between users and the designed product. Thus, we propose that this assessment of compatibility needs to be complete to encompass the whole spectrum of capabilities and should be conducted at a number of levels, based on the distinction between hierarchies of sensory, motor and cognitive levels of human functioning.

The theoretical approach taken in the present framework is similar to that of Kondrasake's Elemental Resources Model (ERM) of human performance (Kondrasake, 2000). It is based on the calculation of compatibility by utilising resource-demand constructs for a set of basic functions at a low hierarchical level such as visual acuity, contrast sensitivity, working memory capacity, force generation and movement capabilities. These functional capabilities are characterised by thresholds or human

performance envelopes for users with variable or reduced functional capacity. The exclusion or difficulty on a given task can be determined simply by comparing the task demands to the demanded capabilities.

Preliminary results suggest that the framework is capable of capturing the critical exclusion regions of interaction between product features and capability ranges. However, in order to perform these experiments, it was necessary to operationalise the demand of product features such as a vacuum cleaner or mobile phone in terms of a set of tasks and with measured variables such as forces, time and perceived difficulty. In order to validate the effectiveness of this framework, an experimental approach compared human performance and measures of perceived difficulty with actual measurements of forces, cognitive and perceptual demand for a range of products and users with varying capability ranges. We examine how this was achieved and assess the effectiveness of the experimental approach.

Methodology

An empirical study was conducted using four consumer products used in activities of daily living : a clock-radio, a mobile phone, a food blender and a vacuum cleaner. Older and disabled users were recruited, and their sensory, cognitive and motor capabilities were evaluated using objective capability tests.

Participants were contacted from organisations and charities such as U3A, CAMSIGHT and the Hester Adrian Centre (Papworth Trust). Nineteen participants in total were recruited and participated in the study. Ethical approval was obtained from the Cambridge Psychology ethics committee. The study was conducted in the Usability Lab in the William Gates Building at the University of Cambridge.

Users performed a number of standard tasks with each of the products while being videotaped. Tasks were randomly assigned to avoid order effects. After performing a task, subjective difficulty ratings were collected for selected actions. The results were analysed to determine how well the capability measures collected prior to task performance predicted difficulty in actual use, using the capability-demand model.

Tasks performed:

(a) Clock radio – setting the time to 4.30 PM
(b) Mobile phone – taking the ringer off via the menu
(c) Blender – blend a banana and water as if making a smoothie/drink
(d) Vacuum cleaner- vacuum a piece of carpet till clean

Participants first signed a consent form (large print forms were available). They were then asked background questions to gather demographic, medical and product experience information. Participants were also asked to rate their experience with four consumer products and to describe how they would go about using these products to perform tasks (one task per product). Data was recorded on a questionnaire sheet and via an audio recorder.

Secondly, a battery of capability tests was administered. These tests included sensory tests (visual acuity, contrast sensitivity, hearing level: Colinbrander, 2003;

Figure 1. Users performing common tasks with the range of products.

DeBonis et al., 2004), cognitive tests (working memory, speed of processing, long term memory: Baddeley, 2000) and motor tests (force exertion in various positions, balance: Kanis, 1993). It was emphasised to the participants that they were not being tested; rather information was being collected to investigate how well the products matched their capability. Participants had a short break after the sensory and cognitive capability assessment was performed.

Some of the participants took other breaks during the capability testing session if they became tired, while other participants chose to continue straight through without a break till all of the capability tests were completed. All capability testing data was recorded on a pre-designed testing sheet. Data from the computerised cognitive capability tests was stored in a computer database and later exported for analysis. Finally, participants were asked to perform one task each with four consumer products: a clock-radio, a mobile phone, a blender, and a vacuum cleaner. The product order and four tasks were randomised for each participant. At the start of the tasks, participants were informed that they could stop at any time for any reason. While tasks were performed, their performance was videotaped. On completion of each of the four tasks, participants rated the level of difficulty and frustration experienced for selected actions on the task with the aid of a graphic difficulty scale.

Data analysis

There were 175 variables recorded during the study and during one task for each product. The relationship between a number of variables for one product only, the blender, are shown as an example of the ongoing analysis. These show the correlation scatterplots and linear regression lines for Rated overall physical and mental demand and a number of variables expected to be related to exclusion or difficulty of use, such as Right Hand Maximum Lift capacity and Short Term Memory Span (Figure 2).

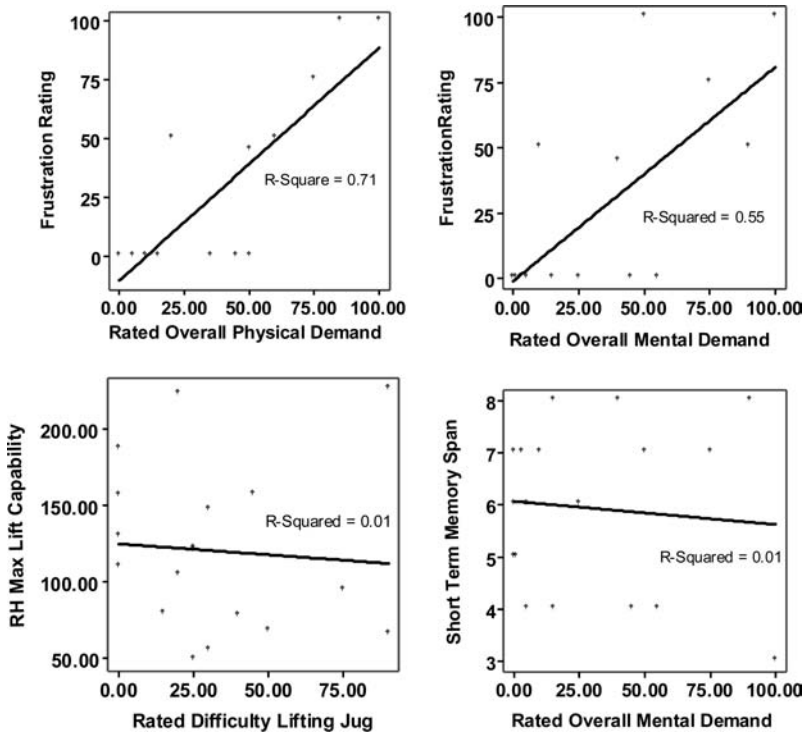


Figure 2. Correlation and regression lines for rated mental and physical demand and difficulty in strongly related variables (upper pair) and weakly related variables.

The upper two graphs in Figure 2 are examples of strong relationships between data variables, in this case, between the users frustration rating with the blender and the overall physical demand. A linear model accounted for significant amounts of the variance of frustration with physical demand $R^2 = 0.707$, $F = 19.7$, $df = 18$, $p < 0.001$ and Mental Demand: $R^2 = 0.547$, $F = 20.5$, $df = 18$, $p < 0.001$). The lower two graphs in Figure 2 are examples of weak relationships between data variables, in the first case between the users RH maximum Lift capacity and the rated difficulty of lifting the Jug made during the task. A linear model failed to account for any of the variance $R^2 = 0.01$, $F = 0.09$, $df = 17$, n.s. The second case shows the relationship between users Short Term Memory Capacity as measured by a digit span task, and the rated overall mental demand of the Blender task. As before, a linear model failed to account for any of the variance $R^2 = 0.01$, $F = 0.144$, $df = 18$, n.s.

Exclusion results

For each task, either: (1) The task was not attempted, (2) The participant successfully completed the task (3) The participant failed to complete the task.

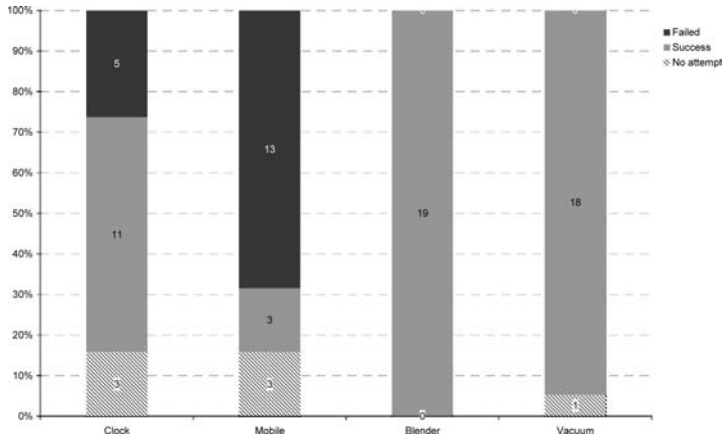


Figure 3. Graph of proportions of levels of exclusion by product.

Discussion

In order to operationalise the demand of a range of products and the capabilities of a number of users in product interactions, a number of sampling issues were addressed. In particular it was necessary to match the range of capabilities of the user sample to the demands of the product such that around 50% of users were able to complete the tasks. It was necessary to ensure this for sensory tests; such as visual acuity, contrast sensitivity, hearing level; cognitive tests such as working memory, speed of processing, long term memory, and for motor tests such as force exertion in various positions and balance. On this basis the study was designed so that most people would not fail outright. Despite this a significant number failed to complete a task or made no attempt at tasks (Figure 3). This was particularly true for the mobile phone and clock radio but was less true for the blender and the vacuum cleaner. This suggests that the experiment successfully operationalised difficulty of use and exclusion in a naturalistic interaction and that the product range and task difficulty was such that individuals were both excluded and included from use.

The users' ratings of overall physical and mental demand were strongly related to user frustration for the lender. This was true despite the fact that seven users rated frustration at zero. Of these three to four users rated the mental or physical demand as very low and the remainder clearly captured people whose difficulty with the product reflected their capability ranges. In order to perform the blender task, the rotary switch required little physical strength, grip or coordination but the removal of the lid required all three. The mental demand of the blender largely lay with interpreting the use of the rotary switch and pulse control but also in operating the lids and the jar unlocking mechanism. The difficulty lifting the Jug was not directly related to the RH lifting capacity, indicating that the physical demand was not due to lifting and was therefore probably due to the large forces needed to remove the safety lid. The STM capability did not predict mental demand suggesting that the cognitive demand of the blender was not limited to large time-limited or information

rich demands in its operation sequence but was probably related to reasoning about the lid and locking mechanism.

We conclude that operationalisation of the task demand of these daily living products was successful in targeting the range of physical and cognitive demands of the products and matched them well to the users' range of capabilities. The preliminary analysis of 127 variables suggests that the measurement and ratings taken were, in principle and practice, capable of differentiating the exclusion relationships between product demand and user capability.

References

Baddeley AD (2000) The Episodic Buffer: A New Component of Working Memory? Trends in Cognitive Sciences 4: 417–423.

Colenbrander A (2003) Aspects of vision loss – visual functions and functional vision. Visual Impairment Research 5: 115–136.

DeBonis DA, Donohue CL (2004) Survey of audiology: fundamentals for audiologists and health professionals. Pearson/ Allyn and Bacon, Boston, Mass.

Persad U, Langdon PM, and Clarkson PJ (2006), Inclusive Design Evaluation Utilising A Model Of Capability And Demand, Universal Access in the Information Society, V6, Number 2/September, 2007, pp. 119–135 Springer–London.

Kanis H (1993) Operation of Controls on Consumer Products by Physically Impaired Users. Human Factors 35: 3.5–328.

Kondraske GV (2000) A working model for human system-task interfaces. In Bronzino JD (ed.), The Biomedical Engineering Handbook, 2nd edn, CRC Press vol. 2, pp. 147–118.

THE ASSISTIVE EYE: AN INCLUSIVE DESIGN APPROACH

Emilie Moellenbach, Alastair Gale & John Paulin Hansen

AVRC, Loughborough University, Loughborough, LE11 3TU
Innovative Communication, IT University of Copenhagen, Denmark

In everyday modern living Assistive Technologies (AT) have become increasingly commonplace, serving to enhance the quality of life for users by providing a wide range of environmental controls. Smart homes, which are truly ubiquitous, require complex systems to be managed by the user which provides the challenge of providing adequate and appropriate user-centric interfaces. This is particularly important where the user is either elderly or suffers from a physical impairment which restricts their limb movements.

Introduction

This paper discusses the issues in providing suitable gaze interactive AT interfaces for severely disabled individuals and presents research which addresses issues concerning the implementation of an interface which supplements Type to Talk (TTT) systems by allowing the user to interact directly with their environment through a real-time virtual representation on screen. Ultimately, this would allow for direct communication with carers and autonomous interaction with everyday household appliances. The context of the user is examined in terms of their environment and their need to communicate, both socially and practically, which enables the establishment of an interaction taxonomy for limited mobility. The issues raised by this ongoing research with disabled users and possible solutions are discussed along with the benefits and pitfalls of gaze interaction.

The fundamentals of gaze interaction

The eye

The eye is a complex organ connected to an even more complex visual cortex. There are only a few aspects of sight which directly effect gaze interaction and which therefore need some explanation. The eye is basically a light sensitive organ, which absorbs light impulses and directs them to the brain. The retina consists of light sensitive cells that differ in both density and function, they are called cones and rods. Rod cells are responsible for our peripheral vision and for high definition viewing the cone cells are used. In the centre of what is called the macula is the fovea, which contains the highest density of cones, and is responsible for foveal vision, which represents an angle of circa 1,5° of high definition vision and it is this small visual area which is very relevant when designing for gaze interaction [14].

Whilst there are many different types of eye movements, there are two basic temporal states in which the eye moves that are of interest here; fixations and saccades. Fixations are moments that last between 200–600 ms in which the eye is relatively still and the central, foveal vision is used to view an object clearly [9]. Between fixations the eye moves in what is called saccades. Saccades are ballistic in nature, which means, that the motion is predetermined, and the eye moves directly from one point to the next [10]. During this motion, which can cover between a 1° and 40° angle, most of the visual input ceases; this is called saccadic suppression. With saccades lasting between 30–120 ms, this suppression ensures smooth visual input. Thus saccades are the eyes' way of making up for not being entirely made up of cone cells. As an object cannot be perceived clearly unless it is seen by the limited area of the fovea, the position of a person's eyes is a general indication of that individual's point of interest [14].

Eye movements

Eye movements have been subject to research for more than 100 years. Initially the process of tracing eye movements was an intrusive procedure employing mechanical objects being placed directly on the cornea. In the 1950's eye-tracking was conducted using head mounted equipment and in the 1970's the big leap came with: '*the discovery that multiple reflections from the eye could be used to dissociate eye rotations from head movement*' [11].

At present the invasive equipment has all but vanished and substituted with increasingly precise and fast eye-trackers. One implementation projects infrared light from several different points, creating a stable reflection, essentially red-eye, then an infrared sensitive camera collects the information allowing a computer to track the direction of the gaze [13]. This type of equipment is still quite expensive and therefore has accessibility issues; however, attempts have been made to make eye-trackers as accessible as possible. Low resolution gaze trackers have been created with standard digital video cameras [5]. These systems are not as robust as the commercially produced ones. However, if the accompanying software applications are designed to tolerate noisier input, the systems work well.

Diagnostic and interactive gaze interaction

There are two prevalent methods used when dealing with eye tracking; diagnostic and interactive. Diagnostic is the passive registration and analysis of eye movements which is often employed in the fields of neuroscience, psychology and marketing. Some of the earliest diagnostic work was done by Fitts et al. in the 1950's where the eye movements of 40 pilots were examined while landing a plane and the technique has been seen as promising since then [9]. '*While there has been considerable use of eye tracking in usability engineering over the 50+ years since Fitts' pioneering work [. . .]. We see however, that just in the past ten years, significant technological advances have made the incorporation of eye-tracking in usability research much more feasible.*'[9].

The technological advance, mentioned by Jacob, enabled the second method of eye tracking; interactive gaze. Here eye movements are used as a real-time input device. Since the beginning of the 1980's this has been a subject for visionaries in human computer interaction. Bolt was one of the first who began to conceptualize how the use of eye-tracking equipment could be implemented in real time. In 1981 he wrote: '*While the current application of eyes at the interface is in a sense a special case, eyes-as-output has a promising future in interactive graphics, especially where the graphics contemplates the possibilities arising out of the orchestrated use of videodiscs.*' [2] Over the last two decades two interactive applications have seen great advances. Firstly 'gaze contingent displays' which attempt to solve the bandwidth problems by only rendering the user's area of interest in high definition and secondly 'on- screen selection'[3] which has been the focus of many studies concerning AT.

The challenges of eye tracking

Employing gaze exclusively and not as a supplementary input device creates some problems which are linked to the single modality of gaze. The eye-tracker recognizes that someone is looking at a point in space, not the intention behind the gaze, in a sense the eyes are 'always on'. This results in the so-called Midas Touch problem. In Greek mythology Midas wished that all he touched would turn to gold. When this wish was granted, he quickly found a flaw in his plan, as all he touched, including food and drink, indeed turned to gold and he almost starved to death. Because of gaze input functions as a single modality there is the risk of selection and activation when this is not required or wanted '*Everywhere you look, another command is activated; you cannot look anywhere without issuing a command. The challenge in building a useful eye-tracker interface is to avoid this Midas' Touch problem.*'[9]

One solution to this problem is dwell time activation, which is intended to ensure that the user is actually fixating on a specific object and not just passing it by with a glance [7]. There is also a problem in using the eyes as both an input and output modality, both looking at feedback from the system and directing interaction on the interface [1] Finally, we are not used to having objects, onscreen or otherwise, react to our sight and we rarely need to employ the type of precision required for controlling information spaces with gaze, as it is relatively difficult to control eye position precisely as a pointing device [9].

Eye tracking in assistive technology

Mobility and interaction taxonomy

Severe impairments, such as; MND (motor neuron disease), Strokes or other forms of LIS (locked in syndrome) can be aided by AT which can then provide this user group with self-directed living. Often this group of users has very little consciously controlled movement, except for their eye movements. Consequently, people with LIS have been instrumental in eye tracking being developed as an AT navigational

Table 1. Interaction Taxonomy.

Abilities	User Context	Interaction Capability
Non-restricted Manual Dexterity	Initial fatigue symptoms are noticeable. Slight mobility restrictions	Capable of using keyboard and mouse in conjunction
Restricted Manual Dexterity	Severe fatigue. Lack of communicative ability. Restricted mobility in either one or both hands.	Capable of using a mouse or similar input restricted
Switch Control	Classic LIS. Loss of mobility in extremities. No communicative ability.	Capable of using a switch through various muscle control
Non-restricted Ocular Inspection	Incomplete LIS. Full Eye movements are still possible.	Full use of eyes and therefore gaze equipment.
Restricted Ocular Inspection	Close to complete LIS. Limited eye movements.	Capable of eye gestures, either vertical of horizontal.

tool as it directly enables the user with a new level of communicative autonomy and environmental control, which otherwise would be impossible.

Advances in eye tracking technology has moved it from cumbersome and intrusive head-mounted systems to miniature and remote systems which can be built into systems so as to provide a high degree of usability. This is exemplified by research carried out within COGAIN EU Network of Excellence [15].

There are two approaches to understanding the capabilities of the user groups and how these effect their ability to interact with equipment. First of all there is the need to gain an understanding of the concept of LIS and secondly there is the need to understand the progression of degenerative diseases such as ALS and MND, so as to create an interface which adapts to the different stages.

A three part classification of LIS can help to understand the fundamentals of limited mobility. First of all there is classic LIS where patients can only move their eyes vertically and upper eyelids. Secondly there is incomplete LIS, here other eye and eyelid movements are possible and finally there is complete LIS: Total paralysis where the patient has no movement capabilities [4]. The understanding of the progression of degenerative neuron diseases is an important factor when defining the modes of interaction the interface and subsequent system needs to adapt. The following interaction taxonomy (Table 1) is inspired by the work of Hansen et al. [6] on defining the 5 stages of ALS, as well as conversations with experts in the field, i.e. Mick Donegan from the ACE centre who works closely with people with complex diseases.

Applicability of eye tracking

Advances in eye tracking technology in recent years has improved its usability. However there is still sufficient scope for further development. The three key areas

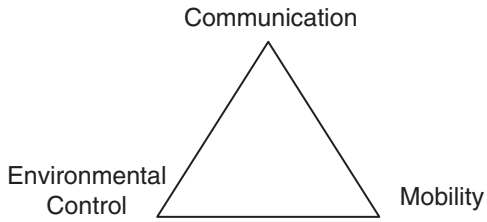


Figure 1. The three keys of interaction for Assitive Technology.

which need addressing when developing assistive technology and gaze interaction systems for people with mobility problems are; communication, environmental control and mobility (Figure 1).

Communication facilitated by gaze, the so-called 'Type-To-Talk' (TTT) – systems, allow users to communicate solely by eye movements. This area has seen some recent advances, GazeTalk and Dasher [8] are two examples of TTT systems which are giving people with limited mobility the possibility to independently interact with the world, some for the first time. GazeTalk not only supports typing but includes other options such as online browsing. Symbol based communication systems also exist, e.g. The Grid 2 [16]. An environmental control system using eye gaze directly on objects has also proved to be possible, e.g. the ART system (Attention-Responsive Technology). The idea is to locate the user's gaze within a three dimensional environment and then use this information to interact with everyday objects, simply by looking at them. For example, if a user wants to switch on an electric fan or a lamp they simply have to look at the item in order to activate the appropriate interface [12]. Building upon the principles of the ART project, the concept of the assistive eye is intended to supplement Type to Talk (TTT) systems by allowing the user to interact directly with their environment through a real-time virtual representation on screen. By using object recognition the user is able to interact with the object or simply point at areas of interest, in order to speed up communication with carers. There is a definite need for further development within the field of gaze interaction in assistive technology. If applied effectively the potential user group which could benefit from improved environmental and mobility control stretches from beyond ALS and MND patients to the elderly and users of all ages who suffer from mobility limitations.

When dealing with users who have very specific needs it is vital that these needs are fully understood and that the entire context of the user's environment is taken into account. The aim is not only to support the user directly by allowing a higher degree of autonomous control; but also to enable those in the immediate environment to better help and support the user. Inclusive design and participatory design is considered an expendable luxury in many commercial design processes. However, in the field of assistive technology it is a necessity. Many technological boundaries are being pushed within this field not because they can, but because they have to.

Acknowledgements

We would like to thank the EPSRC for supporting this research. We also thank the COGAIN EU Network of excellence for providing funding for this research.

References

R. Bates. Have patience with your eye mouse! eye-gaze interaction with computers can work. Proc. of the 1st Cambridge Workshop on Uni.Acc. & As. Tech. (CWUAAT), 2002.

R.A. Bolt. Gaze-orchestrated dynamic windows. Proc. of the 8th an. Conf. on Comp. graphics and interactive techniques, pages 109–119, 1981.

T. Andrew, Duchowski. Eye Tracking Methodology, Theory and Practice. Springer-Verlag London Limited, second edition, 2007.

A.G. Gale and Shi, F. Look in, turn on, drop out. In P.D. Bust (Ed.) Contemporary Ergonomics 2006. Taylor & Francis, London, 2006

D.W. Hansen, D.J.C. MacKay, J.P. Hansen and M. Nielsen. Eye tracking off the shelf. Proc. of the 2004. Symp. on Eye tracking res. & app., pages 58–58, 2004.

J.P. Hansen, H. Lund, H. Aoki and K. Itoh. Gaze communication systems for people with ALS. pages 35–38, COGAIN, 2006.

J.P. Hansen, A.S. Johansen, D.W. Hansen, K. Itoh and S. Mashino. Command Without a Click: Dwell Time Typing by Mouse and Gaze Selections. Proc. of INTERACT 2003, pages 121–128, 2003.

K. Itoh, H. Aoki and J.P. Hansen. A comparative usability study of two Japanese gaze typing systems. Proc. of the 2006 Symp.on Eye tracking res. & app., pages 59–66, 2006.

R.J.K. Jacob. The use of eye movements in human-computer interaction techniques: what you look at is what you get. ACM Trans. on Info. Systems (TOIS), 9(2): 152–169, 1991.

R.J.K. Jacob. HCI: Input Devices. ACM Comp. Sur., 177–179, 1996.

R.J.K. Jacob and K.S. Karn. Commentary on Section 4. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements. Oxford: Elsevier Science. 2003

F. Shi, A.G. Gale and K.J. Purdy. Eye-centric ICT control. In P.D. Bust (Ed.) Contemporary Ergonomics 2006. Taylor & Francis, London, 2006.

Tobii. Product Description Tobii 50 Series.

Semir Zeki. A Vision of the Brain. Blackwell Scientific Publications, 1993.

COGAIN Network of Excellence: http://www.cogain.org/

The Grid 2, Sensory Software International Ltd: http://www.sensorysoftware.com/

CALIBRATING CAPABILITY LOSS SIMULATORS TO POPULATION DATA

Sam Waller[1], Pat Langdon[1], Carlos Cardoso[2] & P. John Clarkson[1]

[1]*Engineering Design Centre, Cambridge University, Trumpington St, Cambridge, UK, CB2 1PZ*
[2]*Faculty of Industrial Design Engineering, Delft University of Technology, Delft, The Netherlands*

When designing mainstream products, issues related to accessibility will often have to trade-off with other factors that relate to commercial success. To make such compromises in an informed manner, it is helpful to compare how many users would be affected by alternative design choices. For example, supposing a proposed design improvement means the text on a product is now readable when viewed with simulated reduced vision ability, it would be useful to know how many additional people would potentially be able to use that product. This paper discusses the challenges associated with obtaining such numbers, and presents an example of a wearable vision loss simulator that might be mapped to available user data.

Introduction

In order to support the design of products that are accessible to people with the widest range of abilities, designers need information about the variation of people's capabilities, and the impact that capability loss has on user-product interaction. However, common ways of presenting capability data in written forms can often 'data overload' the reader, and can be hard to relate to product use.

When considering data about users, it is helpful to define the following terms. Various medical conditions, diseases or ageing can affect the physiology of one or more parts of the body, which causes reduced functional ability. Disability and exclusion may occur when the demands required to perform everyday tasks exceed the ability of the user. For example, cataracts (a medical condition) primarily cause the lens to become cloudy (a physiological change), which reduces the ability of the eye to see fine detail (a functional ability), which means the person becomes unable to read ordinary newspaper print (a disability). Functional ability ranges from high to low in a continuous manner across the entire population. A product may be easy, frustrating, difficult or impossible to use, depending upon the match between the user's ability level, and the level demanded by the product in order to use it. Reducing the capability level required to use a product can improve the experience for a broad range of users, and this paper demonstrates how capability loss simulators can provide a simple and objective way of comparing the capability demand levels of different products or concepts.

Simulation of reduced ability presents a valuable way of helping designers to empathise with end-users and appreciate the impact of capability variation on product use (Nicolle and Maguire, 2003; Poulson *et al*, 1996). It involves demonstrating the effects of physical and sensory impairments, either by using special wearable kit to reduce the designers' own capabilities (Hitchcock and Taylor, 2003; James, 2003) or by modifying image or sound files to show how they would appear to someone with sensory impairments (Neubig *et al*, 2005).

Simulated ability loss is intended to complement, not replace, other methods for assessing usability, such as expert opinion (Poulson *et al*, 1996), and user trials (Jordan *et al*, 1996; Aldersey-Williams *et al*, 1999). Such simulators only approximate true effects, they cannot yet account for the effects of cognitive impairments, and they do not capture the experience of actually living with disability (Poulson *et al*, 1996). Cardoso and Clarkson (2006) provides a thorough review for various different types of capability loss simulators, and their appropriate use in the design process.

Using simulators to measure demand levels

Capability loss simulators are often designed to represent the effects of certain conditions, such as cataracts, enabling the wearer to better understand the condition, and empathise with those who have it (VINE, 2005). However, this paper considers the implementation of capability loss simulators as a way of assessing the capability level demanded by a product in order to use it. Products that remain usable despite a simulated loss of a particular functional ability place a low demand on that functional ability. Simulating ability loss with various degrees of severity therefore enables the capability demands of alternative products to be compared, now demonstrated using an example for vision.

The glasses shown in Figure 1 have been designed so that various numbers of translucent filters can be inserted, thereby changing the severity of the simulated acuity loss. The glasses are used to provide an objective measure of the visual acuity demand associated with reading the text on two different pill bottle tops, also shown in Figure 1. The assessment is carried out in typical indoor lighting conditions, when the pill bottles are resting on the desk and the reader is sitting upright. The glasses could also be used with other filters to simulate loss of central or peripheral field, or loss of binocular vision, although only acuity demands are considered here.

The wearer starts off attempting the task with a large number of translucent filters in place, then continues to remove filters one at a time until the task becomes possible. The pill bottle with the black text first becomes readable with two filters in place, while the white text is first readable with zero filters (A high number of filters indicates a low visual demand). The apparent size of the detail to be detected, together with the brightness and colour contrast difference between foreground and background, and the ambient illumination will primarily affect the acuity demand level. This model assumes that the user's eyes can correctly focus given the viewing distance, which is usually achievable through prescription glasses, or even combinations of different types of glasses (reading glasses, distance glasses,

Figure 1. Vision capability loss simulator with variable numbers of translucent filters, and two example pill bottle tops being assessed for being able to read the text.

bifocal or varifocal glasses). Colour blindness is important to consider when assessing visual demands, although this best accounted for using software impairment simulators.

Calibrating capability loss simulators

Having assessed the capability demand required to use a product, it is desirable to calculate the proportion of the UK adult population that would not be meet that demand. Of particular interest is to understand how this number can be broken down according to the demands made on different capability categories, such as vision, hearing, thinking and dexterity. Informed choices can then be made regarding alternative design concepts, with due consideration for the difference they make to the levels of exclusion. Such calculations require a dataset that is generalisable to UK population and covers all of the capability categories required for product interaction. The 1996/97 Disability Follow-up Survey (DFS) remains the most recent dataset to satisfy these requirements (Grundy *et al*, 1999). Within the framework of this survey, the number of translucent filters through which a task is just possible for a fully sighted person has been approximately calibrated to the number of people that would be unable to perform this task, according to the results for the DFS. This calibration was performed by a small group of participants with 'normal eyesight', in typical office lighting conditions.

It is important to consider that the disability levels described in the DFS could have been the result of different types and combinations of medical conditions and physiological changes. For instance, one of the disability levels (on the DFS) states that people "cannot see well enough to read a newspaper headline". However, one condition, such as age-related macular degeneration, could cause several effects, such as blurred vision and/or a blind spot in the middle of a person's field of vision.

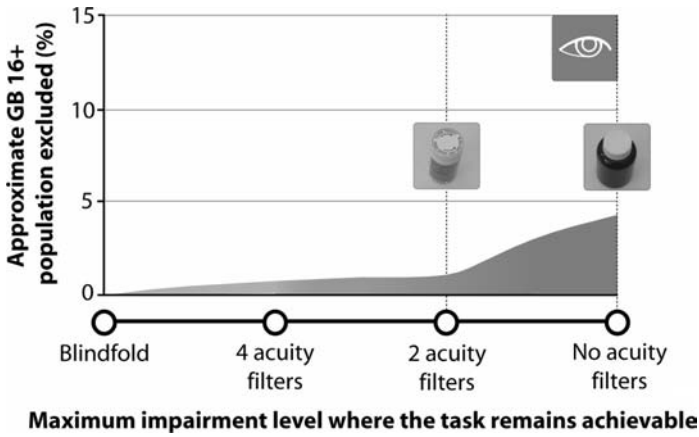


Figure 2. Graph showing an approximate calibration for the visual loss simulators.

Both of these effects could cause difficulties "reading a newspaper headline", yet only one of them is represented by the translucent filters. However, the purpose of this calibration is to demonstrate the manner in which it allows user capability data to be presented, rather than to claim a particular level of accuracy for the population numbers. Nevertheless, the rough calibration is expected to be valid for comparing orders of magnitude for the numbers of people who would be excluded from using a product. Further research by the authors will investigate ways to achieve a more rigorous calibration, including designing and undertaking a population survey specifically for this purpose.

The results of the approximate calibration are shown in Figure 2, which also shows the levels of demand and exclusion associated with reading the text on the two pill bottle tops considered earlier. Note that if a task is possible while wearing a blindfold, then there is no associated visual demand, in which case no-one would be excluded on the basis of vision. The primary difference between the two bottle tops is an additional skim of black paint, which will clearly have a cost impact. Presenting the information in this form allows this additional cost of this process to be evaluated against the additional number of people who are able to read the text, thereby enabling an informed design decision.

The difficulties in using capability loss simulators to measure levels of demand and exclusion are more severe for the other capabilities involved with product interaction. The decline in motor ability can be simulated by attaching stiffening straps to various parts of the body, but the resulting effect will inevitably vary depending upon the strength of the wearer. For tasks involving dexterity, the functional ability of the wrist and each finger is difficult to correlate with the ability to perform a particular task, because that task can typically be performed using many different combinations of fingers, thumbs and palms from one or both hands. For fully able people the maximum force and comfortable force exertions are typically similar, but conditions such as arthritis cause large difference between these two

measurements. Difficulties with motor abilities are often caused by motor control issues rather than strength issues, adding further difficulties to valid simulation. The decline in cognitive ability cannot yet be meaningfully reproduced by simulation, while valid simulation of hearing loss requires an active system that modifies the sound received from a microphone, which limits their practicality (Moore, 1998).

Clarkson et al (2007) details approximate methods for measuring the demand levels in all these categories, which can then be combined with the vision demand level to provide an overall level of exclusion.

Conclusions and further work

Capability loss simulators provide a useful, exciting and interactive tool to help a designer understand capability variation, and its impact on product use. For vision capability loss, using a variable number of translucent filters can provide an objective measure for the acuity demand associated with a particular task, which can be converted into an approximate proportion of the UK adult population who will be unable to perform that task.

Calibration can help to understand the relative significance of different levels of capability loss, thereby helping to make informed decisions when aspects related with usability have to trade-off with other factors, such as aesthetics, cost, performance etc. This calibration has been performed in a approximate manner for one aspect of vision capability loss, in order to provide an example for how such data could be presented and used. Further research is required to develop a more accurate calibration for the vision capability loss simulators, and for the other capabilities involved with product interaction. A national survey to measure these capability levels is one outcome of the current i∼design research project, due for completion around 2010.

However the maximum achievable accuracy for such calibration will always be limited, because of the complex relationship between various conditions, their physiological affects, and the corresponding deterioration in functional performance. Difficulties with product use are often caused by a combination of multiple ability losses, such as vision, thinking and dexterity, but the effects of cognitive impairment cannot yet to be simulated. Capability loss simulators should therefore be used to complement, not replace, other methods for understanding users, such as expert appraisal and user trials.

References

Aldersey-Williams H., Bound J., and Coleman R. (1999) *The methods lab: User research for design*. Design for Ageing Network. Available at www.education.edean.org (11/09/07)

Cardoso C. and Clarkson P.J. (2006) Impairing designers: using calibrated physical restrainers to empathise with users. In proc. *2nd International Conference for Universal Design*. International Association for Universal Design, Kyoto, Japan.

Clarkson P.J., Coleman R., Hosking I., Waller S.D. (2007) *Inclusive design toolkit*. Engineering Design Centre, Cambridge, UK. Available at www.inclusivedesigntoolkit.com (11/09/07)

Grundy E., Ahlburg D., Ali M., Breeze E., Sloggett A. (1999) *Disability in Great Britain*. Department of Social Security, Corporate Document Services, London, UK.

Hitchcock D., Taylor A. (2003) Simulation for inclusion . . . true user-centred design? In Proc. *Include*. Royal College of Art, London, UK.

James K. (2003) *Instant aging with a wearable time machine*. DW-World. De Deutsche Welle, Bonn, Germany.

Jordan P.W., Thomas B., Weerdmeester B.A., McClelland I.L. (Eds.) (1996) *Usability evaluation in industry*. Taylor & Francis, London, UK.

Moore B.C.J. (1998) *Cochlear hearing loss*, Whurr, London, UK.

Neubig G., Bender F., Gorczowski K. and Dickey R. (2005) *Visual impairment simulators*. College of Applied Life Studies, University of Illinois at Urbana/Champaign, IL, US.

Nicolle C., Maguire M. (2003) Empathic modelling in teaching design for all. In Stephanidis C. (Ed.) *Universal Access in HCI: Inclusive Design in the Information Society*. Lawrence Erlbaum Associates, London, UK.

Poulson D., Ashby M., Richardson S. (1996) *USERfit: A practical handbook on user-centred design for rehabilitation and assistive technology*. HUSAT Research Institute for the European Commission.

VINE (2005) Visual Impairment North-East. North Shields, Tyne & Wear, UK. www.vine-simspecs.org.uk (11/09/07)

MEASURING PRIOR EXPERIENCE IN LEARNED PRODUCT INTERACTIONS

Pat Langdon, Tim Lewis & John Clarkson

Engineering Design Centre, Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, UK

In a earlier experiment we addressed the extent to which prior experience of a product affect its use as measured by the time taken to complete a set of tasks and by an protocol analysis of the number and type of errors made during interactions. An uncontrolled variable from the previous work was that of the degree of experience of relevant interface elements. Metrics of experience were devised to measure this but depended on self report of amount of experience weighted by product, interface and brand similarity. Hence, we have addressed prior experience directly with a training transfer experiment. A preliminary analysis of the data suggests that learning during the training phase occurred quickly over a few trials. Transfer to a second product with identical functionality and interface sequence but different layout and appearance of controls gave rise to a significant (4 minute) increase in time to complete set tasks. A third product with the same functionality but different controls, layout and appearance roughly doubled the times and errors.

Introduction

In 2020, more than half the population of the UK will be over the age of 50 and the over-80 range will be growing the most rapidly (Keates and Clarkson, 2004). In order to meet the ideals of inclusive design, a designer must create a product that minimises the number of people excluded or who experience difficulty with a product (Nicolle and Abascal, 2001). Whilst inclusive products reduce the difficulties suffered by the elderly and those users with disabilities, the products often attract those without either.

Training transfer research has looked at the relationship between the similarity of a user's training to the actual task to effectiveness of training. For example, a flight simulation will provide an improved performance over studying a video of a pilot at work (Lintern et al, 1990). This can be counter productive if the training product is too similar to the actual product and acquired accepted behaviour on the training product represents an error on the actual product. In the product design world, the designer may change the function of a button on an interface from one model to another causing experienced users proactive interference, or failure to learn the function of that button when using the later model.

Previous work has shown that ubiquitous, existing symbologies across product families are only noticed by some generations (Lewis and Clarkson, 2005). Studies in the Netherlands have explored this further by outlining technological generations that appeal most to the learning processes of those 25 and younger (DoCampo, 2000). On this basis, the electro-mechanical era can be considered pre-1928, 1928–1964 sees the remote control era, 1964–1990 is dominated by displays and post 1990 layered menu systems are generally popular.

In previous work (Langdon, Lewis, Clarkson, 2007) we have addressed the issue of the extent to which prior experience of a product affect its usability as measured by the time taken to complete a set of tasks and by an protocol analysis of the number and type of errors made during interactions. Over several experiments, a number of products including motor cars, digital cameras and microwave ovens were tested with users from stratified ranges of ages and levels of ability. This data was then compared with users' performance on several detailed experience questionnaires and a number of non age-corrected tests of general and specific cognitive capabilities. The principal finding was that similarity of prior experience to the usage situation was the main determinant of performance, although there was also some evidence for a gradual, age-related capability decline. Users of all ages adopted a means-end or trial and error interaction when faced with unfamiliar elements of the interaction. There was also a strong technology generation effect such that older users were reluctant or unable to complete the required tasks for a typical modern consumer electronic product: a digital camera. When product complexity and function was subsequently kept constant and interfaces varied, a simple traditional interface was found to be more effective that buttons and display panels type for all age groups despite general overall reduction in performance with age and cognitive capability.

An uncontrolled variable from the previous work was that of the degree of experience of relevant interface elements. Measures of experience were devised to measure this but depended on self report of amount of experience with weightings for the similarity of product, interface and brand similarity. Here we addressed prior experience directly by ensuring all users were trained to a criterion of zero error with a common product interface and then examined the changes in performance that result from testing on two different interfaces that varied only in a known set of properties from the training product.

Methodology

In the present study, 20 participants were selected in the 40–60 years and 60–80 years old age bands and three levels of cognitive capability as scored by CCS20 (<95, 95–105 , >105).

All were informed the trials were being conducted to examine the products' performance with a range of users rather than to compare individual users' performances. They were informed they could stop the trial at any point and all their data could be deleted on demand. The detailed methodology followed the codes of practice of the British Psychological Society (BPS Ethics, 2006).

Figure 1. DAB radios: Roberts1, Roberts2 and Uno.

Each product user carried out a 15-minute combined cognitive test (CCS20, 2006). The overall score contained a normalising age-correction factor,. However, as trial performance was expected to vary with age, the effect of this factor was removed scoring all users as if they were aged 20 and renamed combined cognitive score 20 (CCS20).

Products were required that were similar enough to vary only in terms of visual features and sequence of activity rather than functionality. Hence, all were DAB digital radios two from the same manufacturer and a third chosen with identical functionality but different interface components. Various products and tasks were trialed to arrive at a set that presented both a learning challenge and yet still be possible to achieve a criterion of zero errors.

Users were trained to use a DAB radio (Roberts1) to a set criterion of no error. After performing each training trial they were shown any errors they had made on video and trained how to perform the task with no errors. After reaching criteria the same task set was presented with a second test DAB radio (Roberts2) and times and errors recorded as before. A third test radio (Uno) was then carried out. The order of the presentation of the two test radios was alternated to avoid an order effect.

Results

A preliminary analysis of the data shows that learning occurred quickly over twenty users, all of whom reached criterion within four trials. The training data for all users shows substantial improvement in performance by the second trial. Figure 2 shows times to complete tasks and errors for comparison showing that error rates were consistent with task times. Although starting capabilities and rates of learning vary, all but one user (14) reached criteria in 4 trials.

Test trials

Transfer to a second product with identical functionality and interface sequence but different layout and appearance of controls gave rise to a significant (4 minute) increase in time to complete set tasks and an accompanying increase in the number of errors.

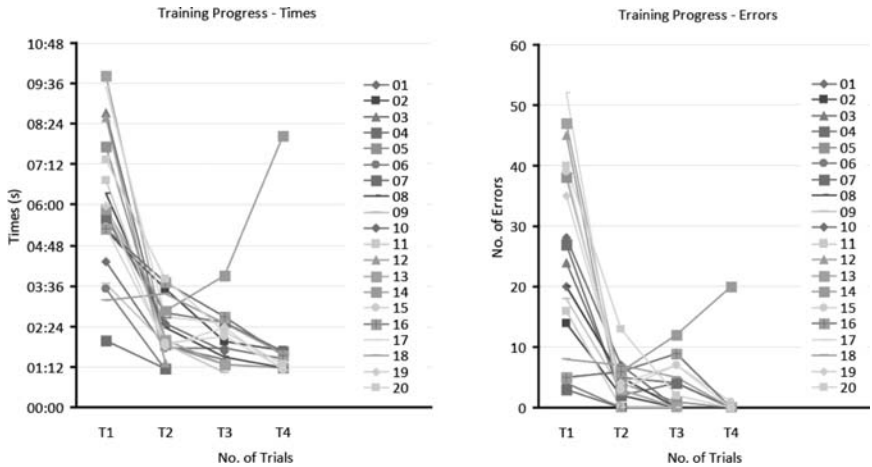


Figure 2. Average time and error to criteria over 4 training trials (radio "Roberts1").

A third product with the same functionality but different controls, layout and appearance roughly doubled the times and errors.

Discussion

Of twenty users who carried out the training trials on BAB radio 1, all but one were able to reach a criteria of zero error in 4 trials. Learning was extremely rapid between the first and second trials although rates varied between individuals, some reaching criterion after two or three trials. There was a greater variability in times on the first trials (T1) of between 2.0 and 10 minutes and this is presumed to reflect the additional variable time arising from tackling differing errors, possibly with trial and error approaches. This is consistent with our previous findings (Langdon et al, 2007) and suggests that despite the design of the radio to be similar to iconic designs of the 1950's, most users did not possess the appropriate knowledge necessary to operate the underlying DAB/FM functions.

The resulting response-selection errors would be expected to give rise to multiple operation of alternative controls in a search for an anticipated system response, incrementing the overall timings. Comparing the T4 trials to the times resulting from the transfer to the similar radio (Roberts 2, Figure 1) suggests that users were able to transfer many aspects of the interface use to the new radio as mean times were below that of T1 (Figure 3) and showed relatively low variability (Figure 4). This is presumed to reflect the fact that the underlying functional model of the radios were identical and that many of the interface controls such as the tuning dial and selection buttons were identical in appearance and labelling but were differently laid out (Figure 1). Hence, it can be assumed that increased times were due to the necessity of locating and operating the controls in their new position and layouts

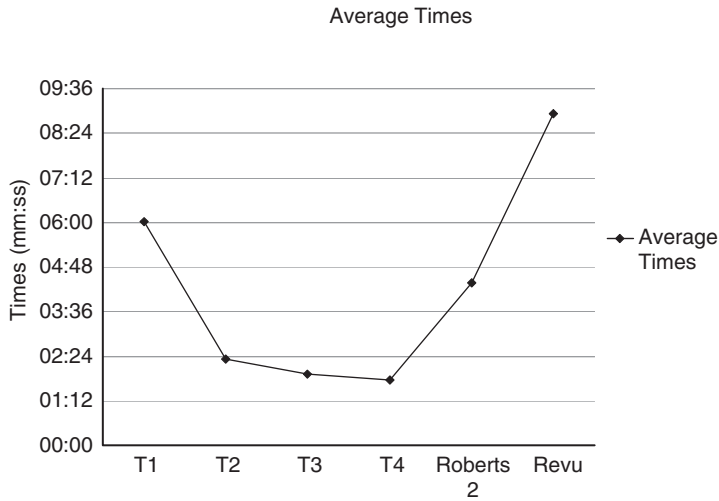


Figure 3. Average times to criteria across 16 users training on Roberts1 (T1–T4) and on transfer to Roberts2 and Uno.

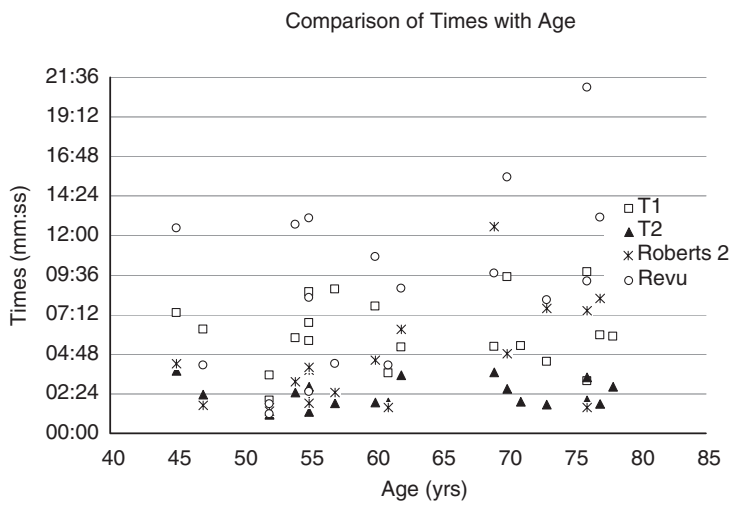


Figure 4. Average time and error to criteria by ages of users.

and, in particular, distinguishing them from a number of additional alternative controls that were not present on the training radio. Completely consistently, the times for the Uno radio with the same functional model but a completely different control interface; both in the nature of the elements and their layout, were greatly increased, with considerably greater variability and a greater volume of errors recorded. It appears that that these errors and times reflect, primarily, response

selection errors and a trial and error selection of controls as well as a significant perceptual component required to learn the appearance of the different interface elements. Further work is in progress analysing the nature of the errors made during these test radio trials and their relationship to the timing data.

There was little evidence of any relationship between age (Figure 4) and times suggesting that neither age-related decrements of performance or generational effects were responsible for timing data. Finally, a weak correlation was found between measures of cognitive capability CCS20 and well-learnt or poorly learnt product interactions. Further work is in progress analysing the nature of the errors made during the test radio trials and their relationship to the timing and cognitive score data.

References

BPS Ethics (2006), http://www.bps.org.uk/the-society/ethics-rules-code-of-conduct.

CCS20 (2006), www.intelligencetest.com, Accessed July, 2006.

Clarkson, J. and Keates, S. (2004) Countering design exclusion an introduction to design inclusion, Springer – Verlag: London.

Docampo, R.M. (2000) Technology Generations handling complex User Interfaces. PhD Thesis, TU Eindhoven. ISBN 90-386-0913-2.

Langdon, P., Keates, S. and Clarkson, P.J. (2003) Developing cognitive capability scales for inclusive product design. In Proceedings of ICED, Stockholm, Sweden, August 2003.

Langdon, P.M., Lewis, T. and Clarkson, P.J. (2007), The effects of prior experience on the use of consumer products, Universal Access in the Information Society, Volume 6, Number 2 September, 2007, pp 179–191 Springer-London.

Lewis T., Clarkson P.J. (2005) A User Study into Customising for Inclusive Design. In: Include 2005. London

Lintern, G., Roscoe, S.N. and Sivier, J. (1990), Display principles, control dynamics, and environmental factors in pilot performance and transfer of training. Human Factors, 32, 299–317.

Nicolle, C. and Abascal, J. (eds) (2001) Inclusive Design Guidelines for HCI. Taylor & Francis,

Persad, U., Langdon, P.M. and Clarkson, J.P. (2006) 'Inclusive design evaluation and the capability-demand relationship'. In Clarkson, P.J., Langdon, P.M. and Robinson, P. (2006). Designing accessible technology. Springer-Verlag.

TACTILE PAVING A NECESSARY INTERVENTION, BUT DOES IT SUIT EVERYONE?

Mohammad Faruk[1], Marcus Ormerod[1], Rita Newton[1], Hamish MacLennan[1] & Mohamed Yusoff Abbas[2]

[1]*SURFACE Inclusive Design Research Centre, SOBE (School of the Built Environment), 430 Maxwell Building, The University of Salford, The Crescent, Salford, UK M5 4WT*
[2]*Faculty of Architecture, Planning and Survey., University Tek. MARA, Malaysia*

The University of Salford as part of the research consortium for I'DGO TOO is currently investigating tactile paving as an indicator of access hazards for vision impaired people as part of an EPSRC funded project. The question that this project poses is the extent to which tactile paving is inclusive within itself as far as other vulnerable population groups are concerned (e.g. older people). The project requires full characterization of the pedestrian and access hazard environment, taking into account other contextual issues affecting the performance of the pedestrian. This paper presents the proposed project research outline in brief and the cognitive model developed from extensive literature reviews, pilot studies and observations that represent a pedestrian's experience in negotiating each of the access hazards (pedestrian crossings and stairs). The paper concludes with a series of specific recommendations for the completion of the next stage of the project.

Introduction

I'DGO TOO (Inclusive Design for Getting Outdoors) is a collaborative research project undertaken by the I'DGO consortium, funded by EPSRC (I'DGO, 2007). The overall aim of the consortium is to addresses the demands, consequences and impacts of new policies and design strategies on older people's environments, and make specific recommendations to make them inclusive thereby improving quality of life for older people, in such a way that it will become general practice in the years to come. I'DGO TOO's research domain covers both private and public open spaces. As part of the project, the University of Salford research team (SURFACE Inclusive Design Research Centre and Centre for Rehabilitation and Human Performance Research) is currently investigating whether tactile paving is inclusive as an indicator of access hazards for vision impaired people and other vulnerable population groups (e.g. older people). In doing so, the team is reviewing existing knowledge of people's reaction on pedestrian journeys in terms of their interaction with tactile paving as a hazard indicator; specifically looking at pedestrian crossings and stairs. This also calls for characterization of the pedestrian and access

hazard environment and any other contextual issues that may affect performance of the pedestrians.

This paper presents the proposed project research outline in brief and the cognitive model developed from extensive literature reviews, pilot studies and observations that represent a pedestrian's experience in negotiating each of the access hazards. The pilot study was found to be representative of the larger demographic sample in terms of pedestrian experience, which justified the cognitive model for the real world; nevertheless there are linkages that require further study on the basis of further laboratory/field studies. These linkages would specifically relate to biomechanical issues and the anatomy of the risk of falling at various surface transition points within the hazard. The paper concludes with a series of specific recommendations for the completion of the next stage of the project.

Tactile paving and I'DGO TOO objectives

Tactile paving began to be introduced in the UK in 1990s to provide a tactile warning for vision impaired individuals to aid their independent mobility. In the years since its introduction tactile paving has become a key design feature in improving the accessibility of public spaces and city centres. The advent of the Disability Discrimination Act (DDA, 1995) with its requirements for inclusive design and equality of access to services is likely to increase the rate at which tactile paving is being installed. Tactile indicators are primarily intended for vision impaired people and it may be less suitable for other vulnerable groups such as older people, potentially posing a trip hazard (Loo-Morrey. M, 2005).

The criteria for the design, siting and laying of tactile paving is currently embodied in various guidelines (DFT, 1998; 2007a & 2007b; LTNZ, 2004) and for many this represents the benchmark for pedestrian crossing design and construction. As various local authorities around the United Kingdom may interpret the benchmarks differently, there could be differences between the various sites in each local area that needs to be taken into account by this study while fulfilling the research objectives. Research objectives of I'DGO TOO are, (1) To examine how blister and corduroy paving is designed, sited and laid; (2) To examine older people's perceptions and approach in using tactile paving; (3) To quantify the relationship between tactile paving design parameters and the biomechanics of ambulation and the risk of falling. Although the objectives are relatively straight forward and yet the factors that determine the risk of falling and the manner in which pedestrians tackle stairs and pedestrian crossings as are quite complex. The design of the research therefore decided to adapt a prototype approach and conduct a pilot study as part of the research methodology that will provide more cohesion in terms of focus between method and the objectives.

The project plan

This research involves gathering data from real world sites (54 sites); conducting a suite of controlled studies of human movement and performance in the laboratory;

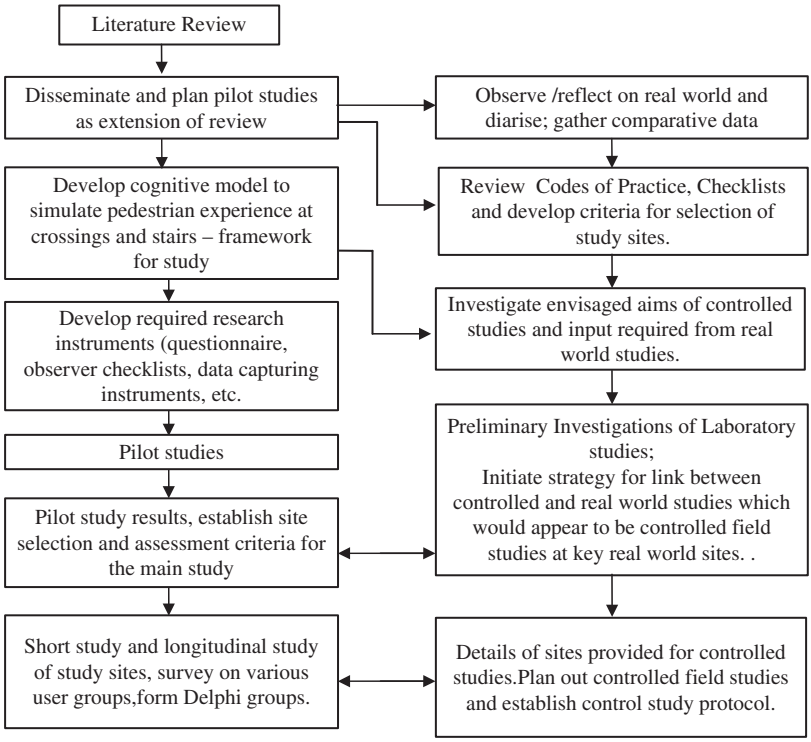


Figure 1. Preliminary project outline to establish knowledge base & main research.

and linking the two in terms of risk of falling and/or injury via further more detailed and semi-controlled studies at key sites. The summarised project plan is as follows:

- Establish Knowledge Base and Recommendations for Conduct of Main Research Project
- Complete Initial Real World Studies and Design of Laboratory Experiments/Controlled studies at key real world sites.
- Finalise Controlled Study Protocols and Experiments/Conduct Controlled Field Studies
- Complete Controlled Study Experiments and Disseminate Data/Establish Protocols for Integrated Analysis of Real World and Controlled Study Data.
- Analyse Data and Present Initial Findings and Dissemination

The pedestrian experience as a driver of the research

There has not been a study that integrates the process of pedestrian navigation through a pedestrian crossing or stair in the form of a model logically linking

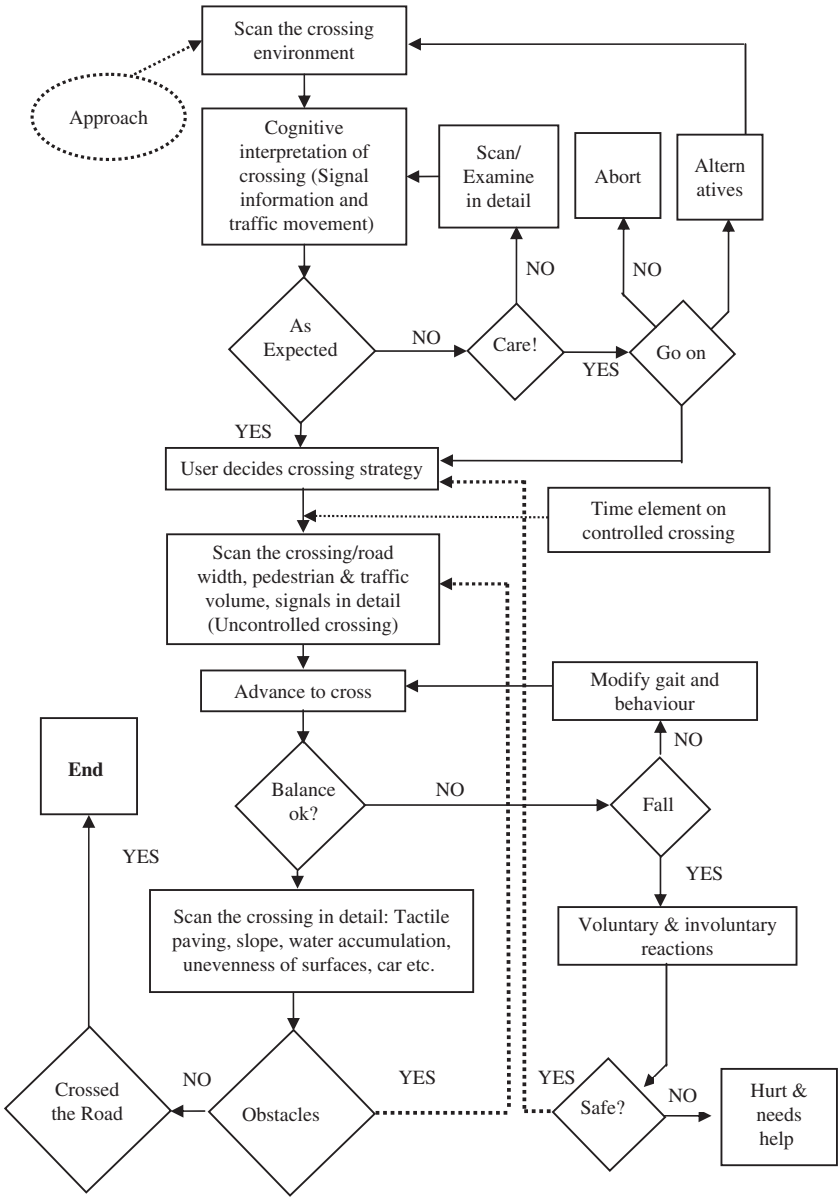


Figure 2. Pedestrian crossing cognitive model.

various acts and decisions involved (Maclennan, 2007). Both pedestrian crossings and stairs are seen as a hazard, which is confirmed by literature reviews (DFT, 1998; 2007a & 2007b; LTNZ 2004; Abbas, M. Y, 2008, Forthcoming). From this work a cognitive model has been prepared for navigation of pedestrian crossings (Figure 2).

Although there may be no direct concrete comparison between stairs and crossings they are both hazards and they have in common the presence of distractions. For this reason that the Cognitive Model for the Stairs (Hale and Glendon, 1987) was adapted by the research team. This model will be subject to continuous review throughout the project by all the various advisory groups. And it can be changed over the course of the study so that it will be immense value to designers and those educating the pedestrian in the safe use of pedestrian crossings.

Recommendations and conclusions

The risk of falling is something that spans across the entire project. The focus is whether or not it is the tactile paving on its own that causes the falling, or whether there are other contributing factors (site and its environment, impact of Code of Practice, pedestrian's physical characteristics). It is here that the outcome of the study as a whole could be improved through the introduction of controlled site studies using human accelerometry techniques and methods (Menz, 2003). For the next phase of the research the site selection and assessment criteria have been established (Faruk, M., Ormerod, M., Newton, R., Maclennan, H. A., 2008, Forthcoming) and 54 pedestrian networks consisting of pedestrian crossings and stairs from England and Scotland have been selected, studied and observed. Longitudinal case studies for the next two years will be carried on a short list of 18 sites out of these 54 pedestrian networks. The findings will be cross matched and simulated in the laboratory to understand the tactile paving associated risk of falling among the older population group.

References

Abbas, M. Y., Ormerod, M., Newton, R., Faruk, M., MacLennan, H. A. (2008), Pedestrians: Just a simple flow?, *Ergonomics and human factors annual conference*, 1–3 April 2008, University of Nottingham. (Forthcoming)

DDA (Disability Discrimination Act). 1995, Available at, http://www.dwp.gov.uk/dda.

DFT (Department for Transport). 1995a, Local Transport Note 1/95, *The assessment of pedestrian crossings*, HMSO.

DFT (Department for Transport). 1995b, Local Transport Note 2/95, *The design of pedestrian crossings*, HMSO.

DFT (Department for Transport). 1998, *Guidance on the use of tactile paving surfaces*, The Scottish Office, DETR, 2007 updates available at, http://www.dft.gov.uk/transportforyou/access/tipws/guidanceontheuseoftactilepav

Faruk, M., Ormerod, M., Newton, R., Maclennan, H. A., 2008, Site selection criteria for I'DGO, *CIB W70 International Conference in Facilities Management*, Heriot Watt University, Edinburgh 16th–18th June, 2008. (Forthcoming)

Hale, A. R., Glendon, A. I., (1987), Stair user cognitive model, *The Staircase*, MIT Press.

I'DGO. 2007, http://www.idgo.ac.uk/, Visited on December 1, 2007.

Loo-Morrey, M. 2005, *Tactile paving survey*, Health and Safety Laboratory, Harper Hill, Buxton.

LTNZ. 2004, *Pedestrian network planning and facilities design guide*, Land Transport New Zealand, Available at, http://www.landtransport.govt.nz.

Maclennan, H. A., Ormerod, M., Faruk, M., Newton, R. 2007, Pilot study Report, *I'DGO TOO working paper*, Surface inclusive design research centre, University of Salford.

Menz, H., Lord, S. R., Fitzpatrick, R. (2003), Acceleration patterns of the head & pelvis when walking on level and irregular surface, *Gait and posture*, Vol. 18, pp. 35–46.

TIME IS MARCHING ON. WHAT REALLY DETERMINES THE TIME IT TAKES US ALL TO GET FROM A TO B?

Hamish MacLennan[1] Marcus Ormerod[1] Rita Newton[1], Mohammad Faruk[1], & Mohamed Yusoff Abbas[2]

[1] *SURFACE Inclusive Design Research Centre, SOBE (School of the Built Environment, The University of Salford, Maxwell Building, The Crescent Salford M5 4WT*
[2] *Faculty of Architecture, Planning and Surveying Univ. Tek. MARA, Malaysia*

Long accepted beliefs, data and formulas that have evolved from the late 1960's in pedestrian planning and design are continually challenged by researchers. Pedestrians occupy more personal space than they did before. The number of distractions in the pedestrian journey has increased along with the pedestrian's functional ability. The old beliefs, data and formula need to be challenged and a way forward proposed. The objective of this paper is to establish whether or not the long accepted beliefs, data and formulas are still realistic in the light of the changing anthropometric and functional ability profile of populations generally. This paper presents work in progress that establishes a way forward for inclusively determining the time that it takes to get from A to B.

Introduction

Much of the data used in the design of circulation routes in open and enclosed spaces was collected over thirty years ago (Fruin, 1987) (Proulx, 2002). Certain Codes have recognized the increasing incidence of obesity impacting on the amount of space people occupy and their circulation characteristics (Rouphail et al, 1998). People are still quite often treated as 'ball bearings' when designers estimate the time it takes to travel from A to B. This approach is termed the macroscopic approach. The microscopic approach removes the ball bearings and replaces them with human beings with very distinct physiological and cognitive characteristics (Fujiyama, 2005). Leake (1991) showed the increased range in walking speeds and behaviour across a carefully selected sample of people that represented the total population profile. The microscopic approach therefore demonstrates that is not just 'a matter of a simple flow' in traveling from A to B, but rather one that relies on the full characterization of the pedestrian (LTNZ, 2004), or building occupant and the associated space (Boyce, 1999) (LTNZ, 2004).Distractions such as the simple use of the mobile telephone make the approach even more complex (Hatfield 2007). The microscopic approach is now more widely accepted with decision based simulation being used in conjunction with real world observation (Castle, 2007)(Kerridge, 2007). The question that needs to be asked though is whether or not this approach is inclusive.

Macroscopic Density Model and Demographic Change

There is concern about the ongoing demographic change that deals with the functional ability of the population inclusively (i.e. ageing, obesity, and impairment) in both pedestrian and evacuation facility design, but it is primarily the approach that needs to be re-examined, especially for crowds (Konnecke, 2005). Such an approach is challenged via simple cognitive models for circulation elements such as stairs and crossings (MacLennan, 2007), where the concept of designed versus required traversal time places the user under stress, because the required traversal time exceeds that of the design. There is still a need for a simplified approach but this needs to be coupled with an acceptable microscopic model (Castle, 2007) with inclusively developed input. This input is vital for use with a new design process along the lines of that suggested by MacLennan (1998) for enclosed spaces. It requires advancing to take account of demographic, environmental and technological changes. Coupling this with pedestrian/building occupant characterization provides the opportunity for more realistic estimating of how long it takes us all to get from A to B.

Population Profile and Occupant Characterization – Functional Ability and Behaviour

A snapshot of the UK population profile shows up those sections of the population whose functional ability will vary their circulation ability:

- Ageing and associated health disorders where approximately 10% of the population are over 65 years and that this will grow to 11.5% in 2025.(US Census Bureau, 2007)
- 19.7% of the population is considered disabled, with impairments that would affect their functional ability to circulate being 15.77%. (DRC, 2007)
- Approximately 29% of the population has a BMI > 30 and 5% > 40. (The Information Centre, 2006)

The above represents an overall population profile where 1 in every 5 persons could be at risk, and these need to be analyzed in more detail to assist with pedestrian and building occupant characterization (Boyce, 1999). The framework is into which this information is set is vital (LTNZ 2004) as follows:

Characteristic Behaviour > Resulting in > Impacting Upon

e.g. Reduced rate of joint motion, obesity and aged related conditions is the characteristics resulting in a slower walking speed impacting on longer traversal time for the pedestrian journey or the evacuation which can increase stress.

In this way most of the pedestrian characteristics concerning functional ability can be analyzed to determine whether they will impact on the time taken to get from A to B.

This applies to other characteristics or behaviours e.g.

- Vision impairments – increased wayfinding effort – longer traversal time
- Talking on the mobile phone – distracted – longer traversal time (Hatfield, 2007). Talking to friends whilst walking- distraction (MacLennan, 2007)

The complexity of occupant or pedestrian characterization therefore confirms that a microscopic approach is required in circulation design (Abbas, 2007) (Castle, 2007).

The Impact of Space

Density Models that depend on the old data need to change. The impact of the increase in critical personal space due to obesity is important. This critical personal space is termed the 'body ellipse' (Fruin 1987) (Rouphail, 2007). The body ellipse was originally set at 600 mm transversal and 500mm in the direction of walking (Fruin, 1987). This has been increased to a circle with a diameter of 600 mm (Rouphail 2007).

Crowd Dynamics (2007) have developed a table for the human anthropometrical footprint across a number of cultures. The projected plan area for the UK Male is 0.26 m^2. This will still fit within the ellipse proposed above 0.28 m^2 but could be challenged for the normal buffer zone allowance. This still does not cater for a BMI > 40 (5% of the population and increasing) when compared with the ellipse of 0.44 m^2derived from measured data from an abdominal CT scan coupled with arm dimensions (Geraghty, 2003)(Ostchega, 2006). The chance of such a footprint from the population profile above is 1 in 20. Allowing for a defined space such as stairs or a crossing the body ellipse would need to be increased further. A frequency of 1:20 can result in a single very obese person reducing the descent speed of an entire platoon of people in an evacuation to 0.4 m/sec which was a reduction of some 33% from the other stairwells (Proulx, 2006). The spatial arrangement of people in platoons and/or clusters is supported by other observations in the field (MacLennan, 2007) (LTNZ 2004).The cluster which is smaller than the platoon is formed by some kind of interactive behaviour or group relationship. Clusters can also be created by slow movers due to functional ability or being distracted by a phone or some other situation (Hatfield, 2007) (MacLennan, 2007) (Proulx, 2006). It is also likely that clusters will not merge with others trying to enter the flow. The traversal time for this type of crowd movement is dictated by the critical walking speed (slowest mover). In larger crowd scenarios the density issues need to be revisited and the various densities quoted for each level of service challenged. This will also affect the flow rates and is part of an ongoing study (MacLennan, 2007).

Impact of Critical Walking Speed

The critical walking speed is often seen as being driven by a goal. People's response will be constrained by their functional ability and can be challenged for people with

Table 1. Walking Speeds.

Leake, 1991 – Leeds Ped. Precinct					I'DGO TOO		
Fastest	10%	50%	90%		10%	50%	90%
Wheelchair Users	1.45	0.86	0.56				
Vision Impairment	1.25	0.86	0.50	Overall	1.57	1.2	0.6
Cane Users	0.91	0.56	0.37		The 1 in 20 was 0.67		
Ambulant Unaided	1.25	0.86	0.45	Overall Stairs			
Older people 65yrs+	1.20	0.86	0.45	Treads 300+	0.94	0.62	0.45 Min 0.35
Non-disabled people Bodied	1.20	0.91	0.77				

impairments who make up 19.7% of the population. The distance that they can safely cover is also of concern. A comparison of two data sets (MacLennan, 2007) (Leake 1991) provides a useful snapshot of the situation in Table 1.

(LTNZ, 2004) set the speed at some 0.7 m/sec based on a comprehensive review of a large number of studies as compared with design speeds of 1.2 m/sec+. Research on walking speeds on stairs (Fahy, 2001) (MacLennan, 2007) (Fujiyama 2005) (Hamel, 2004) require the same review. Walking speeds also vary somewhat with step geometry. Foot placement is critical so that side on placement on narrow treads can generate falls. This occurs in descent when compared to ascent and increases the use of the handrail (Hamel, 2004). There are also many other stair safety issues which are not addressed in this paper that can also affect walking speed. Taking into account the impact of the slow mover the critical walking speed could vary from 0.45 m/sec to 0.35 m/sec. There are people with further limited functional ability whose walking speed would be less (Fahy, 2001). Ramps are the other option. This challenges the design speeds often used for evacuation of 1.0–1.3 m/sec. in free descent which will be modified substantially by density (Nelson, 2002).

Conclusion

The time it takes to get from A to B is not just a simple flow (Abbas, 2007). Demographic, environmental and technological change now dictate that it is the characteristics of the pedestrian and/or building occupant (behavioural) and the pedestrian route that need to be considered. Using the microscopic approach will allow the designer to cope with the entire population profile, hopefully within a design process that is inclusive and reliable.

References

Abbas M.Y., 2007, Pedestrians: Just a Simple Flow, *Paper to be presented at the Annual Conference of the Ergonomics Society, Nottingham 2008,* Ergonomics Society UK.

Boyce K.E., Shields T.J. and Silcock G.W.H., 1999, Toward the Characterization of Building Occupancies for Fire Safety Engineering: Prevalence, Type and Mobility of Disabled People, *Fire Technology, Vol 35, No.1.*

Castle Christian J.E., Guidelines for Assessing Pedestrian Evacuation Software, *Working Paper Series, Paper 115,* University College London, March.

Crowd Dynamics, 2007, Graphic Levels of Service, www.crowddynamics.com, site visited on 6th December 2007.

Disability Rights Commission, 2007, Disability Briefing, May Disability Rights Commission.

Fruin John J., 1987, Pedestrian Planning and Design, *Revised Edition,* Elevator World Inc., Mobile, Alabama.

Fujiyama, T., 2005, *Investigating Use of Space of Pedestrians. Working Paper*. Centre for Transport Studies, University College London.

Fujiyama, T. and Tyler N., 2005, Pedestrian Speeds on Stairs, An Initial Step for a Simulation Model, *Working Paper, University College of London.*

Fahy R.F and Proulx G., 2001, Toward creating a database on delay times to start evacuation and walking speeds for use in evacuation modeling, NRCC 44758, NRC,CNRC.

Geraghty Estelle M. and Boone J., 2003, Determination of Height, Weight, Body Mass Index and Body Surface Area with a Single Abdominal CT Image, Radiology, September 2003, 857–863.

Hatfield, J. and Murphy, S. 2007, The effects of mobile phone use on pedestrian crossing behaviour at signalized and unsignalised intersections, *Accident Analysis and Prevention*, **39**, 1, 197–205.

Hamel K.A. and Cavanagh P.R., 2004, Stair Performance in People Aged 75 and Older, *Journal of the American Geriatric Society (JAGS)*, 52, 563–567.

Konnecke R., and Schneider V., 2005, Evacuation from underground railway stations; Available and safe egress times from different station types and general evaluation criteria, Pedestrian and Evacuation Dynamics, Springer Berlin Heidelberg, 363–368.

Kerridge J., Armitage A., Binnie D., Lei L., and Sumpler N, 2004, Monitoring the Movement of Pedestrians Using Low Cost Infra Red Detectors: Initial Findings, Transportation Research Record.

Land Transport NZ/Opus Consultants, 2004, Pedestrian Network Planning and Facility Design Guide, Land Transport New Zealand.

Leake G.R., May A.D., and Parry T., 1991, An ergonomic study of pedestrian areas for disabled people, *Institute of Transport Studies, University of Leeds,* Transport and Road Research Laboratory, Department for Transport.

MacLennan H.A., Ormerod M., Faruk M. and Newton R., 2007, Pilot Study Report, *Working Paper, I'DGO TOO Project,* Surface IDRC, University of Salford.

MacLennan H.A., Regan M. and Ware R., An Engineering Model for the estimation of occupant premovement and/or response times, *Fire and Materials*, 23, 6, 255–263.

Nelson H.E. and Mowrer F.W., Emergency Movement, (third Ed.) The SFPE Handbook on Fire Protection Engineering, NFPA, Quincy, 367–380.

Ostchega Y., Dillon G., Prineas RJ, McDowell M. and Carroll M., 2006, Tables for the selection of correct blood pressure cuff size based on self reported height and weight and estimating equations for mid arm circumference: data from the US Health and Nutrition Examination Survey, *Journal of Human Hypertension*, 20, 15–22.

Proulx Guylene, 2002, Movement of People: The Evacuation Timing, (third Ed.) The SFPE Handbook on Fire Protection Engineering, NFPA, Quincy, 3–342.

Rouphail N., Hummer J., Milazzo J. and Allen P., 1998, 'Literature Synthesis Pedestrians' Chapter 13 of the *Highway Capacity Manual*, Federal Highway Commission, Washington D.C.

US Census Bureau, 2007, *International Data Base, www.census.gov./ipc/…, last visited 6th December 2007.*

PEDESTRIANS: JUST A SIMPLE FLOW?

Mohamed Yusoff Abbas[1], Marcus Ormerod[2], Rita Newton[3], Mohammad Faruk[4] & Hamish MacLennan[5]

SURFACE Inclusive Design Research Centre, BuHu Research Institute, The University of Salford, Maxwell Building, Salford M5 4WT
[1]*Assoc. Prof., Dr., Fac. of Arch., Plan. and Survey., Univ. Tek. MARA, Malaysia*
[2]*Professor & Director, SURFACE Inclusive Design Research Centre*
[3]*Prog. Dir. PGCert/Dip/MSc Accessibility & Inclusive Design, The Univ. of Salford*
[4,5]*PhD Candidate, SURFACE Inclusive Design Research Centre*

In the design of facilities for pedestrians, it seemed to be a general practice amongst designers/road safety practitioners to regard pedestrians very simply as just one simple flow of human beings. This study looked into the implications of such a simplistic design approach. More recent studies pertaining to pedestrian behaviours were reviewed. Results revealed that variations of composition amongst pedestrians necessitated different fundamental ergonomic requirements considerations. Failure to consider and acknowledge those requirements not only caused unnecessary inconvenience amongst pedestrians but more seriously affected their safety. Hence, a much more localised inclusive design approach in acknowledging the various fundamental ergonomic requirements should be adopted.

Introduction

It has been made known for decades that behaviours amongst humans vary. Designers in particular had been reminded in not to make assumptions about human behaviour because recurring mistakes can be avoided if they understood and take into considerations those aspects in their design process (Deasy and Lasswell, 1985).

In the provision of facilities for pedestrians, it seemed to be a general practice amongst designers/road safety practitioners to regard pedestrians very simply as just one simple flow (termed in this study as the 'macroscopic approach') – devoid of peculiar walking characteristics and behaviours (Fujiyama, 2005). Such an approach assumed that all the various users that make up the pedestrians need the same fundamental ergonomic requirements, have similar characters, and display similar behaviours on the street. This too simplistic approach clearly did not heed earlier reminders to designers against making assumptions about human behaviours in their design process.

This study looked into the implications of such a very simplistic design approach. More recent studies in relation to pedestrians' behaviours were reviewed and

analysed. Findings were categorised according to the various attributes of the pedestrians.

Arguments against the macroscopic approach

Several studies (e.g. Fujiyama, 2005) argued that while the macroscopic approach is suitable for planning/designing mass pedestrian facilities, such as a railway platform for commuting people, however it is not appropriate to obtain basic knowledge for the creation of a much convenient walking environment on the street. By adopting such a simplistic approach, pedestrian behaviours seemed to be poorly understood by designers/road safety practitioners. In particular are about the needs of the more vulnerable group – children, older people, and disabled people – which consequently are rarely assessed beyond their fundamental ergonomic requirements.

Others (e.g. Daff and Cramphorn, 2006) had called for those concerned to put themselves in the position of the users, prior to the design and provisions of road safety facilities for the pedestrians. The non acknowledgement on pedestrian behaviour by those professionals responsible often led to the non-conformity amongst pedestrian to the facilities provided can be fatal. Studies on vehicle-pedestrian crashes (e.g. Hughes, 1998) revealed that in majority of those conflicts, it was the pedestrians who were normally at fault. However, pedestrians would not commit faults if the facilities provided met their basic fundamental requirements. LTNZ (2004) cautioned that by assuming the average pedestrian to be fit, healthy, and non-disabled misrepresented significant proportion of the population. Localised pedestrian behaviours should be taken into account in the design process for the provisions of their facilities.

Understanding pedestrian behaviours

Pedestrians constitute a diverse group of users, whom vary in terms of age, gender, race, physical capabilities and limitations. Age is normally categorized into three groups:- 'children' (16 years old), 'adult' (16–65 years old, and 'older people' (above 65 years old). In terms of physical capabilities and limitations, pedestrians can be regarded as 'non-disabled' or 'disabled people'. The 'disabled people' constitute those who are partially sighted, blind, deaf or people with mobility impairments (on wheelchairs). Those varied attributes of pedestrians would naturally affect upon how they conduct themselves on the street. Findings about the variations and differences in pedestrians' behavioural patterns which call for attention towards their different fundamental ergonomic requirements had been reported in the literatures as discussed below.

Gender difference

In relation to gender differences, Steenbekkers, et al. (1999) found marked differences in the physical variables between the genders. They noted the difference

in body build and the amount of force exerted. At the same time, while differences were also found in the psychomotor variables, however few differences were found in the sensory and cognitive variables.

In terms of patience and conformity to pedestrian control signals, Holland and Hill (2007) revealed that women were less likely to intend to cross than men and perceived more risk. They thus were most likely not to violate the traffic signal.

Other different characteristics between the genders involved the use of hand phone while making the crossing. Females crossed more slowly and revealed more dangerous road safety behaviour such as less likely to look at traffic before and while crossing. Males were slow in crossings at unsignalised crossings (Hatfield and Murphy, 2007).

Age difference

In terms of the different age groups, many of the studies done seemed to revolve around issues that concerned older people (e.g. Oxley, et al., 2005). Findings revealed that older pedestrians have a higher risk of collision with road vehicles due to the perceptual, cognitive and physical deterioration associated with ageing. The failure of initial judgement combined with failure to modify behaviour increases the risk of collision, particularly in complex traffic situation. Further differences within the older people group, were attributed to differences in health and physical abilities rather than to differences in age. In addition, older pedestrians were observed to be more conforming and not violating the signal control. They showed more cautious behaviour in specific traffic situations. and more often choose to walk up to a pedestrian crossing if they can see one.

Amongst the few research that involved children as pedestrians (e.g. Lupton, et al., 2002) findings revealed that the children's main concern was the non compliance of drivers at road crossings or their non-reliability to stop at designated crossing points. Some of the children's choice of crossing point was influenced by perceptions of distance traveled or by social interaction. In terms of their decision making capabilities on various circumstances (safe and unsafe options) in making road crossings, children were not as smart as the adults pertaining to safer crossings and in the choice of safer gaps chosen. Large individual differences amongst the children were observed which suggested that some children were more at risk than others.

Disabled pedestrians

Different requirements are needed by disabled pedestrians.. For example blind people or visually impaired people require indicators such as kerbs, street furniture, audible signals, tactile paving to detect where they are and to orientate themselves. Thus, Hans Monderman's Shared Space concept should be implemented with cautious, for the sake of blind people or visually impaired people. As such, Pey (2006) recommended the creation or re-instatement of a pavement with a kerb, regular dropped kerbs and properly laid tactile paving. It should also be noted that sound cues are masked on crowded footways for visually impaired people and blind people. For people with mobility impairments, the width of footway should

allow the passing of wheelchairs with ease. There also should be proper care in the choice of surface materials as specifying the wrong material could result in slippery surfaces when wet, and much friction when dry.

While the provisions in facilities for the disable pedestrians are applauded, these should not however cause unnecessary inconvenience for non-disabled people. Kobayashi, et.al. (2005) argued that facilities such as tactile ground surface indicators installed on sidewalks help visually impaired people walk safely however, these indicators sometimes cause the non-disabled people to stumble. In order to solve the problem, they designed a new recessed tactile surface that may help both visually impaired people and older people with normal vision to walk safely.

Cross-cultural differences

Cross-cultural differences are central in the study of human behaviour (Abbas, 2000). Even though not much research has been done on racial differences amongst pedestrians, however evidence of differences was reported in Maryland, USA by Reed and Sen (2005) who found that African American people and Asian people were more aware of pedestrian safety and complied with laws as compared to their Caucasian counterparts. In another study which involved the comparison of parental risk perceptions on childhood pedestrian road safety in Sydney, Lam (2005) found that Chinese and Arabic speaking parents or caregivers perceived the road environments as significantly less hazardous for their 4–12 year old children as pedestrians when compared with the Vietnamese and English speaking parents.

Waiting time

In road crossings, Holland & Hill (2007) found that the intention to cross amongst the different age groups seemed to be due to differences in perceived value of crossing rather than differences in perceived risk. The probability for a pedestrian to cross the road, when it is unsafe, varies with waiting time, and age group.

In general people normally do not want to wait too long to cross streets. As signal waiting time increases, they get impatient and they either violate the traffic signal which places them at increased risk of being struck by a motor vehicle, or tended to cross at other locations thus, reducing the efficiency of the crossing for both people and vehicles. Tiwari, et.al. (2007) suggested that by reducing the waiting time for pedestrians not only decreases the probability of pedestrian crossers being hit by a motor vehicle, but also increases the efficiency of the crossing.

At times, pedestrians often experienced waiting unnecessary in making the road crossing despite the absence of threatening vehicles. Daff and Cramphorn (2006) blamed the signal designers' 'safe side conservative practice' approach for the unnecessary delay experienced by pedestrian. The approach make assumption for the benefit of slower pedestrian, or to give more time for speeding approaching vehicles to make the stop. This has often led to pedestrians not trusting the signal and hence violating it.

Walking speed

According to LTNZ (2004), walking speed has been shown to be affected by four factors:– pedestrian characteristics (such as age, gender, physical condition), trip characteristics (such as walking purpose, route familiarity, route length, etc.), route characteristics (such as width, gradient, pedestrian density, etc.) and environmental characteristics (such as weather conditions). The vast majority walk at speeds between 0.8 m/s to 1.8 m/s. The general speed of a fit healthy adult at around 1.2 m/s, while older pedestrians and those with mobility impairments between 0.9 m/s to 1.0 m/s.

Summary and conclusion

This study had shown that the variations of pedestrians in terms of age, gender, race, physical abilities and limitations influenced their behavioural patterns and characteristics on the street. In the provisions of facilities for the pedestrians, the general practice in the design process of considering pedestrians very simply as just one simple flow is too a simplistic a design approach. Failure to understand and acknowledge different user requirements would consequently not only result in the inconvenience and inaccessibility for pedestrians but more seriously detrimental towards their safety. This study suggests for the adoption of a much more localised inclusive approach when dealing with pedestrians rather than treating them as just one flow.

References

Abbas, M.Y. 2000, *Proxemics in Waiting Areas of Health Centres*: *A Cross-cultural Study*, PhD Thesis, The University of Sheffield, U.K.

Daff, M., and Cramphorn, B. 2006, Signal design for pedestrians as if they were thinking adults, *The 7th International Conference on Walking and Liveable Communities*, October 23–25, Melbourne, Australia

Deasy, C.M. and Lasswell, T.E. 1985, *Designing Places For People: A handbook on human behaviour for architects, designers and facility managers*, Whitney Library of Designs, New York

Fujiyama, T. 2005, *Investigating Use of Space of Pedestrians. Working Paper*. Centre for Transport Studies, University College London

Hatfield, J., and Murphy, S. 2007, The effects of mobile phone use on pedestrian crossing behaviour at signalised and unsignalised intersections, *Accident Analysis and Prevention*, **39**, 1, 197–205

Hughes, C.J. 1998, Pedestrian safety at signalised junctions, *Road Transport Information and Control*, 1998, 21–23 April, 50–54.

Holland, C., and Hill, R. 2007, The effect of age, gender and driver status on pedestrians' intentions to cross the road in risky situations, *Accident Analysis and Prevention*, **39**, 2, 224–237

Kobayashi, Y., Takashima,T., Hayashi, M., and Fujimoto, H. 2005, Gait Analysis of People Walking on Tactile Ground Surface Indicators, *Transactions on Neural Systems and Rehabilitation Engineering*, **13**, 1, 53–59

LTNZ. 2004, *Pedestrian Network Planning and Facilities Design Guide*, Land Tranport New Zealand, http://www.landtransport.govt.nz

Lam, L.T. 2005, Parental risk perceptions of childhood pedestrian road safety: A cross cultural comparison. *Journal of Safety Research*, **36**, 2, 181–187

Lupton, K., Colwell, J. and Bayley, M. 2002, Aspects of children's road crossing behaviour. *Proceedings of the Institution of Civil Engineers: Municipal Engineer*, **151**, 2, 151–157

Oxley, J.A., Ihsen, E., Fildes, B.N., Charlton, J.L. and Day, R.H. 2005, Crossing roads safely An experimental study of age differences in gap selection by pedestrians, *Accident Analysis and Prevention*, **37**, 5, 962–971

Pey, T. 2006, "Nine out of ten cars will stop . . . What about the tenth?" *Building Engineer*, **81**, 10, 14–15

Reed, R., and Sen, S. 2005, Racial Differences and pedestrian safety. Some evidence from Maryland and implications for Policy, *Jnl of Public Transportation*, **8**, 2, 37–61.

Steenbekkers, L.P.A., van Beijsterveldt, C.E.M., Dirken, J.M., Houtkamp, J.J., Molenbroek, J.F.M. and Voorbij A.I.M. 1999, Design-relevant ergonomic data on Dutch elderly, *Injury Control and Safety Promotion*, **6**, 3, 99–115

Tiwari, G., Bangdiwala, S., Saraswat, A., and Gaurav, S. 2007, Survival analysis: Pedestrian risk exposure at signalized intersections, *Traffic Psychology and Behaviour*, **10**, 2, 77–89

PATTERN AND TREND OF PEDESTRIAN CASUALTIES IN THE UK: A CASE OF MISUNDERSTOOD ERGONOMICS?

Mohamed Yusoff Abbas[1], Marcus Ormerod[2], Rita Newton[3], Hamish MacLennan[4] & Mohammad Faruk[5]

SURFACE Inclusive Design Research Centre, BuHu Research Institute, The University of Salford Maxwell Building, Salford M5 4WT
[1]*Assoc. Prof., Dr., Fac. of Arch., Plan. and Survey., Univ. Tek. MARA, Malaysia*
[2]*Professor & Director, SURFACE Inclusive Design Research Centre*
[3]*Prog. Dir. PGCert/Dip/MSc Accessibility & Inclusive Design, The Univ. of Salford*
[4,5]*PhD Candidate, SURFACE Inclusive Design Research Centre*

The UK's national policy in walking unveiled in 2000, is towards a more convenient and safer road environment. This study investigated how far those objectives have been met five years after it was launched. Data based on the Department for Transport's (DfT) pedestrian casualties statistics was used in the analyses. Overall findings showed a downward trend in pedestrian casualties. However, peculiar pattern and trend which involved pedestrian casualties were also observed. The common factor that linked with such pattern and trend seemed to involve conforming to pedestrians' requirements. Despite the availability of several guidelines and best practices, could it be that the ergonomic requirements of the pedestrians been misunderstood?

Introduction

The UK's national policy on walking is towards the creation of a much more easier, pleasant and safer environment (DETR, 2000). How far has this been successful since it was launched in the new millennium? Have pedestrians' safety been improved? What about provisions pertaining to the more vulnerable road users – children, older people, and disabled people? Are the provisions more convenient and accessible? Have pedestrians' requirements been understood and acknowledged in practice?

This study investigated the pattern and trend of pedestrian casualties within the succeeding five year period after the policy was launched. The purpose being to gauge the successfulness of the policy that had been put into practice. The Department for Transport's statistical data on pedestrian casualties from 2001 to 2005 was used as the basis for the analyses. Focus was on peculiar pattern and trend of pedestrian casualties which were against the interest of pedestrian safety. Salient features of the analysis were then highlighted. These were then categorized

into main pattern and trend. Related issues were then identified in determining the possible causation factors, with suggestions for possible remedial measures.

The analyses

Categorisation

The analyses were categorised into four sections:- A) User Type; B) Location and Circumstances; C) Age Band, Location and Circumstances; and D) Location and Pedestrian Crossing Type. The analyses for the 5-year period were based on the yearly per 100,000 populations for all accumulated severities.

Pattern and Trend (P&T)

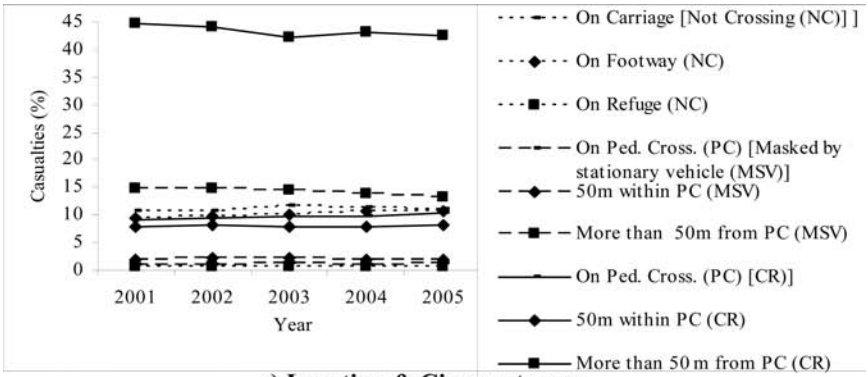
The findings are as shown in Figure 1. Salient features of the findings are categorized under 3 main pattern and trend as summarized below:-

P&T 1) Higher pedestrian casualties together with an upward trend
These involved casualties on the pedestrian crossings as compared to a distance within 50 m from the crossing, which particularly involved older people (see Figures 1a and 1b).
P&T 2) Upward trend of pedestrian casualties
These involved locations at light controlled junction, footways, and distance more than 50 meters away from a pedestrian crossing, which particularly involved adults (see Figure 1b).
P&T 3) Higher pedestrian casualties
These involved:-
a) Pedestrians as second highest casualty users after the car users.
b) Locations at:-
- i) more than 50 m from a pedestrian crossings; esp. children in the masked circumstances (see Figure 1b)
- ii) carriageway and on footway involved the adult (see Figure 1b)
- iii) signal-control pedestrian crossings as compared to zebra crossings (see Figure 1c)
- iv) non-junction pedestrian light control (see Figure 1c).

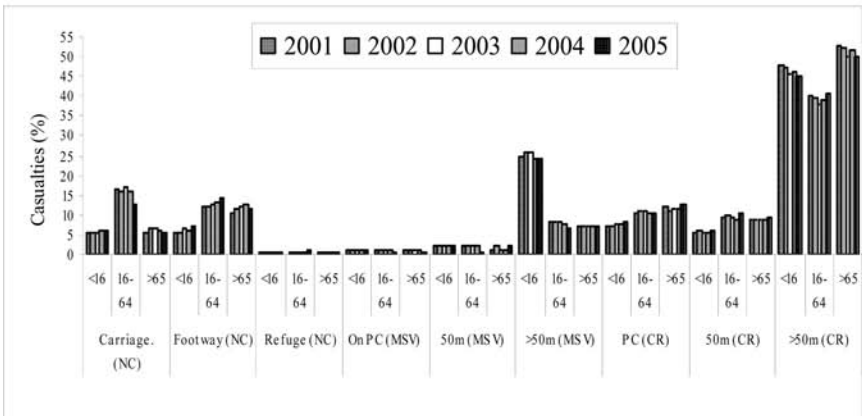
Discussion

Casualties higher on pedestrian crossings

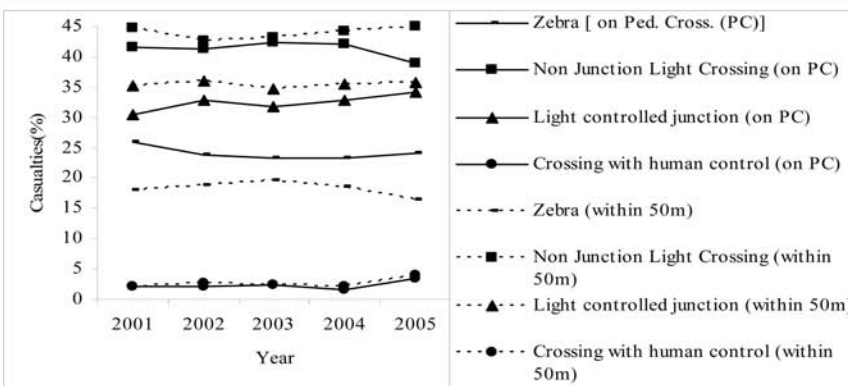
P&T 1 above does seemed alarming as pedestrian crossings were provided for the convenience and safety of pedestrians. In fact, similar findings were also reported in the Netherlands (SWOV, 2006). Other studies (e.g. Hughes, 1998) have shown that in vehicular-pedestrian conflicts, majority of the faults were caused by pedestrian behaviour. With the availability of several guidelines and recommended good practices (e.g. DFT, 2000; LTNZ, 2004) those findings do seemed odd. As



a) Location & Circumstances



b) Age Band, Location & Circumstances



c) Pedestrian Crossing Type & Location

Figure 1. Pattern and trend of pedestrian casualties per 100,000 population in UK – 2001 to 2005.

adult pedestrians seemed to be minimally involved, could it be that provisions of pedestrian facilities were based entirely on requirements of adult pedestrians, forgoing fundamental ergonomic requirements of the more vulnerable users?

Those findings seemed to further increase the suspicions that provisions of the pedestrian facilities have not taken into considerations requirements of the pedestrians fully. The reason being if those requirements had been considered, the provisions of facilities would have accommodated the pedestrians behaviour, thus maximizing pedestrians safety. Practices of regarding a typical average pedestrian as healthy, fit and able will misrepresent significant proportion of the population had been cautioned by LTNZ (2004, p.21).

Providers of pedestrian safety should realize that pedestrians are of different age groups, gender, race and physical abilities. Their behavioural pattern involving risks of crossing and walking speeds greatly varied. In relation to disabled people, there seemed to be contradictory opinions regarding the designs of road crossings with flush kerbs and the development of the tactile paving system to benefit visually impaired people. Pey (2006) called for the re-instatement of a pavement with a kerb, regular dropped kerbs and properly laid tactile paving. Kobayashi, et al. (2005) argued that facilities such as tactile ground surface indicators installed on sidewalks help visually impaired people walk safely however, these indicators sometimes cause the non-disabled people to stumble. In order to solve the problem, they designed a new recessed tactile surface that may help both the visually impaired people and older people with normal vision to walk safely. In addition to other safety countermeasures that involved visibility, orientation and clarity, Leden and Johansson (2006) suggested that speed cushions be situated at a longer distance from the marked crosswalks. This could reduce the actual and perceived threat of traffic amongst pedestrians. Hence, those fundamental ergonomic requirements should be understood, acknowledged and taken into account in the provision of facilities for them.

Time factor

Pedestrians, especially the older people often complained of not enough time allocated by the signal-control for them to make the crossing. In addition, it should be understood that walking rates are also influenced by a variety of other factors which included vehicle volumes, number of pedestrians crossing in a group, the signal cycle length, and the timing of the various pedestrian-signal phases.

Tiwari (2007) linked the pedestrian behaviour factor with the time factor (waiting time to cross, and the length of time to make the crossing) as causal factors that might have led to the fatalities. Being human, pedestrians can be impatient at times to make the crossing, especially more so in the absence of vehicles threatening the crossing, even though the signal-control does not allow them to cross. Thus, pedestrians often do not conform to the requirements of the signal-control, hence risking their lives in making the crossing. Daff and Cramphorn (2006) blamed the safe side conservative practice approach as the reason for such unnecessary delay in waiting time. The approach make assumption for the benefit of slower pedestrian, or to give more time for speeding approaching vehicles to make the stop. This

has often led to pedestrians not trusting the signal and hence violating it. Halden (2005) noted that due to the long waiting time, pedestrians tended to cross at other locations, which tended to place them into further risks, thus making the signal control inefficient for both drivers and pedestrians.

Zebra crossings much safer?

Despite the improvements and ingenuity of the more recent light-controlled type of pedestrian crossings being used, the "older' zebras seemed to have recorded a much lower proportion of casualties compared to the former as highlighted in P&T 2 and P&T 3. Even though such findings seemed to be contradictory to other findings (e.g. LTNZ, 2004), however Halden (2005) argued otherwise. Not only that he found delays to pedestrians at pelican crossings were five times more than at the zebras, he also opined that it is not necessary to adopt pelicans at higher vehicle flows to reduce vehicle delays although those might still be the better option, for example the refuge islands. Such provisions not only could solve the problem of slower pedestrians but also increase their confidence in making the crossing. In fact, Fidler (2007) concurred with this, and suggested that those be made wider. Seats as a resting place should also be considered at the refuge islands for the convenience of older people.

Summary/Conclusion

This paper had analysed the pattern and trend of pedestrian casualties in UK over the period from 2001–2005, based on the Department for Transport's (DfT) statistics. Overall, a downward trend on pedestrian casualties was observed during the five-year period. However, peculiar pattern and trend on pedestrian safety were also revealed. These were all linked to one common factor – fulfilling the requirements of pedestrians. This study opines, despite availability of good practices and guidelines, objectives in maximising pedestrian safety and convenience would be difficult to achieve if pedestrian requirements are still not truly understood and fulfilled. Provisions of pedestrian facilities and crossings should be truly ergonomically and inclusively designed to accommodate requirements of the pedestrians, particularly the vulnerable users.

References

Daff, M., and Cramphorn, B. 2006, Signal design for pedestrians as if they were thinking adults, *The 7th International Conference on Walking and Liveable Communities*, October 23–25, Melbourne, Australia

DfT. 2000, *Tomorrow's Roads Safer for Everyone*, Department for Transport, The Stationery Office, London

Fidler, E. 2007, A Dance with Death: Pedestrian Safety on Route 1, *Rethink College Park*, http://www.rethinkcollegepark.net

Halden, D. 2005, Civilising Pedestrian Crossing Facilities, *European Transport Conference*, October, Strasbourg

Hughes, C.J. 1998, Pedestrian safety at signalised junctions, *Road Transport Information and Control*, 1998, 21–23 April, 50–54

Kobayashi,Y., Takashima,T., Hayashi, M., and Fujimoto, H. 2005, Gait Analysis of People Walking on Tactile Ground Surface Indicators, *Transactions on Neural Systems and Rehabilitation Engineering*, **13**, 1, 53–59

LTNZ. 2004, *Pedestrian Network Planning and Facilities Design Guide*, Land Tranport New Zealand, http://www.landtransport.govt.nz

Leden, L.,Garder, P., and Johansson, C. 2006, Safe pedestrian crossings for children and elderly. *Accident Analysis & Prevention*, **38**, 2, 289–294

Pey, T. 2006, "Nine out of ten cars will stop. . . What about the tenth?" *Building Engineer*, **81**, 10, 14–15

SWOV. 2006, *Institute for Road Safety Research, Fact Sheet, Pedestrian*, September, SWOV, Leidschendam, Netherlands

Tiwari, G., Bangdiwala, S., Saraswat, A., and Gaurav, S. 2007, Survival analysis: Pedestrian risk exposure at signalized intersections, *Traffic Psychology and Behaviour*, **10**, 2, 77–89.

SAY WHAT YOU SEE? THE EFFECTS OF BEING ABLE TO SEE WHAT'S HAPPENING ON VOCABULARY DESIGN FOR SPEECH INPUT

Alex Stedmon & Manjit Aujla

Human Factors Research Group, University of Nottingham, Nottingham UK

This paper considers the way that vocabulary is used depending on the ability of users to see what they are doing. Blind, blindfolded and able-bodied participants took part in a user trial to complete a speech based task. A 'Wizard of Oz' paradigm was used to mimic the use of a speech recognition system that allowed user to speak freely and still be recognised without any training. The results indicated that the different user groups used speech in different ways. The paper will present these results and develops a number of recommendations for speech input.

Introduction

Speech input is used in a wide range of applications from simple text entry to hands free interfaces for mobile technologies in robotics and virtual reality (Stedmon, et al, 2005). In many automated processes (such as the use of robots in industry or non-transparent interfaces) it may not always be possible users to physically see what is happening. As a result they may have to work 'blind' in order to conduct a task outside their line of sight or when using speech input in advanced interfaces where it is not possible to 'see' that processes are taking place within the software.

Speech input software has many limiting factors (for a detailed investigation please refer to Stedmon, 2005) of which one area is the vocabulary used to develop command structures (the words people are allowed to speak, based on what the system is able to recognise). When the vocabulary is restricted, users have to learn specific words in order to interact successfully. If they say anything outside of the recognised vocabulary the system will fail in some way.

A more user-centred approach is to develop vocabularies that are easier to learn and use. The concept of free-speech therefore allows users to speak any commands they wish rather than rely on a prescribed set of commands which can be hard to remember (Baber, et al, 1996). The underlying principle is that free-speech does not have to be learnt in the same way as a restricted vocabulary and that is more intuitive by exploiting commonly used words.

The idea of developing intuitive vocabularies is still quite new. Previous technical development has focussed on expanding the vocabulary sets that speech interfaces refer to in order take account of more words that might be used for similar instructions. A well chosen vocabulary is still seen as the best method of reducing errors

(Rodman, 1999) rather than a 20,000 word vocabulary that is more likely to have words that sounds very similar and therefore easily confused by the system (Gold & Morgan, 2000). With this in mind, it is necessary to investigate the language that potential users will adopt and incorporate this into the design of vocabularies for future use. This provided the basis for the following user trial which sought to address the way that people use language depending on the degree to which they could see what they were doing.

Method

Participants

19 participants took part in the user trial: 5 blind users (mean age = 55.2 years), 7 blindfolded users (mean age = 22.4 years), and 7 able-bodied users (mean age = 21.0 years). The blindfolded, able-bodied and one of the blind participants were recruited from the student population at the University of Nottingham. The remaining blind participants were recruited from the Nottinghamshire Royal Society for the Blind (NRSB). The blind users were familiar with speech input and screen reading software, whilst the other participants had limited experience of either.

Apparatus

A Sony VAIO laptop was used running JAWS v.8 screen reading software and Microsoft Word. A digital recording device was used to record the speech of each participant so that it could be analysed. A stopwatch was used to log performance on the task and a checklist was used by the experimenter to record data.

Design

The task as based on editing a Microsoft Word document using speech input. A 'Wizard of Oz' paradigm was employed where the experimenter carried out the actions that participants instructed the interface to perform. The experimenter therefore performed the role of a free-speech recogniser without the technical problems associated with using speech recognition software. The screen reading software then communicated the performed actions back to the participants. This paradigm has been used successfully in previous speech input research and meant that participants did not have to train speech recognition software prior to the trial or contend with any problems due to mis-recognition errors (Stedmon, 2004). Measures were taken for task completion time, number of instructions repeated, assistance required, number of errors, mean number of words used to complete the task, total number of words used to complete the task, level of difficulty and a rating of speech input.

Procedure

Participants conducted the speech task once only. The task actions required were to: open a file, request the computer to read it aloud, insert a specific graphic file

Table 1. Summary results for the different participant groups.

	Blind	Blindfolded	Able-bodied
Task completion time (mins)	13.04	9.21	7.15
Mean instructions repeated	1.6	0	0.14
Mean times asked for assistance	2	1	1.14
Mean errors made	0.2	0.43	0
Mean words used	120	121.29	104.43
Total words used	108	96	90
Level of difficulty (0 = difficult, 10 = easy)	8.4	7.0	5.9
Rating of speech input (0 = negative, 10 = positive)	8.8	3.43	5.43

from clipart, edit the title of the text, edit the header and footer, and finally save and close the file. The trial took approximately 30 minutes to complete and participants were paid for their time.

Results

A summary of the task based data are presented in Table 1 above.

Discussion

The blind participant group took the longest time to complete the task as they tended to give longer commands. The able-bodied users tended to give one word commands which made the task quicker to perform. Also as the mean age of the blind participants was much higher than the other participant groups they may have had trouble understanding the task instructions which appears to be reflected in the number of times they asked for them to be repeated. The blindfolded participants also took longer than the able-bodied participants to complete the task. Like the blind participants they used longer commands and often took longer to think about what they needed to do next which compounded the task completion time. The blindfolded participants were the only participants operating in an environment they were not used to whereas the blind and able-bodied users were familiar with their modes of interaction (even if perhaps they were not so used to performing the task). The able-bodied participants completed the task in the quickest time but as they could see the screen and use this for feedback in addition to the screen reader software they did not need to visualise the screen and could use shorter commands to complete the task.

The blind participants asked for more instructions to be repeated and as already stated this could have been due to them being considerably older than the other two user groups. It was surprising that the blindfolded users did not require any instructions to be repeated as it was expected that with them performing the task in an unfamiliar situation they might need further assistance. This finding is supported by the number of times participants requested assistance. The blind participants required more assistance than the other participants but again the blindfolded users

required less than the able-bodied participants. However, when taken with the number of errors, blindfolded participants made the most mistakes and may have made less mistakes if they had asked for more help. The number of words used illustrated that the able-bodied participants used less words in total to complete the task which translated to shorter commands. Both the blind and blindfolded participants used more words as they gave more detailed commands to compensate for the lack of visual cues.

Participants rated how difficult they felt the task was and all groups rated it as being easy to perform. The findings illustrated that the users who could see what to do in the task rated it as more difficult than either of the participant groups who could not see the task interface. From these results it would appear that being able to see the screen made it more difficult to rely only on speech. The able-bodied and blindfolded participants said they found it frustrating not being able to use more conventional interaction devices such as a keyboard and mouse. The highest rating came from the blind users which, as they were used to speech input and screen reading software, illustrated that they found the task easier to complete than the other participant groups. However, the blind participants took longer, required more instructions to be repeated and asked for assistance more than the other participant groups. This could indicate that whilst they found the task easy to complete they were more careful to perform it properly by taking their time and getting as much information as possible as they went through the different stages. The ratings for speech input also illustrated that the blind participants preferred it the most. This is probably a reflection of their familiarity with the technology which might also explain why the other participant groups did not rate it very highly. Many of the able-bodied and blindfolded participants were said they would prefer to use conventional interaction methods rather than speech input even though the trial had been designed to mimic a system that allowed them to use free-speech. From this study a number of observations and recommendations are summarised in Table 2 below.

One of the aims of this study was to investigate the use of intuitive and uncommon commands for speech input and how users from different backgrounds might use different vocabularies to complete the same task. The language used by the blind and blindfolded participants tended to be more informal and relaxed than able-bodied participants who used more formal commands. This would seem to indicate that when participants could not directly see the task interface their speech was different to those that could see it. The language participants used tended to be similar and many words were based on the instructions provided (such as ‘open …’, ‘close …’, or ‘insert …’) which were then used as part of the following command participants spoke. As such, it can be argued that these were not freely spoken in the strictest sense however they provide an insight in the convergence effect that can happen when users begin to take on the existing vocabulary a system uses.

Conclusion

This investigation recruited blind, blindfolded and able-bodied participants in order to investigate issues of task performance and vocabulary sets when using a speech

Table 2. Observations and recommendations.

Observations	Recommendations
– Blind and blindfolded participants tended to use multiple word commands and took longer to perform the task	– If users do not have visual contact with a task and/or interface, anticipate longer command structures and/or task completion times
– Able-bodied participants tended to use shorter commands (single words) and completed the task more quickly	– Where time is critical it is beneficial for users to have direct visual contact with the task and/or interface they are using
– Blind participants and blindfolded participants used a more informal vocabulary	– When visual cues are missing, anticipate that users will use larger vocabulary sets
– Blind participants requested instructions to be repeated more than the other participants	– If users are unfamiliar with a task provide more instruction in order to support their understanding
– Blindfolded participants made the most errors	– If users are operating outside of their usual experience be prepared for them to make more mistakes and not request information
– Able-bodied participants judged the task to be more difficult than the other participants	– In order to simplify the interaction it might be useful to break visual contact with the task and/or interface
– Blind users were more supportive of the idea of speech input	– If users have already had experience of speech input they be more prepared to used it for new tasks and/or interfaces

interface and screen reader as a task interface. Whilst there were clear differences between the performance of blind and able-bodied participants, blindfolded participants presented characteristics of both blind users (such as using longer command structures and finding the task easier) and able-bodied users (such as shorter task completion time, fewer repeated instructions or requests for assistance). It is apparent that all user groups dealt with the task differently and that this provides an insight into developing suitable vocabularies for speech input systems according to different user needs.

Acknowledgements

The authors would like to thank the NRSB who for their assistance with this study.

References

Baber, C., Mellor, B., Graham, R., Noyes, J., & Tunley, C. (1996) Workload and the use of ASR: the effects of time and resource demands, *Speech Communication*, 20, 37–53.

Gold, B., & Morgan, N. (2000) Speech and audio signal processing. *Processing and Perception of Speech and Music*. John Wiley & Sons: London.

Rodman, R.D. (1999) *Computer Speech Technology*. Artech House, Inc: New York.

Stedmon, A.W., Nichols, S.C., Patel, H., & Wilson, J.R. (2004) Free-speech in a virtual world: A user-centred approach for developing speech commands. In, P. McCabe (ed). *Contemporary Ergonomics 2004.* Taylor & Francis Ltd: London.

Stedmon, A.W. (2005) *Putting Speech In, Taking Speech Out: Human Factors in the Use of Speech Interfaces.* PhD Thesis: University of Nottingham.

INCLUSIVE COMMUNICATION – THE OTHER BREAKTHROUGH

Robert Chesters[1], Richard Baines[2] & John Mitchell[3]

[1] *Intelligent Health Network Manager, Medilink West Midlands*
[2] *Director of Sustainable Development, Black Country Housing*
[3] *Principal, Ergonova*

The 'Five Keys to inclusion' (1) were developed as a means of integrating and addressing the various barriers that disable people with different impairments.

It proposes a coherent, common sequence that everyone appears to follow when they are interacting with people, places, products and processes (2).

Before anyone can interact with any of these, they must first find them, then reach them, understand how to interact with them, relate to them without being intimidated and finally, engage in two-way interaction with them. Non-inclusive items prevent many users, including people with different impairments, from 'turning' one or more of these 'Keys'. After establishing this sequence, the *i*-House team was asked firstly to produce a guidance specification for sustainable, inclusive home refurbishment (3) and then to use this for refurbishing a Victorian end-of-terrace house in West Bromwich, Birmingham.

Disabled home users in Dudley (3) and in Sheffield (4) had been consulted about the barriers in homes that they would like to see 'refurbished out'. The 'Triple Walkthrough' (5) was then developed to enable disabled people to test out the house for themselves and find where inclusion had and had not been fully achieved. Surprisingly, only two barriers were exposed in this way, non-inclusive physical access and non-inclusive communication. While the effects of non-inclusive access are already widely recognised, the pervasive and far-reaching effects of non-inclusive communication have still to be fully appreciated and addressed. This paper uses the 'Five Keys' and the 'Triple Walkthrough' to reveal the areas in which non-inclusive communication limits the functions of disabled home users, makes them dependent on other people and sometimes also puts them at risk.

Background

People spend a large proportion of their lives in their homes and use them for a whole series of activities that are vital for their wellbeing, productivity and fulfilment (Bennington (8), Autonomy). This paper is one of three submitted to the conference,

please also see Inclusion – blending user and professional insights (6) and 21 Building Blocks for 21st Century Inclusion (7).

Many older homes contain substantial disabling barriers that disrupt these activities. Although young, agile and unencumbered people may not even be aware of them, the restrictions they place on people with impaired sight, hearing, movement/dexterity, understanding and mental health can be very heavy indeed.

Historical barriers and building technology

Many home barriers date from the 19th and 20th centuries when contemporary building technology was powerless to overcome them. Since then, knowledge, building technology and ICT has moved on considerably. Barriers which could not have been overcome in the past could now be avoided quite easily.

In response to this the *i*-House team in the West Midlands has developed a network of disabled people and local industries to find and develop technologies that can ensure homes, both old and new can become fully inclusive and sustainable. The task of identifying the barriers that must be 'designed-out' was carried out in Sheffield (3,5) and in Dudley (4). The team asked groups of home users with various impairments about the barriers in their present and previous homes and what their priorities would be for inclusive refurbishment.

Integrating and addressing barriers and their effects

It was not easy to analyse disabled peoples' feedback on these barriers because, with the exception of access barriers which mainly affect people with physical impairments, there did not seem to be many obvious links between particular barriers and particular impairments, such as impaired hearing, understanding, mobility or mental health. Further consideration suggested that there might be a common sequence of use that people generally appear to follow when they are trying to use things.

The sequence, known as the 'Five Keys to Inclusion', suggests that before people can interact with other people, places, products or processes, they must first find them, then reach them, understand how to interact with them, relate to them without being intimidated and finally, engage in two-way interaction with them.

The sequence effectively integrates barriers against people with different impairments. For example, people with impaired sight, hearing and understanding can have difficulty in finding their way, as can people with impaired mobility if information and signage is placed outside their line of sight. Barriers such as this can prevent many users, including people with different impairments, from 'turning' one or more of these 'Keys'. A useful example of this can be found when looking at common consumer products such as the mobile phone illustrated in Figure 1.

Identifying the barriers that disrupt the 'Five Keys'

The 'Five Keys' provided a functional sequence that all users, regardless of any impairments, appeared to follow whenever they were using things. This sequence

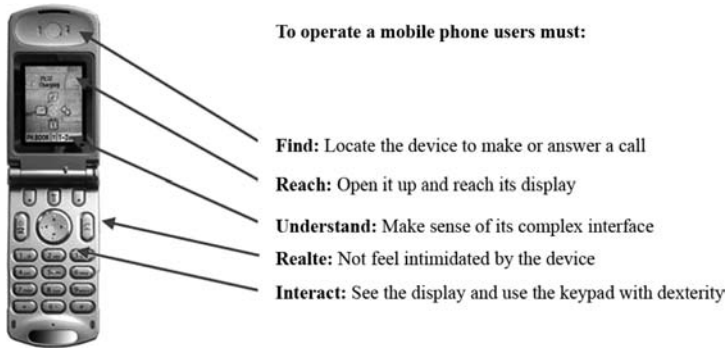


Figure 1. Turning the five keys to operate a mobile phone.

was also helpful because it raised the question about the barriers that prevented people from turning one or more of the keys. What were these barriers and was there any commonality between them? Surprisingly, only two barriers were exposed in this way, non-inclusive physical access and non-inclusive communication. While the effects of non-inclusive access are already widely recognised, the pervasive and far-reaching effects of non-inclusive communication have still to be fully appreciated and addressed.

Integrating inclusion into home refurbishment

On the basis of their work so far, Urban Living, the Government Housing pathfinder for Sandwell and Birmingham commissioned the *i*-House team produce a guidance specification for sustainable, inclusive home refurbishment (3). Subsequently, Advantage West Midlands the regional development agency for the West Midlands has provided funding to apply the standard to a dilapidated end-of-terrace house in West Bromwich as a demonstrator of the principles involved.

The team then required a structured, systematic approach for exposing the barriers in the demonstrator home and for deciding how to overcome them. A similar structured approach will also be required by disabled people when they visit the refurbished home on completion and evaluate its inclusively for themselves. To enable them to accomplish these tasks, the team then developed the 'Triple Walkthrough' (5).

The 'Triple Walkthrough'

The 'Triple Walkthrough' aims to group key areas of home function together into coherent systems to make them easier to evaluate and address. The first walkthrough is concerned with finding, reaching and getting into, around and out of homes. The second walkthrough is concerned with communicating and controlling systems within the home regardless of user's impairments. The third walkthrough looks at using the facilities within a home, such as those found in the kitchen, the bathroom, the living room and the bedroom.

Applying the 'Triple Walkthrough'

The selected demonstrator house typifies a standard Victorian terrace with inherent physical, sensory and cognitive barriers. Based on the leaning from the Five Keys research the *i*-House team developed a design proposal for the demonstrator house. Attention was given to the design of every aspect from entering the property from the street, interaction with the front door to gain access right through to the information and environmental control needs within the house. The walkthrough methodology was then used with the end user groups to test the design and identify any significant issues. Each user group was presented with a PowerPoint presentation and taken through the building design proposal on plan at the 3 walkthrough levels.

The walkthrough model was crucial in verifying the design, in particular the physical access. However, the most notable findings came about from the communication walkthrough research. The team found that communication between the user and the systems within the home was an equal barrier to physical access. From this research significant investment has been made in developing a smart house which can provide audible and visual feed back to the user and a level of built in intelligence. Central to the house is an integrated network of sensors and actuators which are controlled by a central computer and a remote control unit.

Some key systems include a concierge unit on the front door with bio metric finger print reader replacing the key. When a registered friend, carer or relative activates the biometric lock an audible message is relayed to the occupant via ceiling mounted speakers in each room to say who is entering the house. This also logs when they entered and exited the property. The computer driven system can also activate lighting automatically, control remote opening of doors and windows, and control the central heating all with audio and visual feedback to the user. The system can announce and act on numerous situations, for example window in bedroom 1 open, closing down the radiator valve to conserve heat.

This central control facilitates a 2 way communication between the occupant and the house. It is possible to profile rooms within the property against personal requirements. For example in the kitchen a biometric finger print reader is being used to enable users to 'Log in' to elements within the kitchen. For example the system can allow certain users access to the cooker and kettle, for example a carer. These appliances may not be safe for children in the house or an elderly occupant with dementia.

In addition to this system the house features a programmable water management system. This system provides temperature controlled water to sinks and the shower via touch activated taps. This was another feature welcomed by the user group participants; the system removes the need to manually activate taps which can be difficult for those with poor dexterity. The built in intelligence also ensures that once the sink is full the water supply is cut off removing the risk of a tap being left on. This is a common problem with users who have dementia or learning difficulties. The bathroom and en-suite also feature hands free flush on the toilets, the user activates the flush by passing a hand over the sensor plate. Both of these systems also promote hygiene by minimising contact surfaces.

All of these environmental control options will be activated from a hand held remote control unit and through standard telephone key pads.

This is an exciting new development which is intended to illustrate the potential for housing which is flexible to changing need. The team are exploring options to roll out the principles to a larger number of homes, perhaps within a housing association. This would enable a larger scale evaluation of the design standard. More information on the house can be found at http://www.medilinkwm.co.uk/ihealth

References

1 Mitchell, J and Chesters, R, 2005, *'The Sheffield Formula for Inclusion'*, Include, the Helen Hamlyn Foundation, Royal College of Art, London, ISBN 1-905000-10-3.
2 Mitchell, J, 1980, '*Problem-solving in Ergonomics*', Keynote presentation at Ergonomics in Action', IEA Conference, Oslo.
3 Mitchell, J, Chadwick, A and Hodges, B, 2004, '*Why can't we have a Home with Nothing Wrong in It?* – disabled peoples' aspirations and barriers in inclusive living at home', / Report to the Sheffield Housing Department, Ergonova, Sheffield.
4 Baines, R, Chesters, R and Mitchell, J, 2007, '*A design standard for sustainable, inclusive housing refurbishment*, urban Living, West Midlands.
5 Mitchell, J, Barraclough, K, Davis, G, Gamage, Sean, Hodgkinson, C and Smethurst, J,2006, *'Innovation and Collaboration between users and providers'*, Contemporary ergonomics, Taylor and Francis and the Ergonomics Society, London and Loughborough, ISBN 10 0 415 39818 5, pp. 395–398.
6 Baines, R, Chesters, R and Mitchell, J, 2007, '*Inclusion – blending user and professional insights'*, Contemporary Ergonomics 2007, Taylor and Francis and the Ergonomics Society, London and Loughborough.
7 Mitchell, J, Baines, R and Chesters, R, 2007, *'21 Building Blocks for 21st Century Inclusion'*, Contemporary Ergonomics 2007, Taylor and Francis and the Ergonomics Society, London and Loughborough.
8 Bennington, J, McClenahan, J and Mitchell, J, 2000, *'Autonomy for Disabled People – Mutual Problems, Mutual Solutions'*, Report to the King's Fund, London.

INCLUSION – BLENDING DISABLED USER AND PROFESSIONAL INSIGHTS

Richard Baines[1], Robert Chesters[2] & John Mitchell[3]

[1]*Director of Sustainable Development, Black Country Housing*
[2]*Intelligent Health Network Manager, Medilink West Midlands*
[3]*Principal, Ergonova*

The project team is seeking to graft inclusion and sustainability into existing methods of home refurbishment (1,2,3,4,5).

New insights can sometimes meet resistance if they are not seen as being clear, valid, easily verifiable, technically feasible and easy to apply. To achieve this, the team began their task by producing a performance specification for sustainable, inclusive housing refurbishment (5). The effectiveness of the specification is currently being demonstrated/tested by refurbishing a dilapidated 19th century, 'end of terrace' house in West Bromwich. The project team combines the expertise of a disabled ergonomist, a building engineer and a product designer.

Much of the environmental and product information that they required was drawn from first hand development of future practice, existing good practice and from publications (5). However, detailed user feedback is harder to come by, particularly for people with sensory or cognitive impairments (1,2,3,4). Accordingly, the team consulted disabled home users in Sheffield (1,2,3,4) and in Dudley (5) about the barriers that they face in their homes.

Gathering and using user feedback tends to be a rather circular process in which participants and researchers work together to pool and make sense of their insights through questions and the development of answers. Initial consultations (1,3) had suggested that barriers against people with different impairments could be understood and addressed through the 'Five Keys to Inclusion'. Subsequent consultations (5) suggested that non-inclusive access and information are the two underlying causes of exclusion in homes.

In this workshop, the project team and five project participants will explain how they are working together to:

- Expose the barriers that restrict them in their homes
- Achieve inclusive, sustainable home refurbishment
- Test for full inclusion in homes

Background to the project

The project team's insights are drawn from the three separate but linked areas of:

- Environmental sustainability, architecture and building technology/engineering
- Design development, design technology and industrial regeneration
- Usability and inclusion

These insights have been pulled together over the past 4 years to address the opportunities presented by developing inclusive, sustainable housing.

In 2006 the team were commissioned by Urban Living (the West Midlands Housing Market Renewal Pathfinder) to develop a design standard for sustainable, inclusive housing refurbishment (5). Urban Living wished to maximise the human, environmental and economic value of home refurbishment. They therefore asked the team to produce a specification/guide for housing developers that would explain:

- Why refurbishment should be fully inclusive and sustainable
- Where the major human and ecological problems occur at present
- How these problems are caused
- How they could be overcome with better design, components, products and systems

Subsequently, Advantage West Midlands have asked the team to refurbish a dilapidated, 19th century 'end of terrace' house in West Bromwich as a demonstrator and test of the specification.

Gathering the environmental and design information

Information on design is already widely available and understood. The team had already developed an in-depth knowledge of sustainability. The principle task was therefore to review the 'state of the art' in each of these two areas. This was done without revealing any unexpected priorities for further development. This is hardly surprising because developments in both design and building technology are entirely driven by factors that lie completely outside of the project. The team therefore waited with interest to see what priorities would emerge from disabled home users.

Gathering user information

There is little uncertainty about physical barriers in homes. These are already well reported and documented. Unfortunately, the same cannot be said for barriers against people with sensory or cognitive impairments; about which very little evidence currently exists. As a result, there is no clear consensus about how barriers against the full range of impairments can be integrated and even less about how they could be systematically addressed.

To overcome these limitations, the team agreed to use the findings of an earlier user consultation in Sheffield (Mitchell et al, 2004). During this study, home users with impaired sight, hearing, movement, understanding and mental health had explained the barriers they encounter in their present and previous homes and neighbourhoods. Of particular interest were their descriptions of what a 'perfect' barrier-free home would be like from their own points of view.

While this study exposed many barriers against people with the full range of impairments, it did not suggest how they could be integrated or how a common strategy for addressing all of them simultaneously could be developed. This is an

important deficiency because there is little chance of removing barriers with a 'piecemeal', one-by-one approach.

The following year, Ergonova (Mitchell et al, 2005) developed the 'Five Keys to Inclusion' to offer a means of:

- Integrating barriers against the full range of impairments
- Addressing Barriers against all impairments simultaneously
- Evaluating the quality of any resulting solutions

The team wished to check that these starting points were valid and applied to people in the West Midlands as well as in Sheffield. With the help of the "Access in Dudley Group", consultations were set up with local disabled home users with all types of impairment. As well as endorsing the 'Five Keys to Inclusion' and the barriers that had already emerged, they also contributed their own experiences and priorities.

A key part of the methodology was to present design proposals for a refurbishment project to disabled people and to explore where residual barriers might exist.

In effect, the 'useability' part of the project had now produced:

- A list of the barriers that disabled people face in their homes
- A means of integrating them and of working out how they were caused
- An 'innovation agenda' for overcoming them

The effects of inclusion

If homes were fully inclusive, their users, regardless of any impairment, would be able to make full use of them, along with their neighbourhoods and all their domestic products and systems as well.

The 'Five Keys to Inclusion' had been used to expose the points at which disabling barriers had prevented them from doing as they wished.These points were useful in pinpointing the places at which designs were not fully inclusive and which therefore required further consideration. Surprisingly, it appears that just two major barriers can disrupt every stage of the 'Five Keys' by preventing some users from finding, reaching, understanding, relating to and interacting with the things they wished to use. For example, non-inclusive access prevented many people with physical impairments from reaching their homes or the things they wished to use in and around them. It could also prevent them from interacting with and using many systems inside their homes, such as their central heating or their alarms.

Non-inclusive communication had similar consequences for people with impaired sight, hearing, understanding or mental health. It prevented them from finding, understanding, relating to and interacting with anything that requires either one or two way communication. Examples of these problems include using way finding information, visual and audible displays, many forms of instructions, information and guidance and communication systems such as telephones, sound media and alarm or intercom systems.

Integrating the team's areas of expertise

The user problems which were exposed in this way provided the team with plenty of unexpected perspectives and starting-points for combining their particular areas of expertise. Seeing problems from unfamiliar points of view enabled them to re-examine familiar approaches and to ask how inclusive or appropriate they actually are. Several examples of this joint working are given here and others will be discussed during the workshop. Some examples of this integrated working include:

(1) **Inclusive escape:** The team had first addressed the problem of access in terms of overcoming differences in floor levels and in using various methods for overcoming restricted spaces in and outside homes. They had not expected to find that one of the important areas of access is escaping from the upper floor of a home during a fire or other emergency.

Escape is not inclusive at present. Only those who have their full physical, sensory and cognitive capacities can be expected to climb through a window, take hold of the cill and then drop to the ground safely. These are formidable barriers for anyone to overcome. The chances of disabled people surviving must therefore be appallingly low.

After consideration, several potentially fully inclusive methods of escape, such as friction tubes, slow-release 'Peter Pan Waistcoats' and others were rejected. Instead, the team decided to specify an existing sprinkler system that is widely used in high rise buildings but which is virtually unknown in smaller homes. By delaying the spread of a fire, this system allows the fire brigade to help everyone in a home to escape, not just those who have the agility to escape unaided.

(2) **Inclusive communication and control**: All users must communicate with, interact with and sometimes control the people, products and systems within their homes. An inclusive, 'homewide' communication system is needed to make sure that everyone can respond effectively to domestic events as they occur. It might be that someone comes to the door, the phone rings, the central heating must be reset or an emergency such as a fire, a burst water pipe or an electricity failure happens.

An inclusive information system is therefore required that can:

- Reach them wherever they happen to be
- Provide them with information in formats that they can understand and use
- Receive their instructions in whatever format they prefer

The team has chosen the best available ways of achieving users' requirements for inclusive, 'homewide' communication and control and is also exploring the prospects for developing the system further. This application also dovetails neatly with environmental considerations. Refurbishment can require the complete removal of power and other systems when the carcass is stripped down to its basic elements. This enables inclusive communication systems to be built into the panels and insulating materials that are used to upgrade the home's thermal performance.

References

[1] Mitchell, J, Chadwick, A and Hodges, B, 2004, '*Why can't we have a Home with Nothing Wrong in It?* – disabled peoples' aspirations and barriers in inclusive living at home', Report to Sheffield's Housing Department, Ergonova, Sheffield.

[2] Mitchell, J and Chesters, R, 2004, '*Designing out the cell of Exclusion*' Contemporary Ergonomics 2004, Ed Paul T McCabe, General Ergonomics, pp 445–449, ISBN 0 8493 2342 8.

[3] Mitchell, J and Chesters, R, 2005, 'The Sheffield Formula for Inclusion', Include, the Helen Hamlyn Foundation, Royal College of Art, London, ISBN 1-905000-10-3.

[4] Mitchell, J, Beever, A, Mortby, P, Sanders, M, Smethurst, J and Townsend,C, 2006, '*Barriers against people with different impairments in their homes and neighbourhoods*', Contemporary Ergonomics, Taylor and Francis and the Ergonomics Society, London and Loughborough, ISBN 10 0 415 39818 5, pp 392–394.

[5] Baines, R, Chesters, R and Mitchell, J, 2007, '*A design standard for sustainable, inclusive housing refurbishment*', Urban Living, West Midlands.

21 'BUILDING BLOCKS' FOR 21st CENTURY INCLUSION

John Mitchell[1], Richard Baines[2] & Robert Chesters[3]

[1]*Principal, Ergonova*
[2]*Director of Sustainable Development, Black Country Housing*
[3]*Intelligent Health Network Manager, Medilink West Midlands*

Barriers and exclusion become matters of choice rather than inevitability when the technology and resources for eliminating them are available but are not used to achieve inclusion (1). The regeneration of manufacturing industry in the West Midlands has stimulated the search for new, commercial products and systems (2). Inclusive, sustainable home refurbishment has a high priority within this context because commercial markets are inevitably restricted when disabled people are excluded from them. Housing new build and refurbishment also offer fruitful areas for user-based innovation and development (3,4,5,6,7).

After developing a guidance specification for sustainable, inclusive home refurbishment (8), the i-House team has been asked to demonstrate the effectiveness of its specification by applying it to a refurbishment project. This work and the programmes that have led up to it have been based on sustained partnerships between home users with different impairments; the disciplines within the i-House project team' and, local manufacturing industries. These partnerships have enabled the team to bring together and further develop a series of concepts for integrating, understanding and addressing the barriers that exclude people with impairments from mainstream living at home. Disabling barriers are often regarded as complex, impenetrable, unyielding and unprofitable, attitudes that can derail any attempts to eradicate them. The authors hope to counteract these myths by integrating the simple realities of exclusion as a means of alerting their partners to the gains they could make by pursuing inclusion. The 21 'building blocks for inclusion' have emerged over the last three decades and are clustered as follows, inclusion, barriers to inclusion, testing inclusion at home, targets for inclusion, integrating barriers to inclusion.

21 'Building Blocks' for 21st century inclusion

Parts of the natural world can be unavoidably hostile for people with impairments because its barriers are largely outside human control. The 'man-made-world' is quite different. Its elements are designed and constructed by people using the

materials and technologies available to them. Whilst earlier technologies could not overcome every barrier to inclusion, their modern equivalents can potentially overcome most, if not all, disabling barriers. Barriers and exclusion become matters of choice when the technology and resources for eliminating them are available but are not used to achieve inclusion (1).

Inclusion in new build and home refurbishment

These issues are particularly relevant in the West Midlands where the regeneration of manufacturing is stimulating the search for new, commercial products and systems (2). Within this context, inclusive, sustainable home refurbishment has a high priority because commercial markets are inevitably restricted when disabled people are excluded from them and because housing new build and refurbishment offer a fruitful area for user-based innovation and development (3,4,5,6,7). Initially, Urban Living had asked the *i*-House project team to develop a guidance specification for sustainable, inclusive home refurbishment (8). Subsequently, Advantage West Midlands asked the team to demonstrate the effectiveness of the specification by refurbishing a dilapidated 19 century, 'end-of-terrace' house in West Bromwich.

Using available technology to achieve inclusion

Modern design and technology offer a potentially huge field of new and rapidly developing options. It would be almost impossible to choose, integrate and develop the most promising of these options without, clear priorities from disabled users, well established collaboration with industry and an understanding of the 'building blocks for inclusion'. The first two requirements have been met through close partnerships between members of the I-health team, home users with different impairments and West Midlands manufacturing industries. These partnerships have also helped the team to meet the third requirement by enabling them to bring together and further develop a series of concepts for integrating, understanding and addressing barriers in the home and elsewhere. Disabling barriers are often regarded as complex, impenetrable, unyielding and unprofitable, attitudes that can derail attempts to eradicate them. The authors hope to counteract these myths by integrating the simple realities of exclusion as a means of alerting their partners to the gains they could make through inclusion, the 'Inclusion Dividend' (2). Six of these concepts are set out below. Each has one or more components that combine to form the 21 'Building Blocks for Inclusion'.

The 'Building Blocks for inclusion'

The 'Building Blocks' are an interlocking set of concepts and principles that are intended to explain what barriers are how they are caused and how they can be eliminated. The six sets of concepts have emerged over the last three decades and can be set out as follows:

1. Assembling the 'Building Blocks':- Inclusion
   It could be said that one of the marks of civilisation is the liberation of its members' and their potential by removing any barriers that restrict their productivity

and participation. Many disabled people see the 'man-made-world' as being as hostile and impenetrable as the most threatening parts of the natural world. Nevertheless, it is potentially changeable through design by humans and, if their knowledge and skills are used to full effect, barriers could be 'designed-out' and full inclusion could become a reality (1).

2. *Assembling the 'Building Blocks':- Barriers to inclusion*
   The team have consulted disabled users about the barriers that prevent them from making full use of their homes, their domestic products and systems and their neighbourhoods (4,5). To their surprise only two fundamental barriers emerged from this consultation, non-inclusive physical access and non-inclusive communication. It would appear that if both physical access and communication were barrier-free, people with impairments would be able to access, communicate with, interact with and control all the essential elements of their homes (10).
3. *Assembling the 'Building Blocks':- Testing inclusion at home*
   Effective inclusion is only likely to be achieved if each new development is rigorously evaluated to find out if it is fully inclusive and, if not, where fresh efforts are required. The need to answer these questions in the West Bromwich demonstrator refurbishment led the team to develop the 'Triple Walkthrough' (12). This is intended to enable disabled users to check that they can, a) find, reach and get around the home being tested, b) communicate with and control home functions, regardless of users' location and c) use all specific functions, such as kitchens and bathrooms
4. *Assembling the 'Building Blocks':- Targets for inclusion*
   Full inclusion requires that people, places, products and processes (the 4Ps) are barrier-free (12). Any barriers that emerge can then be addressed through either retraining or redesign.
5. *Assembling the 'Building Blocks':- Integrating barriers against inclusion*
   The huge array of barriers against people with different impairments has helped to create the impression that barriers are essentially complex and intractable. The 'Five Keys to Inclusion'(13) offer a straightforward series of stages that appear to apply to anyone who is trying to use or interact with people, places, products or processes. These sequential functional requirements have the effect of drawing together and integrating barriers that are normally thought of as being separate and linked only to particular impairments. It is hoped that a coherent, everyday sequence of use will help to make barriers easier to understand and address. The concept underlying the 'Five Keys' is that, whenever anyone wishes to interact with people, places, products or processes, they must, find them, reach them, understand how they work, relate to them without being intimidated, and finally interact with them and use them.
6. *Assembling the 'Building Blocks':- The dynamics of inclusion*
   Since the eradication of barriers is a human activity, it is bound to reflect and depend on the knowledge, aspirations, motivations and skills of those concerned. The 'Cell of Exclusion' assembles these dynamics into six negative values represented by the cell's floor, its ceiling and its four walls. These comprise of am acceptance of exclusion by disabled people, acceptance of exclusion by other stakeholders, non-inclusive design/development, lack of evidence about

barriers themselves, lack of evidence about the penalties that barriers impose and lack of involvement of disabled people in design development. Fortunately, these values are only negative for historical reasons and each of them could be transformed into positive values. If this was done, the dynamics of the 'Cell' could be reversed. Instead of combining to keep disabled people out of mainstream living, they could then provide the liberating power that could deliver full inclusion (1).

References

1. Mitchell, J and Chesters, R, 2004, '*Designing out the cell of Exclusion*', Contemporary Ergonomics 2004, Ed Paul T McCabe, General Ergonomics, pp. 445–449, ISBN 0 8493 23428
2. Middleton, J, Chesters, R and Mitchell, J, 2006, '*What's in it for me? Altruism or self-interest as the "driver" for inclusion*', Contemporary Ergonomics, Taylor and Francis and the Ergonomics society, London and Loughborough, ISBN 10 0 415 39818 5, pp. 382–386
3. Chesters, R and Mitchell, J, 2004, *'Consumers, Evidence and Innovation'*, Contemporary Ergonomics 2004, Ed. Paul T McCabe, General Ergonomics pp. 450–454 ISBN 0 8493 2342 8
4. Mitchell, J, Beever, A, Mortby, P, Sanders, M, Smethurst, J and Townsend, C, 2006, *'Barriers against people with different impairments in their homes and neighbourhoods'*, Contemporary Ergonomics, Taylor and Francis and the Ergonomics Society, London and Loughborough, ISBN 10 0 415 39818 5, pp. 392–394
5. Mitchell, J, Barraclough, K, Davis, G, Gamage, Sean, Hodgkinson, C and Smethurst, J, 2006, *'Innovation and Collaboration between users and providers'*, Contemporary ergonomics, Taylor and Francis and the Ergonomics Society, London and Loughborough, ISBN 10 0 415 39818 5, pp. 395–398
6. Mitchell, J and Chesters, R, 2005, *'Models for Inclusion: Evidence for choice and innovation'*, Contemporary Ergonomics, The Ergonomics Society, Loughborough, ISBN 0-415-37448-0, pp 297–301
7. Chesters, R, McClenahan, J, Middleton, J and Mitchell, J, 2005, *'Empowering design by empowering disabled consumers'*, Include, Helen Hamlyn Foundation, ISBN 1-905000-10-3
8. Baines, R, Chesters, R and Mitchell, J, 2007, *'A design standard for sustainable, inclusive housing refurbishment'*, urban Living, West Midlands
9. Mitchell, J, McClenahan, J and Bennington, J, 2000, *'Autonomy for Disabled People: the need for systematic choice and innovation'*, Contemporary Ergonomics 2000, pp. 375–380, Eds McCabe, PT, Hanson, E and Robertson, SA, Taylor and Francis and the Ergonomics Society, London
10. Chesters, R, Baines, R and Mitchell, J,2008, *'Non-inclusive communication -, the other disabling barrier?'*, Contemporary Ergonomics 2008, Taylor and Francis and the Ergonomics Society, London and Loughborough

11. Baines, R, Chesters, R and Mitchell, J, 2008, *'Inclusion – blending user and professional insights'*, Contemporary Ergonomics 2008, Taylor and Francis and the Ergonomics Society, London and Loughborough
12. Mitchell, J, 1980, *'Problem-solving in Ergonomics'*, Keynote presentation at Ergonomics in Action', IEA Conference, Oslo
13. Mitchell, J and Chesters, R, 2005, *'The Sheffield Formula for Inclusion'*, Include, the Helen Hamlyn Foundation, Royal College of Art, London, ISBN 1-905000-10-3

MEDICAL ERGONOMICS

TASK ENVELOPES: THE MISSING LINK

Sue Hignett & Patrick Waterson

Healthcare Ergonomics and Patient Safety Unit, Dept of Human Sciences, Loughborough University, Loughborough, LE11 3TU

This paper offers a theoretical exploration of a new model for ergonomics practice and research in healthcare systems. A new concept for team (small group) working is being proposed, the task envelope (TE), as a conceptual framework for the perceived missing link between the workspace ergonomic envelope (micro) and organisational (macro) systems. It is suggested that healthcare task envelopes may offer a new way to integrate complex contextual information from the multi-disciplinary clinical team in the design process of technology and systems that could be used at the definition and test stages, and then again as the context to measure the benefits.

Introduction

Previous research has suggested that the meso level (team or group working) has received little attention with respect to theoretical ergonomics models (Hignett, 2001). As an increasing number of papers have been published on team working in healthcare, there is a need for an integrated theoretical model that can be applied by both practitioners and researchers. To address this perceived gap, we propose a new conceptual framework as the missing link between the workspace ergonomic envelope (micro) and organisational (macro) systems: the task envelope (TE). The workspace envelope was previously defined as dynamic reach activities in 3-dimensions (kinetosphere; Pheasant, 1996) usually relating to a single worker interface. At the organizational level, macro-ergonomics was developed in the USA to move ergonomics practice from a micro to a macro level, whereas in Europe this was achieved many years earlier with the integration of sociotechnical systems theory (Hignett, 2001). Hendrick (2007) called the meso level the 'middle-out', '*an analysis of subsystems and work processes both up and down the organizational hierarchy from intermediate levels*.' Some authors have included team work as part of the microsystem (Johnson and Barach, 2007).

In order to develop the healthcare TE model, theories from ergonomics and patient safety and socio-technical systems have been considered in this preliminary exploration.

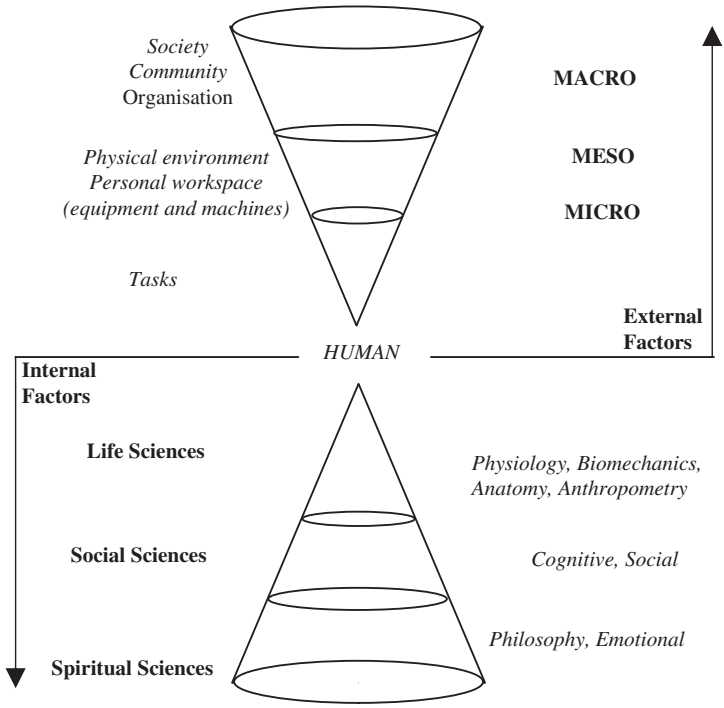


Figure 1. Dynamic interactions model for ergonomics (Hignett, 2001)

Conceptual models from ergonomics, patient safety and S-T systems

Ergonomics

A dynamic model of ergonomics was proposed in 2001 (figure 1) to show the human as a link point (interactions) between internal and external factors with a clear identification of the meso (small group) level. The internal and external factors are portrayed as a dynamic 3-dimensional model with the conical rings moving through the whole vertical range (internal and external). For example a macro level ring (incorporating issues of society, community and organisation) moves through the model to interact at the spiritual sciences level.

Patient safety

In the patient safety literature Bogner (2004) proposed an 'artichoke' model, based on a concentric rings framework. She clustered factors contributing to errors into five categories to represent: the context of care: ambient conditions; physical environment; social environment; organizational factors; and legal – regulatory – reimbursement – national cultural factors. The factors in each of the categories care

are interrelated; they affect and are affected by the other factors in the specific category, as a dynamic model. The three basic care-providing systems are embedded in the context of care as the provider, means of providing care, and the care recipient.

This framework has been used as the basis for a representation of resuscitation with a fluid (constantly changing) overlay for the TE (figure 2). Resuscitation is a complex team operation with the key interaction (means of providing care) between the care giver/provider and care recipient at the centre. In resuscitation this will usually be the first person attending the scene. They will immediately call for assistance (following organizational procedure) at both the local level (within earshot) and from a wider source (e.g. switchboard to the hospital resuscitation team). The immediate workspace (kinetosphere) will be rearranged (bed pulled away from the wall) and new equipment will be collected from the immediate physical environment and possibly from the wider physical environment. As the task continues additional professionals may attend and increase/diversify the team; this will vary with the clinical requirements, time of day/rosters and staff availability. So the team will not have a consistent membership. The exact components of the task will also vary with each repetition due to the different ambient, physical and social conditions for each resuscitation. Within the organization there will be policies that may be interpreted within the context of different professional guidelines (e.g. Royal College of Nursing, Royal College of Anaesthetists), any known patient requests (e.g. living will), and guidance from government organizations (e.g. HSE, NPSA).

Sociotechnical systems

Sociotechnical systems theories generally invoke a set of boundaries between subsystems within the overall system in order to explain where and when miscommunication or larger malfunctions between these levels may occur. Typically, these boundaries cut across organisational (macro), social (meso) and individual (micro) levels of analysis. Within the domain of patient safety there are a number of gaps in the research with regard to these levels of analysis. Hoff et al (2004) for example, have shown that patient safety research has so far focused on teams or leadership without considering the role played by organisational dynamics (i.e., macro and meso levels). Schutz et al (2007) also examined research on patient safety improvement strategies and classified these according to structural components and sociotechnical characteristics (figure 3). Although offering a more comprehensive model that sought to integrate systems models with patient safety improvement initiatives, Schutz et al's model does not include the physical interface between the care giver and recipient and so 'peters out' towards the centre of the artichoke model (the meso and micro levels). In short, the sociotechnical models that exist within patient safety lack integration. The coverage offered by these models is patchy and there is a need to 'join up' macro, meso and micro levels. This is particularly found in healthcare settings where accidents and errors occur at the interface between system boundaries e.g., handover between healthcare teams. We believe that the TE concept is taking a step forward in providing a framework for these links.

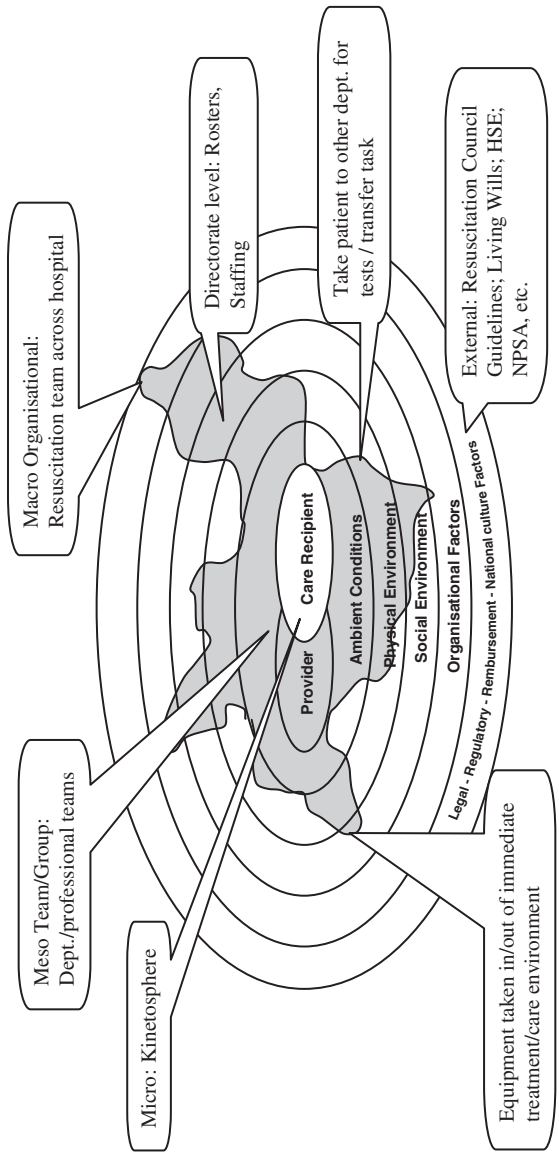


Figure 2. Fluid healthcare task envelope for resuscitation.

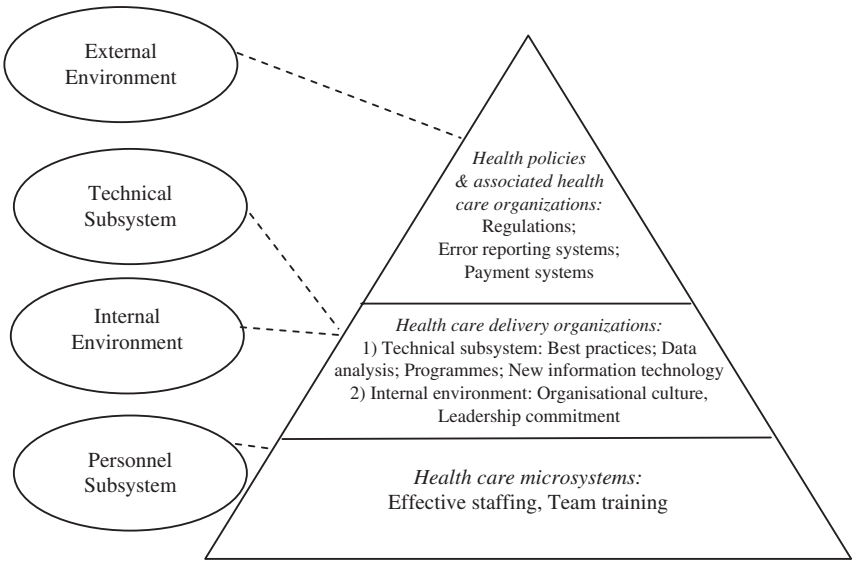


Figure 3. Classification of patient safety improvements (adapted, Schutz et al, 2007).

Conclusion

The TE offers a more complex theoretical representation to incorporate multiple activities, workers and interfaces. The envelope defines the boundary of resources; these can be physical (e.g. spatial constraints), temporal and organisational. The TE concept can support the design process through the provision of information about the incompressible spatial requirements for functional activities rather than prescribing specific dimensions and layouts. It is anticipated that TEs will facilitate the creative process in design, encourage innovation and may offer a speedy response to changes in working practices, equipment and technology. In healthcare, TEs may offer a new way to integrate complex contextual information from the multi-disciplinary clinical team in the design process that will be used at the definition and test stages of development and then as the context to measure the benefits of the new technology.

References

Bogner, M.S. 2004, Understanding Human Error. Chapter 1 in Bogner MS (Ed.) *Misadventures in Healthcare. Inside Stories*. (Mahwah, New Jersey: Lawrence Erlbaum Associates).

Hendrick, H.W. 2007. A historical perspective and overview of macroergonomics. Chapter 36 in Carayon, P. (Ed.) *Handbook of Human Factors and Ergonomics in Healthcare and Patient Safety*. (New Jersey: Lawrence Erlbaum Associates Inc.), 41–60.

Hignett, S. 2001. *Using Qualitative Methodology in Ergonomics: theoretical background and practical examples*. Ph.D. thesis, University of Nottingham.

Hoff, T., Jameson, L. Hannan, E. and Flink, E. 2004, A review of the literature examining linkages between organizational factors, medical errors and patient safety. *Medical Care Research and Review,* **61**, 1, 3–37.

Johnson, J.K., Barach, P. 2007. Clinical Microsystems in healthcare. Chapter 6 in Carayon, P. (Ed.) *Handbook of Human Factors and Ergonomics in Healthcare and Patient Safety*. (New Jersey: Lawrence Erlbaum Associates Inc.) 95–106.

Pheasant, S. 1996. *Bodyspace. Anthropometry, Ergonomics and the Design of Work*. (London: Taylor & Francis).

Schutz, A.L., Counte, M.A., Meurer, S. 2007. Assessment of patient safety research from an organisational ergonomics and structural perspective. *Ergonomics,* **50**, 9, 1451–1484.

REDUCING ERRORS IN SURGICAL CARE: THEORY AND EVIDENCE

Ken Catchpole, Ami Mishra, Simon Kreckler & Peter McCulloch

Quality, Reliability, Safety and Teamwork Unit, Nuffield Department of Surgery, The John Radcliffe, Headington, Oxford, OX3 9DU

The last 5 years has seen considerable advancements in our understanding of the frequency, nature and causes of adverse events in healthcare. However, progress in reducing errors has been slow, in part because significant cultural change is difficult, and in part because empirical evidence about beneficial safety interventions is limited. This evidence is especially important in surgical care, where changes in practice are primarily led through scientific study. In this paper, we summarise the results of 5 years of work in evidence-based improvements in the quality of surgical care, firstly by investigating incident reporting systems, secondly by measuring and understanding the mechanisms of error in the operating theatre; thirdly by evaluating two interventions based on industrial best-practice (Team Resource Management from Aviation and Formula 1 pit-stops); and finally by reporting early the results of an intervention to improve quality of care on surgical wards.

Introduction

Surgery is at the apex of the healthcare system, where diagnostic processes, technology, logisitics, infrastructure, culture, training, expertise, clinical judgement and teamwork combine in limited time and space to provide the best outcome for the patient with the least possible cost. Unfortunately, approximately 1 in 10 patients are unintentionally injured as a result of their healthcare treatment with surgery often regarded as the most prone to error (Vincent et al., 2001). Recognition of this problem in the US in the early 1990's, and in the UK in the latter half of that decade lead to the recognition of the systems approach to error in healthcare, and to increased efforts to understand and reduce the causes of error. In particular this lead to the creation of the National Patient Safety Agency (NPSA), which has responsibility for encouraging better patient safety practices in the UK National Health Service (NHS).

Despite these advances in acknowledging, understanding, and advocating the need for improved safety in healthcare, few genuine improvements can be demonstrated. Since changes to practice in healthcare are largely driven by scientific evidence, if progress is to be made towards a safer system – and if doctors particularly are to be engaged in safety and quality improvements – convincing and clinically relevant evidence needs to be gathered and published in style, language

and forums that are read and respected by clinicians. This requires a theoretical underpinning that, though perhaps derived from other industries or domains, reflects the complexity and unique demands of the healthcare system. This paper describes our progress toward that end.

Incident reporting systems

Most nations have adopted a similar pattern when developing their own patient safety programmes. Following an initial retrospective case note review of a sample of patient notes to record the frequency of incidents (Vincent et al., 2001), an incident reporting system is then developed to help understand the type and causes of medical error and in order to identify underlying systemic causes. This approach has been particularly successful in Australia, where relatively small and specialist incident databases have lead to evidence-based changes in safety practice in Anaesthesia. In the UK, the National Reporting and Learning System (NRLS) developed by the NPSA is the first comprehensive national reporting system for patient safety incidents anywhere in the world. Beginning limited operation in January 2004, and achieving full functioning in 2006, this database now holds over 2 million reports of patient safety incidents.

Our own work with the NRLS (Catchpole et al., 2008) examined 12,606 incidents related to anaesthesia from an early sample of approximately 500,000 reports. The wealth of data available, particularly given the ability to describe an incident in free text, was considerable, and illustrated the large range of causal factors that can contribute to patient harm. In particular, the NRLS data supported the need to improve technologies and procedures, information and documentation, pre-operative communication, continuity of care, and the recognition and prevention of potentially harmful situations. However, with any reporting system there are problems with ensuring data integrity, level of detail, appropriate analytical techniques, and avoiding hindsight and response bias. Perhaps most well documented in the safety literature is that only between 5% and 20% of incidents are reported. Our own finding of 58 NRLS reports of awareness under anaesthesia in an estimated 10 million anaesthetics, compared with that of prospective studies which find an incidence of about 1 in 1000, suggests that some incidents may be reported once for 170 occurrences. Other studies conducted by our group have shown that surgical complications which should also be considered as patient safety incidents are very unlikely to be reported (Kreckler et al., 2008), and that the rate of adverse events may be as high as 40% in some clinical areas (Kaul & McCulloch, 2007).

Incident reporting systems have helped change attitudes toward safety and error, and the range of national projects to develop such systems should be commended. However, the data they produce is currently of limited value for improving surgical care, as incidents are vastly under-reported, particularly by doctors, but also influenced by the incident type, severity of harm, and the view that safety incidents may just be more traditionally viewed as “complications”. Furthermore, a failure to use the data to affect clinical benefit can encourage a view that safety is purely the concern of management, and poor implementation of safety solutions

can antagonise the difficult divide between manager and practitioner. In order to be more effective, incident reporting systems will need to be considered as one component of a safety management system that also includes enhanced mortality and morbidity meetings – the historical, but often ineffective, method for reviewing and addressing complications.

Improvements in the reporting, collection and use of incident reporting systems are therefore needed, and are now underway. However, given the suggested frequency of safety-critical events, it may be possible to observe errors at the point of care, and so understand how they can be prevented. In the next sections, we describe our studies that exploit this feature of healthcare.

Mechanisms of failure in successful operations

The operating theatre is particularly suited to prospective observation of errors and the systemic deficiencies which cause them because the expected course of events can be relatively clearly delineated, and all events happen in a precisely indefinable time and space. Previous observational studies of one particular paediatric cardiac procedure (de Leval et al., 2000) had shown that minor problems can have an effect on patient outcome, and was able to suggest behaviours of surgeons that contributed to those effects (Carthey et al., 2003). Our early studies began using similar techniques in a broader range of paediatric cardiac surgery, in an attempt to describe the mechanisms by which these effects occurred. These first examined the number and type of minor problems that occurred in 24 operations, and observed the escalation of these seemingly innocuous, recurrent, and otherwise ignored problems into far more serious "near miss" situations (Catchpole et al., 2006). We conducted similar studies in orthopaedic operations, and adaptation of a behavioural marker system from aviation to healthcare (Catchpole et al., 2007c) allowed the superimposition of the influence of non-technical skills on this process of escalation. Even though no event led to identifiable patient harm, the more serious "near-miss" events were not reported in any incident reporting system, nor were they reviewed by the team themselves after the operation. However, the observation of these safety-critical situations in otherwise successful operations enabled us to propose a model where small problems originated from deficient components of the system largely independently of the type of operation being performed; the more minor problems there were in an operation, the more likely a "near miss" was; this escalation was more likely in the higher risk operations; and good teams were less likely to have minor problems, or "near misses" (Catchpole et al., 2007a). Furthermore, there was a relationship between minor problems and operative duration, which can have both clinical and efficiency implications. Thus, reducing minor problems – either directly at the source, or by improving resilience through teamwork – might have both safety and efficiency benefits.

Our second set of studies directly followed this theme, examining the impact that teamwork training might have on operating performance and outcomes in Laparoscopic Cholecystectomy and Carotid Endarterectomy operations. We were particularly interested in whether the experience of this type of intervention in

aviation would be beneficial in the operating theatre, and despite considerable opinion that it should, were unaware of any previous studies that could provide convincing evidence that this was the case. The methods used previously were developed to evaluate surgical technical error, minor problems, and non-technical skills. For the latter, we adapted the previous technique to form the Oxford NOTECHS system, which could reliably measure non-technical performance in surgical, anaesthetic and nursing sub-teams, and demonstrated validity against a range of measures, including surgical technical error (Mishra et al., 2007). The teamwork training programme consisted of 9 hours of classroom teaching (including role-play, and a specially made DVD) followed by a 1 month period of in-theatre coaching. We also encouraged our participating teams to conduct a briefing before each operation, and found significant improvements across nearly every dimension measured (attitudes, non-technical skills, non-operating problems, technical error, and length of stay), with brefings being particularly beneficical. To provide convincing evidence that this type of intervention is cost-justifiable and can improve performance, these studies will be extended to a broader range of operations with a larger sample size. We also hope to examine sustainability, and to introduce debriefings to surgical procedures, which should allow better learning and consequently reduce the number of recurrent minor problems.

Improving handover from surgery

Once an operation has been completed, the next stage in the care of a surgical patient is their transfer from the operating theatre to or post-anaesthetic care unit (PACU) or the intensive care unit (ICU). This involves the transfer of information and a plan of care from the surgical team to the new care team, and the transfer of all monitoring, ventilation and infusion equipment to the new care location. This process had previously been observed to be problematic, particularly in high risk paediatric cases. Generally, there are few defined protocols or accepted standards, with each consultant anaesthetist conducting the process in a different way, and expecting the nursing teams to compensate for these differences. After examining the process, we realised an analogy could be made between handover and a motor racing pit-stop, so we enlisted the help of the Ferrari Formula 1 team to understand and improve this process. By building a protocol that improved the design and allocation of tasks, encouraged teamwork and information sharing, and used equipment and information transfer checklists, we were able to measure improvements in all aspects of handover performance. Importantly, we also reduced the opportunity for the escalation of errors by de-coupling equipment transfer errors from information transfer omissions (Catchpole et al., 2007b). This work is now being extended to other areas in the hospital, and to shift-to-shift handovers.

Continuous quality improvement on surgical wards

Our most recent studies examine adverse events and quality of care on a surgical ward, with the aim of affecting change through a quality improvement intervention.

Using the methods advocated in the Toyota Production System (sometimes known as "Lean"), we are working closely with the ward staff to understand and resolve their minor problems. At a basic level, this involves organising the workspaces on the wards to make access to equipment, drugs and fluids more easy and visible. Working from the bottom-up engages staff (since it directly improves their day-to-day work), and helps to addresses the genuine complexity of the healthcare system in a way that "top-down" management instructions do not. While we are in the early stages of this process, one anecdote may suffice to demonstrate the value of this approach. Staff reported that they would sometimes spend up to an hour looking for the intravenous fluids they required. We realised that this was because storage on the ward was badly organised, making it difficult to find the required blends, and difficult to see when stock was running short. Improving the system at the level of the ward exposed the deficiencies in the supply system, where it became apparent that re-stocking was unsystematic, and was based on judgement not need. Some fluids delivered were not even on the stock-list. Neither pharmacy (who placed the order) nor the delivering contractor felt in charge of the process, and a lag of 3 days between the weekly stock-take and delivery, coupled with poor re-stocking, meant that both were overloaded with emergency re-orders from the wards as they ran out of stock. By examining the problem in depth, and engaging all the stakeholders, we are consequently able to resolve this minor problem. Addressing a range of issues in the same way should bring about measurable improvements in safety and quality.

Summary and future directions

A top-down approach to safety alone will not work, since incident reporting data is of poor quality, and solutions that do not engage practitioners or address the complexity of the system will be ineffective. Indeed, such approaches not only overburden the system, but also discourage the view that safety considerations are important for improving quality of care. Though not fortunate for patients, visible systemic deficiencies in healthcare provide us with an opportunity to understand error and its effect on patient care. Indeed, taking this bottom-up approach, the causes of patient harm are relatively easily identified. However, effective solutions and sustainable solutions may not be, and though we can learn from other industries, the unique demands of healthcare must also be taken into account. We must address the complexity and uncertainty in the system, and support our findings with scientific rigour. We can only do this by working directly with, and on behalf of, those directly delivering and receiving care.

References

Carthey, J., de Leval, M. R., Wright, D. J., Farewell, V. J., & Reason, J. T. (2003). Behavioural markers of surgical excellence. *Safety Science, 41*, 409–425.

Catchpole, K., Giddings, A. E., Wilkinson, M., Dale, T., Hirst, G., & de Leval, M. R. (2007a). Improving patient safety by identifying latent failures in successful operations. *Surgery, 142*, 102–110.

Catchpole, K. R., Bell, D. M., & Johnson, S. (2008). Safety in anaesthesia: a study of 12,606 reported incidents from the UK National Reporting and Learning System. *Anaesthesia, In Press*.

Catchpole, K. R., de Leval, M. R., McEwan, A., Pigott, N., Elliott, M. J., McQuillan, A. et al. (2007b). Patient handover from surgery to intensive care: using Formula 1 pit-stop and aviation models to improve safety and quality. *Paediatr.Anaesth., 17*, 470–478.

Catchpole, K. R., Giddings, A. E., de Leval, M. R., Peek, G. J., Godden, P. J., Utley, M. et al. (2006). Identification of systems failures in successful paediatric cardiac surgery. *Ergonomics, 49*, 567–588.

Catchpole, K. R., Giddings, A. E., Hirst, G., Dale, T., Peek, G. J., & de Leval, M. R. (2007c). A method for measuring threats and errors in surgery. *Cognition, Technology and Work, Epub ahead of print*.

de Leval, M. R., Carthey, J., Wright, D. J., & Reason, J. T. (2000). Human Factors and Cardiac Surgery: A Multicenter Study. *Journal of Thoracic and Cardiovascular Surgery, 119*, 661–672.

Kaul, A. & McCulloch, P. (2007). Patient harm in general surgery – a prospective study. *Journal of Patient Safety, 3*, 22–26.

Kreckler, S., Catchpole, K. R., Handa, A., & McCulloch, P. (2008). Factors influencing incident reporting in surgical care. *In Preparation*.

Mishra, A., Catchpole, K., Dale, T., & McCulloch, P. (2007). The influence of non-technical performance on technical outcome in laparoscopic cholecystectomy. *Surgical Endoscopy, In Press*.

Vincent, C., Neale, G., & Woloshynowych, M. (2001). Adverse events in British hospitals: Preliminary retrospective record review. *British Medical Journal, 322*, 517–519.

RADIOLOGIST PERFORMANCE IN THE 21st CENTURY: THE IMPACT OF WORKSTATION TYPE AND MSD'S

Siân Taylor-Phillips[1], Prof. Alastair G. Gale[1] & Dr. Matthew Wallis[2]

[1]*Applied Vision Research Centre, Loughborough University, LE11 3UZ*
[2]*Cambridge Breast Unit, Addenbrookes Hospital, Cambridge, CB2 2QQ*

In the UK Breast Screening Programme there is a growing transition from using X-ray film to digital mammography, and consequently a change in mammography workstation ergonomics. This paper investigates the ergonomics of the changing workstation: from the film workstation currently employed; through a hybrid workstation displaying both film and digital mammograms which will be implemented during the transition period; to a fully digital workstation. Important questions of workstation design are considered, alongside whether prior film mammograms should be digitised to avoid the problems highlighted with the hybrid workstation. Workstation ergonomics is considered not only in terms of user comfort, but also the effects on user behaviour and performance, a key consideration in the safety critical arena of breast cancer detection.

Introduction

In the UK women aged between 50 and 70 years are routinely invited for breast screening mammograms every three years as a part of the National Health Service Breast Screening Programme (NHSBSP). Screening involves radiographers taking two mammograms (X-ray images) of each breast, along cranio-caudal and medio-olateral oblique projections. These mammograms are then reviewed for signs of early cancer by specialist breast screening radiologists or advanced practitioners – experienced radiographers with masters or diploma level training in interpreting mammograms. For simplicity from here onwards these people will be referred to as 'readers'. Those women whose mammograms exhibit possible signs of cancer are recalled for further tests. This process is estimated to save 1,400 lives every year in the UK through early cancer detection and treatment, (NHSBSP, 2006).

Detecting cancer through screening

The methodology which readers employ to determine whether mammograms exhibit signs of cancer is well researched and there is considerable evidence that making comparisons between the current mammograms (from the present screening round), and the prior mammograms (from the screening round three years

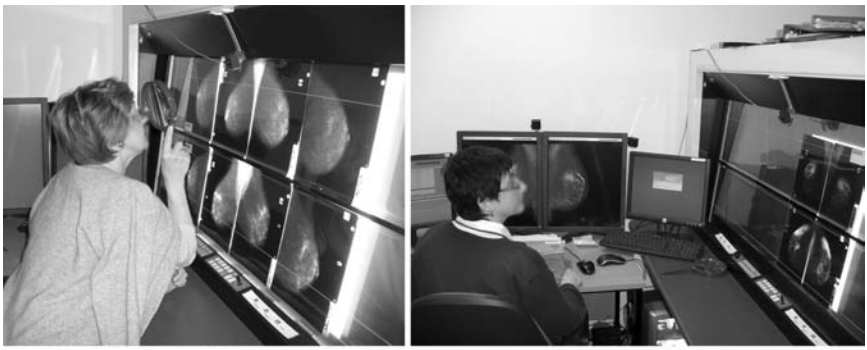
previously) improves detection performance in terms of specificity (Pritchard, 1989; Thurjfell et al., 2000; Sumkin et al., 2003). Therefore, using prior mammograms reduces the false positive rate of readers, so reducing the number of healthy women who are recalled for unnecessary tests. Additionally more recent studies have shown that the use of prior mammograms improves detection performance in terms of overall accuracy as measured by ROC (Receiver Operating Characteristic) curves (Varela et al., 2005; Roelofs et al., 2007), which demonstrates that overall performance is improved in terms of a combination of sensitivity and specificity regardless of the decision criterion for recall chosen. Roelofs (2007) compared performance reading digitized mammograms under three conditions; with priors, without priors, and with priors available upon request. The findings were that performance, measured by area under ROC curve, was improved when priors were always present rather than present upon request, ($p = 0.001$). This suggests that readers underestimate the usefulness of the prior images to their own detection performance.

The transition to digital mammography

Currently, the majority of Breast Screening Centres in the UK take mammograms using X-ray film, which is then developed and displayed on a 'roller' viewer, which illuminates these images from behind. UK trials of digital X-ray technology, both to acquire the images and display them on LCD screens for inspection, have been ongoing and in the future the NHSBSP will move to implementing such an approach in all Breast Screening Centres, (Cush, 2006). Importantly, during the first three years of this transition the current mammograms will be digitally acquired and displayed, but the prior mammograms will have been generated in film format. This presents a choice: either display the prior film mammograms on a roller viewer adjacent to the digital display screen; or digitize the prior mammograms and display them on the screen with the current mammograms, herein referred to as the hybrid workstation and the digital workstation respectively.

There are some ergonomic issues with all three workstations. The film workstation requires repetitive use of a magnifying glass, weighing up to 5 kgs, and the height of the current mammograms encourages postures with the neck in flexion, as shown in figure 1. The height of the seat is adjustable, but if raised the thighs can come into contact with the shelf, and the reader would have to stoop to view the prior mammograms (displayed on the lower row of the viewer) with a magnifying glass.

The hybrid workstation requires repetitive head movements through 90 degrees to compare current to prior mammograms. An alternative layout option is to put the roller viewer parallel to the LCD screen, so instead of repeated neck movement, repeated torso or whole body movement would be required. This however, would require more space per workstation and so would increase costs, and would increase effort to look at the prior mammograms further, without necessarily any ergonomic benefits. Ratib et al. (2000) through a 3D modeling procedure suggest the provision of a light box to display the prior film mammograms above the workstation. This has two distinct disadvantages. Firstly repetitive head movements upwards to a light box are likely to put more stress on the neck than repetitive twisting of the head to the



(a) (b)

Figure 1. (a) A reader using the film workstation. A magnifying glass is used to see tiny micro-calcifications which can indicate cancer. (b) A reader using the hybrid workstation. The head is turned to view the prior mammograms. If closer inspection of the film is required then often the reader will twist and lean their torso.

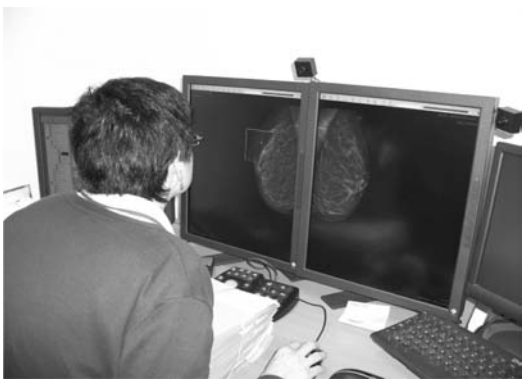


Figure 2. A reader at work at a digital workstation. The magnification tool is being used here but the reader still leans forward to get closer to the screen, putting the torso in extension and the neck in flexion.

side towards a roller viewer. Secondly, a roller viewer allows prior mammograms for a whole set of mammograms to be hung at the same time, so the reader only need scroll through them to the relevant case, a light box would require the reader to hang each set of priors as they review each case, which would use up valuable time and discourage the reader from using the prior mammograms at all.

There are also practical difficulties with the digital workstation, as although the provision of onscreen digital magnification removes the need to hold a magnifying glass, the readers still tend to move their heads close to the screen to see even the magnified views. This results in both poor posture as displayed in figure 2, and proximity of the eyes to the screen viewed has been shown to be a causative

factor for eye strain (Owens and Wolf-Kelly, 1985; Jaschinski-Kruza, 1991). This behaviour may be the result of inadequate magnification tools, or perhaps habit after many years of reading film mammograms.

NHS readers are required to inspect at least 5,000 mammograms per year (NHS-BSP, 2005) with readers working exclusively in breast screening actually reading many more. This is important because it demonstrates the very repetitive nature of the work, and if a movement causes stress then with so many repetitions potential injury can result.

Discussion

The difficulties highlighted in the hybrid workstation where the current mammograms are viewed on LCD screen, and the priors on a roller viewer could result in the reader making less use of the prior mammograms as it takes additional time and effort to view them in a time pressured environment. Considering that Roelofs et al. (2007) have shown that readers underestimate the difference that use of priors makes to their own performance, a self chosen reduction in use of prior mammograms could potentially reduce reader performance in the NHSBSP for the three years of transition to digital technology. If prior mammograms are used less then specificity will possibly decrease, so more women would be recalled for further assessment unnecessarily, with increased financial cost, emotional cost to the women, and increased unnecessary irradiation of the breast. If this is the case then investigation into whether digitizing the prior images and displaying them alongside the current images would encourage the readers to make more use of them is necessary. However, there is no more than anecdotal evidence currently that readers do actually make less use of the priors when they are displayed on a different medium to the current images. Additionally this is based on the assumption that prior images are required just as much when the current image is digital rather than film or digitized, which is not known as all previous research in the area has used either film or digitized images and not digitally acquired images which are of a very different appearance.

Another key question is whether the process of digitizing the images would reduce the quality to an extent to render them undiagnostic. This is unlikely as Roelofs et al. (2006) showed no difference between performance reading film images on a roller viewer to digitized images on an LCD screen.

It is logical to assume that if the prior mammograms were digitized and displayed on the LCD screen alongside the current images they would be utilized by the readers more than the film priors displayed on a roller viewer, but this is not empirically demonstrated. If it is found that digitizing the priors is beneficial to the film readers then the cost of the process of digitization must be considered. In the year 2005/06 1,891,408 women were screened in the NHSBSP (Patnick, 2007), and this figure is increasing year on year, with a large increase projected after plans were recently announced to increase the screening age range from 47 to 73. Therefore, digitizing a set of four prior mammograms for each woman screened would involve digitization of over 7.5 million images which would involve considerable cost. The practical

benefit to offset this would be that staff would not need to be employed to hang the film mammograms on the roller viewer for reading and file them away after viewing.

Cancer detection is a safety critical field where workstation design can affect not only comfort, but also performance. Therefore research attention should be directed at optimizing workstation design for cancer detection in a similar way to past research into optimizing control panels in safety critical industries such as nuclear power.

Acknowledgement

The authors would like to thank the radiologists, advanced practitioners, and staff at University Hospital Coventry who have contributed to this paper.

References

Cush, S., 2006, Evaluation of digital mammography: update on the UK position, in Symposium Mammographicum 2006, *Breast Cancer Research,* **8**, p. 1

Jaschinski-Kruza, W., 1991. Eyestrain in VDU users: viewing distance and the resting position of ocular muscles. *Human factors,* **33**(1), pp. 69–83

NHS Breast Screening Programme, 2005, Quality Assurance Guidelines for Breast Screening Radiologists, Pub. No. 59, (NHS Cancer Screening Programmes, Sheffield)

NHS Breast Screening Programme, 2006, Screening for Breast Cancer in England: Past and Future, Pub. No. 61, (NHS Cancer Screening Programmes, Sheffield)

Owens, D.A., and Wolf-Kelly, K., 1987. Near work, visual fatigue, and variations of oculomotor tonus. *Investigative ophthalmology & visual science,* **28**(4), pp. 743–749

Patnick, J. (ed.), 2007, NHS Breast Screening Programme Annual Review 2007, (NHS Cancer Screening Programmes, Sheffield)

Pritchard, J., 1989. Quality assurance guidelines for mammography (Pritchard report), (NHS Cancer Screening Programmes, Sheffield)

Ratib, O., Valentino, D.J., McCoy, M.J., Balbona, J.A., Amato, C.L. and Boots, K., 2000. Computer-aided design and modeling of workstations and radiology reading rooms for the new millennium. *Radiographics,* **20**(6), pp. 1807–1816

Roelofs, A.J., van Woudenberg, S., Otten, J.D.M., Hendriks, J.H.C.L., Bodicker, A., Evertz, C.J.G., Karssemeijer, N., 2006, Effect of Soft Copy Display Supported by CAD on Mammography Screening Performance, *European Journal of Radiology*, **16**, pp. 45–52

Roelofs, A.A., Karssemeijer, N., Wedekind, N., Beck, C., van Woudenberg, S., Snoeren, P.R., Hendriks, J.H., Rosselli del Turco, M., Bjurstam, N., Junkermann, H., Beijerinck, D., Seradour, B. and Evertsz, C.J., 2007. Importance of comparison of current and prior mammograms in breast cancer screening. *Radiology,* **242**(1), pp. 70–77

Thurjfell, M.G., Vitak, B., Azavedo, E., Svane, G. and Thurjfell, E., 2000. Effect on Sensitivity and Specificity of Mammography Screening with or Without Comparison of Old Mammograms. *Acta Radiologica*, **41**(1), pp. 52–56

Sumkin, J.H., Holbert, B.L., Herrmann, J.S., Hakim, C.A., Ganott, M.A., Poller, W.R., Shah, R., Hardesty, L.A. and Gur, D., 2003. Optimal reference mammography: a comparison of mammograms obtained 1 and 2 years before the present examination., *American Journal of Roentgenology*, **180**(2), pp. 343–346

Varela, C., Karssemeijer, N., Hendriks, J.H.C.L. and Holland, R., 2005/11. Use of prior mammograms in the classification of benign and malignant masses. *European Journal of Radiology*, **56**(2), pp. 248–255

HUMAN RELIABILITY ASSESSMENT IN HEALTHCARE – WHERE NEXT?

Melinda Lyons

Department of Engineering, Cambridge University, CB2 1PZ

Whilst high reliability industries have accepted human reliability assessment (HRA) as part of their safety regime, healthcare has been slow to adopt this as normal practice. The clinical culture, practicalities of carrying out safety analysis in a continuous working environment and the question of how to guarantee robust discrimination between human error and the inevitable progression of human ill-health all challenge the use of HRA in healthcare. These issues are additional to those experienced when applying HRA techniques in other industries - such as the need for validation, hybridization and practical application. This paper discusses the main drivers influencing the past development of HRA and suggests possible future directions.

HRA development in a market environment

In attempting to predict "where next?" for human reliability assessment (HRA), it is useful to examine how it has developed in the past, comment on the environmental influences and try and extrapolate from these a possible direction for the future. Ansoff's (1957) growth matrix (see Figure 1), is a common management tool for evaluating and planning strategies for product growth based on the market drivers – and can be used to evaluate the "products" of HRA in a similar way.

	New Product	Existing Product
Existing market	Market penetration: Innovative HRA techniques, New computer-based tools (Kirwan, 1992a,b, 1996, 1997a,b, 1998);	Product development: Further refinement of existing techniques, automation of tools; validation (Humphreys, 1988, Kirwan, 1992a,b, 1996, 1997a,b)
New Market	Market extension: Tailor-made industry-specific techniques (e.g. TraceR for Air traffic management, Shorrock *et al,* 2003)	Diversification: Initial Emergence of HRA; Development of "resource flexible" techniques, (JHEDI / HRMS; ASEP / THERP) to suit the needs of different users. (Kirwan, 1994)

Figure 1. HRA growth using Ansoff's Product Market Growth Matrix (after Ansoff 1957).

From its humble beginnings, HRA could be considered as first emerging as "diversification" of existing established "products" of the field of reliability assessment applied to a "new market" of human failure. Over time it became recognised as an independent field; though it could still be debated whether its composite parts (e.g. task analysis) are HRA techniques in their own right and whether use of retrospective accident investigation and non-human safety assessment techniques for HRA should allow them to be included within its scope. Everdij (2004) identified 520 safety assessment techniques for use in Air Traffic Management. This collection could be considered incomplete for other industries; notable omissions including FRANCIE (Haney, 2000), HACCP (Hyman, 2003), HFMEA (Linkin *et al*, 2005) and OCHRA (Tang *et al*, 2004), the last three being used in healthcare. Though with such large numbers of techniques available, a broad range of factors are likely to influence practitioner choice as well as product growth in the market.

If one assumes that "markets" will drive HRA along similar themes in the future, there is an expectation that the older techniques will be refined and further validated, whereas new HRA techniques will continue to emerge, subject to "Darwinian selection" in terms of their acceptability. The biggest change is likely to be the areas of application. New markets will be identified – such as new users in existing industries and new industries, particularly those that are focussed around people rather than technology.

To instigate change in healthcare requires at least a convincing evidence base (Grol, 1997). In a sector that demands rigorous clinical trials for introducing new medication, it should be expected that HRA must convince this market that it can provide objective measurable improvements in performance. This implies that validated existing techniques would be better accepted. However, healthcare also has a culture of exclusivity and a reluctance to accept practices simply because they had been successful elsewhere (Hadridge, 2005). To date, HRA within healthcare can be observed to show a similar trend to its progress in industry – from the highly technological aspects (e.g. surgery) to the more human based aspects (Lyons *et al*, 2004). Therefore it is necessary to consider the environmental forces that would suggest the direction for HRA within healthcare.

Environmental influence on application of HRA

The key environmental influences can be considered at different depths of focus – from safety, to human factors / ergonomics to HRA.

The most influential driver for safety could be considered the customer perception of its importance to products, processes and services; whether these "customers" are members of the public or the workforce. However, being that safety is largely opaque to customers, it can only establish itself as a weak driver, acting through marketing and influence on public image. To the internal management, the influence of safety is often perceived as less critical in comparison with its more quantifiable "by-products", such as efficiency and expense reduction.

As safety issues battle for recognition against stronger drivers such as productivity and financial gain, the field of ergonomics and human factors must battle for

recognition within the subset of safety activities. In this, human factors, like the social sciences that underpin many of its techniques, may be viewed with scepticism by safety areas largely underpinned by the "basic" sciences. The safety issues of handling of dangerous chemicals, biohazards and radioactive substances as well as the safe lifting of heavy loads are more transparent. Thus, safety initiatives are easier to justify for physical concepts than the uncertain and hidden human factors. With financial motivations a key influence, it is reasonable to expect organisations to be more willing to invest in physical-based safety as opposed to human-based. Only where the public perception of the consequence of human error is so significant that it could hamper operations would such investments normally be supported at a high level, such as the nuclear and aviation industries.

Following a similar pattern, HRA must battle for recognition within the subset of ergonomics techniques, challenged both by the skills requirement of the techniques of the ergonomist (often including advanced understanding of psychology and mathematical formulae) and the opportunities to apply this within the organisational culture.

Whilst high reliability industries have accepted HRA as part of their safety regimes, these could be some of the reasons why healthcare lags behind. Both hard and "soft" factors are used to justify this reluctance – from the intricacies of the clinical culture to the practicalities of a continuous 24/7 non-stop working environment.

Validation versus Hybridization

Validation and hybridization could illustrate contradictory aims for HRA. Either there is an intention to provide a scientifically robust technique, outputting data of a quality a scale equivalent to the more easily validated physical data; or there is an intention to provide a user-friendly toolkit that takes the best aspects of other HRA techniques and tailors them to the given context. There are challenges from both perspectives. In many cases, validation has been limited to the application of techniques only (Kirwan, 1997a,b). There is, as yet, no assessment and validation of the skills of HRA experts in each technique. Thus, there can be no guarantee that there is any consistency in the way these techniques are applied. There is also not always assessment of the reliability of the data inputted – whether integral to the technique or obtained from subject matter experts and there is no specification on the limitations on the technique's scope. In the case of some of the cognitive techniques, full validation would be impossible.

Based on the large number of techniques that have been developed, HRA users have perceived the need to disregard the available techniques and develop their own. Because many HRA techniques follow a similar template (task analysis, error identification etc), they lend themselves to hybridization of the best parts of previous techniques. Thus, whilst this may result in challenges to scientific validity, it may result in greater user acceptability. For example, the concept of "performance shaping factors" can be used without commitment to an intensive technique such as THERP. Though with hybridization, validity is lost – and in the hands of a

novice, there is a risk that techniques that are "too usable" may be misunderstood and become opaquely unreliable.

Given that the systematic evaluation of "quality" has been recognised only recently in the clinical world (Komashie *et al*, 2007), it is likely that HRA would need to demonstrate its scientific validity in a manner analogous to a new medication or medical device to justify its use. Such transparency is challenging even within high reliability industries, – where there is a benchmark of technological performance for most systems and a process of replacing weak components. Therefore, it is difficult to see HRA would fare in an environment where the artefacts – "Humans" – are not designed to a "standard" and replacement of "weak components" is often hazardous, if possible at all. "System failure" from iatrogenic harm cannot be always distinguished from the evaluation of the "standard" medical ability to deal with the problem or the inevitable progression of human ageing. Even where there is retrospective analysis (for example, autopsies) of incidents that could be used to validate HRA, there would be significant variation in patients, clinical staff skill as well as the quality of the retrospective data. The human-oriented nature of healthcare provides its greatest challenges where reliance is placed on those outside the workforce to guarantee their own safety (e.g. Patient compliance with drug regimes). Some of the processes may result in contradictory reliability, in that refusing to prescribe a patient a particular medication may be deemed correct in terms of procedural safety, but may damage the relationship with patients to the extent that they would turn to individuals and organizations less scrupulous to obtain the medication they perceive they need. All these factors challenge the potential usability and utility of HRA in healthcare and risk the loss of possible investments in HRA training for clinicians and risk managers. Nevertheless, if patient safety is to experience a step change rather than local or short-term fluctuations in performance, HRA and comprehension of the human contributions to medical error would make the critical contribution.

Where next?

From the discussions above, HRA can be seen to progress according to market needs – with all the social, technological, economic, ecological and political drivers. In recent years, the emergence of radically new HRA techniques has slowed and it is likely that development will continue including hybridization and refinement of individual techniques for usability and tailoring them for industry-specific use. Research is likely to pursue these goals together with concerted efforts towards validation as well as supporting the categorization of large numbers of techniques to provide toolboxes.

In terms of application, HRA has extended from its traditions in analysing physical tasks through cognitive operational tasks to more organisational issues, such as safety culture. This is a trend that is likely to continue – to consider organizational errors at strategic level – for example, management and policy. At the same time, this is likely to advance into new areas to reflect these skills from the more physical human-machine interactions of manufacturing to human-human interactions of the

services industries, but also incorporating areas of political interest, for example, environmental management and sustainable development.

In healthcare, these changes are likely to be reflected at a granular level. User-friendliness will be a key selector for techniques used in the domain, within the pressures of the non-stop working environment. Research studies will move from the primarily physical tasks of surgery to those with increasingly cognitive emphasis such as the challenges posed by the mental services. Together with this, there would be a requirement to extend the data of error quantification to include the patient populations (children, older adults and those with disabilities) to aid in supporting the reliability of home-based care.

Acknowledgements

P.J. Wild, A.M. Flintham, J.G. & P.M. Lyons, Prof. Clarkson, IMRC.

References

Ansoff I. 1957, Strategies for diversification. *Harvard Business Review*, **35**, 5, Sep-Oct, 113–124.

Everdij M. 2004. *Review of Techniques to support the EATMP Safety Assessment Methodology*, Volume II Annex. EUROCONTROL, EEC Note 1/4, Brussels.

Grol R. 1997, Personal Paper: Beliefs and Evidence in Changing Clinical Practice. *BMJ*, **315**, 418–421.

Hadridge P. 2005, Leading Safety "Ridiculous" Lessons from oil and other industries. *Journal of the Royal Society of Medicine*. **98**, 539–541.

Haney L.N. 2000, Framework Assessment Notorious Contributing Influences for Error (FRANCIE). *Aviation, Space and Environmental Medicine*, **71**, 9, Section II, September, A145–A150.

Humphreys P. 1988, *Human Reliability Assessors Guide*. (NCSR, AEA Technology, Warrington) October.

Hyman W.A. 2003, The Application of HACCP in Clinical Engineering. *Journal of Clinical Engineering*, July/August/September **28**, 3, 158–162.

Kirwan B. 1992a, Human error identification in human reliability assessment. Part 1: overview of approaches. *Applied Ergonomics*, **23**, 5, 299–318.

Kirwan B. 1992b, Human error identification in human reliability assessment. Part 2: detailed comparison of techniques, *Applied Ergonomics*, **23**, 6, 371–381.

Kirwan B. 1994, *A Guide to Practical Human Reliability Assessment*. (Taylor & Francis, London)

Kirwan B. 1996, The validation of three human Reliability quantification techniques – THERP, HEART and JHEDI: Part 1 – technique descriptions and validation issues. *Applied Ergonomics*, **27**, 6, 359–373.

Kirwan B. 1997a, Validation of Human Reliability Assessment Techniques: Part 1 – Validation Results. *Safety Science*, **27**, 1, 25–41.

Kirwan B. 1997b, Validation of Human Reliability Assessment Techniques: Part 2 – Validation Results. *Safety Science*, **27**, 1, 43–75.

Kirwan B. 1998, Human error identification techniques for risk assessment of high risk systems – Part1: review and evaluation of techniques. *Applied Ergonomics*, **29**, 3, 157–177.

Komashie A., Mousavi A. and Gore J. 2007, A review of historical developments of quality assessment in industry and in healthcare. *Proceedings of the 10th Quality Management and Organisational Development (QMOD) conference*, Sweden, 18–20 June.

Linkin D.R., Sausman C., Santos L., Lyons C., Fox C., Aumiller L., Esterhai J., Pittman B. and Lautenbach E. 2005, Applicability of healthcare failure mode and effects analysis to healthcare epidemiology: Evaluation of the sterilization and use of surgical instruments. *Clinical Infectious Diseases*, **41**, 7, 1014–1019.

Lyons M.N., Adams S., Woloshynowych M. and Vincent C., Human Reliability Analysis in Healthcare: A Review of Techniques. *International Journal of Risk and Safety in Medicine*, **16**, 223–237.

Shorrock S.T., Kirwan B. and Smith E. 2003, Individual and group approaches to human error prediction – a tale of three systems. *OBC Conference on Preventing Human Errors and Violations*, London, 27–28 February.

Tang B., Hanna G.B., Joice P. and Cuschieri A. 2004, Identification and Categorization of Technical Errors by Observational Clinical Human Reliability Assessment (OCHRA) During Laparoscopic Cholecystectomy. *Archives of Surgery*, November **139**, 11, 1215–1220.

ANYTIME, ANYWHERE MAMMOGRAPHIC INTERPRETATION TRAINING

Yan Chen, Alastair Gale & Hazel Scott

Applied Vision Research Centre, Loughborough University, Loughborough, LE11 3TU

Mammographic interpretation is a very difficult task and in the UK, a national self-assessment scheme (PERFORMS) helps to address this issue. Accordingly, further training is suggested to improve performance. Ideally, such training would be whenever and wherever an individual decides to undertake it. To use a portable device for such a purpose would be attractive on many levels, although currently it is not known whether such technology could be used effectively for such high resolution mammographic images. Previous studies have shown possibilities of viewing low spatial resolution medical images (e.g. CT) on PDAs. Thus, it is important that we set out to investigate possible factors affecting feasibility of a PDA as a training technology for examining large high resolution mammographic images.

Introduction

Globally, cancer is the leading cause of death, being responsible for 13% of all deaths worldwide. Breast cancer is of particular interest here and is the most common type of cancer amongst woman. Breast cancer is responsible for 7% of all cancer deaths and almost 1% of all deaths worldwide in 2005 (World Health Organization, 2006). Furthermore, breast cancer is the most common cancer in England. In 2004, there were around 36,900 new cases of breast cancer diagnosed. Additionally, around 10,300 women died from breast cancer in 2004, a rate of 28 deaths per 100,000 women (National Statistic Online, 2007).

The National Health Service Breast Screening Programme (NHSBSP) was initiated in 1988 with the aim of reducing the number of women dying from breast cancer. This scheme provides free breast screening every three years for all women in the UK aged 50 and over. With screening introduction, the death rates began to fall as a result both of the early cancer detection along with improved treatments.

Screening involves the examination of mammograms for the presence of abnormalities that are indicative of cancer. This task is undertaken by specialist, trained mammographic film-readers (radiographers or radiologists). After a mammogram is examined, the film-reader must decide whether or not to recall that woman for further investigation, depending upon their interpretation of mammogram. However, this examination task is especially difficult. This is partly due to the rarity of abnormalities within the screening population (circa 7:1000) and partly due

Table 1. Screening activity – women aged 50+ (NHSBSP, 2007).

	2005/06
Total number of women invited	2,381,122
Acceptance rate (50–70)	74.9%
Total number of women screened	1,891,408
Number of women recalled for assessment	87,469
Number of benign biopsies	1,751
Number of cancers detected	14,851

to the various subtle ways in which an abnormality can present. Consequently, mistakes can occur. For instance, a film-reader might recall a woman for further investigation unnecessarily, because a cancer is suspected where there is none (a false positive). Alternatively, a film-reader might fail to recall a woman for further investigation because a cancer is not suspected, even though a cancer is present (a false negative). It is shown in Table 1 (based on data from NHSBSP, 2007) that the total number of women invited to breast screening in England reached 2.3 million in 2005–06. Some 75% of these (approximately 1.9 million) were screened and over 87,000 women were recalled for further assessment. The figure also shows that around 16,000 True Positive (Correct Recall Percentages) and True Negative (Correct Return to Screen Percentages) decisions were made (14,851 cancers & 1,751 benign biopsies) within the whole recalled group. All these figures indicate the difficulty of mammographic interpretation and the importance of further training further to improve film readers' everyday performance.

Film-readers can improve their skills in a number of ways. Reading a high volume of cases is an important practice as this increases exposure to a wide range of normal mammographic appearance, which aids film readers to develop their skills of identifying abnormality. Also, appropriate experience of attending to specific abnormal features has been shown to develop film readers' ability of recognising early abnormality (Gale, 1997).

Thus, accruing the necessary experience and expertise is a high priority for film-readers. The PERFORMS (PERsonal PerFORmance in Mammographic Screening) self-assessment scheme aimed to address this issue. It is a free and anonymous educational exercise for film readers in the UK undertaken bi-annually where they interpret recent difficult known screening cases. The scheme identifies specific mammographic features that individual film-readers find problematic and suggests further, tailored training, as appropriate. It provides a participant the opportunity to access a range of abnormal appearance within a short time period. PERFORMS provides immediate feedback on their performance as well as subsequent detailed feedback on their respective performance based on a radiological 'gold standard' (Gale, 2003).

Given the inherent high workload associated with the task of mammographic film-reading, it can be difficult for film readers to find time to undertake additional training. This problem is exacerbated by the fact that because screening uses mammographic film then any training undertaken has to take place where mammographic auto-alternators (or 'roller viewers') are, which is at breast screening

Figure 1. Mammogram auto-alternator.

centres (figure 1). Ideally, such training would be *on-demand*; that is 'whenever' and 'wherever' an individual decides to undertake it.

Digital mammography

For the past 100 years, radiographic images have been captured, archived and displayed exclusively on radiographic film. In recent years an increasing use of digital imaging within radiology has been introduced where the original radiological images are digitally captured and then displayed to the radiologist on computer monitors. Furthermore, related modern information technologies have been introduced across the health service which allows images to be distributed nationally and viewed on monitors in any networked location. Thus, digital images can be effortlessly transmitted for a second opinion and these images are easy both to archive and retrieve.

At the beginning of 2005, the NHSBSP set up a steering group for digital mammography to address all kinds of issues on the whole process of introducing digital mammography (NHSBSP Digital Imaging Technologies Steering Group, 2005). This has led to various trials of digital mammography and digital screening being undertaken. Within the next few years it is expected that digital mammographic breast screening will be fully applied across the UK.

The widespread use of digital mammography could be exploited by extending the PERFORMS self assessment scheme to provide dedicated and individuated training. Instead of the training being restricted to viewing physical film on large analogue workstations (based in screening centres themselves), mammographic interpretation training could be undertaken on a digital workstation, laptop, or even on a Personal Digital Assistants (PDAs). This then enhances the possibility of having the training delivered both whenever and wherever it suits the individual.

Of particular interest here is whether or not mammographic training, which has typically relied upon large, high definition mammographic films, could be

pursued, in a digital format, in a useful manner on a relatively small, cheap, and low resolution screen. This type of format would facilitate a high degree of flexibility and mobility in training. An extreme example of such a format is the display of a PDA.

Personal Digital Assistant (PDAs)

Personal digital assistants (PDAs) have become more popular in people's day-to-day life due to their portability, and growing functionality. Increasingly, they are also accepted in professional life.

In recent years, there has been widespread adoption of PDAs within medicine, primarily by medical students and residents who use them as training aids by accessing appropriate text and graphical images. Based on the published literature, it appears that approximately 60% to 70% of medical students and residents use PDAs for educational purposes or patient care, such as storing electronic textbooks; accessing medication reference databases, medical calculators & patient-tracking software; collecting research data; planning ward rounds etc., as well as the more usual general functions such as using the Internet, e-mail, database and spreadsheet software (McLeod *et al*, 2003; Flanders *et al*, 2003; Kho *et al*, 2006).

With reference to the usage of PDAs in radiology, according to a recent survey, 45% of radiologists in North America use a PDA (Boonn *et al*, 2005). However, just what applications can radiologists expect from the PDAs may be limited to some very basic usage, e.g. sign off on reports, access to teaching materials. The potential for radiological images to be viewed on PDAs has not yet been comprehensively researched, mainly due to the fact that radiology inherently requires the highest possible quality image for diagnosis. With regard to PDA usage then a key question would be whether they can provide sufficient image quality to allow primary diagnoses to be carried out directly from the small PDA display. 'Common sense' would possibly imply this not to be an achievable reality. However, the actual resolution of a PDA screen can equate to that of a PC monitor – the main difference then only being one of physical display size. Previous radiological research indicates that PDAs show promise for specific applications, such as: 'neurosurgical emergency consultation'; 'access to senior staff'; 'bedside access to patient data'; 'radiologist in contact with modality', and 'access to specialist' (Schweitzer *et al*, 2002). Recently, a few studies have demonstrated that PDAs can be used successfully to report on CT images, whose small physical size and resolution is adequately handled by the PDA screen resolution and size (Toomey *et al*, 2007).

PDAs in breast screening training

Is there a potential for mammographic training to be undertaken on a PDA? The obvious answer is no: a typical PDA screen resolution would only accommodate

around 1/10th a single mammogram – and in screening each woman must be represented for inspection by four such images. However, information from the PERFORMS scheme offers another answer. When individuals undertake PERFORMS they report cases using a tablet PC which utilizes small mimic displays of the real mammographic images (as shown in figure 1). These individuals also provide additional information via questionnaire from which 43% report that the thumbnail images displayed on the tablet PC (in order to record the location of suspicious features) were rated either 'somewhat' or 'very useful' to them in aiding their decisions about the X-ray mammographic images.

Thus small sized mammographic images may well be useful to the radiologist for showing mammographic features. This information can then be used to facilitate the development of on-demand training in breast screening which could be offered through the use of a portable device. Ongoing research is currently investigating the performance of experienced screening radiologists in making screening judgments using a PDA which displays a single view mammographic image, as well as further elaborating whether the limits of such PDA usage lies in the challenging small PDA display size or within the limits imposed by the display's pixel resolution. The fundamental point then becomes one of how to effectively interact with the display. These HCI issues are currently under investigation, with the aim of delivering a PDA based on-demand training resource for breast screening interpretation within a couple of years.

References

Boonn, W.W., Flanders, A.E., 2005. *Survey of Personal Digital Assistant Use in Radiology*. RadioGraphics 25, 537–541.

Flanders, A.E., Wiggins, III R.H., Gozum, M.E. 2003. *Handheld Computers in Radiology.* RadioGraphics 23, 1035–1047.

Gale, A.G. 1997. *Human response to visual stimuli*. In Hendee, W. & Wells, P. (Eds) *Perception of Visual Information*, 2nd edition. (Springer Verlag), 127–147.

Gale, A.G. 2003. *PERFORMS-a self assessment scheme for radiologists in breast screening*.

Kho, A., Henderson, L.E., Dressler, D.D., Kripalani, S. 2006. *Use of handheld computers in medical education*. Journal of General Internal Medicine; 21(5), 531–537.

McLeod, G.T., Ebbert, J.O., Lymp, J.F. 2003. *Survey Assessment of Personal Digital Assistant Use among Trainees and Attending Physicians*. Journal of the American Medical Informatics Association 10, 605–607.

National Statistics Online 2007. *Breast Cancer: Incidence rises while deaths continue to fall*. Retrieved on 1st Dec, 2007.

NHSBSP Digital Imaging Technologies Steering Group 2005. *NHSBSP Equipment Report 0502: Review of the Literature on Digital Mammography in Screening*.

NHS Cancer Screening Programmes 2007. *NHS Breast Cancer Annul report 2007.*

Schweitzer, T., Engelmann, U., Schröter, A., Boraelv, E., Meinzer, H.P. 2002. *Teleradiology on a Personal Digital Assistant*. Second Conference on Mobile Computing in Medicine, Workshop of the Project Group MocoMed, 11.

Toomey, R.J., Ryan, J.T. McEntee, M.F., Evanoff, M., Ryan, P., Manning, D.J., Chakraborty, D.P., Curran, K., Brennan, P.C. 2007. *The Use of Personal Digital Assistants in the Reporting of Radiological Images*. Medical Image Perception Society Meeting.

World Health Organization 2006. *Fact sheet No. 297: Cancer*. Retrieved on 1st Dec, 2007.

HEAD BANGING: THE ROLE OF ERGONOMICS IN THE INTERPRETATION OF NEURORADIOLOGICAL IMAGES

Lindsey H.K. Cooper & Alastair Gale

Applied Vision Research Centre, Loughborough University, Loughborough, LE11 3TU

In the UK, radiology is now almost completely digital with medical images being inspected on high resolution monitors in dedicated reading environments. Digital imaging procedures today commonly include PET, CT and MRI which produce varying 2D, 3D and 4D images. However, much of the research into factors affecting the interpretation of radiological images is based on 2D images and tends not to encompass such multidimensional imaging. Here, the domain of neuroradiology is addressed with key ergonomic factors pertinent to the efficient reporting of neuro-MRI images being examined. It is argued that data extrapolation from 2D images is not fully adequate to encompass complex visual inspection domains and ongoing research is introduced which aims to address these issues.

Introduction

Ergonomics has an important role to play in contributing to patient safety within the domain of radiology. This is with regard both to factors pertinent to acquiring the relevant medical image of the patient (with respect to the health, safety and comfort of patient and radiographer) and factors that relate to the appropriate and accurate interpretation of the resultant image by the relevant radiological specialist. The latter is the subject of concern here.

Diagnostic errors in the interpretation of medical images have been reported since the 1940s and abnormalities have been either missed, or over-read, with various rates across numerous experimental studies. Most of these investigations have concentrated upon the chest X-ray which is a large static 2D image but with the introduction of breast screening in many countries in the last 20 years, research has focused more upon the interpretation of the breast mammogram.

Radiology in the UK is now almost completely digital with images being inspected on very high resolution computer monitors (rather than film displays on illuminated light boxes) in dedicated reading environments. Other radiological imaging procedures commonly available today include ultrasound, PET, CT and MRI. These techniques can produce varying 2D, 3D and 4D images, which need to be examined thoroughly in an efficient and timely manner by a specialist. Despite this, much of the actual research into factors affecting the interpretation of radiological imaging tends not to encompass such multidimensional imaging.

Medical imaging, accuracy and interpretation

Images of the human musculoskeletal system and anatomy have been accessible since Roëntgen's cathode ray discovery in 1895, when he first visualised the bones in his wife's hand. Since then, cathode rays have been used with great success in the medical field to diagnose bone fractures, skeletal abnormalities and signs of disease. The 1940's sparked a surge of investigation into perceptual psychophysics with 'radiologic physicists characterising imaging systems in terms of contrast rendition, spatial resolution, and the radiation dose required to make an image' (Kundel., 2006). Following on from this period of intense work, inter-rater reliability and diagnostic errors in the interpretation of medical images came to the fore of perception research.

Accurate image interpretation is a challenging but important task. Should an abnormal mass or structure be overlooked, there are dire consequences for the patient (missed diagnosis and/or suspension of treatment) and the medical team involved (potential disciplinary action or litigation). As Manning et al. (2005) pointed out, a number of variables impact upon decision-making: perceptual (the accuracy of image inspection), cognitive (medical knowledge, experience and decision-making) and ergonomic factors (seating, lighting and the environment used for interpretation). Specialists must examine the image and make prompt decisions, usually within seconds, about which aspects, if any, require further attention or action. To complicate matters further, simple 2D images depict intricate 3D anatomical structures within the human body, which radiologists have to disambiguate what might be indicative of a complex problem from what might be a concatenation of overlapping structures.

Previous research has found mixed results with regard to the 2D-3D image problem. Krupinski (2000) reported that more errors tend to represent false negatives rather than false positives with search, recognition, and decision errors occurring when scanning anatomical images. Monitoring how observers appraise anatomy (using eye-tracking technology), whilst giving them feedback has been used to improve performance with varying degrees of success. Whilst Krupinski et al. (2000) reported a 16% increase in pulmonary nodule detection using feedback, other work found mixed results. For instance, circling an area of the scan to focus attention did not produce a clear improvement in observer performance (Krupinski et al., 1993), which also appeared dependent upon radiographer expertise (Donovan., 2007).

Research has allowed insights into contributory factors surrounding visual search, abnormality detection, and cognitive interpretative processes with models being subsequently developed to offer further explanations and guide future studies. This work facilitates the deployment of appropriate training to improve the accuracy of the radiological process. This applied research is being drawn and combined from psychology, cognitive science, statistics, physics, and ergonomics with a clear medical focus to influence overarching frameworks, which subsequently filter down to impact upon clinical practice. In addition, the recent move towards a blame-free culture (DoH., 2000) can only assist in the prevention of clinical adverse events.

Medical imaging and the digital era

As clinical conditions and new challenges (e.g. cancer & chronic health conditions) have emerged in the last few decades, medical research and practice has had to change dramatically to keep pace. In radiology, the move away from film and into digital has been prompted by technological advances specifically in the last two decades. From an interpretation stand point, it is important that radiologists can move towards using the new technology quickly and effortlessly to ensure that patient care and safety are not compromised (Krupinski., 2000). Much work has been conducted on how an image is displayed (e.g. alterations to image presentation and display technology; Krupinski., 2007) and the environment of the observer e.g. average viewing box brightness, and ambient light levels (McCarthy & Brennan., 2003); to prevent diagnostic performance in viewing areas and wards from being affected. Sakas (2002) suggested that the use of enhanced imaging purely for diagnostic purposes is no longer appropriate and technology must go beyond traditional 2D imagery to be useful and efficient. Computer visualisation systems have been developed to 'build up' CT, MRI, and Ultrasound data in 3D and 4D representations with multidimensional images being collated to gather more information than ever before. The need exists for observer performance studies to thoroughly encompass these complex image types.

The radiological Picture Archiving Communication Software (PACS) works in collaboration with the digital imaging equipment and can assist in developing radiology training packages, small or large scale clinician conferences, in preoperative planning, for rare or highly complex procedures. With the use of the internet, digital images can be shared in real-time between specialists in different countries, even when using mobile equipment (Sakas., 2002). It appears then, that there has been a recent shift away from using images in isolation and towards systems that guide further interventions for the patient and/or act as preoperative 'virtual playgrounds' for medics with one shot at getting a procedure right, especially in the field of neuroradiology. It could be hypothesised that providing the experts with practice in complex scenarios may also reduce the incidence of surgical complications, further surgical care and patient burden. However, is clinical technology advancing at a speed that medical professionals might find overwhelming? As further practical changes are implemented within the health service, further questions arise: How has the switch from static to digital imagery affected efficiency and efficacy? What impact will further alterations have on the radiology workforce? How will the shift from 2D to 3D and 4D affect performance and clinical errors? How can researchers measure and predict the consequences?

MRI imaging

MRI imaging is a challenging interpretation task for even the expert radiologist and is the focus of the ongoing research described here, which concentrates upon brain imaging – the head banging of the paper title refers to the rhythmic noise generated during MRI brain scanning. MRI is still undergoing great technical changes and will become more important and widespread in the future (Sakas., 2002). MRI

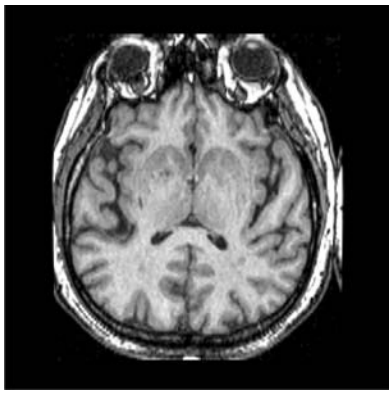


Figure 1. This horizontal image shows a 'healthy' brain.

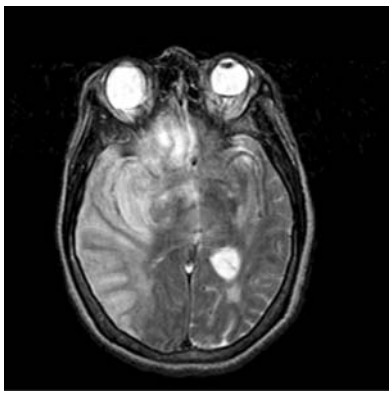


Figure 2. This image was taken prior to a fatal Stroke.

data allows information to be gathered on water-based tissues (e.g. organs, muscle mass, ligaments, and abnormalities including tumour growth and structural deficits) in contrast to radiographic imaging which is based on imaging opacity to X-ray radiation. A single MRI examination can of itself produce hundreds of images which each need to be examined thoroughly, in an efficient and timely manner. The technique permit good visualisation of a large number of conditions. For instance; vascular diseases such as stroke, and degenerative diseases such as Alzheimer's, present themselves quite differently (localised or generalised) in a scan and pose different accuracy and interpretation problems for the observer (figures 1–4). Of particular interest here is what constitutes radiological expertise in the area of neuro-radiology and how does it differ, if at all, from other specialties?

Eye-tracking technology provides an objective, experimental insight into how radiographers and radiologists appraise a complex medical image and how errors might occur. Radiology studies using eye tracking technology have examined duration and location of observer gaze (scan paths and foveal fixations) from novice to expert, to inform skill acquisition and expertise. Qualitative research enhances this

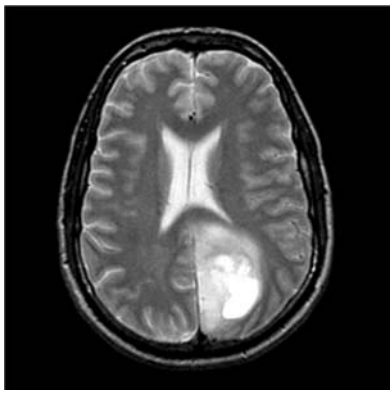


Figure 3. This image shows the presence of a Glioma (tumour).

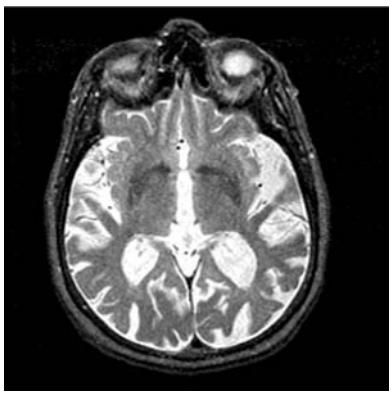


Figure 4. This image highlights the brain shrinkage/atrophy and hypoperfusion that characterises Alzheimer's Disease.

data by providing radiologist accounts, which could be combined and compared with reported clinical errors in domains such as neurology and neurosurgery.

Conclusions

Medical interpretation and decision errors are endemic in any inspection task and radiology is no different to many industrial inspection tasks. More demands have been made of the radiology workforce as technology has switched from 2D static analogue imagery to digital, with 3D and 4D images being produced regularly. Despite this scenario, there appears a lack of research into factors affecting the interpretation of multidimensional imaging and subsequent effects on performance (efficiency and efficacy), clinical errors, and patient safety. Previous literature has focused on chest and breast screening to identify the presence or absence of malignant or benign nodules. Few studies have examined complex interpretative

tasks such as brain MRI. It has been argued that there is an overall 'lack of published research into radiographic interpretative accuracy within the context of clinical practice and decision-making' (Donovan., 2006, pp10) that requires more attention.

MR images combine perfectly the investigation of complex digital imagery with the examination of the accuracy and interpretation of neuroradiological images. Our ongoing research is exploring interpretative accuracy and decision-making using a combination of qualitative (exploration of radiologist accounts) and quantitative (eye-tracking) methodologies. It is intended that the proposed research will inform practice in the field of 3D and 4D neuroradiological imaging with subsequent benefits in terms of clinical outcomes.

Acknowledgement

MRI images reproduced with permission from The Whole Brain Atlas http://www.med.harvard.edu/AANLIB/home.html

References

Department of Health. 2000, *An Organisation with a Memory.* (London).

Donovan, T. and Manning, D.J. 2006, *Successful reporting by non-medical practitioners such as radiographers, will always be task-specific and limited in scope.* Radiography, **12**, 7–12.

Donovan, T., Manning, D.J., Phillips, P.W. and Crawford, T.J. 2007, *The effect of perceptual feedback in fracture detection and lung nodule perception.* (LaMINeR conference proceedings).

Krupinski, E.A., Nodine, C.F. and Kundel, H.L. 1993, *A perceptually based method for enhancing pulmonary nodule recognition.* Investigative Radiology, **28**, 289–294.

Krupinski, E.A. 2000, *The Importance of Perception Research in Medical Imaging.* Radiation Medicine, **18**, 329–334.

Krupinski, E.A., Williams, M.B. Andriole, K. Strauss, K.J. Applegate, K. Wyatt, M. Bjork, S. and Seibert, A. 2007, *Digital Radiological Image Quality: Image Processing and Display*. American College of Radiology, **4**, 389–400.

Kundel, H.L. 2006, *History of Research in Medical Image Perception*. Journal of the American College of Radiology, June, **3**, 402–408.

Manning, D.J., Ethell, S. Donovan, T. and Crawford, T. 2006, *How do Radiologists do it? The influence of experience and training on searching for chest nodules.* Radiography, **12**, 134–142.

Manning, D.J., Gale, A. and Krupinski, E.A. 2005, *Perception Research in Medical Imaging*. The British Journal of Radiology, **78**, 683–685.

McCarthy, E. and Brennan, P.C. 2003, *Viewing conditions for diagnostic images in three major Dublin hospitals: a comparison with WHO and CEC recommendations.* The British Journal of Radiology, **76**, 94–97.

Sakas, G. 2002, *Trends in medical imaging: from 2D to 3D*. Computers & Graphics, **26**, 577–587.

METHODS AND TOOLS

LAPTOPS IN THE LECTURE THEATRE: AN ERGONOMIC FOCUS ON THE CRITICAL ISSUES

Rachel Benedyk & Michalis Hadjisimou

UCL Interaction Centre, University College London, Remax House, London WC1E 7DP, U.K.

A previous survey showed evidence that the casual use of laptops by students in university lecture theatres causes a significant level of discontent for both students and staff. Typical complaints related to sightlines, poor seating, and environmental intrusions. Lecture theatres are characterised by fixed furniture arranged in tiered rows, often in an arc around a low dais for the lecturer, with some form of information projection behind, for the audience. Whereas most teaching in such theatres is didactic in style, there remains a critical need for direct interaction between audience and speaker; does the student laptop compromise this interaction in a poorly integrated theatre environment? This study set out to use two complementary ergonomic methods (Soft Systems and the Hexagon Ergonomic Model) to identify in a structured manner the range of issues and conflicts that influence the integration of the student laptop into the lecture theatre. The implications for future lecture theatre use and design are discussed.

Introduction

A previous paper (Benedyk et al, 2007) discussed the recent trend for university students to bring personal laptops onto campus to support their study and for private use. Despite the fact that universities provide dedicated computer suites for student use both on campus and in halls of residence, and despite the fact that ergonomics would advise against such intensive use of laptops, students find the portability and immediacy of laptops to be advantageous. Universities need to set themselves up to support their use.

The survey reported by Benedyk et al (op cit) identified 19 different settings in which students use their laptops, including libraries, classrooms, bedrooms, cafes, bars and parks. Only 6% of the respondents reported using a laptop in lecture theatres; yet lectures are the mainstay of university study activity. Why is it that students seem wedded to their laptop almost everywhere except in the lecture theatre? This study set out to evaluate the student-laptop-lecture theatre dynamic from an ergonomic perspective.

While university teachers are encouraged increasingly to use laptops themselves for demonstration, data projection and using on-line resources during teaching sessions, they do not often show the same encouragement to their students. The

survey by Benedyk et al uncovered a mainly negative attitude by teachers to student use of laptops. The two main objections were environmental intrusion (flashing screens and noise from key-tapping, start-up tunes and fans) and disruption to engagement between teacher and student: either apparent (eg lack of eye contact) or real (15% of those students who do take a laptop to lectures admitted they use it for recreational activities not connected with the lecture!). Three teachers even admitted they had banned the use of laptops in lectures. 10% of students reported that other students had also voiced objections to laptop use in lectures, mostly because of noise. These negative attitudes may well be part of the reason for the lack of laptop use in lecture theatres, but not the whole story.

29% of students surveyed in fact reported that use of a laptop would be helpful in lectures – what are the various factors that are hindering them?

The lecture theatre

The main use for a university lecture theatre is to project the spoken word from a teacher, together with illustrative material such as slides, to a large audience of students; typically, they are designed more for information flow from the teacher to the student than from student to teacher, or student to student, than might be the case in other teaching spaces such as classrooms. Architecturally, the acoustic arrangement thus facilitates this direction of sound projection, and the seating arrangement (which is typically fixed) facilitates the sightline of students towards teacher and projection screens. Current trends (JISC, 2006) would suggest design of teaching spaces is moving towards much more flexible furniture and facilities, to reflect the increase in active and collaborative learning etc, but lecture theatres are still the near future on many campuses.

Ergonomically, then, we often see fixed multiple seated workstations, arranged in tiers, sometimes arced, around a centre-frontal teaching station. Lighting often excludes daylight, and prioritises dimming facilities to enhance projection slides, and sometimes downlights for student note-taking and spotlights for the speaker. Power sources tend to be concentrated at the dais, for use with projection equipment. The teaching station frequently is organised to accommodate a laptop computer for the speaker, which can be connected in to the theatre data-projector.

The student workstation, however, facilitates only the tasks of listening, watching, and usually note-taking – for which a narrow fixed table or a chair arm-plate is provided at each seat. The seat itself is often of the tip-up type, to allow access and egress of the multiple occupants of each row. These workstations are not inherently designed to accommodate student laptops; as recently as 2006, university minimum standards for the technical equipping of lecture theatres make no mention of facilitating student laptops (Victoria University, 2006).

Ergonomic evaluation of laptop use in lecture theatres

This study used two complementary ergonomic evaluation techniques to identify in a structured manner the factors in lecture theatres likely to come into effect when

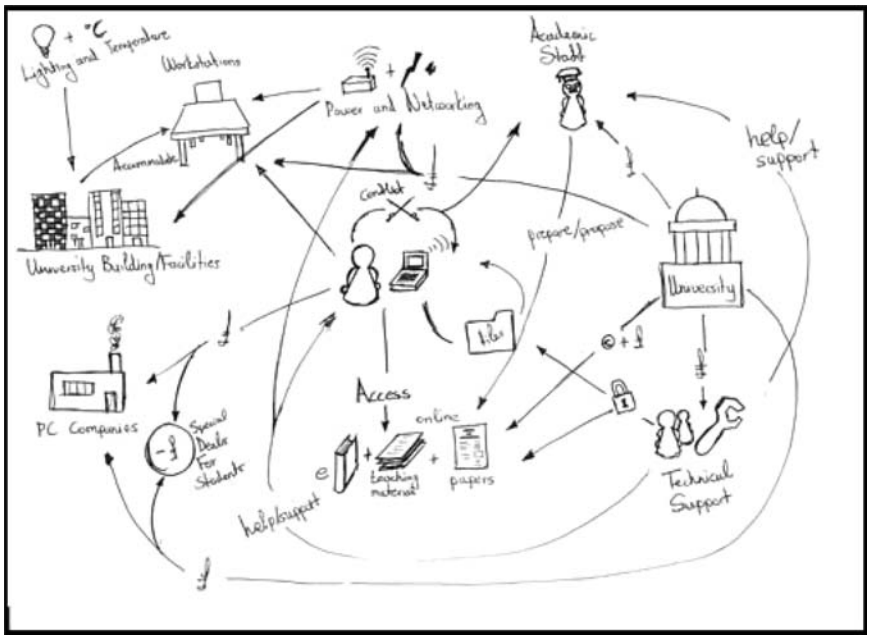


Figure 1. Rich picture of the student-laptop-university system.

students try to use laptops there. First, the known information on university students using laptops was used with the Soft Systems Methodology (SSM) (Checkland and Holwell, 1998) to help identify comprehensively the key actors involved in this complex scenario, and the possible conflicts between them. Then the Hexagon model of educational ergonomics (Benedyk, Woodcock and Harder, 2006 and in press) was applied to focus in on the lecture theatre and to extract the ergonomic issues arising from student/laptop interaction in a lecture theatre setting.

SSM summary

The survey data from Benedyk et al (2007) was used to create a rough rich picture, in the middle of which the student is represented working on a laptop, which itself can cause ergonomic conflict (Heasman, 2000). The student interacts with the workstations in university settings (both when using and not using the laptop) and with the PC companies (buying a laptop). The student also interacts and shares files with other students and with the academic staff, with the possibility of further conflicts.

The university facilities are characterised by their configuration, space, lighting, temperature, ventilation etc and they accommodate the workstations. The university has a budget allocated for buying these workstations and for making power and networking available for the students using them. The university also employs the lecturers and the technical support people. Finally, the university has to pay to get copyright for the e-Books and papers accessed by its students.

	Organisational sector		Contextual sector		Personal sector	
	Management factors	**Infrastructure factors**	**Task factors**	**Tool factors**	**Individual factors**	**Social factors**
External Level	Government funding to universities; DfES policy	Developments in integration of IT facilities into buildings	Best practice in university teaching and use of computing	Trends in use of computers in university teaching and learning	Student income level (for buying a laptop); peer pressure	Cultural and family expectations of university facilities
System Setting Level (University)	Allocation of budget for teaching space; costs and benefits of having lecture rooms suitable for laptop users	Provision of wireless connectivity; facilities spec for lecture theatres; technical support	Syllabus and content of lectures	Age and antiquity of university facilities, flexibility of use of space	Attitudes of students at this university to needing to own and use a laptop in lectures	Attitudes of lecturers to student use of laptops; trust that student will not use laptop for leisure when in lecture theatre
System Workplace Level (Lecture Theatre)	Procurement and Planning choices for design and facilities in this Lecture theatre; room size and Layout	Connectivity and power supply in overall space; accessibility, temperature, acoustics, lighting	Style of delivery of teaching; type of learning	Suitability of workplace configuration; environmental issues eg key-tapping noise; adjustability	Crowding factor; relation of this workstation to Neighbours	Attitudes of neighbouring students and of this lecturer to the use of this student's laptop; rake of tiered seating
System Workstation Level (Laptop use)	Procurement choice of individual workstation eg adjustability and size – suitability for supporting a laptop	Connectivity and power supply in individual workstation; task lighting	Task factors influencing frequency of, and reason for, use of a laptop in the lecture	Laptop design; fit to this work-station; potential for docking station; intrusion eg loudness of fan, volume of sound	Posture and flexibility of posture factors imposed by current configuration	Sightlines and engagement between lecturer and student
Interaction Level						
System User Level (Student)	Personal characteristics, abilities, preferences, expectations and knowledge of particular student population in relation to laptop use in lecture theatres					

Figure 2. Hexagon model analysis table.

The academic staff prepare or propose online teaching material, as well as delivering lectures. The technical support people are responsible for securing the shared files and the online material and are responsible for maintaining the networks, but not the laptops.

After drawing this rich picture of the system it became obvious that the key actors are the students and lecturers, with technical support services playing a secondary role. The conflict representation indicates that student–student and lecturer–student interaction might be problematic when the student uses their own laptop in a lecture theatre not designed to accommodate or support laptops. This picture was then examined in more detail for the lecture theatre setting, using the Hexagon model.

Hexagon model analysis

This analysis (Figure 2 shows part of this) pointed up the likely factors, in all sectors and at all levels, potentially intrusive to successful student/laptop interaction in lecture theatres. From the organisational sector, we see the design and facilities of the teaching space potentially limiting the interaction through poor connectivity, unsuitable workstation design, or lack of power sockets. From the contextual sector, the style of teaching and learning will encourage or discourage the use of the laptop, and the laptop itself will contribute potential ergonomic limitations. The personal sector highlights attitudes and expectations from the lecturer and other students, and we see clearly how apparent disengagement between lecturer and student might occur during laptop use.

If lecture theatres are to accommodate students using laptops successfully, there are a host of ergonomic issues that evidently need to be addressed. For example, what is the right lighting for a room where students either use a laptop or take notes? How can all laptop users have network and power supply? How can a laptop user maintain real engagement with the lecturer, and is any such disruption perhaps related to the rake of the seating? Will a large number of laptops in a lecture room cause too much heat and noise? Is the workstation configuration suitable for both taking notes and using a laptop? Can lecture theatre workstations integrate laptop docking stations? Can the crowding of students into fixed workstations be relieved to allow sufficient flexibility of posture?

Purpose-designed flexible teaching spaces can happily address all these ergonomic points and more, and there is ample evidence that students can interact with laptops in a supportive way given the right facilities and positive attitudes (JISC, 2006). However, many current lecture theatres fall sadly short of this on multiple fronts; it is unsurprising therefore that many students do not use their laptops in lectures, despite wishing to do so.

References

Benedyk, R., Harder, A., Hussain, R., Ioannidis, A. and Perera, M., 2007, The ergonomic impact of student use of laptops in university settings. In P. Bust (ed.) *Contemporary Ergonomics 2007*, (Taylor and Francis, London).

Benedyk, R., Woodcock, A. and Harder A., 2006, Towards a new model of Educational Ergonomics. *Pro. 2nd Conf on Access and Integration in Schools*. (Ergonomics Soc).

Benedyk, R., Woodcock, A. and Harder A., in press. *The Hexagon model of educational ergonomics*.

Checkland, P. and Holwell, S. 1998 *Information, Systems, and Information Systems: Making Sense of the Field*. (Chichester: John Wiley).

Heasman T., Brooks A. and Stewart T., 2000 *Health and safety of portable display screen equipment*, (HSE Contract Research Report 304/200).

JISC, 2006 *Designing spaces for effective learning*, JISC e-Learning programme, (HEFCE).

Victoria University Australia, *Minimum acceptable standards for teaching spaces*, July 2006 htp://tls.vu.edu.au/SLED/ETSU/TSAG/classroom%20standards-%20Rev%20Doc%20Draft%201-July%2006. PDF accessed November 2007.

DEVELOPMENT OF SOFTWARE SIMULATION TO ENABLE HF ASSESSMENT OF SOME ASMGC's FEATURES

Kee Yong Lim[1], Dora Tanuj[2], Ranieri Yung, Ing Koh[1], Samuel Zhen Yu Ow[1] & Dan Ruey Chour Han[1]

[1]*Center for Human Factors & Ergonomics School of Mechanical and Aerospace Engineering Nanyang Technological University, Singapore*
[2]*Dept. of Industrial Engineering & Management Indian Institute of Technology, Kharagpur W. Bengal, India*

This paper describes the development of a to-scale simulation to enable assessment of the efficacy and Human Factors implications of proposed functions of an Advanced Surface Movement Guidance & Control System (ASMGC's). To enable efficacy assessment, out of cockpit views simulating the field of view of the pilots' during aircraft taxiing were created and projected onto an Elumens™ dome to enable an immersive experience to enhance fidelity during testing. The A380 aircraft and the aerodrome of Singapore's Changi International Airport were modeled as the case-study platform. Comparative assessment of the efficacy of newly proposed/specific functions of a proposed ASMGC's may then be made against traditional wayfinding systems. Feedback from pilots may also be elicited using the simulation and safety concerns (if any) may then be identified and suitable modifications proposed.

Introduction

The ICAO (An-Conf/11-IP/4) manual describes the general goal of an ASMGC's as the provision of enhanced support of aircraft operations to meet the required traffic capacity and still maintain a high level of safety especially under low visibility conditions. To this end, an ASMGC's exploits modern technologies, advanced automation and a high level of integration between various systems to provide a flexible solution that can be adapted to accommodate future changes in traffic density, surface congestion and aerodrome layouts. 11 high level goals were defined in the ICAO manual including the provision of the following key functional support:

- Surveillance to provide controllers (eventually pilots and vehicle drivers) with enhanced situational awareness of the movement area (i.e. a surveillance display showing the position and identification of all aircraft and vehicles), taking into consideration visibility conditions, traffic density and aerodrome layout.
- Control to provide runway conflict detection and alert (and eventually the whole movement area). An automated linkage between surface and terminal and

enroute airspace traffic shall be established to enable seamless operation with reduced controller and pilot workload.

- Routing designated manually (eventually automatically) to specify the most efficient route for each aircraft or vehicle.
- Guidance to pilots and drivers during taxiing in the aerodrome. Comprehensive guidance shall be provided to ensure safe surface operations on the aerodrome by directing in an unambiguous and reliable way, pilots and vehicle drivers along their assigned routes. Aerodrome visual aids to guide surface movement shall be included as a part of the integrated system.

An ASMGC's is thus designed to suit the needs of a particular airport and might differ in function between locations. The efficacy of specific ASMGC's functions proposed for a particular airport should therefore be assessed before system implementation. Further, with increased automation an ASMGC's might introduce new human factors concerns, such as task-induced stress on pilots and controllers, reduced human–human/human–computer interaction and deskilling with increased automated aids for guidance and control. To enable these human factors assessments, computer simulations of various features and characteristics of a proposed ASMGC's may be created conveniently and economically. Safety concerns (if any) may then be identified and suitable design modifications proposed to address them. As aerodrome displays remain a crucial navigation aid for pilots after ASMGC's implementation, the simulation can also be used to verify the adequacy and efficacy respectively of existing and new aerodrome displays (see also Lim and Khoo, 2005). To address these wayfinding concerns, a simulation has been developed for an A380 aircraft taxiing in the aerodrome of Singapore's Changi International Airport. An account of the steps involved in the development of the simulation follows.

Development of the Simulation

Steps entailed by the development of the simulation were as follows:

1. *Generating a model of a section of Changi Airport.*

   A section of Changi Airport was modeled using AutoCAD™ 2004 and Discreet™ 3DS Max 6. Markings and displays on the runway and taxiways were created exactly as per the standards mentioned in *ICAO Annex 14 – Aerodromes (Volume 1).*

2. *Selecting a taxiing route in the aerodrome to implement and assess specific traffic control features of an ASMGC's.*

   A possible taxiing route of an A380 aircraft from Terminal 1 to Runway 20R-02L is chosen for the simulated implementation of selected functions of an ASMGC's. The route selected is an intersection at which the selected functional features of an ASMGC's may be exemplified and tested for their efficacy. The route also includes a set of traditional aerodrome displays (see Civil Aviation Authority, 2005) to enable an assessment of the visibility from the cockpit of an A380 aircraft. The simulation comprises 3 parts, namely: a night simulation of the taxiway with and without the ASMGC's and a day simulation to assess the

adequacy of the existing set of aerodrome displays to satisfy visibility requirements in relation to the taxiing speed, turning radius, pilot seating positions, nose cone and cockpit configuration of the A380 aircraft (see Singapore Airlines, 2007).

3. *Modeling the pilot's field of view from the cockpit of an A380 aircraft.*

   To simulate the pilot's field of view (FOV), two virtual cameras are placed at the normative eye level of the pilots in the aircraft cockpit (i.e. seat adjusted in accordance with the eye reference height for the aircraft). As an initial simplification, the simulation assumes that the pilots maintain their heads in the neutral position looking in the forward direction as they taxi the aircraft. For the same reason, a central vision cone of 45 degrees in the horizontal plane and a 10 degree downward tilt in the vertical plane is specified. These visual cones will be modified subsequently to address anticipated characteristics of aircraft navigation and control tasks and the implied position and direction of the pilots' heads and eyes. The camera settings and so the resulting FOV, may be adjusted easily in the simulation after a task analysis of the pilots' tasks has been completed.
4. *Animating and rendering the simulation.*

   The intention of the animation is to create an approximation of the taxiing speed of the aircraft and so determine what may be seen dynamically from the cockpit, i.e. the pilots' FOV at each stage of aircraft taxiing and the duration of which specific aerodrome displays remain visible. The simulation assumes a 15 mph maximum taxiing speed limit at linear sections of the taxiway and a 10 mph maximum turning speed limit. A rate of 10 frames per second is thus set for the animation. This is found to be sufficient for smooth animation and yet would not take too long to render. Turning of the aircraft in the day simulation is assumed to be performed by wing-gear tracking.
5. *Projection of the simulation on an Elumens™ Visionstation to enable an immersive experience.*

   The Elumens™ Visionstation enables immersive viewing by projecting an adapted image onto a hemispherical dome (1.5 m in diameter) which supports a vertical and horizontal FOV of up to 60° and 160° respectively. The wide FOV creates a sense of space and depth without the need for 3D goggles or glasses (see Figure 1). Nevertheless, it is planned that the display system shall be upgraded later to support stereoscopic 3D viewing.

As indicated earlier, the simulation comprises 3 sets of animations. The first two night-time animations (corresponding to a traditional aerodrome and an ASMGC's equipped aerodrome) depict an incursion event in which two aircrafts are about to collide into each other as their taxi paths intersect (see Figure 2). The efficacy of specific ASMGC's functions may then be assessed comparatively in relation to their effectiveness in warning the pilots of the incursion and so avert a collision of the aircrafts. ASMGC's functions tested include the use of automated 'follow the green' taxiing guidance, supported by surface radar and ground sensors. When an incursion event that might lead to aircraft collision occurs, a flashing red light in the cockpit is activated together with a verbal alarm to alert the pilots to the critical situation that is about to happen. A more speculative water screen stop bar/display (see Figure 3)

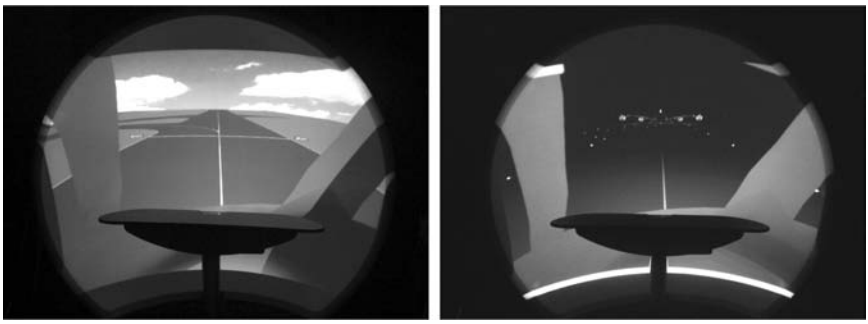


Figure 1. Running the simulation on the Elumens™ Visionstation.

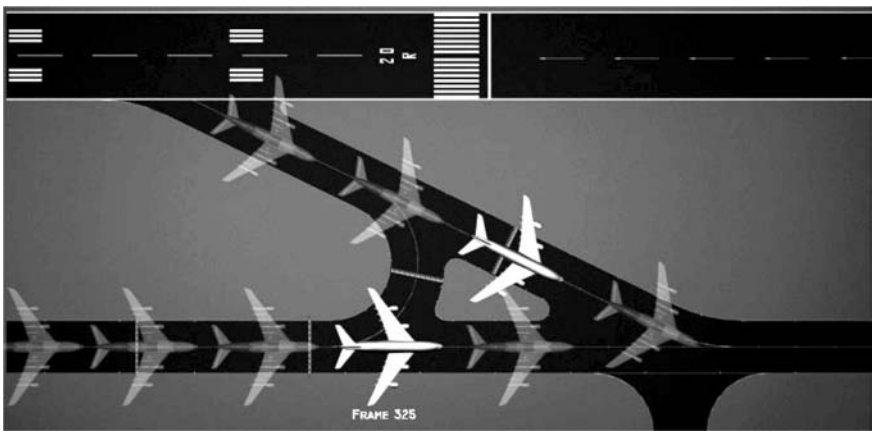


Figure 2. Route of the two aircrafts in the simulated incursion event.

Figure 3. View from the cockpit with the water screen stop display activated.

recruited from a novel traffic display for land transportation (LaserVision, 2007), is also examined for its efficacy in alerting the aircraft pilot in comparison with conventional stop bars. Appropriate performance measures shall be worked out with pilots and civil aviation officers to ascertain the efficacy. Subject testing can only begin after these measures have been specified.

The third day-time animation is created to supplement the first two. This shows the pilots' FOV as they taxi the A380 aircraft along the route chosen for this particular simulation (i.e. from Terminal 1 to Runway 20R-02L). As indicated earlier, the simulation is used to assess the adequacy of existing aerodrome displays.

Concluding summary

The animations developed in this project illustrate the usefulness of software simulations in supporting the assessment of various human factors concerns in relation to: the introduction of new large aircrafts into an existing aerodrome and the efficacy of an ASMGC's. Safety concerns may then be identified and suitable design modifications proposed. Plans for future work include the following:

- Modification of the visual cones of the pilots' to address anticipated aircraft navigation and control tasks and the implied position and direction of the pilots' heads and eyes. A stereoscopic 3D simulation would be created if funds are available to upgrade the capability of the Elumens™ Visionstation.
- Definition of appropriate performance measures to ascertain the efficacy of an ASMGC's. These indicators should be defined with pilots and civil aviation officers. Subject testing can only begin after these measures have been specified.
- More extensive modeling of the aerodrome of Changi International Airport to support the assessment of traditional aerodrome displays in relation to the introduction of the A380 aircraft.

References

Civil Aviation Authority (2005). *WSSS AD 2.9 Surface Movement Guidance and Control System and Markings.*

ICAO. *Manual of Advanced Surface Movement Guidance and Control Systems (A-SMGCS).*

ICAO. *Annex 14 – Aerodromes (Volume 1).*

LaserVision (2007). *Softstop – Sydney Harbour Tunnel.* Retrieved September 13, 2007, from: http://www.laservision.com.au/page.asp?lid=1&sec=Projects&subsec=Permanent+Attractions&subsubsec=Softstop

Lim, K.Y. and Khoo, A.M. (2005). Using an Animated Simulation to Assess the Visibility of Aerodrome Information Displays from the Cockpit. *In: Contemporary Ergonomics, Proceedings of the Ergonomics Society's 2005 Conference, Hatfield, UK, Taylor and Francis.*

Singapore Airlines (2007). *Singapore Airlines A380 Technical Specifications.* Retrieved September 13, 2007, from: http://www.a380.singaporeair.com/content/aircraft/index.html

ASSESSMENT OF PHYSICAL ACTIVITY AT WORKPLACES

Britta Weber[1], Ingo Hermanns[1], Rolf P. Ellegast[1] & Jens Kleinert[2]

[1]*Institute for Occupational Health and Safety of the German Social Accident Insurance (BGIA), Sankt Augustin, Germany*
[2]*Institute of Psychology of the German Sport University Cologne, Cologne, Germany*

The Institute for Occupational Health and Safety of the German Social Accident Insurance (BGIA) has developed a mobile measuring system for long-term activity monitoring. The CUELA activity system (computer-based measurement and long-term analysis of activity behaviour) provides information on the type, intensity, frequency and duration of physical activity (PA). In this paper an evaluation of the intensity determination during a number of office activities is presented. The comparison between PA intensity predicted by heart rate (HR) and by the CUELA activity system indicates that the system offers a good estimation for physiological demands for the investigated activities.

Introduction

Labour-saving devices and the widespread use of computer applications have led to a significant decrease in PA at a growing number of workplaces. A lack of PA at the workplace is a key aspect in the increasingly inactive lifestyle of people in industrial countries. Static body postures and physical inactivity constitute a serious health hazard. The consequences include diseases of the cardiovascular system and musculoskeletal complaints, as well as mental problems. These health risks can be reduced by preventive strategies which promote PA and the variation between sitting, standing and moving at workplaces. Measures promoting PA must be evaluated by a method which provides objective and reliable information on PA in field studies. For this purpose, the sensor-based CUELA activity system was developed. It automatically classifies various activities, such as standing, sitting, lying, kneeling, walking and climbing stairs, and estimates the energy expenditure (EE) in terms of METs (1 MET = average EE during static sitting or $1\,\mathrm{kcal} \cdot \mathrm{kg}^{-1} \cdot \mathrm{h}^{-1}$ or 3.5 ml oxygen uptake $\cdot\, \mathrm{kg}^{-1} \cdot \mathrm{min}^{-1}$). The objective of the present pilot study was to compare MET measurements during various activities involving the CUELA activity system and HR records. HR has been shown to be an accurate predictor for the EE in consideration of age and fitness (Strath *et al*, 2000).

Table 1. Activity protocol.

Activity	Duration [min]	Comments
Static sitting	5	
Dynamic sitting	6	Standardised office tasks
Static standing	2	
Dynamic standing	4	Standardised office tasks
Slow walking (4 km/h)	3	200 m
Moderate walking (5 km/h)	3	250 m
Fast Walking (6 km/h)	3	300 m
Walking downstairs	1	88 steps
Walking upstairs	1	88 steps
Total duration	**28**	

Methods

Four healthy people (two men, two women) participated in the study. The average age was 26.2 (SD 2.3). The test subjects were each monitored for about 30 minutes by the CUELA activity system with synchronised HR recording. The subjects were instructed to perform a number of activities, whose sequence and duration are listed in Table 1. The subjects were required to perform a number of standardised office tasks at a desk (reading, writing, typing, filing) and whilst standing (sorting files in an office cabinet, using the telephone). In order for the duration and sequence of tasks to be maintained constant for all test subjects, a video was presented parallel to the activity. The entire procedure was explained prior to performance, and the video fast-forwarded through. In addition to the video, the measurement was accompanied by written and verbal instructions. In order to standardise the speed during walking and climbing stairs, the test subjects were accompanied by a person who set the pace with a clock.

CUELA activity system

The system is derived by the standard CUELA system, which is used since several years in ergonomic field analysis (Ellegast and Kupfer, 2000). The CUELA activity system measures the motion behaviour by means of sensors, each consisting of a 3D accelerometer and a 1D gyroscope. Elastic and breathable straps are used to attach a total of seven sensors to different points on the body under the clothing: on the upper arm of the dominant arm, on the back at the level of the thoracic spine and the lumbar spine, and on the thighs and calves of both legs (see Figure 1).

The system is set to a default sampling rate of 50 Hz. The measured values are stored by the data logger on a flash card, and subsequently imported into the associated analysis software on the PC. The software provides a range of analysis functions, such as 3D animation of the movements performed, graphic display of angle/time characteristics, automatic detection of activity, estimation of the intensity of activity in METs, and statistic analyses. In addition, the values can easily be exported for further processing in other applications (e.g. MS Excel or SPSS).

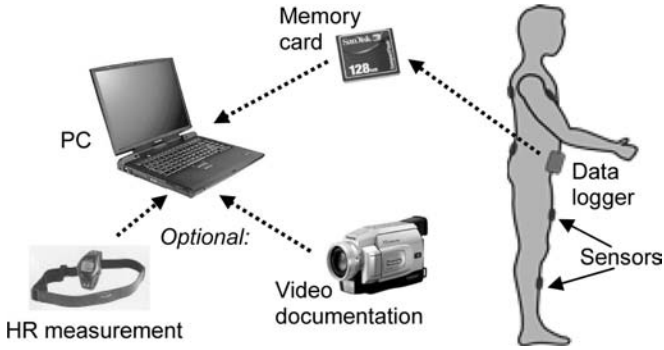


Figure 1. The CUELA activity system.

HR measurement

Wrist-worn monitors with recording function (Polar® Vantage NV) were employed for measurement of the HR. The heart rates were stored at five-second intervals and imported into the CUELA activity measurement as a supplementary data stream. The resting HR (HR_{rest}) was determined by averaging the five lowest measurement values (Janz, 2002). The maximum HR (HR_{max}) was estimated by means of the formula [220 – subject age].

Estimation of cardiorespiratory fitness (CRF)

The CRF (maximum MET level (MET_{max}) or maximum oxygen intake ($\dot{V}O_{2\,max}$) was determined for each test subject by means of the non-exercise prediction equation of Jurca *et al* (2005). The formula takes account of sex, age, BMI, HR_{rest} and a physical activity score determined on a five-stage scale.

Data analysis

The data were analysed by the CUELA software and MS Excel. The MET intensity of activity was determined from the recorded HR and the measured kinematic data.

Determination of MET values from HR (MET_{HR}): With reference to Strath *et al* (2000), the activity HR (HR_{act}) was used to estimate the EE in METs. The procedure followed is shown in Figure 2. The percentage of HR reserve (%HRR) in each case was first calculated [1]. Swain et al. (1997) demonstrated that %HRR and percentage of reserve ($\%\dot{V}O_{2R}$) are closely linked and numerically similar. Based upon this finding, the $\%\dot{V}O_{2R}$ was used to determine the absolute $\dot{V}O_2$ ($\dot{V}O_{2act}$) [2], with $\dot{V}O_{2\,max}$ being determined by means of the non-exercise prediction equation of Jurca *et al* (2005). Division by 3.5 [3] yields the estimated EE (MET_{HR}).

Determining of MET values by means of the CUELA activity system (MET_{CA}): The system automatically detects the type of activity performed by the person (for the programme performed: sitting, standing, walking, walking upstairs and downstairs). Based upon the MET compendium (Ainsworth *et al*, 2000), the

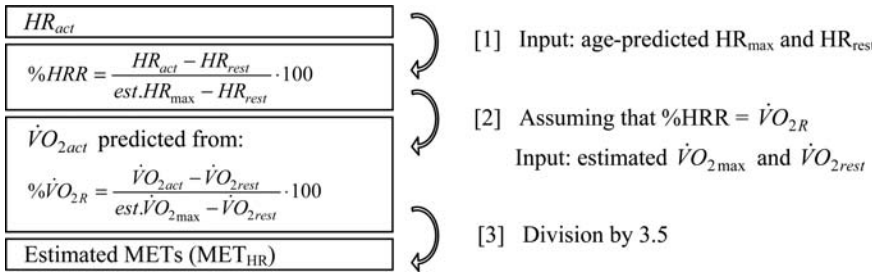


Figure 2. Flow diagram demonstrating the use of HR_{act} to calculate the MET_{HR} (modified according to Strath *et al*, 2000).

corresponding base MET values (MET_{base}) are stored. The latter are combined with the physical activity intensity (PAI), which is determined from the acceleration data. For each body segment fitted by a sensor, the vector magnitude of the 3D acceleration vector is calculated. Subsequent high-pass filtering with a cut-off frequency of 0.1 Hz removes the constant signal portions. To obtain the current movement intensity $PAI_{segment}$ a moving root mean square (RMS) is calculated for the high-pass filtered vector magnitudes (VM_{filt}) across $T = 150$ readings:

$$PAI_{segment_t} = \sqrt{\frac{1}{T}\int_{t-\frac{T}{2}}^{t+\frac{T}{2}} VMfilt_{segment_t}{}^2(t)dt} \tag{1}$$

According to the distribution of segment masses assumed in biomechanical models the segment values are compiled as follows to a value for the entire body (PAI_{total}):

$$\begin{aligned} PAI_{total} = {} & 0.4 \cdot (0.5 \cdot PAI_{thoracic\ spine} + 0.5 \cdot PAI_{lumbar\ spine}) \\ & + 2 \times 0.2 \cdot (0.65 \cdot PAI_{upper\ leg} + 0.35 \cdot PAI_{lower\ leg}) \\ & + 0.2 \cdot PAI_{upper\ arm} \end{aligned} \tag{2}$$

Tables 2 and 3 show the formulae for calculation of MET_{CA}. The values for seated and standing activities are derived from the sum of MET_{base} and a value (MET_{PAI}) calculated from PAI_{total}. The constants were determined using linear regression. In the measurement of intensity for walking tasks, the sum is corrected by $C_{fitness}$. This correction makes allowance for the impact of the test subject's fitness upon the EE during the performance of dynamic activities.

Results

The results of the comparison between METs from HR and the CUELA activity system are presented in Figure 3. For every subject the means of MET_{HR} and

Table 2. Calculation of MET_{CA} for sitting and standing tasks.

$MET_{CA} =$	$MET_{base} =$		$MET_{PAI} =$
$MET_{base} + MET_{PAI}$	Sitting Standing	1.0 1.2	$w_{stat} \times PAI_{total} + o_{stat}$

Table 3. Calculation of MET_{CA} for walking tasks.

$MET_{CA} =$	$MET_{base} =$		$MET_{PAI} =$	$C_{fitness} =$
$MET_{base} + MET_{PAI} + C_{fitness}$	Walking: Downstairs: Upstairs:	3.5 3.0 8.0	$w_{dyn} \times PAI_{total} + o_{dyn}$	$0.1 \times (10 - MET_{max})$

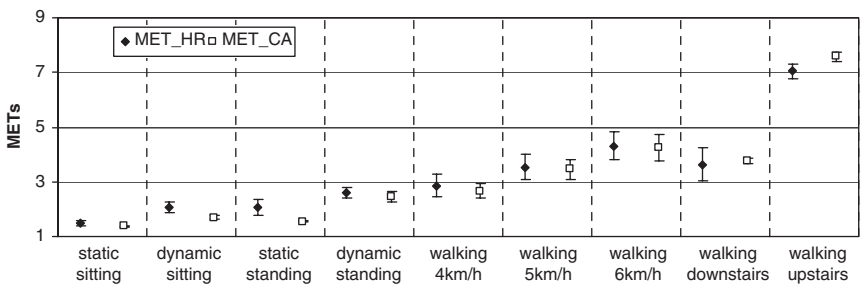


Figure 3. Comparison between MET_{HR} and MET_{CA} (Mean and SD; n = 4).

MET_{CA} for each activity category were calculated. The means per category were then averaged across all subjects respectively.

For all categories the means of MET_{HR} and MET_{CA} are very similar. For *dynamic sitting* and *static standing* the CUELA activity system estimated slightly lower values than the MET prediction via HR. Whereas during *walking upstairs* MET_{CA} is averaged at a higher level than the MET_{HR}. For the remaining categories both MET means are nearly of the same value. In addition, the standard deviations of mean MET_{HR} and mean MET_{CA} are very similar for most activities. Only for *walking downstairs* a wider range of variation occurs for the METs predicted by HR.

Discussion

The METs determined with the aid of the CUELA activity system were shown to yield a close estimation of the EE. Compared to the METs determined from HR, the values were very similar during all activities studied. However, further measurements are to be conducted on a larger scale to examine the occurring differences. In order to apply the gold standard of EE determination, indirect calorimetry should be employed.

Higher variations for mean MET_{HR} (e.g. for *walking downstairs*) arise from the fact that the identical level of movement intensity does not result in the same cardiovascular reaction. Metabolic demands are not only influenced by factors like fitness level, age or sex, which are partly considered in the calculation of MET_{CA}, but also by additional (e.g. emotional) factors. For this reason HR might be included in the MET determination by the CUELA activity system. The use of rating scales to measure the perceived exertion is also conceivable.

References

Ainsworth, B.E., Haskell, W.L., Whitt, M.C., Irwin, M.L., Swartz, A.M., Strath, S.J., O'Brien, W.L., Bassett, D.R., Schmitz, K.H., Emplaincourt, P.O., Jacobs, D.R. and Leon, A.S. 2000, Compendium of physical activities: an update of activity codes and MET intensities. *Medicine and Science in Sports and Exercise*, **32** (Suppl.), 498–516

Ellegast, R.P. and Kupfer, J. 2000, Portable posture and motion measuring system for use in ergonomic field analysis. In K. Landau (ed.) *Ergonomic Software Tools in Product and Workplace Design* (Institut für Arbeitsorganisation, Stuttgart), 47–54

Janz, K.F. 2002, Use of heart rate monitors to assess physical activity. In G.J. Welk (ed.) *Physical Activity Assessments for Health-Related Research*, (Human Kinetics, Champaign, Ill), 143–161

Jurca, R., Jackson, A.S., LaMonte, M.J., Morrow J.R., Blair, S.N., Wareham, N.J., Haskell, W.L., Van Mechelen, W., Church, T.S., Jakicic, J.M. and Laukkanen, R. 2005, Assessing cardiorespiratory fitness without performing exercise testing. *American Journal of Preventive Medicine*, **29**, 185–193

Strath, S.J., Swartz, A.M., Bassett, D.R., O'Brien, W.L. King, G.A. and Ainsworth, B.E. 2000, Evaluation of heart rate as a method for assessing moderate intensity physical activity. *Medicine and Science in Sports and Exercise*, **32** (Suppl.), 465–470

Swain, D.P. and Leutholtz, B.C. 1997, Heart rate reserve is equivalent to $\%VO_2$ Reserve, not to $\%VO_2$max. *Medicine and Science in Sports and Exercise*, **29**, 410–414

A FOLLOW-UP STUDY USING PARTICIPATORY ERGONOMICS TO REDUCE WMSD's

Jason Devereux & Robert Manson

Robens Centre for Public Health, University of Surrey, Guildford, Surrey, GU2 7TE

An intervention within a site of a pharmaceutical company (GSK) using participatory ergonomics was conducted to reduce the incidence of work-related musculoskeletal disorders. Significant reductions in musculoskeletal disorders has been observed and substantial cost savings have been realised. The use of participatory ergonomics is suggested as an intervention for reducing work-related musculoskeletal disorders.

Introduction

Participatory ergonomics has become increasingly used for implementing organizational change and a framework (the PEF) has been developed to assist in the initiation and support of such programmes (Haines et al., 2002). Participatory ergonomics requires ergonomic design, analysis and end user involvement in both the design and analysis stages (Devereux and Buckle, 1999).

Participative change strategies have been shown to improve productivity and comfort (Vink et al., 2006). However, long term follow up studies evaluating changes in the incidence of musculoskeletal disorders have not been conducted. A five year follow-up of the outcomes from a participatory ergonomics strategy within a pharmaceutical organisation is described.

Structure of the participatory ergonomics programme

In 2001, managers at a large pharmaceutical manufacturing site implemented its participative ergonomics programme to investigate the increasing numbers of lost time illnesses and injuries related to work-related musculoskeletal disorders (WMSDs), identify and assess associated risk factors and achieve operational excellence. All lost time incidents in the previous year were caused by WMSDs at this particular site. The site employed approximately 1400 staff working in product laboratories and offices, site maintenance, warehousing and plant rooms. A detailed description of the approach taken is provided elsewhere (Devereux and Manson, 2005).

Table 1 describes the structure of the participatory ergonomics group. Referring to permanence in table 1, GSK saw the importance of integrating the programme

Table 1. The design using the Participatory Ergonomics Framework.

	Case study 1
Permanence	ongoing
Involvement	partial direct representation
Level of influence	organization
Decision-making	delegative participation
Mix of participants	operators, senior management, technical staff
Requirement to participate	voluntary
Topics addressed	physical design of equipment, workplaces, work tasks, formulation of policy or strategies
Brief: problems	identification, solution development, implementation change, monitor process
Role of Ergonomics Specialist	initiates and guides, consults, audits

into existing operational excellence strategies in order to increase the likelihood of having a more lasting impact.

Partial direct representation was used. Participants represented a typical subset of a larger group and their views did not necessarily represent the views of their peers in the first instance. This potential bias was minimized by engaging with the larger workforce on specific projects to gather end user data and test assumptions made. This allowed systematic data to be collected that could be analyzed objectively by members of the ergonomics programme.

For level of influence, the organizational/workplace improvements made were shared across work sites around the world via intranet to exemplify best practice.

Senior management gave members of the participatory ergonomics program discretion and responsibility to adapt the work to individual users within the work areas that they were responsible for.

It was considered important to have a senior manager recognized as a team champion who was highly motivated and respected across business units within the organization. Their leading role was to support and advise team members when developing ergonomics solutions.

The appropriate number of team members had to be large enough to ensure representation of all parts of each organization. The maximum number of members did not exceed fifteen in order to allow effective communication. This helped to achieve full participation from all team members. A representative from health and safety, engineering and operations was sometimes present to ensure that core expertise was brought to each team when necessary. Other expertise such as procurement staff were made available to attend specific meetings to address issues requiring their support.

Due to the autonomous culture of the organization, recruitment of members on the participatory ergonomics team was voluntary. It was important to select highly motivated and interested individuals willing to undergo ergonomics training and take responsibility for process changes. Time was needed to alter beliefs and attitudes towards the objectives of the programme.

To address the problems, an ergonomic improvement team, comprising 12 experienced manual handling assessors, was set up to enable employees to participate in

the improvement process by identifying ergonomic hazards and seeking solutions to reduce the risk of WMSDs.

Responsibilities of the team included:

- learning about work systems in the organisation;
- evaluation of workstations and tasks for risk factors;
- ranking and selection of risk factors to control;
- implementation of improvements to reduce risk factors;
- conducting a risk reduction evaluation;
- providing information and assistance to employees and managers to address risk factors;
- document risks, preventative action and information concerning the risk reduction and productivity improvements;
- assist managers and supervisors when investigating potential WMSDs.

The initial focus was the physical redesign of laboratory/process equipment taking into consideration the organization of work, tasks performed, the physical and social environment, interaction with other technology/equipment and end user physical/mental reactions when using pieces of equipment under investigation.

The potential risk for a WMSD was assessed using methods accepted by team members. The risk assessment methods used by the EIT were easy to use, identified the relevant risk factors for the type of ergonomic problems being addressed and allowed accurate identification of risk factors and assessment of exposure levels. The same risk assessment methods used for the initial assessment were used to conduct a post-intervention assessment to verify the effectiveness of the solution.

For employees performing a particular work task, a discomfort form was sometimes used to gather information concerning areas of discomfort, likely causes, and possible solutions. Additionally, the discomfort form was also re-administered after the intervention to verify changes in intensity, frequency and duration of discomfort and symptoms of WMSDs.

The ergonomics expert was responsible for initiating the programme by training participants, developing usable exposure and discomfort assessment tools for the range of occupations covered by the programme. The expert was also responsible for providing scientific data using expert methods such as video observation analysis and compression forces models for identifying high-risk work tasks when complex issues arose. The quality of record-keeping by team members was also audited by the expert to ensure that pre-and post-intervention results were accurately recorded. Once the confidence and skill set up a team members improved, the consultant reduced the level of supervision and monitoring.

Key ergonomic problems included pallet lifting, repetitive work using poor hand tools in labs, cylinder lifting, loading and unloading storing plugs, packaging boxes, working in isolators, hand tool redesign, operation of high rise trucks etc.

Results of the programme

Twelve months after the participatory ergonomics programme was initiated at the site, 31 work system improvements were achieved. One intervention involved

automation and eliminated human involvement, two interventions mechanised the work process, and thus, eliminating the risk but human involvement remained. Three interventions used administrative controls whereby there were changes to job design or work policy. Twenty-five interventions reduced exposure to multiple risk factors (force, repetition and postural load) for WMSDs by introducing new equipment, repositioning or modifying existing equipment and introducing new ways of working.

There was also a 40% reduction in WMSDs attended to on site by the company physician. A reduction in WMSDs was not observed at similar sized manufacturing sites that did not initiate an ergonomics participatory programme.

Within 3 years, the EIT was considered a resounding success and resulted in the following additional improvements:

- 65 ergonomic improvements implemented and evaluated in total
- 160% increase in ergonomic hazards reported
- efficiency savings and reduce cycle times of up to 40%
- best ever employee health and safety performance-3.4 million hours work free of lost time illness and injury reached in 2003
- best global health and safety audit score 94%

The total cost of implementing the ergonomics programme was approximately £20,000 a year at the site. This figure was derived from costs based on meeting attendance times, design course attendance time and fees, materials for training and consultant fees. This was viewed as a significant cost benefit in relation to the costs of incident investigations, claims, compliance etc. In response to these achievements, the site was awarded the Chief Executive Officers Environmental, Health and Safety Excellence Award in 2004.

The potential barriers influencing the change process were analyzed qualitatively using focus groups, questionnaires and reflective diaries by the team champion. Business disconnect, knowledge gaps by team members, time prioritization, fear of involvement and reluctance to change were themes derived from the data.

Since the successful introduction of this participatory ergonomics programme, GSK has taken this model and implemented it at 15 other of their sites in the UK.

Some specific recorded benefits in GSK sites have included:

- Implementation of a robust ergonomic risk assessment which has seen an increase in the identification of high risk tasks
- Increased identification of ergonomic problems by employees
- Reduction of 40% work related musculoskeletal injuries and illness from 2001–2006
- Reduction of 19% certified days lost due to musculoskeletal conditions on sites from 2004–2006. This has been estimated as saving up to £1.5 million.
- Overtime reduced from £60,000 to £5,000 each year since 2001 following redesign of production lines
- Cycle times reduced from 3 hours to 5 minutes on production lines as a result of redesigning process flow.

Discussion and conclusion

This case study showed significant benefits in reducing the risk of work-related musculoskeletal disorders and the subsequent incidence of disorders by using participatory ergonomics.

It is important that a participatory ergonomics programme aligns itself with operational excellence initiatives currently used within a work organization by using similar performance benchmarks, recording processes and communication channels. Using the participatory ergonomics framework may assist in making the appropriate design decisions for the programme. It is also suggested that such an approach be used to assist in redesigning jobs for workers returning to work from injury.

This case study has been awarded for best practice in health and safety at work for both UK and European competitions.

References

Haines, H., Wilson, J.R., Vink, P. and Koningsveld, E. (2002). Validating a framework for participatory ergonomics (the PEF). *Ergonomics*, **45**(4), 309–327

Vink, P., Koningsveld, E.A.P. and Molenbroek, J.F. (2006). Positive outcomes of participatory ergonomics in terms of greater comfort and higher productivity. *Applied Ergonomics*, **37**, 537–546

Devereux, J. and Manson R. (2005). Ergonomic improvement teams at Glaxo-SmithKline. *Occupational Health Review*, **114**, 28–30

Devereux, J. and Buckle, P. (1999). *A participatory strategy to reduce the risk of musculoskeletal disorders.* In M.A. Hanson, E.J. Lovesey and S.A. Robertson (Eds.). *Contemporary Ergonomics (pp. 286–290).* London: Taylor and Francis

A COLLABORATIVE DESIGN APPROACH TO PREVENTING WORK-RELATED MUSCULOSKELETAL DISORDERS

H.K.G. Punchihewa & D.E. Gyi

Department of Human Sciences, Loughborough University, Loughborough, Leicestershire, LE11 3TU

Work-related musculoskeletal disorders (WRMSDs) affect the health and wellbeing of workers and incur staggering costs to industry. This research is being conducted to examine the potential of Quality Function Deployment (QFD), an established collaborative design tool, in reducing WRMSDs.

To this end, a case study was conducted in a cleaning environment to evaluate a floor scrubbing task. The objectives were to investigate user ability to identify and specify user requirements to prevent WRMSDs. A sample of workers (n = 10) and managers (n = 3) took part. MSD prevalence, body discomfort data and REBA scores were compared with the user requirements extracted from the interviews. This paper presents the methodology and the initial findings from the case study. Finally, the QFD collaborative approach is discussed for the work task.

Introduction

WRMSDs account for a significant proportion of all occupational injuries or illnesses and recognition and control of them has become a major concern because of health risks and associated costs (Buckle, 2005). Various attempts have been made in the past to reduce the prevalence of MSDs among the working population. A plethora of intervention programmes, standards and guidelines have been developed to eliminate workplace risk factors (Ruotsalainen *et al*, 2006). Yet, WRMSDs are still commonplace in the industry and further research is necessary in order to provide safer working conditions for the workers.

Aspects of design may be given priority in reducing MSDs in future research (Buckle, 2005; Karwowski, 2005), but a mismatch between the user requirements and what the engineers ultimately produce seems to be a drawback in the design process (Shinnar *et al*, 2004). Therefore, essential user requirements to reduce MSDs are not reflected in the products that are designed and finally being manufactured.

There are various ergonomics assessment methods being used to identify MSD risk factors, but there is a shortage of tools for ergonomists to identify user requirements and then to systematically convey them to the engineers. A collaborative approach to fill this void between the users and the designers may help the designers better understand the user requirements that reduce WRMSDs (Dul *et al*, 2003).

Reduction of vulnerability towards WRMSDs may in turn save labour hours and reduce related costs resulting in increased productivity.

In order to achieve this, a suite of tools that will encourage stakeholder participation in the design of processes and equipment for the industry is proposed in this research. It will examine the potential of QFD, an established collaborative design tool (Chan and Wu, 2002) in minimising WRMSDs by helping to derive design specifications for processes and equipment. To this end, a case study was conducted in a cleaning environment to evaluate a floor scrubbing task. The main objectives of the study were, to assess the ability of the workers to be involved in a participative design approach and to test the developed methods.

Materials and Method

Permission was obtained from senior managers to recruit participants for interviews and observations prior to the study. Data were collected from 10 workers (age: mean 44.5 yrs, SD 13.5 yrs) which include 5 females (height: mean 159 cm, SD 7 cm; weight: mean 65 kg, SD 4 kg) and 5 males (height: mean 174 cm, SD 6 cm; weight: mean 90 kg, SD 20 kg). They had varying levels of experience (median 1 yrs, range 10 yrs) in cleaning floors using scrubbing machines. Also, the line managers from the company (n = 3) with varying levels of experience were interviewed (median 9 yrs, range 9.5 yrs). Informed consent was obtained from every participant and the Loughborough University ethical guidelines for studies involving human subjects were strictly observed.

Initially, the floor scrubbing work task (using two types of scrubbing machines) was observed and work elements were identified using task analysis. The main work elements identified were: filling water and additives; scrubbing open areas; scrubbing edges of the floor; and emptying the dirty water containers. The two types of scrubbing machines had basic differences in their designs, but the main work elements required for the scrubbing task were the same for both types of machines, such that both types of machines were studied.

To collect data from the workers, semi-structured interviews were conducted with the workers to, elicit demographic data (i.e. age, gender, height, weight and ethnic background); acquire job information (e.g. job title, experience, involved work tasks and individual work schedules); identify their awareness of MSDs using the "worker stage of change" questionnaire (Whysall *et al*, 2007); gather perceived user requirements for the different work tasks by encouraging them to reflect on their work; record musculoskeletal troubles using a modified "standard Nordic musculoskeletal questionnaire" (NMQ) (Kourinka *et al*, 1987); explore the involvement in the task design decision (i.e. rating their involvement in the task design decision on a 9-point Likert scale; and knowledge of the involvement of other stakeholders). Probing questions were asked wherever felt necessary to clarify points of interest. The interviews took 15–30 minutes.

Secondly, video recordings were taken while the workers performed their normal work routine to assess MSD risks using REBA (Hignett, and McAtamney, 2000). A camera was set-up to capture information from two directions where recording

angles are perpendicular to the frontal and sagittal planes of the workers. It was directed as much as possible at the origin of the three intersecting planes (i.e. frontal, sagittal and transverse) to record upper and lower body postures during work tasks. These angles were maintained as much as possible to ensure repeatability and reproducibility although it was impossible to achieve full control. Each work task identified was expected to be captured for at least 15 minutes covering all work elements from both directions mentioned above. Workers were also asked to talk through the process while being recorded.

Then, whole body discomfort (WBD) scales (no discomfort = 0 to extreme discomfort = 6) based on the body part discomfort scale (Corlet, 1990) were distributed among participating workers that regularly used scrubbing machines to obtain discomfort due to posture before and after the work. These were distributed just before the commencement and at the end of scrubbing. The machines are used for approximately 1 hour each day. Simultaneous direct observations of work tasks were also carried out, guided by the work element recording checklist (Konz, 1990). These data were expected to aid the assessment of work-related risk factors for MSDs and to supplement the interviews by identifying further user requirements.

Finally, a second set of interviews were carried out with managers to obtain job information (i.e. job title, experience and job assignments); awareness of MSDs using the "manager stage of change" (Whysall *et al*, 2007); involvement in the task design decision (i.e. involvement in the task design decision on a 9-point Likert scale; procedures being followed when deciding on designs; access to and awareness on ergonomic standards and guidelines; knowledge of the involvement of other stakeholders; and opinion on shop floor level workers being involved in task design decisions) from managers. All interviews took 13–16 minutes.

Analysis

To gain an understanding of the ability of workers to provide specifications for work tasks, responses from the interviews were compared with the REBA scores, MSD prevalence and discomfort data. Next, REBA analyses of the scrubbing work elements were performed using video recordings to identify the risks for MSDs. The most common and the worst postures were used in the estimation of the REBA scores (Bao *et al*, 2007). The MSD prevalence and discomfort data were analysed to help identify risk factors. User requirements were then extracted from the interviews and the observations for the work task, to input into QFD in order to assess the feasibility of deriving specifications for the studied work task.

Results

REBA scores for the work elements were calculated (Table 1) for a 7th percentile woman and a 90th percentile man according to stature (Pheasant and Haslegrave, 2006).

Table 1. REBA scores for scrubbing work elements.

Work element	REBA scores for 7th percentile woman	REBA scores for a 90th percentile man
Filling water	3	4
Scrubbing open areas	4	5
Scrubbing edges and corners	5	9
Emptying containers	8	10

MSD prevalence data showed that shoulder, wrist and lower back problems are common among participating workers. Unfortunately, discomfort data could only be collected from 3 workers due to work pressure, but revealed that the scrubbing work task gives rise to moderate discomfort in wrists, hands, fingers and lower back.

All the workers appreciated that the scrubbing machines were a huge improvement to mopping. It was interesting that user requirements to reduce the risks of MSDs and remedies to form the low level design specifications in this QFD process were identified by the majority of workers. They were:

- Eliminate lifting and carrying the dirty water tanks by incorporating hose to empty;
- Allow proper holding of the handle by eliminating the wire cord that supplies power;
- Eliminate simultaneous holding of the handle and pressing the operating lever by separating the operating lever;
- Eliminate bending over and holding the machine tilted simultaneously using a mechanism to lower the brush;
- Eliminate bending and use of hands to clean when the suction tube gets clogged by adding a device to remove debris from the suction tube.

Discussion

The methodology given in this paper is one element of a suite of tools based on QFD being developed to enhance collaboration among the stakeholders of the ergonomics evaluation and design process. It needs to be tested in different environments in order to gain a comprehensive understanding of its capability and then to make changes if necessary. The method discussed extends beyond the mere assessment of risk factors of MSDs. It helps the translation of risk factors into useful user requirements followed by low level design specifications that could ultimately be understood by the designers.

Interestingly, an association between the REBA scores and the suggestions for improvement by the workers was observed, where most of the suggested improvements were for the high risk areas identified using REBA analysis. This provides a glimpse of the worker ability to provide user requirements to reduce WRMSDs in

the workplace. The company practice of testing the equipment that they are planning to purchase with the users also revealed that the workers and managers can collaboratively provide specifications for equipment. However, further research is needed to ascertain the ability of the worker to specify user requirements to reduce WRMSDs.

The methods used to develop this collaborative design tool have been devised and selected to assist free thinking of the participants to help elicit as much information as possible. Dynamics of real life industrial situations were given prominence in selecting the methods. This method may also help the workers orient their thinking towards making their work tasks risk free in terms of MSDs by reflecting on their work and equipment. It will enable future such assessments with the same participants more effective.

The procedure used in the case study was facilitated by one person and the interviews and observations were carried out concentrating on one participant at a time unlike in group techniques such as focus groups, which resulted in minimal interruption to the normal operation of the organisation. Later, the data was compiled to list the user requirements. As a result, this methodology may suit the current pace of the industrial activities where participation of several personnel on a single occasion is impractical. Even though the workers were interviewed separately, they suggested similar improvements to their work, but further work is necessary in this regard.

Furthermore, a collaborative approach such as this may receive acceptance especially by the designers because the required information can be made available to them electronically after collecting and compiling the data. The collaborative environment may be created using server based technology where all the stakeholders in the design process could participate in the design process. This method leads to an opportunity for participatory design of tasks and equipment to reduce the prevalence of WRMSDs even though the stakeholders are physically at different locations.

Future work of this research project will involve refinement of the data elicitation techniques and processes to make them more efficient and effective. In order to refine the proposed suite of tools, further case studies in different work environments with diverse work tasks will be carried out.

References

Bao, S., Howard, N., Spielholz, P. and Silverstein, B. 2007. Two posture analysis approaches and their application in a modified Rapid Upper Limb Assessment evaluation, *Ergonomics*, **12**, 2118–2136

Buckle, P. 2005, Ergonomics and musculoskeletal disorders: overview. *Occupational Medicine*, **55**, 164–167

Chan, L.K. and Wu, M.L. 2002, Quality function deployment: A literature review. *European Journal of Operational Research*, **143**, 463–497

Corlett, E.N. 1990, The evaluation of industrial seating. In: Wilson, J.R. and Corlett, E.N. (eds.) *Evaluation of Human Work: A Practical Ergonomics Methodology 1990*, (Taylor & Francis, London), 500–515

Dul, J., Willemse, H. and de Vries, H.J. 2003, Ergonomic Standards: identifying stakeholders and encouraging participation. *ISO Bulletin*, **September**, 19–22

Hignett, S. and McAtamney, L. 2000, Technical note: Rapid Entire Body Assessment (REBA), *Applied Ergonomics*, **31**, 201–205

Karwowski, W. 2005, Invited Plenary Paper: Ergonomics and human factors: the paradigms for science, engineering, design, technology and management of human-compatible systems, *Ergonomics*, **48**, 436–463

Konz, S. 1990, *Work Design: Industrial Ergonomics*, Third Edition (Publishing horizons Inc. Scottsdale, Arizona)

Kourinka, I., Jonsson, B., Kilbom, A., Vinterberg, H., Biering-Sorensen, F., Anderson, G. and Jorgensen, K. 1987, Standard Nordic questionnaires for the analysis of musculoskeletal symptoms, *Applied Ergonomics*, **18**, 233–237

Pheasant, S. Haslegrave, C.M. 2006, *Bodyspace: Anthropometry, Ergonomics and the Design of Work*, Third Edition, (Taylor & Francis, London), 20–54

Ruotsalainen, J.H., Verbeek, J.H., Salmi, J.A., Jauhiainen, M., Laamanen, I., Pasternack, I. and Husman, K. 2006, Evidence on the Effectiveness of Occupational Health Interventions, *American Journal of Industrial Medicine*, **49**, 865–872

Shinnar, A., Indelicatoa, J., Altimaria, M. and Shinnar, S. 2004, Short Communication: Survey of ergonomic features of supermarket cash registers, *International Journal of Industrial Ergonomics*, **34**, 535–541

Whysall, Z.J., Haslam, C. and Haslam, R. 2007, Developing the Stage of Change Approach for the Reduction of Work-related Musculoskeletal Disorders, *Journal of Health Psychology*, **12**, 184–197

CROSSMODAL CONGRUENCY BENEFITS FOR TACTILE AND VISUAL SIGNALING

P.A. Hancock[1], L.T.C. James[2] & L. Merlo[2]

[1]*University of Central Florida, Orlando, FL 32816*
[2]*United States Military Academy, West Point, NY 10996*

We conducted an experiment in which tactile messages were created based on five common military arm and hand signals. We compared response times and accuracy rates of novice individuals responding to visual and tactile representations of these messages, which were displayed either alone or in congruent or incongruent combinations. Analyses were conducted on trials where tactile and visual signals messages were presented either individually or concurrently. Results indicated beneficial effects for concurrent, congruent message presentations with both modalities showing a superior response time and improved accuracy when compared to individual presentations in either modality. These results confirm the promise for tactile messages to augment visual messaging in challenging and stressful environments where visual messaging may not always be possible.

Introduction

Many operational conditions such as combat, firefighting, or certain law enforcement and/or emergency management situations impose significant demands on operator's sensory capabilities. Noisy (e.g., weapons fire, vehicle engines, etc.) and murky (e.g., smoke, sandstorms) environments, for example, impose great demands on hearing and vision, and can compromise the ability to exchange critical information through conventional communication pathways. To circumvent these environmental difficulties it may be possible to provide a redundant source of information through the modality of touch, by using tactile signaling.

Previous studies have shown tactile systems can produce relatively stable performance improvements across a variety of body orientations even when spatial translation is required (Oron-Gilad, Downs, Gilson, and Hancock, 2007; Terrence, Brill, & Gilson, 2005) as well as in the presence of physiological stress (Merlo, Stafford, Gilson, & Hancock, 2006). Most of human information processing uses multiple sensory inputs, such as the synthesis of visual and auditory cues (Hancock, 2005; Spence & Driver, 2004; Stein & Meredith, 1993). Literature on experiments that involve the use of two modalities of information presented redundantly each show improvement in the areas of accuracy and response time (Spence & Walton, 2005; Gray & Tan, 2002; Strybel & Vatakis, 2004). The present study seeks to show that similar congruency benefits may be achieved for more complex stimuli presented through both the visual and tactile modalities.

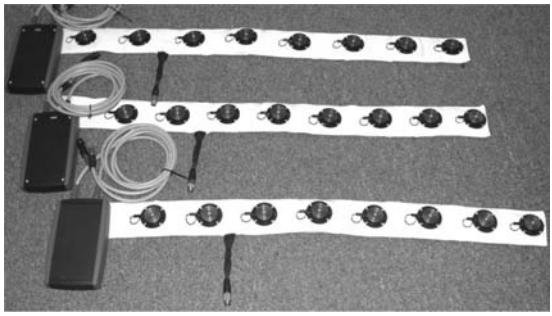


Figure 1. Three tactile displays belt assemblies are shown above along with their controller box.

Experimental method

Experimental participants

To investigate the foregoing proposition, twenty participants (9 males and 11 females) ranging in age from 18 to 48, with an average age of 25 years, volunteered to participate. Each participant self-reported no surgeries, significant scarring or any impediment that might cause lack of feeling in the abdomen or torso area. Additionally, none of the participants had any prior experience with the presented arm and hand signals nor the tactile signals in general.

Experimental materials and apparatus

The vibrotactile actuators (tactors) in our system are model C2, manufactured by Engineering Acoustics, Inc. They are acoustic transducers that displace 200–300 Hz sinusoidal vibrations onto the skin. Their 17 gm mass is sufficient for activating the skin's tactile receptors. The C2's contactor is 7 mm, with a 1 mm gap separating it from the tactor aluminum housing. The C2 is a tuned device, meaning it operates well only within a very restricted frequency range, in this case approximately 250 Hz. The tactile display itself is a belt like device with eight vibrotactile actuators, an example of which is shown in Figure 1. The belt itself is made of elastic and high quality cloth similar to the material used by professional cyclists. When stretched around the body and fastened, the wearer has an actuator over the umbilicus and one centered over his spine in the back. The other six actuators are equally spaced, three on each side, for a total of eight (see Cholewiak, Brill, & Schwab, 2004).

The tactors are operated using a Tactor Control Unit (TCU) which is a computer-controlled driver/amplifier system that switches each tactor on and off as required. This device is shown on the left side of the tactile displays belts in Figure 1. The TCU weighs 1.2 lbs independent of its power source and is approximately one inch thick. This device connects to a power source with one cable and to the display belt with the other and uses Bluetooth technology to communicate with the computer driven interface. Tactile messages were created using five standard Army and Marine

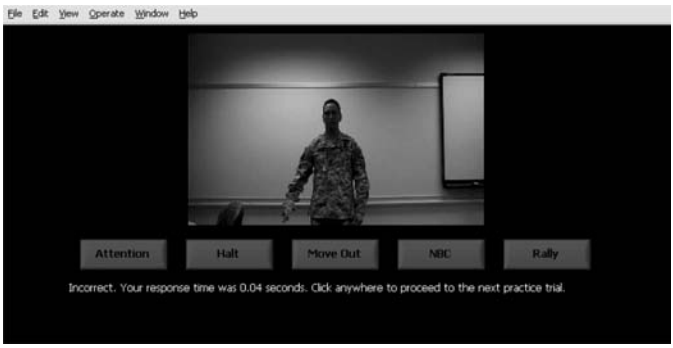


Figure 2. A computer screen shot showing what the participant viewed as the signals were presented. The participant mouse clicked on the appropriate signal name after each presentation.

Corps arm and hand signals (Department of the Army, 1987). The four signals chosen for the experiment were, "Attention", "Halt", "Rally", "Move Out", and "Nuclear Biological Chemical event (NBC)". The tactile representations of these signals were designed in a collaborative effort of scientists at the University of Central Florida and a consultant group of subject matter experts (SMEs) consisting of former US Soldiers and Marines.

Short video clips of a soldier performing the five arm and hand signals were edited to create the visual stimuli. Careful editing ensured the timing of the arm and hand signals closely matched that of the tactile presentations (see Figure 2). A Samsung Q1 Ultra Mobile computer using an Intel Celeron M ULV (900 MHz) processor with a 7″ WVGA (800 × 480) liquid crystal display was used to present videos of the soldier performing the arm and hand signals. This computer ran a custom LabVIEW (8.2; National Instruments) application that presented the tactile signals via Bluetooth to the tactor controller board and captured all of the participant's responses via mouse input. Participants wore sound dampening headphones with a reduction rating of 11.3 dB at 250 Hz to reduce the effects of any auditory stimuli emanated by the tactor actuation.

The display of each message or signal was presented in one of four ways:

- visual only (video presentation of the arm and hand signal)
- tactile only (tactile presentation of the arm and hand signal)
- both visual and tactile simultaneous and congruent (i.e. the same signals were presented both through the video and through the tactile system)
- Both visual and tactile simultaneous and incongruent (i.e. the visually presented signal did not match the presented tactile signal).

Experimental design and procedure

Participants first completed a computer-based tutorial that described each arm and hand signal individually. For each signal, a short description was presented.

Participants then viewed a video of a soldier performing the signal followed by experiencing its tactile equivalent. Finally, the participants were able to play the signals concurrently (both visual and tactile representation). Participants were allowed to repeat the presentation (i.e., visual, tactile, visual-tactile) as many times as desired. Once the participant reviewed the five signals in the two presentation styles, a validation exercise was performed. Participants had to correctly identify each signal twice before the computer would prompt the experimenter that the participant was ready to begin.

Each participant performed two, 60 trial blocks. The blocks had two of each signal presented only visually (10 total), two of each signal with only tactile signals (10 total), four of each signal performed simultaneously with both congruent visual and tactile presentation (20) and four of each tactile and visual signal performed simultaneously but incongruent (20). Each participant performed two blocks of trials, with the 60 trials within the blocks completely randomized for each participant. The entire experiment took less than an hour to complete.

Results

All analyses reported were conducted using SPSS 11.5 for Windows with the alpha level set at .05 for a two-tailed t-test conducted unless otherwise noted. Results were analyzed in terms of the speed of the response and the accuracy of the response under the respective conditions. In the interest of brevity, only the results from the three conditions of visual presentation, tactile presentation and congruent, concurrent presentation are presented in this work.

A one-way Analysis of Variance (ANOVA) was performed on the mean response times across the three experimental conditions of visual presentation, tactile presentation or visual-tactile concurrent and congruent presentation, with the following results: $F(1,\ 19) = 473.45,\ p < .01,\ (\eta_p^2 = .961,\ \beta = 1.00)$. Subsequent a priori pairwise analysis showed, simultaneously presented congruent signals resulted in significantly faster response times than visual signals presented alone $t(19) = -2.25,\ p \leq .04$, see Figure 3. Also, as is evident, the congruent signals were faster than tactile alone $t(19) = -3.98,\ p \leq .01$. Additionally, the visual only presentation of the signal was significantly faster than the tactile only presentation of the signal $t(19) = -2.16,\ p \leq .04$.

Although, there was no significant difference in the accuracy rates observed between the visual and tactile signals to each other when presented alone $t(19) = 2.00,\ p \leq .06$. There was a significant difference in the performance rates when the tactile modality was compared to the concurrent congruent presentation of the signals, $t(19) = 4.03,\ p \leq .01$. The overall lower accuracy rate for the tactile signaling was due to an apparent confusion between the tactile signal for "NBC" and "Halt", which have similar tactile characteristics but low visual similarity. Analysis without the "NBC" tactile signal data removes these differences (it was not significantly different with it included) in the error rate between visual and tactile signals, while not influencing the main effect between congruent modality signaling and single modality signaling at all.

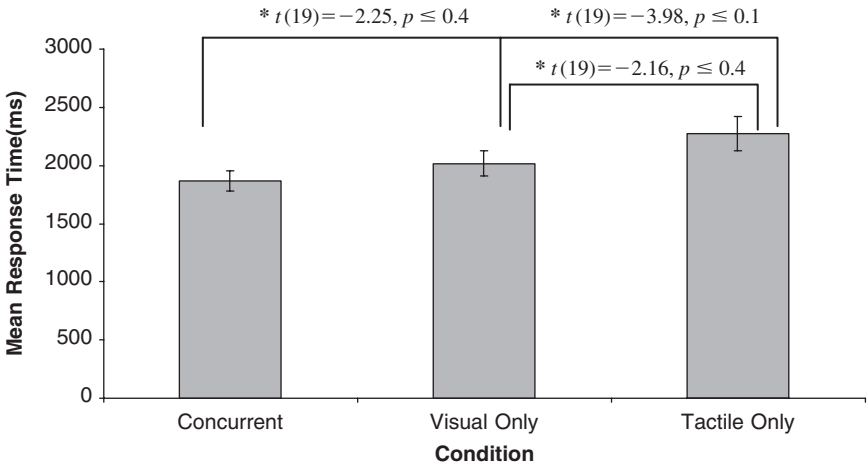


Figure 3. Response Time in milliseconds by signal presentation condition.

Discussion

The overall high accuracy rate displayed by the participants (over 80% in all modalities with fewer than ten minutes of training) is highly encouraging to the current form of tactile display. The accuracy of the messages and the reported intuitiveness with which they were received is also a testament to the utility of the subject matter expert information and the present tactile "language" transformation format. Similarities in both tactile and visual signals that cause confusion among signals are virtually eliminated in concurrent presentation. The rich multi-modal information format seemed to produce faster and more accurate performance.

The question of learning complex tactile communication signals, especially for use in adverse or unusual circumstances is liable to be an important future issue. The tactile system acts as a redundancy gain as the participants now have two means of receiving communication because the visual hand and arm signals would still be available. While initial testing seems to result in superior performance for tactile communication and traditional arm and hand signals combined, the challenge of a universal input device remains a significant hurdle. Stimulus response compatibility will have to be analyzed carefully to maximize performance as different types of inputs are considered for use with tactile displays. However, when individuals are faced with extreme challenges and the traditional sources of information are either diminished of eliminated altogether, the tactile system provides an important alternative communication channel and one that can and should be exploited.

References

Cholewiak, R. W., Brill, J. C., & Schwab, A. (2004). Vibrotactile localization on the abdomen: Effects of place and space. *Perception and Psychophysics, 66,* 970–987.

Department of the Army. (1987). *Visual Signals*. (Field Manual No. 21–60). Washington, DC: Government Printing Office.

Gray, R., & Tan, H. Z. (2002). Dynamic and predictive links between touch and vision. *Experimental Brain Research*, 145(1), 50–55.

Hancock, P. A. (2005). Time and the privileged observer. *KronoScope*, 5(2), 177–191.

Merlo, J. L., Stafford S. C., Gilson, R., & Hancock, P. A. (2006). The effects of physiological stress on tactile communication. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, San Francisco, CA.

Oron-Gilad, T., Downs, J. L., Gilson, R. D., & Hancock, P. A. (2007). Vibro-tactile cues for target acquisition. *IEEE Transactions on Systems Man, and Cybernetics: Part C, Applications and Reviews*, 27(5), 993–1004.

Spence, C., & Driver, J. (Eds). (2004). *Crossmodal Space and Crossmodal Attention*. Oxford; New York: Oxford University Press.

Spence, C., & Walton, M. (2005). On the inability to ignore touch when responding to vision in the crossmodal congruency task. *Acta Psychologica*, 118(1), 47–70.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses.* Cambridge, MA: MIT Press.

Strybel, T. Z., & Vatakis, A. (2004). A comparison of auditory and visual apparent motion presented individually and with crossmodal moving distractors. *Perception*, 33(9), 1033–1048.

Terrence, P. I., Brill, J. C., & Gilson, R. D. (2005). Body Orientation and the Perception of Spatial Auditory and Tactile Cues. *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society, Orlando, FL.*

OCCUPATIONAL HEALTH AND SAFETY

MANAGEMENT OF WORK-RELEVANT UPPER LIMB DISORDERS

A.K. Burton[1], N.A.S. Kendall[2], B.G. Pearce[3], L.N. Birrell[4] & L.C. Bainbridge[5]

[1] *Centre for Health and Social Care Research, University of Huddersfield, UK*
[2] *Health Services Consultancy, London, UK*
[3] *Humane Technology, Rothley, UK*
[4] *Rolls-Royce plc, Derby, UK*
[5] *Pulvertaft Hand Centre, Derbyshire Royal Infirmary, UK*

This review, using a best evidence synthesis, examined the evidence on management strategies for work-relevant upper limb disorders. Articles were found through systematic searching of electronic databases together with citation tracking. Information from included articles was synthesised into high-level evidence statements, which were distilled into key messages. The results covered the following main themes: classification/diagnosis, epidemiology, associations/risks, and management/treatment. The focus was on return to work and took account of distinctions between non-specific complaints and specific diagnoses. Neither medical treatment nor ergonomic workplace interventions alone offer an optimal solution; rather, multimodal interventions show considerable promise, particularly for vocational outcomes. Early return to work, or work retention, is an important goal for most cases and may be facilitated, where necessary, by transitional work arrangements. Successful management strategies appear to require all the players (workers, employers and healthcare) to be onside, and acting in a coordinated fashion. The biopsychosocial model applies: biological considerations should not be ignored, but it is psychosocial factors that are important for occupational as well as clinical outcomes. Implementation of interventions that address the full range of psychosocial issues will require a cultural shift in the way the relationship between upper limb complaints and work is conceived and handled.

Background

The Health & Safety Executive (HSE), in it's guidance *Upper limb disorders in the workplace*, acknowledges that not all upper limb disorders (ULDs) can be prevented. Whilst it seems likely that a mix of physical, physiological, psychological, or social/cultural factors influence the development and impact of ULDs (HSL (Lee & Higgins) 2006), their relevance to overall management is uncertain.

The biopsychosocial model has been shown to be highly applicable to the understanding and management of pain, and has successfully been applied in the management of problems such as low back pain. It may be that a biopsychosocial approach is equally applicable to other musculoskeletal disorders such as ULDs, but before reaching this conclusion it is necessary to consider whether there are issues that might render application of the biopsychosocial model less relevant and useful. The field of ULDs is complicated by the plethora of commonly used diagnoses and classifications, many of which are predicated on specific pathophysiological features, in addition to the accepted existence of non-specific regional pain. Furthermore, so far as work-relevance is concerned, some of the specific conditions are prescribed industrial diseases and eligible (in the UK) for Industrial Injuries Disablement Benefit.

To shed light on this complex topic, the HSE commissioned a review of the literature to determine the nature of ULDs and establish to what extent they should be managed according to biopsychosocial principles. This summary paper presents some key findings and attempts to locate them in an ergonomics perspective; the full report is available in the HSE Research Report Series (Burton et al 2007).

Methods

A systematic search of major electronic databases (mid 2007) was undertaken using appropriate keywords to retrieve articles pertaining to the development and management of upper limb disorders. In addition citation tracking was undertaken, together with searches of personal databases and the Internet. Each article for inclusion ($n \sim 200$) was read and summarised; the original authors' main findings were extracted, checked, and entered into evidence tables. Themes were identified from the evidence tables and the information was synthesised into high-level evidence statements and linked to the supporting evidence, which was graded to reflect the level of support.

The overall methodology, which attempted to bring together a diverse literature on a complex subject, should be viewed as a 'best evidence synthesis', summarising the available literature and drawing conclusions about the balance of evidence, based on its quality, quantity and consistency (Slavin 1995). This approach offered the flexibility needed to handle complex topics, but at the same time took a rigorous approach when it came to assessing the strength of the scientific evidence.

Key findings and context

The results cover the main conceptual themes of classification/diagnosis, epidemiology, associations/risks, and management/treatment, focusing on return to work whilst taking account of distinctions between non-specific complaints and specific diagnoses.

Classification/diagnosis

There is strong evidence that classification and diagnosis of ULDs is particularly problematic; there is a lack of agreement on diagnostic criteria, even for the more common specific diagnoses (e.g. tenosynovitis, epicondylitis, rotator cuff syndrome). Inconsistent application, both in the clinic and workplace, leads to misdiagnosis, incorrect labeling, and difficulties in interpretation of research findings. There is moderate evidence that the scientific basis for descriptive classification terms implying a uniform etiology, such as RSI (repetitive strain injuries) and CTD (cumulative trauma disorders), is weak or absent and they are inconsistently applied/understood.

These diagnostic uncertainties have encouraged some reviewers to discuss ULDs simply as regional musculoskeletal disorders, reflecting the subjective experience and difficulty in determining a specific cause or pathology in the vast majority of cases. Indeed, a considerable number of the articles retrieved for the present review take a 'lumping' approach whereby studies will include a variety of different disorders under labels such as 'work-related upper limb disorder' or simply 'musculoskeletal disorders'. On the one hand it may be argued that a specific diagnosis provides insight into pathogenesis, and therefore to effective treatment. On the other hand, it may be felt that many of the specific diagnoses offered to patients are in reality uncertain, and in any case tell us little about what treatment may be effective.

Epidemiology

The epidemiology of ULDs is essential to understanding how they arise, in whom, and to inform on the natural history. Many older studies of the epidemiology of ULDs relied on cross-sectional observational designs (including surveillance data), which may illustrate an association between a given characteristic (e.g. job) and the existence of symptoms, but does not confirm a causative link. If study design is not carefully considered, along with the criteria for causation – strength, temporality, consistency, specificity, and dose-response of the association, plus biological plausibility – there is a risk of misinterpreting the epidemiological evidence.

There is strong evidence of a very high background prevalence of upper limb pain and neck symptoms in the general population: the 1-week prevalence in general population can be >50% (Walker-Bone et al 2004). Estimates of the prevalence rates of specific diagnoses are less precise, but are considerably lower than for non-specific complaints. Rates vary depending on region, population, country, case definition, and on the question asked.

There is moderate evidence that upper limb pain is frequently experienced in more than one region at the same time (both bilaterally and at anatomically adjacent sites). There is strong evidence that ULDs often lead to difficulty with normal activities and to sickness absence, yet most workers can and do remain at work; this applies to specific diagnoses as well as non-specific complaints.

Associations/risks

The idea that controlling risk at the workplace will result in (some measure of) prevention of injury and ill health is attractive, and has had considerable success for safety outcomes (e.g. reducing major injuries and occupational diseases). However, it has not had the same effect on health outcomes (e.g. pain and disability due to musculoskeletal symptoms, which are sometimes characterised as injuries). This is evident from the high levels of growth in disability and work loss associated with musculoskeletal pain over the very period when industrialised countries have implemented occupational safety and health legislation, and developed inspectorates for compliance and enforcement (Coggon et al 2007).

Large-scale influential reviews published around the turn of the millennium (which included much cross-sectional data) concluded that there were strong associations between biomechanical occupational stressors (e.g. repetition, force) and ULDs: backed by plausible mechanisms from the biomechanics literature, the association was generally considered to be causative, particularly for prolonged or multiple exposures (though a dose-response relationship generally was not evident).

More recent epidemiological studies involving longitudinal designs also suggest an association between physical exposures and development of ULDs, but they report the effect size to be rather modest and largely confined to intense exposures. The predominant outcome investigated (primary causation, symptom expression, or symptom modification) is inconsistent across studies and remains a subject of debate. This is true for regional complaints and (with few exceptions) most of the specific diagnoses.

There is strong evidence that workplace psychosocial factors (beliefs, perceptions, work organisation) and individual psychological factors (anxiety, distress, depression) are associated with various aspects of ULDs, including symptom expression, care seeking, sickness absence, and disability. Similarly, individual psychological factors are associated with various aspects of ULDs, including symptom expression, sickness absence, and disability.

Regardless of the causation debate, the consistent association between upper limb complaints and the physical demands of work shows that ULDs are frequently work-relevant: remaining at work may be difficult or impossible in the face of symptoms. Recognition of this issue is likely to be an important aspect for successful interventions.

Management/treatment

There is moderate evidence that interventions using cognitive-behavioural principles, and multidisciplinary occupational rehabilitation for people with ULDs can improve occupational outcomes. There is moderate evidence that ergonomic work (re)design, directed at equipment or organisation, has not been shown to have a significant effect on incidence and prevalence rates of ULDs, yet ergonomics interventions can improve worker comfort. Although the components of return-to-work interventions vary, there is moderate evidence that integrative

approaches can facilitate return to work; temporary transitional work arrangements (modified work) seem an important component.

Neither medical treatment nor ergonomic workplace interventions alone offer an optimal solution; rather, multimodal interventions show considerable promise, particularly for vocational outcomes. Some specific diagnoses may require specific biomedical treatments, but the components of supplementary interventions directed at securing sustained return to work seem to be shared with regional pain disorders. Early return to work, or work retention, is an important goal for most cases; successful management strategies require all the players to be onside and acting in a coordinated fashion.

Interpretation

The high prevalence of ULDs suggests that the symptoms arise from normal physiological processes and everyday events, such as fatigue or soft tissue strain, rather than some sinister pathology; most cases might best be viewed as regional (pain) complaints. The specific diagnoses implicating pathology or injury are far less common, and many cases will be mislabeled (whether colloquially or by a healthcare professional).

For many people, their symptoms will be work-relevant: their work may be painful or difficult irrespective of the origin of the symptoms. However, even when work is related to the expression of symptoms, that does not mean work was necessarily the underlying cause: it is apparent that work is not the predominant cause of most ULDs.

The biopsychosocial model is certainly appropriate to understand the phenomenon of work-relevant upper limb disorders, and has important implications for their management. Biological considerations should not be ignored, particularly for initial treatment of cases with specific diagnoses, but it is psychosocial factors that are important when developing and implementing return-to-work and stay-at-work interventions, which need to address obstacles to recovery and obstacles to working. Work is beneficial and people need to be helped and encouraged to remain in, or return to, work. This is true both for non-specific upper limb complaints and specific diagnoses, and indicates that cultural shift is required in the way the relationship between upper limb complaints and work is both conceived and handled.

Although early work-return is seen as advantageous, simply sending someone directly back to a job they find painful is counter-intuitive and inappropriate. There is a strong case for using transitional work arrangements as the facilitator, which takes account of both biological and psychosocial obstacles to RTW. The use of temporary modification of activities to support people with regional pain states on their return to work is well established, and there is no clear evidence that the principle cannot or should not be applied to the specific diagnoses.

Just because the epidemiological pattern of most ULDs does not favor ergonomic interventions as a significant primary preventive measure, this does not mean there

is no merit in making work ergonomically acceptable. Jobs, naturally, should be within the reasonable capabilities of the workers; if job tasks are close to, or exceed, physiological limits, a proportion of workers are going to succumb to injury. However, portions of the ergonomics literature and official guidance give the erroneous impression that work is intrinsically the predominant cause of ULDs, and that they will be eliminated by applying a risk identification and control approach. The evidence reviewed here indicates they will not. Furthermore, a possible problem with ergonomic interventions is that they can reinforce workers' beliefs that they are exposed to a serious hazard, and thereby encourage undue reporting of symptoms, inappropriate workloss, and development of disability (Coggon et al 2007). Nevertheless, an ergonomics approach, correctly applied, should improve comfort and efficiency and, importantly, assist in accommodating those with work-relevant complaints or disorders. The adage '*work should be comfortable when we are well and accommodating when we are ill*' (Hadler 1997) is certainly apposite – good ergonomics will not stop all workers' arms hurting, yet it is a necessary, albeit not sufficient, tool for managing the ULD phenomenon.

References

Burton, A.K., Kendall, N.A.S., Pearce, B.G., Birrell, L.N., and Bainbridge, L.C. 2007, *Management of upper limb disorders and the biopsychosocial model*, (HSE Books, London).

Coggon, D., Palmer, K.T., and Walker-Bone, K. 2007, Occupation and upper limb disorders. *Rheumatology* **39**, 1057–1059.

Hadler, N.M. 1997, Back pain in the workplace. What you lift or how you lift matters far less than whether you lift or when. *Spine* **22**, 935–940.

HSL (Lee & Higgins). 2006, *Report on the proceedings of HSE's musculoskeletal disorder and return to work workshop (WPS/06/02)*, (Health & Safety Laboratory, Buxton).

Slavin, R. 1995, Best evidence synthesis: an intelligent alternative to met-analysis. *J Clin Epidemiol* **48**, 9–18.

Walker-Bone, K., Palmer, K.T., Reading, I., Coggon, D., and Cooper C. 2004, Prevalence and impact of musculoskeletal disorders of the upper limb in the general population. *Arthritis & Rheumatism* **51**, 642–651.

GLOVE EFFECTS ON HAND PERFORMANCE CAPABILITIES IN A LIGHT ASSEMBLY TASK

Iman Dianat & Christine M. Haslegrave

Institute for Occupational Ergonomics, University of Nottingham, Nottingham NG7 2RD, UK

The study investigated the effects of wearing gloves on forearm torque strength, error rate and EMG activity of the extensor carpi radialis and biceps brachii muscles when performing light assembly tasks over the duration of a 2 hour work period. Eight male subjects participated in the experiment, being asked to perform a simulated assembly task fitting components together using screws and a screwdriver. The results showed that wearing gloves significantly increased the extensor carpi radialis muscle activity and the error rate. There was also a significant effect of time interval (over the 2 hour duration of the task) on the extensor carpi radialis muscle activity, maximum torque strength and error rate, which suggests that the effects of gloves on hand performance will change over time. Therefore, both task duration and type of task are important considerations in the selection and allocation of gloves to assembly workers.

Introduction

Many industrial tasks require the use of gloves for protection of the worker's hands from different hazards. Although a number of studies have investigated short duration effects of gloves on hand performance measures (for example Mital et al, 1994; Fleming et al, 1997; Bishu and Klute, 1995; Shih and Wang, 1997), longer duration effects are still unknown due to lack of research in this area. Effectively, the studies have used "one-time" tests to evaluate different aspects of hand performance. It is obvious that in most industrial tasks, such as assembly operations, workers have to work with their gloves for extended periods of time (sometimes for 8 hours a day) and it is possible that the length of time alters the effects of the gloves. Therefore, "one-time" evaluation tests may not give a realistic insight into the effects in a work context. Assembly tasks are yet one of the most common parts of many industries such as electronic industry and protective gloves are often needed to protect workers' hands from injury while operating hand tools (Chang et al, 1999). It is important to know the exact effects of gloves on an operator's hand performance capabilities over time so that job demands can be evaluated properly. The purpose of this study was to investigate and compare the short and long term effects of gloves on several aspects of work capacity relevant to industrial tasks in which protective gloves may be required, and to determine the most appropriate duration to be used for

evaluation of each of these performance measures. The physical capacities studied were maximal forearm torque strength, error rate and electromyography (EMG) activity of extensor carpi radialis (ECR) and biceps brachii (BB) muscles, which were selected based on their relevance to the screwdriving task, providing the forearm supination torque force for fastening the screws in a clockwise direction. Error rate, which was defined as the number of falling objects (i.e. screws, components or screwdriver) from the hand while performing the task, can be taken as a measure of dexterity while performing the task.

Method

Eight male subjects aged between 28 and 40 years participated in the experiment. All subjects were healthy with no upper limb injuries and were not professionally involved in working with hand tools, especially screwdrivers. All subjects signed a consent form before the start of the trials and were familiarised with the experiment; their questions were answered by the experimenter and they were told that they are free to withdraw at any stage of the experiment. Their anthropometric data, including height, body weight and hand volume, were recorded at the start of first session. In order to familiarise the subjects with the task, they practiced the simulated assembly task for two minutes before undertaking the actual task. The task was typical of industrial assembly work, consisting of using a screwdriver to join two components together. Each subject performed four experimental sessions in four successive weeks, each session testing a different glove condition (with order of presentation randomized across the subjects). Each experimental session consisted of four 30-minute cycles of the task, where subjects were asked to complete 6 assemblies during each 30-minute cycle, followed by a 5 minute rest period after each cycle. Three commercially available gloves which are commonly used in industrial assembly tasks, as well as a bare hand condition, were tested in the experiments. The gloves were nylon, cotton, and nitrile-coated cotton gloves. These were provided in different sizes (small, medium and large) and subjects selected whichever glove size they felt was the best fit.

Electromyograhic activity (EMG) of extensor carpi radialis (ECR) and biceps brachii (BB) muscles of the right hand was assessed using the MT8-3 Biological Telemetry System from MIE Medical Research Ltd., Leeds, UK. EMG data were recorded using disposable Ag/AgCl surface electrodes (Blue Sensor Electrodes N-00-S, Medicotest Ltd., Denmark), with inter-electrode distance of 2 cm and skin impedance less than 5 KΩ. The preparation of the skin and electrode placement was carried out according to the recommendations of Basmajian and De Luca (1978). EMG signals were amplified differentially (input impedance >10 MΩ; common mode rejection rate >100 dB and gain $\times 1000$), bandpass filtered (20–300 Hz) and converted into root-mean-square (rms) outputs. EMG signals were transferred from the receiver unit to a computer using National Instrument data acquisition DAQ board (model AT-MIO-16E, National Instruments, Chicago, USA) and were sampled at 1000 Hz with software created in LabVIEW (National Instruments Corp., Austin, TX) running on the computer. For normalization of the EMG signals to the

maximum EMG, each subject performed three maximum voluntary contraction (MVC) efforts for 3 seconds for each muscle according to muscle activation methods recommended by Kendal et al (1993) at the start of the first experimental session. The average value of the three readings was used to normalize the EMG data collected during the experimental task.

Torque strength measurements were obtained using an AFG 500N torque meter (Mecmesin, Slinfold, West Sussex, UK) and a T-shape handle similar to that used by O'Sullivan and Gallwey (2005). The T-shape handle consisted of a handle that made an angle of 70° with a shaft in order to provide a neutral wrist position during exertions. Both handle and shaft had a diameter of 25 mm, but the shaft was reduced to a thickness of 8 mm at the joint with the handle for better gripping between the index and second finger. Torque measurements were carried out in a standing position with the upper arm at approximately 0° abduction, forearm horizontal and elbow flexed at 90°. The T-shape handle was attached to the torque meter shaft which was mounted on steel bar and could move vertically along the steel bar to provide the desired height. Subjects were asked to exert maximum voluntary torque over a 3 second period with their dominant hand. Two replications were taken for each experimental condition, with a two minute rest period in between, and the average value was recorded as the maximum forearm torque strength.

The number of objects (i.e. Screws, components or screwdriver) falling from the hand while performing the simulated assembly task during each 30-minute time interval was defined as the measure of error rate. Torque force measurements were recorded before the start of the task and immediately after finishing each 30-minute cycle. Error rate and EMG measurements were carried out during each cycle, while the subjects were performing the simulated assembly task.

Results and discussion

The main results of the study are shown in Table 1. Glove type had a significant effect on the muscle activity of the ECR muscle, which is consistent with other findings (Larivière et al, 2004), and on the error rate. All three kinds of gloves required higher ECR muscle exertion levels (%MVC) compared to the bare hand condition, with increases of 10.9, 10.9 and 14.5%, respectively, for cotton, nylon and nitrile gloves. Higher error rates were recorded when using nitrile gloves compared to the other hand conditions.

Peak forearm torque and ECR muscle activity were both significantly affected by the time interval. Torque capability was reduced by 60 minutes after the start of the task. However, when torque values at different time intervals were calculated as a percentage of the bare hand condition, it was found that the relative level of torque exerted increased steadily over 90 minutes for all three kinds of gloves, with 23, 16 and 29% greater torque exerted relative to that with bare hand with the cotton, nylon and nitrile gloves, respectively (Figure 1). By 120 minutes, a higher muscular effort (%MVC) was recorded, compared to the other time intervals. ECR muscle activity (% MVC) at 120 minutes was 7.8%, 5.7%, and 6% higher, respectively,

Table 1. Mean (standard deviation) of all dependent variables under various test conditions.

Dependent variables	Peak torque (Nm) Mean (SD)	Muscle activity (%MVC) ECR Mean (SD)	Muscle activity (%MVC) BB Mean (SD)	Error rate Mean (SD)
Gloves				
Bare hand	9.0 (2.35)	38.5 (5.75)	39.6 (3.74)	0.4 (0.80)
Cotton	9.9 (2.12)	42.7 (7.21)	42.8 (6.76)	0.8 (0.83)
Nylon	9.1 (1.89)	42.7 (4.69)	41.7 (6.99)	0.6 (0.74)
Nitrile	9.6 (0.92)	44.1 (6.97)	40.1 (6.49)	1.3 (1.11)
Time interval (min)				
0	10.0 (1.86)	NA	NA	NA
30	9.9 (1.91)	40.8 (6.40)	39.6 (3.74)	1.0 (0.96)
60	9.4 (1.72)	41.6 (6.84)	42.8 (6.76)	0.8 (0.78)
90	9.0 (1.82)	41.5 (6.19)	41.7 (6.99)	0.5 (0.84)
120	8.8 (2.08)	44.0 (6.50)	40.1 (6.49)	0.9 (1.08)

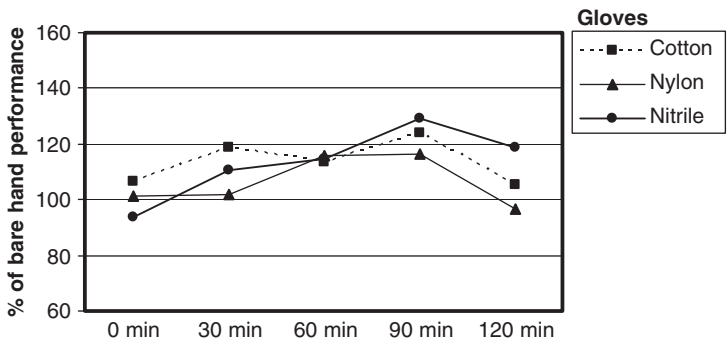


Figure 1. Forearm torque strength (as % of bare hand performance) over time.

than at 30, 60 and 90 minutes. Muscle activity levels as a percentage of bare hand performance did not differ significantly over time.

There was a significant interaction effect between glove condition and time interval for peak forearm torque. This showed higher values at 0 and 30 minutes than at 90 and 120 minutes in the bare hand, cotton and nylon glove conditions, whereas for nitrile gloves there was no torque decrement over time.

Conclusions

The results of this study showed that a higher extensor carpi radialis muscle effort was needed and error rate increased when wearing gloves, especially nitrile gloves, while performing the simulated assembly task. It was also found that the duration was an influencing on glove effects. This means that, if a “one-time” torque force

measurement was made without consideration of the duration of glove use in actual working situations, the reported gloves effect could be misleading. Therefore, it is recommended that, for glove evaluation studies, the duration of using gloves by workers in actual working conditions should be considered and strength tests should last for at least 90 minutes. However, muscle activity and error rate measurements can be done without any time considerations.

There was an interaction between type of glove and duration of wearing. The decrement in strength over time was considerably less for the thicker (nitrile) gloves than with bare hands, (i.e., nitrile gloves in this study which can be considered as an advantage for this kind of glove compared to both bare hand condition and thinner (cotton or nylon) gloves. In contrast, for those tasks that require dexterity and fine movements of hands and fingers, the bare hand condition and cotton and nylon gloves serve better (as indicated by lower error rates and less extensor carpi radialis muscle activity. For torque force exertions, nitrile gloves have better overall performance.

References

Basmajian, J.V. and De Luca, C.J. 1978, *Muscles Alive: Their Functions Revealed by Electromyography*, Fourth Edition, (Williams & Wilkins, Baltimore)

Bishu, R.R. and Klute, G. 1995, The effects of extra vehicular activity (EVA) gloves on human performance, *International Journal of Industrial Ergonomics*, **16**, 165–174

Chang, C.H., Wang, M.J.J. and Lin, S.C. 1999, Evaluating the effects of wearing gloves and wrist support on hand-arm response while operating an in-line pneumatic screwdriver, *International Journal of Industrial Ergonomics*, **24**, 473–481

Fleming, S.L., Jansen, C.W. and Hasson, S.M. 1997, Effect of work glove and type of muscle action on grip fatigue, *Ergonomics*, **40**, 601–612

Kendall, F.P. McCreary, E.K. and Provance, P.G. 1993, *Muscles, testing and function: with posture and pain*, (Lippincott Williams and Wilkins, Philadelphia)

Larivière, C., Plamondon, A., Lara, J., Tellier, C., Boutin, J. and Dagenais, A. 2004, Biomechanical assessment of gloves. A study of the sensitivity and reliability of electromyographic parameters used to measure the activation and fatigue of different forearm muscles, *International Journal of Industrial Ergonomics*, **34**, 101–116

Mital, A., Kuo, T. and Faard, F. 1994, A quantitative evaluation of gloves used with non-powered hand tools in routine maintenance tasks, *Ergonomics*, **37**, 333–343

O'Sullivan, L.W. and Gallwey, T.J. 2005, Forearm torque strengths and discomfort profiles in pronation and supination, *Ergonomics*, **48**, 703–721

Shih, Y.C. and Wang, M.J.J. 1997, The influence of gloves during maximum volitional torque exertion of supination, *Ergonomics*, **40**, 465–475

LONG-TERM STUDY OF MUSCULOSKELETAL PROBLEMS DURING A REORGANISATION IN ROYAL MAIL

Corinne Parsons

Royal Mail Engineering, Royal Mail York, York YO26 4ZZ

Royal Mail introduced new mail delivery patterns nationally in 2003/4, replacing the first and second deliveries with one longer single daily delivery. Musculoskeletal problems and stress related conditions were monitored from sick absence records over a 21 month period in 30 delivery offices (over 3000 employees) to ensure that longer spans did not adversely affect delivery personnel. The results showed a peak in sick absence due to musculoskeletal problems and stress related conditions around the time of the reorganisation in the majority, but not all offices and this returned below the levels prior to implementation before the end of the study. This picture was also reflected in the national sick absence rates for Royal Mail. The level of sick absence was affected by the type of delivery equipment (i.e. trolleys, cycles or foot), used in the offices.

Introduction

New mail delivery patterns were introduced nationally in 2003/4 replacing the first and second delivery with one longer single daily delivery. Instead of the delivery being planned to last 2.5 hours the deliveries were planned to last 3.5 hours. These changes were made to improve the efficiency in delivery, reducing costs to allow an increase in basic pay, and in most offices a change to a five day week from a six day week. It was anticipated that the change in delivery spans would result in a significant reduction in jobs, but in practice the headcount reduction was relatively low.

Prior to the introduction of the new delivery spans a study was carried out at pilot sites which monitored heart rate and subjective measures of fatigue. A second study was commissioned after the introduction of the new delivery span to the monitor health in the longer term. This study was based on the identification of musculoskeletal problems from sick absence data and ill health retirement to ensure that longer spans did not adversely affect delivery personnel in terms of musculoskeletal problems. Recent studies by Devereux (2003) have shown a link between the reporting of musculoskeletal problems and psychosocial factors and so sick absence due to stress was also monitored.

Method

Sick data from musculoskeletal and stress related mental disorders and medical retirements were extracted from the Human Resources database for 30 offices

over a period of 21 months which gave data 5 months before the reorganisation and 12 months afterwards for all of the offices. The offices were selected to give a geographical spread and to represent a range in terms of size of office, size of conurbation, terrain, housing types and equipment used. Several offices were included because their sick absence rate was known to be high. The sample included 3111 members of staff in post in the study group at the beginning of the study and 3070 at the end, falling to a lowest point of 2982 in the middle of the study.

The offices included in the study were not aware of their inclusion until the end of the study when offices were visited, or contacted to discuss changes that had been made in the offices that may have affected sick absence.

Results

Sick absence

The sick absence records on the Human Resources database include a description of the cause of absence, the length of the absence, and the office that the data came from. The descriptions of the cause of absence varied in detail from those giving a specific cause e.g. sciatica, osteoarthritis in knees, to those that were more vague e.g. pain in back, hand operation, which may have been related to musculoskeletal problems or may have had an entirely different cause.

The sick absence data was analysed so that all records that could have been due to work related musculoskeletal or stress conditions were included in the data set for further analysis. Problems described as resulting from none work related causes were excluded e.g football injuries and bereavement.

The sick absence data was analysed in relation to the reorganisation and the results showed that overall there was an increase in sick absence due to musculoskeletal and mental disorders just before and after the reorganisation, which did not return to the pre-implementation level until around 7 months after the reorganisation had been implemented. By 10 months after was below the level in the months prior to implementation.

As can be seen from the Figure 1 all categories of musculoskeletal and stress related sick absence appeared to increase after the reorganisation except for operations, which appeared to reduce. This data was compared to the overall national picture for sick absence for 2003–2005 and was found to show the same general pattern around the reorganisation. This did not show the same trend as the data for the previous years and was not therefore part of the normal seasonal pattern of absence.

Length of sick absence

The sick absence data was analysed by length and the results showed that 60% of staff took no sick absence due to mental or musculoskeletal disorders during this period. 27% of those who did take sick absence lost 5 or less days and 33% of those taking sick absence lost more than 30 days. Individuals absent for more than 30 days, made the greatest contribution towards the overall absence rate with 82% of the days lost being due to those taking more than 30 days absence.

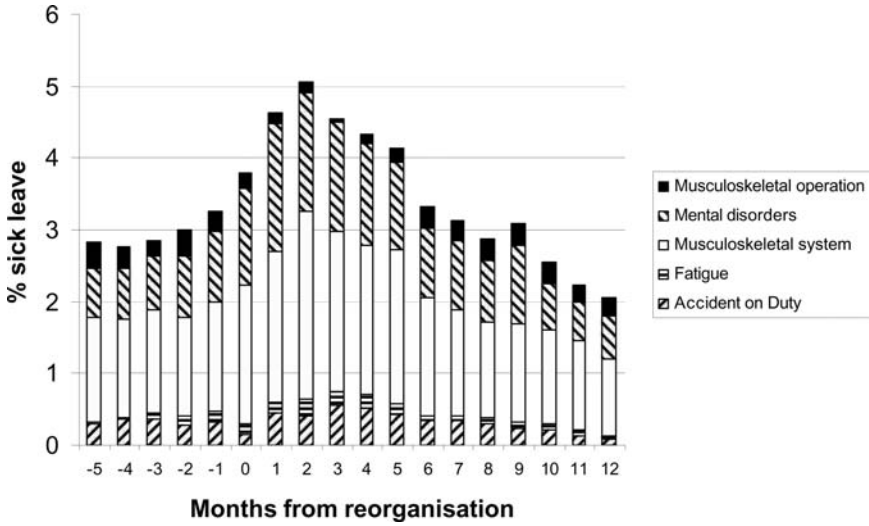


Figure 1. Sick absence in relation to delivery reorganisation.

Further analysis of the results showed that approximately twice as many days were lost due to musculoskeletal disorders than for mental disorders, Four times as many people took absence due to musculoskeletal disorders than for mental disorders but the absences due to musculoskeletal disorders tended to be shorter with 90% taking less than 5 days and musculoskeletal disorders only contributing to 63% of the absences over 30 days.

Sick absence and delivery equipment

The sick absence for the first 18 months of the study was analysed according to the main delivery method in the office. Offices were categorised by the predominant equipment type i.e. pouch on foot, cycle or trolley offices, and where a range of delivery methods were available the office was designated as “mixed” for the analysis.

As can be seen from Figure 2 sick absence was lower in offices where trolleys were used than other equipment. In particular there was less absence due to back, shoulder, foot and knee problems. The differences were significant at the P 0.01 using a paired T-test for trolley offices compared foot offices and P 0.1 for trolley offices compared to cycle offices. Sick absence in offices where cycles were used was not significantly different to foot offices.

Sick absence and terrain

Offices were categorised according to terrain into Steep, Undulating, Flat and also Tenement, which included offices with a high proportion of multi-storey dwellings. The results showed that sick absence for tenement offices was higher than for other offices, this was notable in knees, shoulder and back. These differences were not

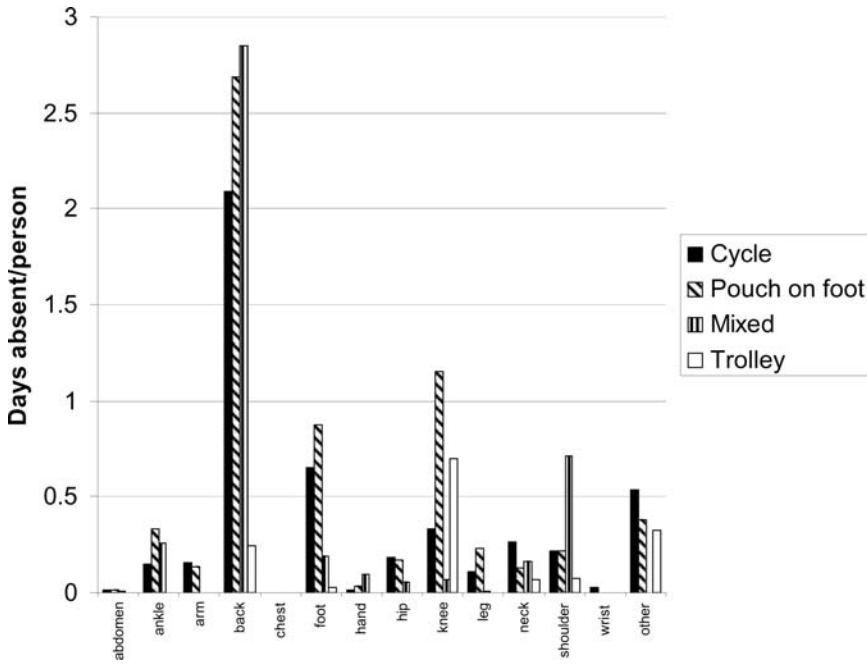


Figure 2. Sick absence in relation to equipment type.

quite statistical significant using a Paired T-test (P 0.11 flat: tenements), but as there was a relatively small number of offices in this category (approx 180 staff) further investigation is required to identify whether there is a higher rate of musculoskeletal disorders in offices with tenements/flats.

Sick absence for offices with deliveries on steep terrain was slightly higher than for flat or undulating terrain but this was not statistically significant. Tenements and steep terrain are less suited to the use of trolleys and cycles to manage weight on delivery and this may be a factor affecting the difference in absence rates.

Ill health severances and medical retirements

During the 03/04 financial year there were 26 retirements/severances due to ill health in the study offices, (prior to the reorganisation), compared to 65 retirements/severances during the 04/05 financial year, (post reorganisation). Details of the ill health problems leading to the retirement/severance were not available to this study and so it is not known whether these were related to musculoskeletal, mental or to entirely different problems.

Discussion

The sick absence levels increased around the time of the reorganisation in the majority of offices and these began to fall 3–4 months after the reorganisations had

been implemented. The sick absence trends were discussed with the managers from approximately 2/3rds of the offices in the study group, in some cases the staff gave views as well.

The delivery revisions resulted in major changes to the routes and the organisation in a number of offices. In offices where the reorganisation when smoothly and staff perceived benefits from the changes there was little difference in the sick absence rates. However, in some cases the reorganisation did not go well, a backlog of work developed and staff became disillusioned and stress levels were high. Ongoing reviews of walks to sort out individual problems and taking pressure off by making "minor tweaks" to duties resolved the problems in some offices.

Other managers believed that the peak in sick absence was due to an initial shock in offices where staff had relatively easy walks prior to the reorganisation. This resolved when staff accepted that the new deliveries were in to stay, and either became used to the walks or chose to leave the business. Medical retirements and resignations of staff, not capable of, or not willing to carry out the new deliveries, will have undoubtedly had an impact on absence rates.

Other factors such as the changes to the absence procedures and the increased importance placed on following them, coupled with incentives such as the "In it to Win it" scheme were also considered to have helped to bring sick absence rates down.

Conclusions

Sick absence records and ill health retirements/severances have been monitored for 30 offices over a period of 21 months between October 2003 and June 2005. This allowed a direct comparison of data for 5 months before and after the reorganisation. The results showed a peak in sick absence due to musculoskeletal and stress related mental disorders around the reorganisation in the majority, but not all offices, which returned below pre-reorganisation levels within a year. This picture was also reflected in the national sick absence rates for Royal Mail.

There was a lower sick absence rate for trolley deliveries compared to foot and cycle deliveries, this was found to be statistically significant. The areas of difference were back, neck, shoulders and knees Sick absence rates appeared to be higher in areas with tenements and steep terrain than in flat or undulating areas, but the difference was not statistically significant.

The greatest contribution to days lost was long term sick absence and an increase in medical retirements/severances post-reorganisation will have had an influence on reduction in sick absence following the post reorganisation peak. Managing weight as well as managing absence processes also appeared to be key to reducing sick absence.

References

Deveraux, J. 2003, Work-related stress as a risk factor for WMSDs, *Contemporary Ergonomics*. Taylor & Francis.

TEACHING ABOUT MUSCULOSKELETAL DISORDERS – ARE WE BARKING UP THE WRONG TREE?

Andrew Baird

University of Derby, Derby, DE22 1GB

Musculoskeletal disorders remain a major health problem. If ergonomics is to play its part in addressing this problem effectively, ergonomists must recognise the true nature of these conditions and the impact of the minority of individuals who develop chronic problems. This requires a move away from simplistic models based on physical risk factors to a broader understanding of pain and health behaviours that will facilitate the application of bio-psychosocial models.

Introduction

Musculoskeletal disorders (MSDs) remain a major health problem across the globe. Dr Bruntland, Director General of the World Health Organisation, has suggested that back pain has reached 'epidemic' proportions as we move into the new millennium, with 80% of people experiencing the condition at some point in their lives. Within Europe, as much as 40% of workers compensation can be attributed to musculoskeletal problems and costs have been calculated to be as much as 1.6% of a country's gross domestic product. To put this in perspective in relation to other health problems, MSDs have a societal cost greater than that of cancer, coronary artery disease and AIDS combined (Thomsen et al, 2002).

In launching this year's 'Lighten the Load' campaign, Vladimír Spidla (EU Commissioner for Employment, Social Affairs and Equal Opportunities) said:

'Tackling MSDs is a priority for the EU if we are to create more and better jobs in Europe. . . We can increase productivity and therefore prosperity in the EU if we manage to improve the situation of days lost to MSDs'.

The primary 'weapon' in the fight against MSDs is typically seen to be ergonomics – in the UK, the Health & Safety Executive (HSE) state in their guidance to Manual Handling legislation, that 'modern medical and scientific knowledge stresses the importance of an ergonomic approach to remove or reduce the risk of manual handling injury'. This paper discusses how best to ensure that ergonomists fully understand the 'medical and scientific knowledge', particularly in relation to the key issue of days lost due to these problems.

The realities of musculoskeletal disorders

For many, the 'first thought' in relation to MSDs would be a specific condition such as a disc prolapse or tenosynovitis – reasonably well understood conditions with

relatively clear pathology and aetiology. In reality however, such conditions are relatively rare, representing only a small fraction of MSDs. It has been recognised for some time that whilst most people will experience back pain at some point in their lives only about 15% can be given a positive diagnosis (Deyo & Weinstein, 2001). Increasingly, it appears that there is a similar picture for upper limb problems, with only about 20% fitting a specific diagnosis (Reneman et al, 2005).

Whilst individuals often want a 'scan' which they hope will show where they are 'damaged', such investigations are of little value in most cases (Waddell, 2004). Indeed, most musculoskeletal problems are essentially self-limiting and individuals will recover almost regardless of what treatment is offered. Even when tissue damage is identified, we are often left facing a 'pathology paradox' – when looking at neck pain, Peterson et al, (2003) could identify no significant differences in either pain severity or disability between those with and those without cervical degeneration. Even in osteoarthritis, a condition that can produce quite profound joint deterioration, a review by Kidd, (2006) describes a very poor correlation between deterioration and symptoms – indeed 40% of those with severe radiographic change are symptom free. It is clear therefore, that pain and tissue damage are not the same thing and the presence of pain should not be taken as indication of tissue damage. Failure of the 'injury model' led to profound changes in MSD management in clinical and rehabilitation settings. Biomedical models have been dropped in favour of more encompassing bio-psychosocial approaches and increasingly researchers are investigating the nature of pain itself rather than simply treating it as correlate of tissue damage.

Whilst MSDs are clearly complex, the majority of individuals do recover without a significant, lasting effect on their physical function or quality of life. Unfortunately, a minority fail to recover. Whilst this chronic, care-seeking and/or work-absent group typically represents less than 10% of individuals who experience MSDs, this minority accounts for the vast majority (as much as 90%) of the cost (Verbunt et al, 2003). Typically these individuals exhibit no more objective physical abnormalities than those who go on to recover and it is often difficult to explain the level of their ongoing symptoms. Biomedical models would suggest that these individuals are either mad or malingering, but their symptoms can be understood by recognising that they don't have a musculoskeletal problem so much as a pain problem – they suffer from chronic pain.

For this group, the best predictors of disability (i.e. a failure to recover) are not physical signs, but rather mental health issues, such as depression and the so-called 'psychosocial factors' – including both workplace factors such as monotonous work and lack of support (Macfarlane et al, 2000) and individual factors such as fear avoidance and catastrophising (Woby et al, 2004). In fact, fear of pain is more disabling than pain itself (Crombez, 1999).

In the UK, HSE figures show that whilst there was a reduction in the number of individuals affected by MSDs between 1995 and 2002, over the same time period the number of days lost increased. This was due to an increase of around 50% in the average number of days lost per case. This increase was not universal however, but reflected the impact of 'outliers' in the form of chronic cases. It follows therefore, that to deal with the recognised issue of 'days lost' ergonomists must address the

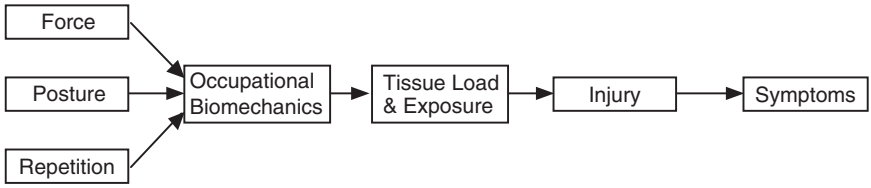


Figure 1. Traditional view of musculoskeletal disorders.

issue of chronicity. To do so they must address psychosocial factors including pain beliefs and behaviours.

Whilst it is common for those involved in MSDs to talk in terms suggestive of acute injuries, in many cases ergonomists do not see individuals in the early, acute phase of their condition. In reality, ergonomists are often called in only once an individual's problems have failed to resolve over a considerable time and/or the individual has had extended sickness absence (often months). Alternatively, an ergonomist may be called-in to look at a process once a pattern of symptom reporting has emerged – again ensuring that by the time the ergonomist looks at the problem, the reported cases are effectively chronic. It is fair to say therefore that not only should the avoidance of chronicity be a fundamental goal of ergonomics interventions, but it is chronic cases that ergonomists frequently have to face.

Traditional approaches

The Ergonomics approach to MSDs has traditionally focussed very much on physical risk factors. As such it reflected a biomedical view and sat within a 'safety' mindset, focused on primary prevention. The paper by Ulin & Armstrong, (1992) is indicative of this, outlining a strategy covering a range of physical factors, but with no mention of psychosocial factors, issues of chronicity or the need to manage those with MSDs. Such an approach was also evident within workplace guidance of the time such as HSG60 (1990) '*Work related upper limb disorders – a guide to prevention*' from HSE. This approach is summarised in Figure 1.

Whilst such a model could be said to have some merit – force, posture and repetition (particularly in combination) can indeed impact upon MSDs – it falls some way short of explaining the presentation of MSDs that are seen in the workplace and can in fact be rather misleading. In 1992, Huskisson, suggested that 'for ergonomists, RSI was the answer to a maiden's prayer. They multiplied and measured, taking an inch off the height of the desk and adding a centimetre to the width of the backrest. But did it help? Most commentators have concluded that it didn't'. Whilst something of a slight on the profession of ergonomics, it may be justified as reviews of the Australian 'RSI epidemic' failed to support the idea of a predominantly physical cause for the reported problems (Caple & Betts, 1991). Lucire (2003) argued that those seeking to manage the problem served only to escalate it and a review by Szabo & King (2000), whilst recognizing the positive impact of ergonomics on comfort, suggested that the 'many proponents of ergonomics assert that the elimination of certain risk factors related to force, repetition, and posture can

prevent or even cure work-related musculoskeletal disorders of the upper extremity. However, there is little scientific support for this position'.

Perhaps one of the biggest problems however, is that a failure to understand the complex nature of pain has led individuals to use the model shown in Figure 1 'backwards' i.e. if symptoms are reported, it is assumed there must be an 'injury' with associated tissue damage and therefore the tissue's capacity must have been exceeded by some biomechanical factor. It is also then assumed that the scale of the reported discomfort is indicative of the scale of the biomechanical insult. We have seen this recently in relation to back pain amongst school children – when rising levels of back pain were recognised, there was a clamour to identify physical causal factors such as the weights being carried, the design of school bags and the furniture provided in schools. In reality, research suggests that other factors such as emotional problems and the presence of other somatic complaints are more important (Jones et al, 2003).

Whilst psychosocial factors are increasingly recognised, they still tend to be seen as relatively 'peripheral' in ergonomics and health & safety settings – an additional issue to be considered rather than a fundamental part of an encompassing bio-psychosocial model. Whilst the term 'psychosocial' is becoming commonplace, very few in the workplace really understand what it means (Baird et al, 2006). Even in clinical settings it is not unusual to find that individuals 'revert to type' and adopt biomedical approaches when faced with specific individuals and symptoms, as opposed to talking in general terms (Thunberg et al, 2001). This suggests that despite the drive to adopt bio-psychosocial models, few practitioners can truly operationalise them.

Teaching about MSDs at Derby

At Derby, the module 'Musculoskeletal Disorders & Ergonomics' has been re-designed in an attempt to both better reflect the research evidence and relate to the realities of the problems students will face once they graduate and begin to practice. It addresses recognised knowledge gaps and its design reflects a number of key criteria:

- A primary aim of MSD management systems must be the avoidance of chronicity – this directly addresses the issue of 'lost days', but also addresses key quality of life indicators (particularly in relation to the 'label' of disability).
- Those working with MSDs must have an understanding of pain – without an understanding of pain it is not possible to fully understand the impact of psychosocial factors and by extension it is not possible to operationalise bio-psychosocial models.
- MSD management systems cannot rely solely upon primary prevention – whilst workplace factors are important, other individual and non-work factors will have as big an impact and cannot be controlled with traditional risk-based approaches.

The module has eight units: Introduction; Pain; Basic Anatomy & Physiology; Biomechanics, Posture & Movement; Common Disorders & Associated Risk

Factors; Ergonomics Interventions; Treatment & Rehabilitation; and MSDs & the Law. A further module 'Human Behaviour & Health' provided later in the programme, covers health beliefs and behaviours, supporting the MSDs module in relation to return-to-work and rehabilitation.

Great emphasis is placed on the understanding of pain. After the scene-setting introduction, the next unit is pain – this provides the 'context' within which later topics can be covered effectively. Pain is discussed from biomechanical, neurological, psychological and socio-cultural perspectives, all of which impact upon MSDs.

The module is assessed using a scenario based within the 'Virtual Workplace' discussed elsewhere within these proceedings (see 'Teaching Ergonomics at a Distance: A Virtual Workplace.' – Richardson & Baird). In addition to an outline description and images of the workplace and workstations, video 'interviews' are provided for an individual with musculoskeletal symptoms and their supervisor. This mimics a real-world scenario with contradictions both within and between each of the accounts. The scripts have been designed to be realistic whilst also alluding to a range of issues – emphasis given to potential issues by the actors however, is not necessarily a true reflection of the 'scale' of the problems and some issues are referenced subtly. Perhaps most importantly, the videos allow for the introduction of 'psychosocial factors' (both organisational and individual) in a more realistic way than could be achieved with a traditional paper-based scenario.

Response of students

The revised module has run for the first time in Autumn Semester 2007. The materials and related discussions have been described as challenging, interesting and enjoyable. Particular praise has been given to the materials on pain – with students engaging in wide-ranging debates and providing many case study examples. The discussion forum related to this module has been very busy in terms of individual posts, numbers of threads (issues raised) and views (indicating that even those not contributing were following the debates).

Conclusions

The MSD module at Derby is designed to better reflect both the research evidence and the needs of industry in terms of chronic conditions resulting in 'lost days'. It does this by ensuring a sound understanding of pain and the psychosocial factors associated with a transition from acute to chronic problems. It utilises novel technologies within its assessment to ensure that psychosocial factors can be addressed effectively. It has also proved popular with students.

References

Baird, A., Haslam, R. & Haslam, C. 2006, 'Life, lifting and luck' – a qualitative analysis of attitudes towards back pain at work . In: *Proceedings of the 16th World Congress of the International Ergonomics Association*. (Elsevier Science)

Caple, D.C. & Betts, N.J. 1991, RSI – It's Rise and Fall in Telecom Australia 1981–1990. In *Proc, 11th Congress of the International Ergonomics Association*. (Taylor & Francis)

Crombez, G., Vlaeyen, J.W., Heuts, P.H., & Lysens, R. 1999, Pain-related fear is more disabling than pain itself. *Pain, 80*(1–2), 329–339.

Deyo, R.A. & Weinstein, J.N. 2001, Low back pain. *N Engl J Med, 344*(5), 363–370.

Huskisson, E. 1992, *RSI – The Keyboard Disease*. (Charterhouse Conf & Comms)

Jones, G.T., Watson, K.D. & Silman, A.J. 2003, Predictors of Low Back Pain in British Schoolchildren: A Population-Based Prospective Cohort Study. *Pediatrics*, 111, 4, 822–

Kidd, B.L. 2006, Osteoarthritis and joint pain. *Pain* 123, 6–9.

Lucire, Y. 2003, *Constructing RSI – belief and desire*. (UNSW Press)

Macfarlane, G., Hunt, I. & Silman, A. 2000, Role of mechanical and psychosocial factors in the onset of forearm pain: prospective population based study. *BMJ*; 321; 676–681.

Peterson, C.K., Bolton, J.E., Wood, A.R. & Humphreys, B. 2003, A Cross-Sectional Study Correlating Degeneration of the Cervical Spine With Disability & Pain. *Spine*. 28(2):129–133.

Reneman, M.F. 2005, Basis for an FCE Methodology for Patients With Work-Related Upper Limb Disorders. Journal of Occupationa Rehabilitation 353–364.

Szabo, R. & King, K. 2000, Repetitive Stress Injury: Diagnosis or Self-Fulfilling Prophecy? *The Journal of Bone & Joint Surgery*. 82:1314–1322.

Thomsen, A. Sorenson, J. Sjorgen, P. & Eriksen, J. 2002, Chronic non-malignant pain patients and health economic consequences. *European Journal of Pain*, 6, 341–352.

Thunberg, K., Carlsson, S. & Hallberg, L. 2001, Health care professionals understanding of chronic pain: a grounded theory study. *Scand J Caring Sci*, 15, 99–105.

Ulin, S. & Armstrong, T. 1992, A Strategy for Evaluating Occupational Risk Factors of Musculoskeletal Disorders. *Journal of Occupational Rehabilitation, 2*, 39–50.

Verbunt, J.A. 2003, Disuse and deconditioning in chronic low back pain: concepts and hypotheses on contributing mechanisms. *Eur J Pain, 7*(1), 9–21.

Waddell G. 2004, *The Back Pain Revolution*. 2nd Edition, (Churchill Livingstone) p2.

Woby, S.R., Watson, P.J., Roach, N.K. & Urmston, M. 2004, Adjustment to chronic low back pain – the relative influence of fear-avoidance beliefs, catastrophizing, and appraisals of control. *Behav Res Ther, 42*(7), 761–774.

EVALUATION AND INTRODUCTION OF ELECTRICALLY POWERED EQUIPMENT FOR MAIL DELIVERY

Adrian Woodcock & Corinne Parsons

Royal Mail Engineering, Dorcan, Swindon SN3 5XX

This paper describes the role the Royal Mail ergonomics team played in the process of purchasing new delivery equipment to aid the manual task of delivering mail. The ergonomics team had a major input into the process starting with the identification of the need for new delivery equipment, concept testing, operational trials, through to writing specifications for tender and selection of the successful supplier. This paper focuses on the electric trolley but the same process was applied to a range of equipment. The testing used an iterative cycle of 'lab' based assessments, modification or redesign and then testing was carried out in a live environment with delivery staff. Feedback was gathered by; observation, group discussions, structured interviews and discomfort body maps, objective measures such as heart rate were also measured. The final stages of the work have included working closely with the procurement team to help create specifications and choosing between suppliers.

Introduction

Royal Mail is in the process of modernising. Delivery staff typically deliver around 70 kg of mail over a 3.5 hour delivery route. Most of the staff carry a delivery pouch which holds a maximum of 16 kg at a time and collect replenishments of mail as they work their way around their route. New trolleys have been introduced over the past 10 years and cycles are also commonly used. A number of internal studies have shown reduced fatigue levels for deliveries carried out using a trolley or cycle rather than carrying a pouch but user preferences and inadequate security of current equipment has prevented many staff from changing their delivery method. The aim of the current delivery strategy is to provide suitable equipment for all walks so that delivery staff can take all of their mail with them and not carry a pouch over the shoulder.

This paper describes the role that the Royal Mail ergonomics team plays in the design, development, assessment and decision making when introducing new delivery equipment and technology. This paper focuses on electric trolleys but similar processes have been applied to other equipment such as electrically assisted bicycles with trailers and tricycles.

The ergonomics team has been involved in a number of projects which have explored the benefits or limitations of different delivery solutions and has developed a method of categorising deliveries according to terrain, housing type, and security

risk level (Parsons *et al*, 2001). This has assisted in identifying the types of deliveries for which the current manual equipment is not suitable and a range of equipment has been tested for these deliveries. All of this equipment has power assistance to allow delivery staff to carry more mail, this allows them to be self sufficient during their round. The aim of the electric trolley is to deliver mail to areas within walking distance of the delivery office on terrain that is too steep for other trolleys to be used.

The following stages of development of the equipment have involved an ergonomics input:-

- Benchmarking and identifying possible delivery equipment
- Concept testing
- Design development
- Operational trials
- Specification for tender
- Testing of equipment for procurement

Method

Benchmarking and identifying possible delivery equipment

Benchmarking studies were carried out to identify what equipment was used by other postal administrations or organisations that may have been suitable for delivery of mail. Any equipment that was thought to have potential for use in Royal Mail was obtained for concept testing.

Concept testing

As with all stages of the testing the concept testing was carried out in two phases; lab based ergonomics testing on a simulated delivery route, and for successful equipment, proof of concept live user testing on delivery routes. A number of trolleys were tested in the lab based environment but only one progressed to live user trials. This trolley is already used by the German Post Office. The initial testing showed that there were aspects of the trolley that would need to be changed to be suitable for Royal Mail deliveries, in particular the compartment for carrying the mail. However, it was considered adequate to test the basic functionality and suitability of the trolley. The ergonomics team were responsible for developing Safe Systems of Work for the trolley, training the users and assessing the suitability of the walks on which the trolleys were used. This was done by accompanying the postman/woman the first time the trolley was used on their delivery.

The initial proof of concept testing was carried at 8 offices, on 10 deliveries, for 2–4 weeks each by 10 subjects, representing 13% ile to 96% ile users. The trolley was evaluated by observation, structured interviews with users, a user workshop, discomfort maps and heart rate comparison with the use of the trolley compared to normal delivery method. The feedback from the majority of users was very positive, particularly for those with steep delivery routes. The study showed a reduction in

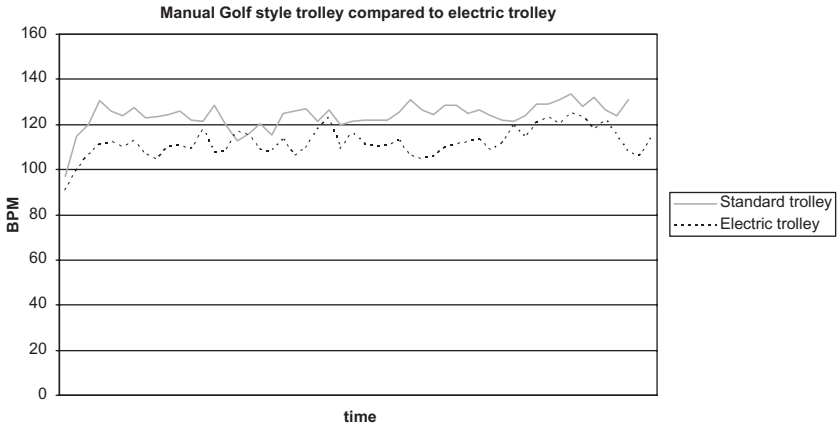


Figure 1. Heart rate comparison between a manual and electrically powered trolley for one participant.

effort levels for the delivery staff with average heart rates being 9 beats/minute lower for the powered trolley than for their normal delivery method. This is illustrated in Figure 1. A number of detailed design changes were identified in addition to the more major changes to the load compartment.

Design development

The data from the questionnaires and feedback gathered by the ergonomists was used to help design concepts for the load compartment of the trolley. These were presented to participants who took part in the trial in short 'one on one' sessions. The concepts were presented using basic models and drawings to give the participants an opportunity to have an input to the development of the design concepts. They were asked to rate the features of each concept. Unsurprisingly simple features and designs that looked familiar were highly rated.

Further development involved a workshop to finalise the shape and size of the mail carrying compartment. This was held with a group of ergonomists, designers and operational managers. An initial design exercise resulted in similar design concepts and themes to those already developed with users. The workshop explored the dimensions of the load compartment using a series of cardboard mock-ups varying in height, width and profile which were mounted on the chassis of the trolley. Loading, unloading and steering the trolley were tested with the mock-ups and further modifications were made until the final dimensions were agreed. The ergonomics team then worked with the designers to produce working prototypes for operational testing from the dimensions and features that had been agreed.

Other design changes were fed back to the manufacturer to improve the design prior to operational testing. A second phase of user testing was carried out at 3 offices with 5 units to obtain more feedback and to verify that the changes made had been successful.

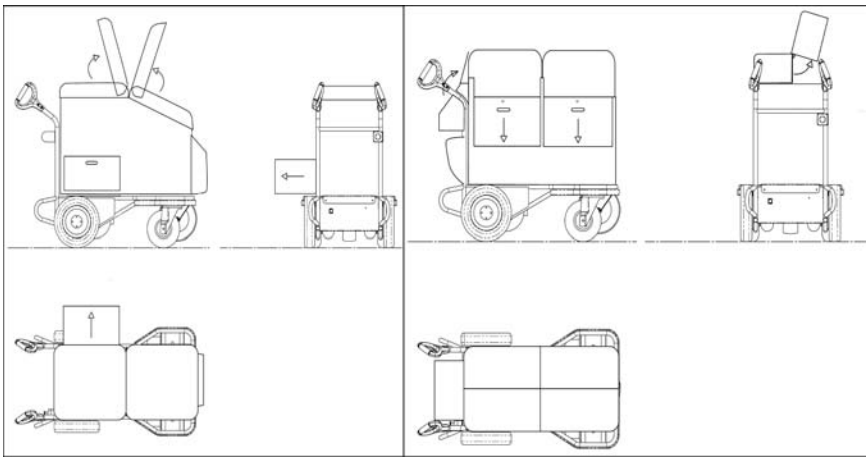


Figure 2. Examples of drawings shown to delivery staff to rate.

Figure 3. Full size card mock up created in design workshop using working electric trolley chassis.

Operational trials

Operational trials were carried out with a range of equipment to identify business benefits from using the new equipment. The trials involved testing of the electric trolleys in addition to manual trolleys, cycles with trailers, electrically assisted cycles, scooters, electrically assisted tricycles and motor vehicles. There were 12 trial sites altogether which represented the full range of terrain, housing and security classifications. The ergonomics input to these trials included assisting in the selection of the trial offices; lab based assessments of equipment to confirm that any changes identified as necessary in the proof of concept testing had been made; working with training teams to develop and deliver training; developing a risk assessment procedure for delivery routes to ensure that they were suitable for

the equipment to be tested and carrying out a user evaluation of the equipment during the trials. In addition to the user evaluation of the delivery methods feedback was also obtained on the training and risk assessment processes.

The electric trolleys were tested at 4 of the trial sites with 16 users representing 50th–95th percentiles. The user feedback was generally very positive although some detailed design changes and improvements to the training were identified.

Specification for tender

Specifications for tender for the various equipment types were developed from findings of the operational trials with operational, engineering and ergonomics leads working together. These specifications were then used for competitive tender to supply the equipment, and to measure the suitability of the equipment produced.

Testing of equipment for procurement

As with other phases of testing, lab based evaluation by ergonomists was carried out over a simulated delivery route to assess suitability prior to use on live deliveries. Equipment considered suitable was then put into a live operational environment for final evaluation, using delivery routes selected as appropriate for each equipment type selected. Evaluation methods similar to those used previously were employed.

Discussion

A lot was learned during the trials including what features delivery equipment has to have to succeed to why some equipment is unpopular. Key features of equipment that were important to users were reliability and simplicity. User feedback and heart rate measurement showed that the introduction of powered equipment to walks reduced effort levels required to complete the delivery.

The trials also showed how difficult some individuals find it to become accustomed to new ways of working and are sometimes reluctant to change their method of delivery to gain full benefits from the equipment provided. Several participants had used manual trolleys for many years in relatively steep areas and were used to pushing hard to initiate movement of the trolley. The electric trolley needed only a relatively light press on the handles to start but as there was a time delay between pressing the handles and the trolley starting, several users exerted high forces on the trolley and even after repeated training found it hard to change this practice. Similar problems occurred with other equipment, in particular tricycles which some individuals found very difficult to get used to if they had previously only ridden bicycles. This led to the development of more a detailed training course which highlighted potential problems to the trainer and longer familiarisation period was allowed for potential users.

Simply collecting the data may have not provided the full picture, achieving a good understanding of the users helped design the equipment and the training requirements for it. This is where the ergonomics team added value to the project.

Conclusions

The ergonomics team had a major role to play in every stage of the introduction of the new technology and equipment. This included; benchmarking suitable equipment, initial testing and evaluation, design development; supporting operational trials by the production of training, working methods, risk assessments and in conducting user evaluation of the equipment; and supporting procurement teams in the development of specifications for tender and the final selection of equipment. The outcome should be more efficient deliveries that are easier for delivery staff to complete.

References

Parsons, C. Langford, J. and Dillon, J., Equipment Selection Process for Postal Delivery Workers, 2001, *PREMUS*, 55.

INITIAL TESTING OF AN INSPECTORS' TOOL FOR THE ASSESSMENT OF REPETITIVE TASKS (ART)

Jeremy Ferreira, Matthew Birtles & Leanne Stanley

Health and Safety Laboratory, Harpur Hill, Buxton, Derbyshire, SK17 9JN, UK

This paper presents user trial results for an inspectors' tool for screening repetitive tasks of the upper limbs for work-related musculoskeletal disorder risks. The main aim of the trials was to inform further improvements to the design of the tool. Data analysis included the extent of scoring agreement between users and specialists, qualitative feedback, and usability measures such as the System Usability Scale. During field trials, 67% of inspector judgements agreed with the specialist's judgement for risk level while 95% of judgements either agreed or were within one category difference of the specialist's judgement. Further results demonstrate sufficient usability and reasonable reliability to allow a further pilot of the tool amongst trained inspectors.

Introduction

In January 2007, the Health and Safety Laboratory (HSL) began work with the Health and Safety Executive (HSE) to develop a tool that health and safety inspectors could use to screen repetitive tasks of the upper limbs for work-related musculoskeletal disorder (WRMSD) risks. The initial concept of the tool, entitled 'Assessment of Repetitive Tasks' (ART), is described elsewhere in this publication (Ferreira *et al*., 2008).

The process used to develop ART was based on previous work by Lee (2005), who outlined an iterative design process for evaluating and improving the usability of paper-based risk assessment tools for musculoskeletal disorders. This process involved a series of non-empirical methodologies (cognitive walkthrough and property checklist) and empirical methodologies (user trials and usability questionnaires) whereby the results were interpreted and used to inform subsequent prototypes of the tool. Peer-review exercises were also carried out amongst ergonomists in HSE/HSL and a small group of external ergonomics practitioners. The purpose of this paper is to present results from the user trials, which were carried out from July to September 2007, and were used to refine the later prototypes of ART.

Methods

The methodology involved two types of user trials: simulated trials, during which inspectors were instructed to assess repetitive tasks shown on video, and subsequent field-based trials.

Simulated user trials

Two events were arranged, with the primary aims of testing the usability of ART and informing further improvements to its design. At each event, information on general upper limb disorder (ULD) risk factors were first presented to participants. Specific training was then provided on how to use ART. This involved an overview of the tool, a demonstration of how ART is applied to assess a repetitive task, and the opportunity for participants to assess a further two tasks themselves, ask questions and receive feedback on their performance. Each participant was then instructed to use ART to assess a further three tasks without assistance or feedback and the score sheets were collected for analysis. The three tasks shown on video and used in the testing of ART involved:

- Picking a piece of chicken and cutting it with a knife;
- Picking product from a conveyor and placing it into a box; and
- Assembling a box, picking a stack of envelopes, tapping the envelop stack on the table and placing them into the box.

Following the assessments, participants completed a System Usability Scale (SUS; Brooke, 1996), a ten-item Likert scale used to give an overall measure of usability and user satisfaction. Qualitative feedback was encouraged with a bespoke usability questionnaire and through further group discussion.

Field based trials

Following the simulated user trials, field-based trials were arranged, during which a visiting inspector and a specialist each used ART to assess a selection of repetitive tasks. Inspectors and specialists were instructed to assess the same repetitive tasks simultaneously, yet to carry out the assessment without assistance from the other person. All inspectors who participated in the field trials had attended the simulated user trials within the previous two months, although most typically for just one or two days prior to the field visits. Three trained HSE/HSL ergonomists acted as specialists during the field trials. These trials allowed specialists to observe how inspectors used ART in the field. In addition, at the end of the visit, the score sheets were collected to allow a comparison of scores between the trained inspectors and trained specialists.

Results

Participants

Overall, 25 inspectors participated in the simulated trials. Of these, 12 participants (48%) were HSE inspectors and 13 participants (52%) were Local Authority (LA) inspectors based in Scotland, London and the south-east of England. Twenty-one inspectors (84%) had received previous training in the use of the Manual Handling Assessment Charts (MAC) and three inspectors were specialists in occupational health.

In total, six inspectors participated in further field trials where ART was used to assess 15 tasks at eight different premises. The types of tasks assessed included: packaging food and other items; manufacturing small parts; feeding linen into ironing machines; assembling boxes, sorting paper and catalogues, and decorating cakes.

Extent of agreement in scores

During the simulated trials, each of the 25 participants was asked to assess three tasks, and for each task this involved making a judgement on 12 risk factors. This provided 888 judgements in total to examine the extent of scoring agreement (one participant only completed an assessment of two of the tasks). Of these, 69% (610/888) of judgements agreed with the expert consensus for risk level while 97% (858/888) of judgements either agreed or scored within one category difference of the expert consensus. However, information on the workpace and duration risk factors was provided to participants, as it could not be obtained from video. These risk factors scored 100% full agreement. If this data is removed from the analysis, 740 judgements are available for comparison. In this case, 62% (462/740) of judgements agreed and 96% (708/740) of judgements either agreed or were within one category difference of the expert consensus. Table 1 provides a breakdown of the extent of scoring agreement for each factor.

For the field trials, the number of judgements available for comparison varied according to the risk factor. This was because the tool allows assessments to be made of the left arm and the right arm separately. This was done for seven of the

Table 1. Extent of scoring agreement for risk factors.

	Simulated trials		Field trials	
Risk factor	Full agreement (%)	Full agreement or 1 category difference (%)	Full agreement (%)	Full agreement or 1 category difference (%)
Frequency	80	100	66	97
Repetition	31	97	73	90
Force	66	100	69	93
Neck posture	64	96	39	94
Back posture	66	92	78	100
Shoulder / arm posture	64	92	62	93
Wrist posture	58	99	52	90
Hand / finger grip	41	71	48	86
Breaks	77	89	67	100
Workpace	*	*	67	100
Other factors	64	99	90	100
Duration	*	*	100	100
Overall	62	96	67	95

* Not included in the analysis as information was provided to participants in written instructions

fifteen tasks assessed in the field, and in such cases, the assessments for each arm were included in the comparison between the inspector and specialist. Thus, the 15 different tasks assessed provided a total of 29 comparisons for risk factors that were specific to each arm (e.g. frequency, repetition, force, shoulder posture, wrist posture, hand grip and other factors) and 18 comparisons for risk factors that were not specific to a single arm (e.g. neck posture, back posture, breaks, workpace and duration). For the field trials, 67% (196/293) of inspector judgements agreed with the specialist's judgement for risk level while 95% (277/293) of judgements either agreed or were within one category difference of the specialist's judgement.

Completeness

During the simulated trials, 98% of assessments (74/75) had a score recorded next to each risk factor. During one of the trials, participants were asked to make a record of the amount of time they required to complete each assessment. Of the 13 participants, 11 provided this information, which suggested that 90% of assessments were completed in 10 to 15 minutes. The average time to complete an assessment was 13 minutes and 42 seconds.

Usability

During the simulated trials, 84% of inspectors reported that ART improved or greatly improved their confidence when assessing ULD risks. Table 2 shows the descriptive statistics for the SUS data. As a measure of overall usability and satisfaction, the SUS measure was favourable throughout the trials (i.e. more than 50%). For both HSE and LA inspectors, the SUS measure steadily improved with subsequent user trials and iterative refinements. However, LA inspectors reported significantly lower SUS scores compared to HSE inspectors (Mann Whitney $U = 33.5$, $p < 0.05$).

Qualitative Feedback

Where possible, qualitative feedback on how to improve the wording and clarity of the tool and associated training material was used to refine ART. Themes that emerged from the feedback were that the tool was practical to use and, for a difficult topic, provided a helpful process to narrow down those aspects of the task that are relevant to WRMSD. Feedback also suggested the flowchart process

Table 2. SUS descriptive statistics.

SUS score (%)	HSE Inspectors	LA Inspectors	Overall
Mean	74.8	63.5	68.9
Median	77.5	65	67.5
Minimum	57.5	50	50
Maximum	87.5	80	87.5
Standard deviation	11.5	8.8	11.5

would be helpful to show employers how repetitive tasks can be assessed. However, although ART was thought to help identify *where* action was needed, it did not suggest *what* action was needed and several users felt that specific information on risk control would be required in addition to training on the use of the tool. A 19 cm × 12 cm document, similar in size to an inspector's notebook, was also requested.

Discussion

The assessment of repetitive tasks using observation-based methods is intrinsically challenging. Thus, it is important that the sensitivity of a tool is appropriately balanced against further requirements for usability and practicality at the work area. The simulated trials were primarily undertaken to inform further improvements to the usability of the ART. Initial trials in the field, although limited in the number of participants and tasks assessed, suggests that trained inspectors will be able to use ART with reasonable reliability. Some difficulties remain when deciding whether localised hand/wrist postures observed are awkward enough to merit scoring. This is an area where further clarification and attention to training will be required. Qualitative feedback, along with observations of inspectors using ART in the field, suggests that ART provides a useful process to screen repetitive tasks of the upper limbs for significant WRMSD risks and to brief employers on the findings of the visit.

The SUS revealed that, compared to HSE inspectors, LA inspectors reported more frequently that they felt they needed to learn a lot of things before they could get on and use ART quickly. LA inspectors also reported that they did not think they would use ART as frequently. It is possible that premises enforced by Local Authorities will undertake work that is less suited to the use of ART. This may help explain the difference in the SUS results between the two groups. Further improvements were carried out with the aim of making ART more usable to assess tasks with less rigid patterns of duration and opportunities for recovery.

Conclusion

The user trials suggest that ART provides a practical process that trained HSE and LA inspectors will be able to use to screen repetitive tasks of the upper limbs for WRMSD risks and share these findings with employers. Initial testing indicates that ART demonstrates sufficient usability and reasonable reliability to warrant a further pilot amongst trained inspectors.

Acknowledgements

The authors wish to acknowledge the support of colleagues in HSE policy and operational directorates who have assisted with and participated in the user trials.

References

Brooke, J. 1996, SUS: a 'quick and dirty' usability scale. In: P.W. Jordan, B. Thomas, B.A. Weedmeester, and I.L. McClelland (eds.), *Usability Evaluation in Industry*, (Taylor and Francis, London)

Ferreira, J., Gray, M., Stanley, L., Riley, D. and Birtles, M. 2008, Development of an inspectors' tool for assessment of repetitive tasks. In P. Bust (ed.), *Contemporary Ergonomics 2008*, (CRC Press, London)

Lee, D. 2005, *Selecting and applying a series of usability methodologies to MAC: an iterative design approach*. MSc Dissertation, (Robens Centre for Health Ergonomics, University of Surrey)

DEVELOPMENT OF AN INSPECTORS' ASSESSMENT TOOL FOR REPETITIVE TASKS OF THE UPPER LIMBS

Jeremy Ferreira[1], Mike Gray[2], Leanne Stanley[1] & David Riley[1]

[1]*Health and Safety Laboratory, Harpur Hill, Buxton, Derbyshire, SK17 9JN, UK*
[2]*Health and Safety Executive, Edgar Allen House, Sheffield, S10 2GW, UK*

A tool has been developed to help health and safety Inspectors screen repetitive tasks of the upper limbs for significant musculoskeletal disorder risks. The tool is based on two observational methods in the literature. These have been incorporated into a flowchart process that examines twelve risk factors grouped into four stages: frequency and repetition of movements; force; awkward postures; and additional factors, which include aspects of duration and recovery. A 'traffic-light' system is used to grade risks. The flowchart is supported with an assessment guide, which provides general instructions on the use of the tool and more detailed information on how to categorise each factor. There are also forms to record information collected as part of the assessment.

Introduction

The prevention and management of work-related musculoskeletal disorders (WRMSD) continues to be a priority for the Health and Safety Executive (HSE) and remains one of the keys to improving occupational health in the UK. HSE and Local Authority Inspectors play an important role in preventing WRMSD, and providing advice and information to employers about assessment and management of WRMSD risk factors. Thus, there is a need for tools that Inspectors can use to screen workplaces for tasks that present significant WRMSD risks. For several years, Inspectors have used the Manual handling Assessment Charts (MAC) to identify high risk manual handling tasks. There is a similar need for a tool to screen repetitive tasks of the upper limbs for some of the common risk factors that contribute to the development of WRMSD.

In January 2007, the Health and Safety Laboratory (HSL) began work with HSE to develop a tool that Inspectors could use to assess repetitive tasks. The purpose of this paper is to outline the concept and initial development of the tool, called Assessment of Repetitive Tasks (ART).

Initial concept of ART

With this work, the intention was not to develop an entirely new assessment method for repetitive tasks, but rather to make some of the more recent developments in this area accessible to health and safety Inspectors. The initial concept was to base the technical content of the tool on the Occupational Repetitive Action (OCRA) methods, developed by Colombini *et al*. (2002) and incorporated into BS ISO 11228-3 (2007): *Ergonomics – Manual handling – Part 3: Handling of low loads at high frequency*. In particular, the OCRA checklist offered promise as an initial concept, as it was simpler to apply than the OCRA index and more suitable for an initial screen of repetitive tasks (Occhipinti and Colombini, 2005). Thus, initial work focussed on careful examination and simplification of the OCRA checklist statements and ensuring the linkage to existing HSE guidance (2002) on upper limb disorders in the workplace (HSG60). An early peer-review exercise with ten HSE/HSL ergonomists indicated that the tool would also benefit from applying many elements of the Quick Exposure Check (QEC, David *et al*., 2008).

The initial concept also included developing a process that was similar in format and appearance to the MAC (Monnington *et al*., 2002). This was believed to offer an advantage, as most Inspectors, when learning how to use the tool, would already be familiar with the underlying process. The result is that ART involves a sequential flowchart process to examine twelve risk factors that have been grouped into four stages: frequency and repetition of movements; force; awkward postures; and additional factors, which include aspects of duration and recovery. For each of the twelve factors, a 'traffic-light' system is used to grade risks, with green representing a lower level of risk and red representing a higher level of risk. At the user's discretion, an assessment can be made of both the left and right arm, or just the arm predominantly involved in the task. The flowchart is supported with an assessment guide, which provides general instructions on the use of the tool and more detailed information on how to categorise each factor. There are also forms to record information collected as part of the assessment.

Selection of risk factors

Stage one: frequency and repetition of movements

Initial examination of the OCRA checklist indicated that, where users had difficulties identifying the sequence and number of technical actions, there was considerable potential to misclassify the frequency of arm movements. In addition, Li and Buckle (1999) have raised the practical concern that, in many tasks, a fixed work cycle either may not exist or may vary periodically, making it difficult to assess the exact frequency and repetitiveness of movements in a short period of time. With this in mind, the QEC was felt to offer Inspectors the most practical system for categorising two risk factors in this stage: the frequency of shoulder / arm movements; and the repetition of the forearm and hand / wrist movements. These aspects of the QEC were shown to be practical and reliable (David *et al*.,

2008). They were previously included in HSG60, and to some extent would be able to take account of the use of job enlargement to provide increased variety of arm movements and speed of work.

Stage two: force

The aim was to develop a graphical approach, which would allow users to evaluate the level of force as well as the duration of force exertion within the representative work cycle. For any actions that require more than a minimal amount of force, users are directed to involve those doing the work in a subjective assessment of the level of individual hand force applied in the task (e.g. light, moderate, strong or very strong). The assessment guide provides a summary of actions that might involve more than a minimal amount of force. For advanced users or specialists, a psychophysical approach using the Borg CR10 scale is recommended to improve the technical quality of the subjective assessment.

Colombini *et al*. (2002) caution that the use of force is rarely perceivable, even to a highly trained observer, and where force is assessed by an external observer there is potential for major errors. However, it may not always be practical for users to communicate effectively with workers about the level of hand force involved in the tasks. For such occasions, the graphical approach also incorporates the QEC categories for individual hand force requirements as a guide. These were based on recommendations from Silverstein *et al*. (1986) that estimated average hand force requirements of more than 4 kg be considered high force jobs, while requirements of less than 1 kg be considered low force jobs.

Stage three: awkward posture

The aim was to develop a process similar to the OCRA checklist that involved a check of both the presence of an awkward posture as well as the timing and duration of the posture. From the OCRA checklist, the tool includes an analysis of shoulder/arm, wrist and hand/finger grip postures. Further analysis of head/neck and back postures are also included, with the aim of expanding the scope of the tool to screen for a wider range of workstation design problems. Descriptive words were used to guide users to the presence of awkward postures. Li and Buckle (1999) have suggested that practitioners prefer the use of descriptive words over specific postures quantified in degrees. It is also unlikely that users would have access to posture measuring instruments when in the field.

Where awkward postures are not present, the factor is coded as green. However, where awkward postures are present, users then consider the duration of the awkward posture or movement. For each body segment, the criteria for the amber category are described consistently as either moving to an awkward posture repetitively or holding an awkward posture for 'a part of the time'. The Department for Trade and Industry's document entitled *'Instruction for consumer products'* (DTI, 1989) suggests that the phrase 'a part of the time' would mean 15–35% of the time to the majority of people. The criteria for the red category are described consistently as moving to an awkward posture repetitively or holding an awkward posture for

'more than half the time'. In the event that the user remains undecided between the categories, or subsequent video analysis reveals an awkward posture to be adopted for either less than 15% or 36–50% of the time, then an amber coding and selection of an intermediate score would be acceptable.

Stage four: additional factors

This stage considers further aspects important to the assessment of repetitive tasks, including the opportunity for recovery, the worker's perceived workload and other psychosocial factors, object and environmental factors, and the duration of the repetitive task.

It is important that work be organised in such a way that there is sufficient opportunity for recovery and a ratio of at least 1:5 between recovery and task duration has been suggested as a basis for assessment (BS ISO 11228-3, 2007). A factor was included in ART to ensure users consider the time that workers perform the repetitive task without a break. This may require information on the frequency; timing and duration of both structured breaks in the repetitive work as well as time spent performing other tasks that do not involve similar repetitive arm movements. A timeline is provided for users to record this information. Where the organisation of breaks is such that workers undertake the task continuously for less than one hour, or recovery can occur within the work cycle, the factor is coded green. Two amber categories are provided for where the task is performed continuously for one to two hours and two to three hours. Two red categories are provided for instances where the task is performed continuously for three to four hours or more than four hours.

A workpace factor was included in the tool using categories originally set out in the QEC (Li and Buckle, 1999). The aim was to help users investigate the extent to which workers perceive they have difficulties keeping up with the work and any reasons for these perceptions. A further aim was to help users begin a dialogue with workers about other psychosocial and organisational factors that can trigger the reporting of WRMSD. Other psychosocial factors are not given a score. However, the assessment guide draws attention to the importance of recording psychosocial factors and cues users to several factors that should be considered through dialogue with the workers. These were drawn from HSG60 and include: little control over how the work is done, frequent tight deadlines, incentives to skip breaks or finish early, lack of support from supervisors or co-workers, monotonous work, excessive work demands, high levels of attention and concentration, and insufficient training to do the job successfully.

The OCRA method lists an assortment of object and environmental factors for which there is evidence of causal or aggravating relationship with WRMSD. They are listed as 'other factors', as each factor may not always be present in the workplace; however, they are important. ART prompts the user to consider factors such as: exposure to hand/arm vibration; localised compression of anatomical structures; the need for fine movements of the fingers or hand, exposure to cold environments or cold surfaces; the use of gloves that restrict movements; and counter-shock actions, such as hammering with a tool.

Finally, a duration factor was included in the tool. The aim of this risk factor was to ensure that users considered the total amount of time that workers spend performing the repetitive task in a day. A further aim was to take account of the benefits of job rotation to non-repetitive tasks and discourage excessive task durations. Duration multipliers from the OCRA method were adopted, with task duration categories of less than two hours, two to less than four hours, four to eight hours and more than eight hours.

Intended use of ART

At this early stage of development, ART is intended for use by HSE and LA Inspectors. Although some users may find ART usable without training, initial indications are that training will be required for improved understanding of the topic, as well as improved accuracy and consistency when using the tool. However, the impact of training on use and reliability will require further investigation.

No attempt has been made to link individual factors or scores to enforcement action, as Inspectors must consider a range of other factors as well. In addition, there is little agreement on how aspects such as repetition, force, posture or other factors should be weighted, or how interactions should be quantified and most scoring systems adopted in the past have been largely hypothetical. However, an overall score can still be useful for estimating the level of exposure, prioritising interventions and assessing improvements when changes have been implemented. For this purpose, preliminary weightings were allocated to each factor. These weightings were based on values suggested in the original methods. At this stage, it was also felt important to ensure that the tool provided a balanced assessment, where none of the four stages individually could come to dominate the assessment and only those repetitive tasks exhibiting a combination of risk factors from several stages would be highlighted through the overall scoring system. The range of total scores possible have been categorised into three proposed exposure levels. However, these exposure levels are hypothetical and proposed only as a starting point for further validation and testing. Users are directed to the colours assigned to each risk factor when exploring risk reduction measures. At this stage, ART only considers individual exposure to a single repetitive task; however, there may be further scope to take account of complex or rotating jobs made up of several repetitive tasks.

The tool was developed through a process that involved peer-review exercises and iterative user trials, the findings of which are reported separately (Ferreira *et al*., 2008).

Acknowledgements

The authors wish to acknowledge the support of colleagues within HSE policy and operational directorates as well as those external ergonomics practitioners who have assisted with the development and review of ART.

References

BS ISO 11228-3. 2007, *Ergonomics – Manual handling – Part 3: Handling of low loads at high frequency*, (British Standards, London).

Colombini, D., Occhipinti, E., and Grieco, A. 2002, *Risk assessment and management of repetitive movements and exertions of upper limbs: job analysis, OCRA risk index, prevention strategies and design principles*, Volume 2, (Elsevier Science, Oxford).

David, G., Woods, V., Li, G., and Buckle, P. 2008, The development of the Quick Exposure Check (QEC) for assessing exposure to risk factors from work-related musculoskeletal disorders. *Applied Ergonomics*, **39**(1), 57–69.

Ferreira, J., Birtles, M., and Stanley, L. 2008, Initial testing of an inspectors' tool for the assessment of repetitive tasks (ART). In P. Bust (ed.), *Contemporary Ergonomics 2008*, (CRC Press, London).

Department of Trade and Industry. 1989, *Instructions for consumer products – Guidelines for better instructions and safety information for consumer products*, (HMSO, London).

Health and Safety Executive. 2002, *Upper limb disorders in the workplace*, HSG60 Second edition, (HSE Books, Sudbury, Suffolk).

Li, G. and Buckle, P. 1999, *Evaluating change in exposure to risk for musculoskeletal disorders – a practical tool*, HSE Contract Research Report 251, (HSE Books, Sudbury, Suffolk).

Monnington, S.C., Pinder, A.D.J., and Quarrie, C. 2002, *Development of an inspection tool for manual handling risk assessment*, HSL Report ERG/02/20 (HSL, Sheffield).

Occhipinti, E. and Colombini, D. 2005, The occupational repetitive action (OCRA) methods: OCRA index and OCRA checklist. In: N. Stanton, A. Hedge, K. Brookhuis, E. Salas and H. Hendrick (eds.) *Handbook of Human Factors and Ergonomics Methods*, (CRC Press, London).

Silverstein, B.A., Fine, L.J. and Armstrong, T.J. 1986, Hand wrist cumulative trauma disorders in industry. *British Journal of Industrial Medicine*, **43**(11), 779–784.

BAGGAGE HANDLING IN NARROW-BODIED AIRCRAFT: FURTHER RISK IDENTIFICATION

Sarah Tapley[1], David Riley[2] & Laraine Oxley[2]

[1]*HSE Wren House, Hedgerows Business Park, Chelmsford, Essex CM2 5PF*
[2]*HSL, Harpur Hill, Buxton, Derbyshire SK17 9JN*

The manual loading and unloading of passenger aircraft continues to be the cause of a high level of reported incidents to the Health and Safety Executive (HSE) in the UK. New handling equipment is starting to become more widely available, driven in Europe by increased regulatory activity however there is limited published literature evaluating these systems. This paper describes some of the new technologies being introduced, reviews relevant new literature, and outlines a collaborative project being run to establish robust guidance for the UK baggage handling industry.

Introduction

Baggage handling activities continue to account for over 40% of all reported incidents at UK airports.

The job of the baggage hander task is often thought to be just loading bags on and off aircraft. This is not the case, the baggage handlers will typically perform the range of turn-around tasks; directing the aircraft to the stand, connecting up power, chocking the wheels, and making safe for the passengers to disembark. Passenger steps in the UK are routinely manually moved into place by ground crew, who also pull and push full and empty baggage carts to and from the hold doors. Hold baggage is then offloaded to carts and taken to the reclaim area and onward bags are loaded from carts onto the aircraft. The bag offload/onload is performed using a variety of methods, the specific method used will vary depending on the plane type and the equipment available.

There are important differences between on- and off-loading. For on-loading, 2 or 3 workers unload baggage from carts into the hold doorway, either direct or via powered conveyor, to two workers in the hold usually knelt or stood stooped. These two work ‘in series’, one near the hold door passes baggage deeper into the hold for the other worker to stack. In some cases there may only be one internal handler, in this case the external handler will lift the bag to the aircraft sill and then push the bag down the hold to the stacker.

For off-loading, the two hold workers pass baggage out to their co-workers either direct, or via a powered conveyor onto flatbed trucks or baggage carts.

It is the internal stacking operations that are accepted to be the highest risk part of the handling task in terms of musculoskeletal ill-health. (Tapley and Riley 2004).

International influences

Aviation and so by default aircraft baggage handling is an international activity, which is highly competitive with narrow profit margins. In Europe, and in particular in the UK ground handling services such as baggage handling are often contracted out rather than being done 'in house' by the carrier in an effort to drive down costs and to meet competition legislation. In Europe, some health and safety enforcing authorities (the Danish and Dutch) are starting to take a stronger line to reduce the total tonnage being moved by baggage handlers.

In the UK there is an increased understanding of the issues although little has been done in practical terms on the ramp to reduce the well established risks. This may be due in part to the lack of influence that a handling company is able to have over issues such as aircraft design, airport layout and carrier demands. New handling technologies are becoming more widely available around the world although there is little in the way of published literature evaluating these systems. Some researchers, in particular Dell (2007) argue that implementations of these new technologies will only come about when the regulators force the issue, as has been the case in parts of Europe.

Updated literature review

Relatively little new material has been identified since the authors original review in 2004 (Tapley and Riley 2004) and there remains little in the way of published material in this field. Dell (2007) in his PhD thesis (obtained by personal communication) present the most definitive review of baggage handling to date covering the work conducted in this area since the 1970's. Several papers reviewed in this thesis are evaluations of mechanical assistance mentioned above.

Stålhammar (1991) observed each of nineteen handlers in a study to lift an average of over 10 tonnes per shift. Culvenor (2004) in a study of Qantas handlers found a cumulative load per shift of 8 tonnes.

The IATA (International Air Transport Association) Airport Handling Manual 2006 recommends a single bag weight limit of 32 kg. British Airways in the UK are introducing a 23 kg single bag weight limit in November 2008. Qantas and Air New Zealand already have 25 kg single bag limits in place. Ryanair have introduced a 10 kg single item system with excess baggage rates applying above this weight. This varied and inconsistent bag weight limit is confusing for passengers.

From an injury reduction view Culvenor (2007) suggests a reduction would need to go as low as 10 kg or under. Dell (2007) applies the NIOSH lifting equation to outside the hold and presents a weight limit of 6 kg. Data gathered by Culvenor on bag weight as part of the Qantas study indicates that only 10% of bags are over 23 kg. Introduction of a 23 kg single bag weight limit will therefore alter the average weight of items very little.

Studies have focused on assessing the task as it is traditionally carried out. Korkmaz et al (2006) report a study investigating the benefits of providing baggage weight indication labels and of a systematic stacking pattern to minimize manual

lifting of the heaviest baggage items. This is an alternative way of stacking from the traditional stacking bags flat. These two factors combine to produce a beneficial effect in reduced spinal loading. This study collected muscle activity using EMG, trunk posture using a Lumbar Motion Monitor, applied force from a floor mounted force plate. There were 12 student subjects and the data was collected in a non-operational context. An EMG-assisted biomechanical model was used to process the data. Although the trial arrangements can be criticised for a lack of realism due to the subjects, the uniform luggage used and the experimental context, the results indicate a clear potential for such a stacking approach to be employed operationally with worthwhile benefits.

New risk reduction technologies

Some of the new equipment available avoids the need for a worker in the hold to perform the function of moving baggage between the hold door and the stacking worker (and vice versa). However the stacking work is acknowledged as presenting the higher risk of injury (Egeskov, 1992, Tapley and Riley 2004). Belt-loader based technologies have been developed which reportedly significantly reduce the amount of manual exertion and lifting and supporting involved in the stacking and un-stacking operation. None of the systems described below have a fundamental effect on the fact that the worker needs to kneel, sit or squat in the confines of the hold. However, it is the combination of heavy handling while in a kneeling posture that presents the greatest risk of back problems developing.

1. Sliding Carpet™ and ACE™ systems are two mechanical systems which have the benefit of eliminating one of the baggage handling operations – that of moving the bags between the doorway and the stacking face, and vice versa. They do not help to improve the posture of the stacking worker, or to reduce the manual lifting and handling operations associated with the stacking and unstacking function. (Jorgensen 1987 and Stokholm 1988). These systems have been available for at least 10 years, but appear to still be relatively rarely encountered, in the UK at least; this may in part be due to the weight that they add to the aircraft.
2. RTT Longreach™ is a belt loader based device designed to work in combination with the 'Sliding Carpet', and to deliver the baggage items near to the position the item has to be stacked, and in theory minimising the amount of manual handling required. Cree and Kothiyal (2006) conducted trials using a mock-up hold and 10 experienced handlers in a comparison with the conventional belt-loader working practices. They used measures of rated exertion (Borg RPE), posture analysis (OWAS), and found significantly reduced ratings of exertion, reduced amounts of pulling and pushing actions, improved postures (increase in good posture, 42% loading, 28% unloading) compared to the conventional use of the belt loader alone.

   However, anecdotally, this combination of equipment is generally used by airlines that self handle, such as Qantas and KLM, rather than low cost carriers who seek to minimise hold equipment to reduce fuel and maintenance costs.

Any of these devices may also be used in combination with modified stacking regimes, such as that proposed by Korkmaz et al (2006). This could result in reduced handling risk (spinal loading) for the stacking worker, as well as eliminating one handler from the team working in the hold. These combinations have not been reported in the literature.

3. Rampsnake™ was designed following regulatory action in Denmark requiring the tonnage being handled by baggage handlers to be reduced significantly in line with the requirement of the EU Manual Handling Directive in 1992 [integrated into UK legislation and implemented as the Manual Handling Operations Regulations (as amended) 1992]. The unit replaces a conventional belt loader and incorporates an extending powered conveyor section that enables baggage items to be transported directly to or from the stack in the hold. At this point it provides an adjustable raising section of conveyor to assist with the transfer between belt and stack. In addition, it features a user positionable adjustable section of conveyor for use at the baggage cart. The aim is to reduce as far as possible the amount of manual lifting and supporting required in the transfer of items between baggage carts and the aircraft, essentially only pushing and pulling level transfer actions remain. Compared with the conventional use of a belt loader, it also removes the need for the worker located at the hold doorway.

   Koelewijn (2004, 2006) reports a study undertaken to evaluate the effectiveness of the Rampsnake handling equipment at Schipol Airport for KLM. The overall conclusion was that the use of the Rampsnake is beneficial and although it does not avoid the need for the manual operation, it significantly reduces the risks to staff. A cost-benefit analysis for KLM suggests that the annual savings in terms of absenteeism costs would be in the range of Euro 0.5 M to Euro 2.1 M.
4. Powerstow™, is very similar to the Rampsnake in function in that is provides an extending powered conveyor from belt loader type device through the hold door right to the working point within the hold. It also has the facility at the end to assist with the raising of baggage items from hold floor level (or vice versa). No studies have evaluated this equipment directly, however, given the functional similarities, Koelewijn's findings for the Rampsnake are considered to be directly relevant. At present it does not have the height adjustable extending conveyor section at the ramp/apron end, for the worker transferring between the baggage cart and the belt, so the risk there is considered to be identical to that for the use of a conventional belt loader. Company marketing material indicates that it is possible to retro-fit the extending conveyor to existing belt loaders.
5. Mallagahan LBT 90™, this is a large driveable belt loader which effectively replaces the baggage cart, the tug to pull them and the belt loader. There are issues around compatibility with aircraft types and especially with airport infrastructure. The baggage hall, its conveyors and its access and roadways all need to be designed to accommodate the Mallaghan, otherwise its effectiveness is significantly reduced, and further risks may be introduced. More work is needed to establish the extent of capabilities of this device in terms of the range of aircraft it can be used with, and in undertaking a more detailed appraisal in relation to the manual handling risks.

Collaborative baggage handling project

It has been recognised that the health and safety issues associated with baggage handling are not owned by any one group but are affected by all those who have input to an aircraft turnaround.

Acknowledging the need for a partnership approach to tackle this issue in the UK, a collaborative team of handling companies, carriers, trade union, an airport operator, CAA and HSE was established to produce robust practical guidance setting out the expected standards for the manual loading and unloading of narrow bodied aircraft at UK airports.

Data has been gathered using an observation based approach using video, as well as measurements and interviews with staff at a UK regional airport carrying out handling activities from a variety of aircraft using a variety of handling techniques. Trials are planned in early 2008 for two handling companies to evaluate the usability and risk reduction offered by Rampsnake and Powerstow. One of the primary questions being posed by the handling companies is whether it is possible to turn an aircraft round, i.e. offload and then onload bags, in the 25 minutes allocated by the low cost carriers. The trails will seek to replicate in part the study carried out by Koelewijn for KLM in 2006, which led to the company purchasing 39 Rampsnake units.

Summary

There is little new published material available since the first study was carried out in 2003, and the industry continues to report high levels of incidents associated with baggage handling activities. The aviation sector is an international, highly competitive industry which, in the UK, is subject to intense price cutting and contracting out of services which make introduction of new technology a challenge for the industry and the regulator.

A collaborative group is working to produce robust guidance, which aims to reduce the risks associated with manual baggage handling activities.

References

Dell, G. 2007 PhD thesis (unpublished).

Egeskov, R. 1992 *Use of conveyor belts and locker loading on B737 Aircraft, Comcare Investigation Report*, Comcare, Brisbane.

Cree, J. and Kothiyal, K. 2006 *Manual handling risks in aviation industry: ergonomics evaluation of the RTT loader*. Meeting diversity in ergonomics, Proceedings of the 16th congress of the international ergonomics association, Maastrict, 10–14 July 2006.

Koelewijn, B.R. 2006 *Rampsnake: a new working aid for loading an aircraft. Meeting diversity in ergonomics*, Proceedings of the 16th congress of the international ergonomics association, Maastrict, 10–14 July 2006.

Koelewijn, B.R. 2004 *Rampsnake intervention study*. KLM Arbo Services bv

Jorgensen, K., Jensen, B., Stokholm, J. 1987 Postural strain and discomfort during loading and unloading of flights. An ergonomic intervention study, Trends in Ergonomics/Human Factors, Elsevier Science Publishers, New Holland.

Stokholm, J. 1988 *A summary of the SAS Ergonomic Report on the use of the Sliding Carpet Loading System on SE-DDT during October–December 1987*, Scandinavian Airlines System, Copenhagen.

International Air Transport Association, 2006, *Airport handling manual*, 26th edition, (IATA).

Culvenor, J. 2004 Baggage Handling: *The effect of baggage weight and other recommended solutions.* Report to Qantas Airways – unpublished, Qantas Airways, Sydney. In Culvenor, J., 2007 and Dell, G., 2007.

Tapley, S. E., and Riley, D. 2004 *Baggage handling in narrow-bodied aircraft: Identification and assessment of musculoskeletal injury risk factors.* Health and Safety Executive.

ERGONOMICS REVIEW OF THE HIGHWAYS AGENCY CONTROL ROOMS

Nigel Heaton[1], Jed Evans[2], Jim McKenzie[1] and Matthew Trigg[1]

[1]*Human Applications, the Elms, Loughborough LE11 1RG;*
[2]*The Highways Agency, Ash House, Exeter EX2 7LB*

The Highways Agency needed to provide advanced DSE assessments for their control room operators, so commissioned a Training Needs Analysis (TNA) to assess the requirements of the control room and to evaluate how much additional skills would be needed by the DSE assessors to teach them how to do suitable and sufficient DSE assessments on the control room staff. The subsequent TNA became a detailed ergonomics report, which helped the Agency to manage its key ergonomics issues and to identify any improvements made in addition to the new DSE assessment process. The work highlighted a number of 'quick wins' and proposed a structured design plan to improve ergonomics management and reduce the likelihood of ergonomics problems occurring. The advanced DSE training ran in parallel with an ergonomics action plan. This paper discusses the approach to undertaking the ergonomics audit, identifies the main findings from the audit and looks at the design and implementation of the training programme that resulted from the work. Key features of the study were extensive user involvement in defining problems, video analysis of potential MSD issues and expert review of the software used. The key learning points from the training, the personnel who were trained and the behaviour changes sought are also discussed. The integration of ergonomics issues into the broader health and safety management system was a key feature for the Highways Agency, allowing it to continue its process and commitment to managing risk into the future.

Introduction

The Highways Agency (HA) operate Regional Control Centres (RCCs) throughout England. A Training Needs Analysis (TNA) was undertaken to provide the background for a training programme for existing HA Display Screen Equipment (DSE) risk assessors to move to an advanced level of DSE assessment. This was required by the specific and unique ergonomic and operational factors in the RCC control rooms.

RCC control rooms exist to manage the HA's responsibilities to the UK motorway network in real time. Operators work in small teams typically with one supervisor

allocated to several operators on a shift. Control centres operate across 24 hours every day of the year. The main activities of the RCCs are:

- Radio despatch: handling calls from police and HA Traffic Officers
- Emergency Response Team call handling: handling calls from road users using Emergency Roadside Telephone boxes
- Traffic management: monitoring CCTV images to anticipate and manage congestion and incidents in the absence of a direct presence at the scene
- Traffic management: setting roadside and overhead signals to control traffic behaviour
- Roadside and rescue liaison: part of call handling that requires telephone contact with roadside assistance and recovery agencies
- Traffic information: liaison with traffic information providers.

The main challenges from an ergonomics perspective was how to ensure that the HA provided the best ergonomics environment that was possible within the constraints of operational requirements and budget.

Methodology

A stakeholder meeting held with various HA representatives to scope the work and to agree the general approach. This meeting gave key personnel from across the HA the opportunity to provide input to the study and to raise those issues currently known to cause concern across the estate. The meeting also gave Human Applications' consultants the opportunity to understand some of the design decisions which had been made and the technological constraints inherent in the roll-out of the RCCs.

Human Applications' consultants visited each of the seven RCCs. At each site, key personnel were interviewed about their usage of the existing systems. Operators and Supervisors were interviewed and observed whilst carrying out normal duties at RCC workstations. Measures of environmental and physical factors were taken where appropriate.

The results of the visits allowed Human Applications to produce a comprehensive report on the training needs of the HA RCC personnel as well as identifying strengths and weaknesses of the existing provision. Human Applications also produced a be-spoke training courses aimed at training local HA RCC personnel in how to conduct DSE risk assessments in an RCC environment.

Results

All operators reported receiving training as part of their initial one-month training to be RCC operators. The HA booklet 'Are You Sitting Comfortably?' had been given to all staff and copies were pinned to modesty panels at some sites. Though the booklet details how and why to make adjustments to the standard chair and to the standard desk, few of the operators interviewed knew how to make the full

Figure 1. Typical working postures in the RCCs.

range of adjustments and none could describe why some adjustments might be advantageous.

In general terms, the provision of furniture (adjustable desks and chairs, etc.) was good but there were concerns about 'fit' from some of the personnel and issues of how the furniture should be used. The most significant challenge to the HA was how to use and get the best out of the monitors and keyboards (the range went from 4–7 monitors; including a touch screen interface on one monitor, plus a variety of input devices).

In terms of the environment, all the RCCs, reported that lighting and temperature was controlled by software maintained by the Facilities Management function. The most common concern raised was that out of core office hours, the system was maintained by a non-specialist (typically the security officer at front of house) and that changes to the system overnight were not made.

Gross control of lighting was available in addition at all RCCs, though the deployment or otherwise of this control seemed to vary both between sites and between operators (some seemed quite happy to change the lighting, others felt it was not fair on their colleagues).

During intensive interaction with the IT systems, almost all operators adopted a bent forwards posture (see Figure 1 provided above). There are four reasons for this, which will be present to varying degrees between operators:

- The viewing distances between the operator and the main screens are too long
- Smaller operators must sit forwards on the chair and must therefore rest their forearms on the worksurface to achieve upper body support
- Regular interaction with the touch screen brings most operators forwards, and they will tend to avoid repeated movements at the trunk
- With very few exceptions, operators do not touch type. The need continually to look at the keyboard brings the neck into flexion (forwards bending), which is uncomfortable to maintain from an erect or reclined posture.

Recommendations

The principal aim of this project was to undertake an audit of the existing situation in HA RCCs with respect to the health, safety and comfort of RCC operators and to use this information to define the training requirements for RCC DSE Risk Assessors. Recognising that RCC control room operators as high risk DSE users needed an enhanced approach over and above the established HA procedures.

A DSE assessment of a computer user will identify risk factors and make recommendations to eliminate or reduce the source of risk to the specific user. However, in reality, within the RCCs, we found that there is a combination of [DSE associated] risk factors for current RCC staff that cannot be addressed *solely* by local changes by DSE Risk Assessors.

Instead, we identified three categories of change required in HA RCCs to reduce the current levels of risk to which RCC operators are exposed and divided our recommendations from the project into three main sections:

- Information and training needs for RCC operators that will provide them with the necessary information and skills to work safely in an HA RCC.
- Changes to the current provisions of workstations, workplaces and working practices in RCCs. Without these changes, the HA may not be providing the most appropriate environment for personnel within the RCCs.
- Training needs to be provide for RCC DSE Risk Assessors that will equip them with the necessary skills, experience and competency to undertake suitable DSE assessments in an environment as complex as an RCC.

Information and training

Additional training should be provided, within each RCC, to demonstrate how and why to make the adjustments for each operator. Subsequent risk assessment should include an opportunity for a risk assessor to validate the use of better set-up. Much of the information and training requirements for RCC operators is common to all DSE users within the HA. However, the HA needs to ensure that DSE information and training for RCC operators be enhanced to include the following issues specific to RCCs.

Changes to provision

The equipment used – the chairs and desks need to be matched to end users and in some instances additional provision will be required. The number of monitors on the desks represents a significant challenge to the HA and efforts need to be made to reduce and rationalise desk layout. Whilst lighting and thermal comfort were not significant challenges, both could be improved via relatively quick win solutions such as the provision of additional task lights and slightly increasing the temperature of the RCCs at night.

DSE risk assessor training

The requirement to conduct a suitable and sufficient DSE assessment would normally be met by training assessors on a 'standard' DSE course providing delegates with the following skills:

- Conducting a workstation risk assessment
- Identifying and ranking risks to users
- Proposing corrective measures
- Providing training to users to improve their use of equipment.

Our analysis demonstrated that, as identified at the outset by the HA, a standard course would not deliver the appropriate assessment skills to operators. The HA course needed to address fundamentally some of the new and different risks associated with the RCCs. The new course need to provide assessor with the ability to support:

1. The removal all footrests, except in special cases.
2. The reduction of the number of keyboards and mice to 2 of each.
3. The replacement all 'ball' mice with 'optical' mice.
4. Training all operators to interact with it using the mouse or reposition the touch screen.
5. Reminding all personnel of the HA eyesight testing policy.
6. Reducing the number of monitors to a maximum of 5 at each workstation.
7. Replacing the CRT screens with a flat screen (we understand that this is an ongoing programme).
8. Providing a range of alternative chairs for assessors to specify for those with special needs.
9. Increasing the light levels to an acceptable minimum.
10. Providing task lights were required.
11. Training operators, were reasonably practicable, in touch typing and keyboard skills.
12. Reviewing and identifying all software short cuts, and providing additional short cuts where appropriate.
13. Providing a range of headsets to suit user needs.
14. Ensuring that the Policy on breaks is followed.

Human Applications and the HA worked to design the course to meet all these requirements. It was decided that an extended 2-day course with internal HA support would be the most effective way of delivering the required number of assessors. It was apparent that all RCCs would require local assessors who would be able to cover all the shifts. It was also clear that a central unit of trained assessors would be able to offer additional support and could also liaise with the external providers of occupational health services.

The course was designed to meet the requirements of the Institution of Occupational Health and Safety's (IOSH's) DSE course and all candidates were required to sit the IOSH exam. The course was designed to be delivered by professional

ergonomists who had a practical experience of the issues associated with working in the RCCs.

Conclusions

As part of its overall strategy to assess, reduce and manage the risks associated with all HA activities, the HA trained specialised DSE assessors and began the process of rolling out a comprehensive risk assessment programme. At the same time, many of the 'quick wins' identified by the project were implemented. A series of specialised training courses were held at the RCCs to train both local operators and central personnel in how to conduct assessments and how to improve the ergonomics of the RCCs.

The assessors then began a process of implementing ergonomics improvements and providing local information, training and assessment of all RCC personnel. In parallel with this work, the HA began to roll-out a much more comprehensive programme of health and safety risk management to ensure that the occupational health and safety of all staff where ever they are employed by the HA is appropriately managed.

PHYSICAL ERGONOMICS

THE EFFECTS OF A MID-MORNING CHOCOLATE SNACK ON MOOD AND GLUCOSE TOLERANCE

Neil Morris

University of Wolverhampton, School of Applied Sciences, Psychology, Millennium Building, Wolverhampton WV1 1SB, UK

Chocolate can be used to 'self-medicate' when energy levels are low and tension levels are high mid-way through a work period (Thayer, 2001). However chocolate consumption may have undesirable effects on blood glucose level and calorie intake. In this study blood glucose level was determined at 10.00 am, 11.00 a.m., and 12.15 p.m. Participants consumed a chocolate bar, either regular milk chocolate or the 'LoCarb' equivalent, at 10.40 a.m. Just prior to each blood sampling Energetic Arousal and Tense Arousal were measured using the UWIST mood scale. The two types of chocolate bar were matched for fat and calorie content but differed in sugar content. Blood glucose rose significantly, in both groups, after ingesting chocolate but returned to normal by the 12.15 p.m. blood test demonstrating good glucose tolerance to this snack. Neither group showed any change in Energetic Arousal but Tense Arousal was considerably reduced in both groups and this reduction was still apparent at 12.15 p.m. It is concluded that a 40 gram chocolate snack can reduce perceived tension in participants with good glucose tolerance without raising blood glucose levels beyond the normal range.

Introduction

Thayer (2001) has argued that we self-medicate our mood states by ingesting mood altering substances, for example, alcohol, coffee, nicotine etc. and we may also use other means to modify our affective state. For example, we may exercise or we might alter our mood by seeking massage to reduce tension (Morris and Wilkes, 2007). The latter, non-substance based approaches are favoured by Thayer because they lack the drawbacks associated with chemically altering mood (for example, 'rebound' as the effect wears off and the danger of addiction). However it is abundantly clear that most of us, particularly in the workplace, will seek a chemical modification of mood. In this study the effects of a fairly innouous mood enhancer, chocolate, is examined. The study seeks to examine the most obvious short-term drawback to ingesting chocolate, namely the undesireable increase in blood glucose level associated with ingesting sweet snacks. Thus an assessment of the mood enhancing properties of chocolate is accompanied by a determination of study participants tolerance to the ingestion of significant amounts of sugar in this form.

Generally, mood is less positive early in the day but improves across the morning and peaks late morning before dipping in the afternoon (Thayer, 2001). Numerous studies have shown that consuming snacks or drinks rich in simple carbohydrates, i.e. sugars, reduces perceived anger, depression and tension (Benton and Owens, 1993; Keith *et al.*, 1991; Martino and Morris, 2003). These mood changes are associated with elevation of blood glucose level. However many foods and soft drinks have a high glycaemic index (G.I.), i.e., they cause a rapid rather than gradual rise in blood glucose level.

In healthy, young individuals blood glucose levels are maintained at around 5 mmol/l. This control is attained via a negative feedback loop. Insulin is released from the pancreas when blood sugar begins to rise above 5 mmol/l and it results in glucose being removed from the circulation and immobilised, as glycogen, in the liver and in muscles. When blood sugar levels drop much below 5 mmol/l the pancreas releases glucagon that mobilises glycogen in the liver. Glycogen is broken down into glucose and released into the blood stream increasing blood sugar until insulin release is triggered. In fact insulin and glucagon are mutually inhibiting so deviations from the 'set point' for blood sugar occurs largely because of absorption of glucose, mainly from the gut, and this is then rapidly regulated. High G.I. food and drink ingestion results in more insulin release and results in glucose level 'spikes' that are not healthy. Thus the benefit of a sugar rich snack, with respect to mood enhancement, is tempered by the insults to glucose regulation.

Ideally a snack would have a low G.I. thus providing slow, smooth blood glucose elevation and good glucose tolerance (i.e. a gradual return to ~5 mmol/l blood glucose within a couple of hours of raising the level). One good candidate for a relatively low G.I. snack rich in sugars is chocolate. The G.I. of a given snack is modified by a number of factors. For example, G.I. is reduced when the pH of the snack is decreased (so adding vinegar can reduce the G.I. of a snack) or by having a significant fat content as this impedes glucose transport across the gut wall thus slowing but not preventing absorption (Pond, 1998). The combination of fat and sugar in chocolate results in a G.I. value of about 43 (where 50 g glucose in a saturated solution = 100) which is considered to be low (Brand-Miller, *et al.*, 2003). However chocolate is still rich in mood enhancing substances (simple sugars, caffeine etc.) suggesting that if a small snack is used to assuage tension and/or hunger then chocolate is perhaps a good candidate.

In this study participants were given a snack consisting either Tesco milk chocolate or Boots 'LoCarb' chocolate (containing much less sugar). Their mood state was established psychometrically using the UWIST mood adjective checklist (henceforth the UWIST – Matthews *et al.*, 1990) prior to chocolate ingestion. Blood glucose level was determined also. Following ingestion of chocolate both mood and blood glucose were again determined after 20 minutes and then both were tested again at the end of the morning. The purpose of the study was to establish the pattern of changes in both mood and blood sugar level. The latter was used to determine if there was good glucose tolerance after ingesting chocolate since it is desirable that the use of chocolate as a mood enhancer does not also result in unhealthy glucose regulation.

Method

Participants

This study employed 88 undergraduate Psychology students. Approximately 80% were female with a median age of 21. None of these participants had fasted immediately prior to the study.

Materials and procedures

In this study mood was measured using the UWIST which is known to be both psychometrically valid and reliable (Matthews, *et al.*, 1990). The UWIST measures Energetic Arousal, Tense Arousal, and Hedonic Tone. It has two other scales – Anger-Frustration and General Arousal but their psychometric properties are not well established. In this study we use only the rigorously validated sub-scales and confine our analysis to the two arousal scales. Detailed descriptions of the properties of these dimensions can be found in Morris, *et al.* (1998). More informally, Energetic Arousal is a measure of the extent to which an individual reports feeling lively and alert, or with a low score, sluggish. Tense arousal is a measure of the extent to which one feels muscular tension. Positive shifts in mood are represented by increases in Energetic Arousal and *decreases* in Tense Arousal. The constituents of the chocolate are shown in Table 1 below. Blood glucose level was tested using BM-Test 1-44 blood glucose test strips, following the manufacturers procedure and then measured with Prestige Medical Healthcare Ltd. HC1 digital Blood Glucometers. The average of two measures was recorded at each testing.

Participants in the study were tested in two groups. The sessions were identical in procedure except that one group consumed Tesco milk chocolate and the other ate Boots LoCarb chocolate. After an initial blood test participants filled out the mood questionnaire. They were then asked to consume either 40 grams of LoCarb chocolate or 38 grams of Milk chocolate. The slightly different portions of chocolate facilitated matching the constituents of the different chocolates. Participants then watched a 15 minute university induction video and this was followed by a memory test and served the purpose of allowing the chocolate to be absorbed. Blood glucose was then determined again and a second mood questionnaire was filled out. Following a further video and memory test a final blood glucose determination was made and mood state was once again measured.

Table 1. Constituents of LoCarb chocolate and milk chocolate. The figures in brackets show the calorific values and sugar content for the amount of chocolate consumed in each condition).

Type of chocolate	LoCarb	Milk chocolate
Kcal/100 g	485 **(40 g ∼ 194 kcal)**	531 **(38g ∼ 195 kcal)**
Carbohydrate	44.0	55.6
Sugars	9.0 **(40 g ∼ 3.60 g)**	54.9 **(38 g ∼ 20.86 g)**

Results

Blood glucose readings were analysed using a 3×2 mixed design ANOVA. The experimental condition (i.e. LoCarb versus milk chocolate) formed the between subjects factor, and the within subjects factor was the blood sampling times (i.e. first, second and third blood test). Results revealed that, overall, mean blood glucose level was significantly higher among those in the milk chocolate group ($F_{(1,86)} = 6.09$, $p < 0.025$). There was a significant main effect of test ($F_{(2,172)} = 8.19$, $p < 0.0001$) that *post hoc* testing revealed to be due to lower blood glucose level at test 3. The interaction was non-significant ($F_{(2,172)} = 1.48$, $p > 0.05$). There were no significant effects derived from the energetic arousal analysis (all F ratios <1). Analysis of the mood data, using an identical anova design, for Tense Arousal revealed no significant main effect for type of chocolate ($F_{(1,86)} < 1$, $p > 0.05$) but there was a significant effect of test number ($F_{(2,172)} = 32.05$, $p<0.0001$) that showed that tense arousal was higher before eating the chocolate. The lack of a significant interaction ($F_{(2,172)} = 2.04$, $p > 0.05$) demonstrated that the reduction in tense arousal after eating chocolate was found with both types of chocolate and persisted to the third test. The means and standard deviations for the blood glucose and mood data are presented in Table 2. The results can be summarised as showing that the chocolate failed to, or only marginally, elevated blood glucose but the level had fallen by the end of the study to a level below the initial value. Energetic arousal did not alter but tense arousal was reduced after either chocolate and this reduction was maintained until the experiment finished.

Table 2. Means and standard deviations (in brackets) for blood sugar levels (in mmol/l) for two groups, one consuming Boots LoCarb chocolate and the other consuming Tesco milk chocolate and mean Mood scores for Energetic Arousal and Tense Arousal.

Condition/Test number	Test 1 1st blood test	Test 2 2nd blood test	Test 3 3rd blood test
LoCarb N = 44	4.85 (1.32)	5.12 (0.99)	4.74 (1.43)
Milk chocolate N = 44	5.36 (1.01)	5.38 (1.88)	4.88 (1.22)
Condition/ Mood dimension	Test 1	Test 2 Energetic arousal	Test 3
LoCarb	20.00 (3.67)	19.97 (4.01)	19.64 (4.12)
Milk chocolate	21.00 (5.71)	20.32 (4.77)	20.30 (4.02)
Condition/ Mood dimension	Test 1	Test 2 Tense arousal	Test 3
LoCarb	18.61 (4.27)	14.66 (3.24)	15.11 (3.14)
Milk chocolate	17.02 (5.47)	14.70 (4.06)	14.86 (4.07)

Discussion

The results show that a modest amount of chocolate can reduce perceived tension. Furthermore it does not matter whether the chocolate is a low sugar brand or a typical milk chocolate. The dosage used in this study was four squares of chocolate and provided less than 200 kcal. Thus a chocolate ration of this size, providing less than 10% of the recommended calorie intake for the average adult woman, is unlikely on its own to create a weight gain problem in the long term. Table 1 shows that there is only a very modest blood glucose level throughout the study. Guidelines on 'healthy' blood glucose levels vary but even the more stringent ones accept <6 mmol/l as unproblematic. These data also show good glucose tolerance in this sample. Blood glucose level was lower at the end of the study, some two hours after the baseline blood glucose level was established even when >20 grams of simple sugar had been ingested.

These data should not be interpreted as implying that workers should eat 40 grams of chocolate mid-morning. Rather one should consider that, just as one might take aspirin when a headache begins, a few squares of chocolate might prove efficacious when tension is first perceived. It would obviously be better to avoid developing tension at work by employing suitable ergonomic measures in the workplace, for example, appropriate seating, regular breaks with exercise etc. However, if tension does begin to develop then a small chocolate snack may be a first, and innocuous, 'drug' treatment. This research forms part of a larger body of work that aims to identify methods for improving well-being, in the workplace and elsewhere, simply, safely and inexpensively. It is not intended that any of these methods are a remedy for complaints of a serious nature. See Morris and Wicks (2007) for a detailed discussion of this.

A number of conclusions can be drawn from these data.

- A 'dose' of chocolate of this magnitude may reduce tension but probably only briefly. It is no substitute for removing sources of tension.
- The tension experienced is not due to sub-optimal blood glucose level.
- This amount of chocolate is not experienced as 'energising'.
- Low sugar chocolate, in 40 gram portions, is not 'healthier' than traditional milk chocolate.
- There is no evidence here for any contra-indications for using chocolate to 'self-medicate' tension but we should not extrapolate to other populations or portion sizes without further empirical evidence.
- A sample of this population rated the palatability of both types of chocolate and the milk chocolate was perceived to be significantly pleasanter. This suggests that the milk chocolate is preferable to low sugar chocolate as a tension reducer.

References

Benton, D. and Owens, D.S. (1993). Is raising blood glucose associated with the relief of tension? *Journal of Psychosomatic Research*, **37**, 1–13

Brand-Miller, J., Foster-Powell, K., and Holt, S. (2003). *Complete Guide to G.I. values*. (Hodder and Stoughton, London)

Keith, R.E., O'Keefe, K.A., Blessing, D.L. and Wilson, D.G. (1991). Alternations in dietary carbohydrate, protein and fat intake and mood state in trained female cyclists. *Medicine and Science in Sports and Exercise*, **23**, 216–218

Martino, O.I. and Morris, N. (2003). Drinking glucose improves mood at the beginning of the working day. In P. McCabe (ed.) *Contemporary Ergonomics 2003*, (Taylor and Francis, London), 226–231

Matthews, G., Jones, D.M., and Chamberlain, G. (1990). Refining the measurement of mood: The UWIST Mood Adjective Checklist. *British Journal of Psychology*, **81,** 17–42

Morris, N., Toms, M., Easthope, Y., and Biddulph, J. (1998). Mood and cognition in pregnant workers. *Applied Ergonomics*, **29,** 377–381

Morris, N. and Wickes, S.B.L. (2007). Mood enhancement by Indian head massage. *Journal of Holistic Health*, **4,** 32–37

Pond, G. (1998). *The fats of life*. (Cambridge University Press, Cambridge)

Thayer, R. (2001). *Calm Energy*. (Oxford University Press, New York)

REDUCED PHYSIOLOGICAL STRAIN IN FIREFIGHTERS USING A PRACTICAL COOLING STRATEGY

Dave Barr, Warren Gregson & Thomas Reilly

Research Institute for Sports and Exercise Science, Liverpool John Moores University. Henry Cotton Campus, Liverpool L3 2ET

The aim of this study was to establish whether a practical cooling strategy reduces the physiological strain during simulated firefighting activities. On two occasions, 12 male firefighters completed two bouts of simulated firefighting activities (separated by a 15-min recovery period) under live fire conditions. During the 15-min recovery period subjects were either cooled via application of an ice vest and hand/forearm water immersion (~19°C) or remained seated with no cooling (control). At the end of the recovery period core temperature and heart rate were significantly lower in the cooling condition compared with control. Core temperature remained significantly lower throughout bout 2 following cooling. There was also a trend for heart rate to be lower following cooling compared to the control condition ($p = 0.053$). We conclude that this practical cooling strategy is effective in reducing the physiological strain during simulated firefighting activities.

Introduction

Firefighters perform strenuous work under conditions of high environmental heat strain (Romit and Frim, 1987; Ilmarinen *et al.*, 1997; Smith *et al.*, 2001). The combined effects of such activity and the addition of protective clothing utilized by firefighters lead to high degrees of thermoregulatory and cardiovascular strain (Rossi, 2003). Reducing this physiological strain in firefighters is therefore important since elevations in body temperature may hinder both physical and mental performance (Hancock *et al*., 2007).

In order to meet the demands of firefighting activity there is a need for practical cooling strategies which possess sufficient capacity to reduce thermal stress in relative short periods of time. Hand and forearm immersion (House *et al*., 1997; Selkirk *et al.*, 2004) and ice vests (Bennett *et al.*, 1995) have been shown independently to be effective in reducing physiological strain during single or repeated bouts of moderate intensity firefighting activity under laboratory conditions. However, little research to date has examined the influence of cooling strategies on the physiological responses during repeated bouts of firefighting activity that more closely simulate the highly demanding conditions that firefighters may be exposed to (Ilmarinen *et al.*, 1997; Ilmarinen *et al.*, 2004). Under these conditions more aggressive cooling strategies may be required to maintain the degree of physiological

strain within safe limits. The aims of the present study were therefore to evaluate whether hand and forearm immersion combined with the use of an ice vest is effective in aiding physiological recovery following a strenuous bout of firefighting activity under live fire conditions and to explore whether any benefits remained during a subsequent bout of work.

Methods

Measurements of core temperature, heart rate, thermal sensation, rating of perceived exertion (RPE), recovery quality and sweat loss were taken on two occasions in 12 male firefighters (age 42 ± 7 years, body mass $= 86.9 \pm 11$ kg, height $= 1.78.1 \pm 0.1$ m, maximal oxygen uptake $= 43.32 \pm 5.4\,\mathrm{ml^{-1}kg^{-1}min^{-1}}$ and body fat of $18.46 \pm 3.7\%$) while performing two bouts of simulated firefighting search and rescue activities in a building containing live fires. Ambient temperatures in the building ranged from 46°C to 574°C. Each bout was separated by a 15-min recovery period during which one of two conditions was administered, Cooling (wearing an ice vest while immersing the hands and forearms in water at ~19°C), and Control (seated rest). During each recovery period the firefighters removed their tunics, helmets, anti-flash hoods, gloves and self-contained breathing apparatus, consumed a controlled amount of water at room temperature ($5\,\mathrm{ml \cdot kg^{-1}}$ body weight) and rested at ~25°C. This study was approved by the Research Ethics Committee at Liverpool John Moores University.

A two-way (condition *x* time) analysis of variance (ANOVA) with repeated measures was used to analyse data for core body temperature, heart rate and recovery quality. Student T-tests were used to analyse data for perceived exertion and thermal sensation. Since trials varied in duration, core temperature and heart rate data were normalized with respect to time by expressing them as a percentage of total trial duration.

Results

Core temperature significantly increased during bout 1 ($F_{1.19,13.16} = 147.01$; $p < 0.01$) and bout 2 ($F_{1.78,19.60} = 233.58$; $p < 0.01$) in both conditions. Core temperature at the start of the recovery period was similar under both conditions (Figure 1). A more pronounced decrease in core temperature was observed during the 15-min recovery period in the Cooling (0.05°C $\mathrm{min^{-1}}$) condition compared with Control (0.01°C· $\mathrm{min^{-1}}$ $F_{1.83,20.13} = 33.45$; $p < 0.01$). At the end of the recovery period core temperature was lower with Cooling compared to Control (37.73 ± 0.27°C vs 38.07 ± 0.26°C). Core temperature remained consistently lower during the second bout of activity in the Cooling condition compared with Control ($F_{1,11.} = 16.27$; $p < 0.01$).

Heart rate increased significantly during bout 1 ($F_{3.41,23.86} = 17.04$; $p < 0.01$) and bout 2 ($F_{3.39,23.75} = 25.06$; $p < 0.01$) under both conditions. Peak heart rates during bout 1 were 166 ± 14 beats $\mathrm{min^{-1}}$ and 169 ± 16 beats $\mathrm{min^{-1}}$ in the Cooling

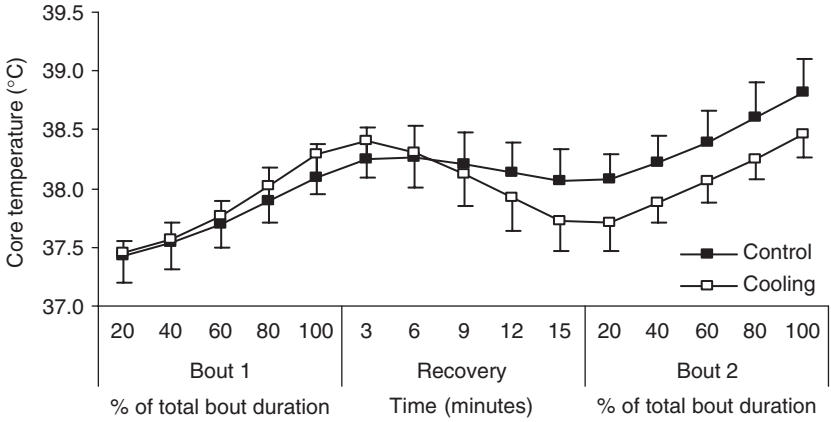


Figure 1. Core temperature responses during repeated bouts of firefighting activity under both Cooling and Control conditions ($n = 412$; mean ± s).

and Control conditions, respectively. Heart rate decreased significantly during the recovery period in both conditions ($F_{3.31,23.21} = 23.28$; $p < 0.01$). This decrease was significantly greater in the Cooling condition compared to Control ($F_{1,7} = 23.45$; $p < 0.01$). Heart rates at the end of the recovery period were 91 ± 11 and 104 ± 11 beats min^{-1} in the Cooling and Control condition, respectively. During bout 2 there was a tendency for heart rate to be lower following Cooling ($F_{1,7} = 5.42$; $p = 0.053$). Peak heart rates during bout 2 were 168 ± 13 beats min^{-1} and 180 ± 14 beats min^{-1} in the Cooling and the Control conditions, respectively. In the Control condition peak heart rate in bout 2 (180 ± 14 beats min^{-1}) was significantly higher than in bout 1 (169 ± 16 beats min^{-1}) ($T_{(7)} = -2.72$; $p = 0.2$). In contrast, in the Cooling condition, there was no significant difference between bout 1 (166 ± 14 beats min^{-1}) and bout 2 (168 ± 13 beats min^{-1}) in peak heart rates ($T_{(7)} = -1.02$; $p = 0.34$).

There was no significant differences in RPE between conditions during bout 1 ($T_{(11)} = 1.75$; $p = 0.24$) or bout 2 ($T_{(11)} = 0.22$; $p = 0.83$). The firefighters perceived bout 1 and bout 2 in both conditions to be 'somewhat hard' and 'very hard' respectively. Thermal sensation increased by a similar magnitude during bout 1 in both conditions ($T_{(11)} = 1.75$; $p = 0.11$), and was regarded as being 'hot'. Prior to bout 2, thermal sensation was lower following Cooling compared to Control ($T_{(11)} = 2.80$; $p = 0.017$). The firefighters perceived themselves to be 'neutral' in the Cooling condition compared to feeling 'warm' to the Control. The firefighters' perceived thermal sensation was similar during bout 2 in both conditions. The firefighters perceived bout 2 to be 'very hot' in both conditions. The firefighters' perception of the quality of their recovery improved significantly ($F_{2.,22} = 46.73$; $p < 0.01$) during the recovery period and was similar between conditions ($F_{1.,11} = 46.73$; $p = 0.53$). Both conditions were perceived to be a 'good recovery'. There was no significant difference in the changes in body mass between cooling and control conditions ($T_{(11)} = 0.46$; $p = 0.65$). The firefighters lost a total of 1.50 ± 0.29 and 1.44 ± 0.42 litres of sweat in the Cooling and Control conditions, respectively.

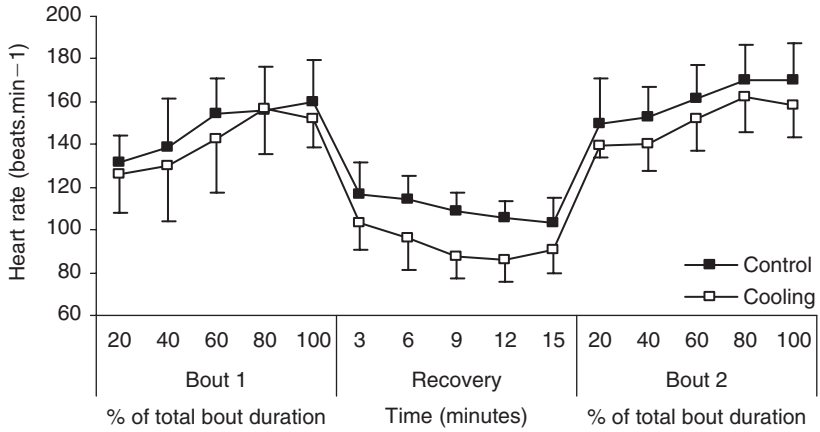


Figure 2. Heart rate responses during repeated bouts of firefighting activity under both Cooling and Control conditions ($n = 8$; mean ± s).

Discussion

The results of the present study confirm previous findings which suggest that firefighter search and rescue activities under hot environmental conditions induce marked thermoregulatory and cardiovascular strain (Romit and Frim, 1987; Smith *et al.*, 2001). The 15-min recovery period in the control condition failed to reduce the magnitude of thermoregulatory strain as indicated by the lack of reduction in core temperature. This finding was observed irrespective of the fact that the firefighters ingested fluids during the recovery period, demonstrating that the current strategy employed by firefighters of seated rest and drinking water is an ineffective means of reducing the physiological strain that occurs during firefighting activities. This observation supports previous investigations that have employed a similar passive recovery period between repeated bouts of strenuous firefighting activity in the heat (Ilmarinen *et al.*, 1997; Ilmarinen *et al.*, 2004).

The cooling strategy used in the present investigation decreased core temperature by 0.7°C during the 15-min recovery period. This reduction is greater than values previously reported for hand and forearm immersion alone (0.3°C decline in rectal temperature) following a single bout of activity (House *et al.*, 1997 and Selkirk *et al.*, 2004). This suggests that cooling via application of an ice vest in conjunction with hand and forearm immersion provides a practical and effective means of mediating decrements in physiological strain during repeated bouts of strenuous firefighting activity. Previous studies have shown that wearing an ice vest per se promotes marked decrements in thermal strain during a single bout of firefighting activity in the heat. Bennett *et al.* (1995) reported that wearing an ice vest for 40 min prior to exercise led to a 0.7 °C reduction in rectal temperature following 30 min of moderate intensity exercise in the heat. Despite ice vests clearly providing physiological benefits to firefighters working in the heat, the cooling strategy in the present study may offer a more practical solution during repeated bouts of strenuous firefighting

activity in real life situations since the timing of incidents is unknown and relatively short periods of preparation may only be permitted prior to initial entry into the heat.

In the present investigation a significantly lower heart rate was observed following the recovery period in the Cooling condition (91 ± 11 beats min^{-1}) compared to the Control (104 ± 11 beats min^{-1}). Selkirk *et al.* (2004) and Barr *et al.* (2007) have previously reported similar finding under laboratory conditions. Lower heart rate during cooling has been reported to correspond with a reduced skin temperature (Selkirk *et al.,* 2004; Barr *et al.,* 2007). These changes in skin temperature reduce the need for peripheral blood flow, thus lowering the cardiovascular strain during any subsequent work that may occur (Rowell, 1974).

The physiological benefits of this cooling strategy in the present study remained evident throughout the second bout of firefighting activity with a consistently lower core temperature observed in the Cooling condition. There was also a tendency for heart rate to be lower throughout the bout in the Cooling condition compared with Control. These results support findings of our research group (Barr *et al.,* 2007) and other studies (House, 1996; Selkirk *et al.,* 2004) performed under laboratory conditions. Cooling reduced peak heart rate during bout 2 to a level similar to peak heart rate in bout 1. This finding was not observed in the Control, suggesting that this cooling strategy is effective at reducing the cardiovascular strain during a subsequent bout of firefighting activities compared to water intake and seated rest.

The firefighters' perception of thermal sensation was significantly lower at the end of the recovery period following Cooling compared with Control. However, no differences in thermal sensation were apparent during the second bout. The firefighters perceived this bout to be 'very hot' in both conditions. There was no difference in the firefighters' perception of recovery between conditions. This lack of difference in these two indices could be due to the restricted range and therefore sensitivity of this scale to detect differences.

Conclusion

It is concluded that a 15-min recovery period of hand and forearm immersion combined with wearing an ice vest accelerates the thermoregulatory and cardiovascular recovery following a bout of simulated firefighting search and rescue activities in a building containing live fires. This cooling strategy also reduces the thermoregulatory and cardiovascular strain in firefighters during redeployment of firefighting search and rescue activities. This suggests the present cooling strategy is effective in promoting marked reductions in physiological strain during extreme occupational challenges regularly experienced by firefighters.

References

Barr, D., Gregson, W. and Reilly, T. 2007, Evaluation of a practical cooling strategy for reducing the physiological strain in firefighters performing work in the heat.

Book of Abstracts, 6th Annual Conference of Sports, Leisure and Ergonomics. Burton Manor: 14th to 16th November 2007 (Liverpool John Moores University, Research Institute for Sport and Exercise Sciences) 7.

Bennett, B.L., Hagan, R.D., Huey, K.A., Minson, C. and Cain D. 1995, Comparison of two cooling vests on heat strain while wearing a firefighting ensemble. *European Journal of Applied Physiology,* **70**, 322–328.

Hancock, P.A., Ross, J.M. and Szalma, J.L. 2007, A meta-analysis of performance response under thermal stressors. *Human Factors,* **49,** 851–77.

House, J.R. 1996, Reducing heat strain with ice-vests or hand immersion In Y. Shapiro, Y. Epstein and D. Moran. (eds). Proceedings of the 7th International Conference on Environmental Ergonomics, Jerusalem, Israel October. 1996. 347–350.

House, J.R., Holmes, C. and Allsopp, A.J. 1997, Prevention of heat strain by immersing the hands and forearms in water. *Journal of Royal Naval Medical Services,* **83**, 1, 26–30.

Ilmarinen, R., Lindholm, H., Koivistionen, K. and Hellsten, P. 2004, Physiological evaluation of chemical protective suit systems (CPSS) in hot conditions. *International Journal of Occupational Safety and Ergonomics*, **10,** 215–226.

Ilmarinen, R., Louhevaara, V., Griffins, B. and Kunemund, C. 1997, Thermal responses to consecutive strenuous fire-fighting and rescue tasks in the heat. In Y. Shapiro, D.S. Moran and Y. Epstein. (eds). Environmental Physiology, Recent Progress and New Frontiers. London. 295–298.

Romit, T.T. and Frim, J. 1987, Physiological responses to firefighting activities. *European Journal of Applied Physiology,* **56**, 633–638.

Rossi, R. 2003, Firefighting and its influence on the body. *Ergonomics,* **46**, 1017–1033.

Rowell, L.B. 1974, Human cardiovascular adjustments to exercise and thermal stress. *Physiological Reviews,* **54**, 75–159.

Selkirk, G.A., McLellan, T. and Wong, J. 2004, Active versus passive cooling during work in warm environments while wearing firefighter's protective clothing. *Journal of Occupational and Environmental Hygiene,* **1**, 521–531.

Smith D.L., Manning T.S. and Petruzzello S.J. 2001, Effect of strenuous live-fire drills on cardiovascular and psychological responses of recruit firefighters. *Ergonomics,***44,** 244–54.

METHODS AND APPLICATIONS OF BODY COMPOSITION ANALYSIS

Thomas Reilly & Laura Sutton

Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Henry Cotton Campus, Liverpool L3 2ET

A range of methods is available for analysis of body composition. These vary from anthropometric measures taken from the body surface to sophisticated imaging systems. The choice of method depends on resources and training of personnel available, use in laboratory or field context and purpose of data collection or examination. The relevance of these measures to health in occupational settings is readily accepted: the consequences for performance have not been elaborated.

The aims of the workshop are:- i) to consider the validity, reliability and relevance of indirect methods of body composition analysis; ii) to illustrate applications of body composition analysis in occupations that include fire fighting, military and professional sport and iii) to highlight limitations of commonly used indices, including skinfold thicknesses, waist-to-hip ratio and body mass index.

Methods include anthropometry, composite indices, densitometry, bio-electric impedance and dual-energy x-ray absorptiometry. The error associated with each approach is addressed.

The methods are illustrated by application to specific populations that include females of different activity levels, firemen, soldiers and professional athletes. The uses of indices for determining overweight in population studies are considered and their limitations addressed.

Introduction

Analysis of body composition is relevant in a range of occupational contexts. Body space requirements may change with an increase in body mass, whether the individual is in a static or dynamic posture. Extra mass as fat acts as dead weight when the body is lifted against gravity as occurs in ladder climbing or locomotion. Added muscle mass is effective in tasks requiring high levels of force. Body composition is relevant also in health-related disorders that contribute to absences from work on a substantial scale. Composite measures such as body mass index (BMI) and waist-to-hip ratio (WHR) have gained widespread acceptance in the public health domain as predictors of cardiovascular risk and predisposing to other morbidities.

A range of methods is also available for the analysis of body composition. Techniques vary from anthropometric variables obtained from the body surface

to sophisticated imaging systems for the whole body. The choice of method may depend on the availability of resources and trained personnel whether in a laboratory or field setting and the purpose of data collection. A further consideration is the number of body compartments being examined, either two, three or four body compartments being distinguished depending on the model adopted.

The aims in this workshop are to:-

i) consider the validity, reliability and relevance of indirect methods of body composition analysis;
ii) illustrate applications of body composition analysis in different groups;
iii) highlight the limitations of commonly used indices, including skinfold thickness, WHR and BMI.

Methods include anthropometry, composite indices, densitometry, bio-electric impedance and dual energy x-ray absorptiometry (DXA). The error associated with each approach should be addressed.

The methods are illustrated by application to specific populations that include females of different activity levels, firemen, soldiers and professional athletes. The use of indices for determining overweight in population studies are considered and their limitations addressed.

Validity and reliability

Approaches to body composition analysis have been divided into anatomical or structural and chemical models. The anatomical approach epitomised by cadaver analysis is the one direct method. Chemical approaches are indirect since they are based on quantitative assumptions. For example, dilution methods for determining body water content are based on assuming a fixed relationship between lean body mass and its water content. Similarly, potassium counting (K^{40}) to determine lean body mass is based on an assumed relationship between cellular potassium isotopes and lean tissue. Furthermore, the method of hydrodensitometry, long accepted as a reference method for comparison with novel approaches, is indirect since it is based on densitometry. Assumptions are that the density of the fat compartment and the fat-free compartment each has a value that is constant. Hence any new method derived from emerging technology that is validated against densitometry is doubly indirect.

Comparison of methods

Dual-energy x-ray absorptiometry is fast becoming the preferred method used in analysing body composition. The approach leads to a division of body compartments into fat, bone and fat-free cum bone-free compartments, the latter being described in commercial software as lean mass. Regional analysis is also provided so that distribution of fat or specific skeletal weaknesses in bone may be identified. Although originally designed for assessment of skeletal health and

Table 1. Levels of analysis in descriptions of body composition.

Level 1	Level 2	Level 3
Cadaver analysis	Hydrodensitometry Dual-energy x-ray absorptiometry Medical imaging techniques Dilution methods Potassium counting	Anthropometry Bio-electric impedance Infra-red interactance

Table 2. Correlations (r) between different methods of measuring percent body fat. Dual energy x-ray absorptiometry (DXA) is used as reference, both total body and sub-total (without the head): n = 84.

Method	DXA total body	DXA Subtotal
Hydrodensitometry (Siri equation)	0.873*	0.910*
Hydrodensitometry (Brozek equation)	0.904*	0.874*
Bioelectric impedance	0.673*	0.683*
Skinfold thickness (Durnin and Wormersly, 1974)	0.796*	0.835*

* indicates $p < 0.001$

risk of osteoporosis, the technique has been promoted as the standard method for determining percent body fat. The radiation dose experienced by the subject is very low and the time for measurement is less than 5 min. The expense of the machine is high which limits its routine application in occupational or sports settings.

Hydrodensitometry was long considered the nearest to a 'gold standard' method, being the reference for classical anthropometric equations such as that of Durnin and Womersley (1974). Despite criticisms of applying densitometry to athletic groups who may have higher bone density and greater muscle mass than average, there is close agreement and strong correlation with measures derived using DXA. Wallace et al. (2006) reported a correlation coefficient of $r = 0.974$ but recommended that residual volume should be measured rather than predicted from vital capacity when hydrodensitometry is employed.

In a further comparative study, bioelectric impedance and skinfold thicknesses were included in an investigation of the relationship between methods. Whilst the various techniques are highly correlated, variability is highest when bioelectric impedance is utilised (see Table 2). The correlation coefficient between percent body fat determined from the sum of four skinfold sites (Durnin and Womersley, 1974) was $r = 0.89$ (Egan et al., 2004) in professional soccer players. The relationship is improved when lower limb sites are included in the prediction equation. This observation supports the recommendation of the British Olympic Association to include the anterior thigh among five skinfolds when access to DXA or hydrostatic weighing facilities is unavailable (Reilly et al., 1996). Anthropometry does appear to be a suitable tool for use in a field setting, provided the assessor is

Table 3. Bone mineral density (g.cm^2) of part time female soldiers (mean ± S.D).

Group	Bone density
Commanders	1.15 ± 0.06
Recruits	1.11 ± 0.06
Sedentary controls	1.11 ± 0.08

Table 4. Bone mineral density and percent body fat for different groups (mean ± SD).

	N	Age (years)	BMD (g.cm^2)	% body fat
Soccer players	35	25.9 ± 42	1.467 (± 0.116)	10.9 ± 1.7
Soccer referees	6	37.5 ± 4.7	1.370 ± 0.34	18.9 ± 3.7
Firefighters	23	41 ± 6	1.302 ± 0.094	20.1 ± 3.8
Soldiers	14	25.9 ± 5.3	1.282 ± 0.126	12.3 ± 2.9
Reference group	24	26.8 ± 5.2	1.292 ± 0.088	17.5 ± 3.6
Wheelchair athletes	10	26.2 ± 7.6	1.160 ± 0.134	21.5 ± 5.4

suitably trained. It would appear also that omission of the head is appropriate since the assumptions about fat content in this area introduce error in use of DXA with both lean and overweight subjects.

A further question is the value of using composite anthropometric variables such as BMI and WHR. The former has been criticised on the basis that an individual with a large muscle mass would be interpreted as being overweight. Nevertheless, in a sample of female subjects drawn from the general population, BMI values were found to be highly correlated with percent body fat as determined using DXA. Both waist and hip girths were correlated separately with percent body fat suggesting that an alternative use of the raw data with correction for body size should improve utility rather than rely exclusively on WHR (Wallace, 2007).

Applications

Body composition analysis has been applied to a number of sports specialisations and occupational workers. In the vast majority of cases the analysis was restricted to assessments of BMI, WHR or skinfold thicknesses. Only recently has DXA been used in the characterisation of different groups with a three-compartmental model.

The data shown in Table 2 imply that military training is effective in increasing bone mineral density. It seems that the impact loading associated with the training of female soldiers has positive benefits for skeletal health. This effect is evident despite the potential consequences of long marches and such activities on risk of stress fracture.

In Table 4, bone density and percent fat are compared for firefighters, soldiers, wheelchair athletes and professional soccer players (Sutton et al., 2007)

and soccer referees (Reilly and Gregson, 2006). The soccer players are clearly separated from the other groups, both in bone mineral density and percent body fat. The occupational groups are within the normal range for body composition in all compartments.

Conclusion

In the choice of method for analysing body composition, the purpose of the investigation should be a deciding factor. The resources available should also determine the method adopted. Body composition variables are employed for characterisation of subject groups, as dependant variables in intervention studies and as predisposing factors in epidemiological investigations. In ergonomics contexts, volume and surface area may also be relevant. Whilst techniques for body mapping are continuously being improved, the sensitivity of the method adopted and its validity are critical considerations.

References

Durnin, J.V.G.A. and Womersley, J. 1974, Body fat assessment from total body density and its estimation from skinfold thickness measurements on 481 men and women aged 16 to 72 years. *British Journal of Nutrition*, **320**, 77–97.

Egan, E., Reilly, T., Chantler, P. and Lawler, J. 2005, Body composition before and after six weeks pre-season training in professional football players. In M. Marfell-Jones, A. Stewart and T. Olds (eds) *Kinanthropometry*, (London: Routledge), pp. 123–130.

Reilly, T. and Gregson, W. 2006, Special populations: the referee and assistant referee. *Journal of Sports Sciences*, **24**, 795–801.

Reilly, T., Gregson, W. and Barr, D. 2007, Evaluation of an ergonomics interaction to reduce the physiological stress in firefighters. Report to Merseyside Fire and Rescue Services, Liverpool John Moores University.

Reilly, T., Maughan, R.J. and Hardy, L. 1996, Body fat consensus statement of the Steering Groups of the British Olympic Association. *Sports Exercise and Injury*, **2**, 46–49.

Sutton, L., Wallace, J., Scott, M. and Reilly, T. 2007. Body composition of international and club-level professional soccer players measured by dual energy x-ray absorptiometry. Communication to Sixth International Conference on Sport, Leisure and Ergonomics, November, Burton Manor.

Wallace, J. 2007, The impact of exercise on human body composition determined using dual-energy x-ray absorptiometry. Unpublished PhD thesis, Liverpool John Moores University.

Wallace, J., George, K. and Reilly, T. 2006, Validity of dual-energy x-ray absorptiometry for body composition analysis. In P. Bust (ed.) *Contemporary Ergonomics*, (London: Taylor and Francis), pp. 513–515.

VIRTUAL ELECTRONIC GAME PLAYING BY CHILDREN CAN BE ACTIVE

Leon Straker[1], Clare Pollock[2], Jan Piek[2], Rebecca Abbott[3] & Jemma Coleman[1]

[1]*School of Physiotherapy, Curtin University of Technology, Perth, Australia*
[2]*School of Psychology, Curtin University of Technology, Perth, Australia*
[3]*School of Human Movement Studies, The University of Queensland, Brisbane, Australia*

The majority of children in affluent communities now play electronic games. This has lead to concerns about increased inactivity and the poor health sequelae. Traditional electronic games using gamepad, keyboard and mouse input have been considered sedentary, but newer virtual game interfaces require active movement. The physical activity demands of these active-input games are unknown. This study compared the movement, muscle activity, energy expenditure and heart rate of 20 children aged 9–12 years playing electronic games with traditional input devices and with a webcam motion analysis (Sony EyeToy®) device. Using traditional input devices children were usually as sedentary as watching a DVD. In contrast, using the active-input virtual gaming device resulted in substantial increases in all measures. These results suggest playing traditional electronic games is indeed a sedentary activity and may thus exacerbate current trends for obesity and other inactivity related disorders. In contrast the new active input virtual technologies offer the potential of allowing children to enjoy playing electronic games without becoming more sedentary.

Introduction

Children in affluent communities have substantial exposure to computers and electronic games at home or at school and use is growing rapidly (Roberts *et al.*, 2005). 92% of Australian 5–14 year olds have used a computer (ABS, 2006) and more than 80% children in the USA have access to a computer and to video games (Roberts *et al.*, 2005). A meta analysis of studies in affluent countries showed mean computer use by children and adolescents of 30 minutes a day in addition to 40 minutes a day electronic game playing on other devices (Marshall *et al.*, 2006). Game playing is the most frequent computer activity for children, with 82% of boys and 59% of girls reporting playing computer games (ABS, 2006).

In a recent review we (Straker and Pollock, 2005) reported that the available evidence suggested computer use targeted on learning areas is associated with enhanced academic achievement (e.g. Harrison *et al.*, 2002), but that electronic game playing in particular may have a negative effect on school achievement (Moseley *et al.*, 2001). Similarly we found that game-related discourse may provide a stimulus for children's social development (Heft and Swaminathan, 2002), although there are concerns about the potential negative effects of violence in electronic games (Villani, 2001).

However, there has been increasing concern over the potential physical health impact of the time children spend playing electronic games. The concerns relate to the impact on posture, muscle activity and related discomfort (Harris and Straker, 2000; Straker *et al.*, 2002) and to suggestions that time spent using computers and playing electronic games not only displaces time spent being physically active (Straker *et al.*, 2006), but may also increase 'inactive' or sedentary time. Whilst the current evidence only supports a weak negative relationship between computer/electronic game use and physical activity (Marshall *et al.*, 2004; Straker *et al.*, 2006), increasing use of computers and electronic games is widely blamed for the rise in body fatness in children (Vandewater *et al.*, 2004).

Yuji (1996) reported evidence that electronic games improved children's fine motor performance. In contrast, gross motor experiences are usually associated with movement of the limbs and torso and there are concerns that computer/electronic game use reduces children's gross motor physical activity (Straker and Pollock, 2005).

Precisely how sedentary children are when playing electronic games is unknown (Vandewater *et al.*, 2004). Watching TV requires no movement and has been used in studies of children (Abbott *et al.*, 2002) to obtain 'resting' levels. Electronic games have traditionally used keyboard/mouse and game pad interfaces which require very little movement. Our research on children's use of computers for educational tasks using keyboard and mouse showed only small movements and low levels of muscle activity (Briggs *et al.*, 2004; Greig *et al.*, 2005) but details of movement and muscle activity during game playing are not available.

In contrast, some new virtual electronic games require 'active-input', and utilize a motion camera (Sony EyeToy®) to capture the user's image and movement and embed this into the virtual game environment. The games require the user to touch or avoid virtual objects.

Recently, three studies have reported energy expenditures whilst children played domestic electronic games. Wang and Perry (2006), Lanningham-Foster *et al.* (2006), and Maddison *et al.* (2007) reported heart rates and energy expenditures whilst children played a traditional input game to be only slightly greater than resting values. In contrast, energy expenditures whilst playing the active input games were found to be greatly increased. These physiological data suggest some electronic games can require substantial movement and muscle activity.

Interventions to reduce the time that children spend being sedentary with electronic media have to date had limited success at a population level; the new active-input virtual games technology may be a more successful approach.

Method

A within-subjects design was used to compare the movement, muscle activity, energy expenditure and heart rate of children using traditional electronic game devices and an active input virtual game device.

Twenty healthy children (12 male, eight female) between the ages of 9 and 12 years were recruited through personal contacts and advertisements placed in community media. All children had played electronic games, but were not experienced with the games provided in this study. Volunteers were excluded if they had a diagnosed disorder likely to impact their movement or electronic game use. The mean (SD) height and weight for males was 141.8 (10.9) cm and 38.0 (9.3) kg, and for females was 141.0 (8.7) cm and 31.6 (4.6) kg. The study was approved by the Human Research Ethics Committee of Curtin University of Technology.

Each child watched a DVD and played games using five different game devices. A Tetris® style puzzle game of cascading blocks was played on the Handheld game device. A car racing game was played using a Gamepad or Keyboard or steering Wheel with pedals connected to a PlayStation 2® (Sony Corporation, Tokyo, Japan) and displayed on a 30 inch LCD screen. The same screen and PlayStation 2® console were also used to run the EyeToy® device. The game used with the EyeToy was called Cascade and required moving hands and/or feet to touch virtual targets shown on the screen.

Heart rate and energy expenditure was recorded using a MetaMax 3B® system (Cotex Biophysik, Leipzig) incorporating a face mask, volume transducer and O_2 and CO_2 sensors.

Motion data were collected using a seven camera Peak Motus® 3D Optical Capture System (Peak Performance Technologies Inc., Centennial, CO, USA). Spherical or semi-spherical, reflective markers were positioned on a cap, right posterior acromial shelf, the midpoint between the right radial and ulnar styloid processes, right thumbnail, over the sacrum and right 3rd metatarsal head. Data were sampled at 50 Hz, then filtered and smoothed using a Butterworth filter (cutoff frequency 4 Hz). The distance travelled by each marker was calculated from the 3D coordinates.

Surface myoelectric activity (sEMG) was collected from right side cervical erector spinae, upper trapezius, anterior deltoid, wrist extensor bundle, lumbar erector spinae and rectus femoris muscles. The electrode sites were prepared by shaving, lightly abrading and cleaning with surgical spirits and pairs of 12 mm diameter Ag-AgCl disposable surface electrodes were placed with 25 mm centre-to-centre distance. Raw sEMG signals were sampled at 1000 Hz via an eight channel AMT-8 EMG cable telemetry system (Bortec Biomedical, Alberta, Canada) with analogue differential amplifiers.

On arrival to the climate and lighting controlled motion analysis laboratory parents and children were informed about the study and provided written consent/assent. Children were then fitted with EKG electrodes, sEMG electrodes, reflective markers, expired gas mask and harness. Following quality control of measures and calibration checks, participants moved to the study area and watched a DVD for 5 minutes. Each child then proceeded to play electronic games in the following order: handheld, gamepad, keyboard, wheel, EyeToy. Children

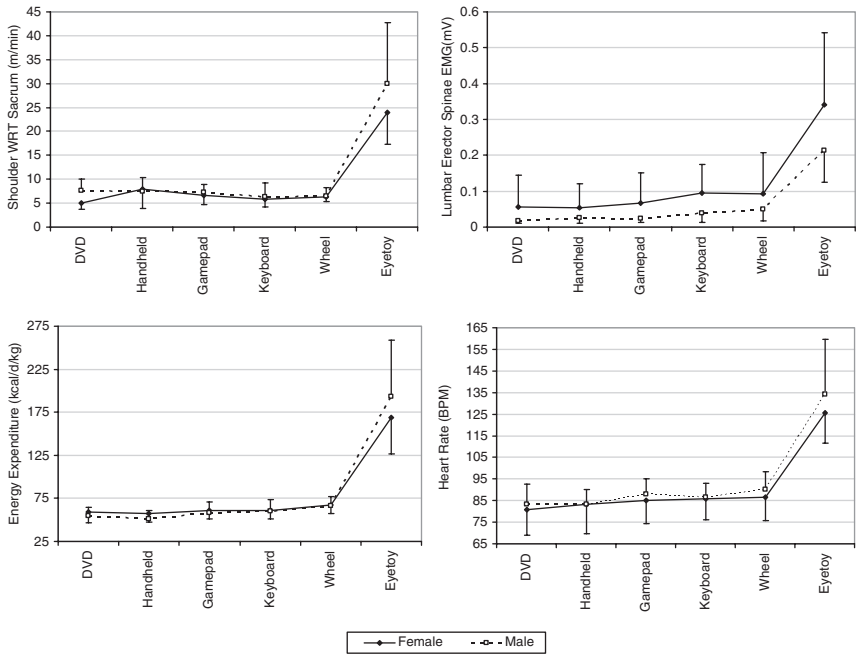


Figure 1. Trunk movement, trunk muscle activity, energy expenditure and heart rate during game play by children.

played with each device for 5 minutes or until a steady physiological state was achieved.

A multivariate mixed model analysis of variance (MANOVA) with device as the within subjects factor and gender as the between subjects factor was performed. Huynh-Feldt epsilon corrections were used as Mauchly's test indicated a lack of sphericity. Within-subjects contrasts were used to compare between devices. A critical alpha level of 0.01 was used. Missing values were replaced with condition means. All analyses were performed with SPSS v13.0 (SPSS Inc., Chicago, USA).

Results

Figure 1 shows the clear trends for similar levels of trunk movement, trunk muscle activity, energy expenditure and heart rate across the traditional devices (DVD, handheld, gamepad, keyboard and wheel) with a marked increase for EyeToy. Univariate tests found a significant effect of device for trunk movement, trunk muscle activity, energy expenditure and heart rate ($p < .001$).

Discussion

These results have shown that children are indeed sedentary when interacting with most electronic game devices. Given the increasing use of electronic games this

gives credit to the concerns expressed regarding the increasing sedentary activity of children and the effect this might have on their health. However, this study has demonstrated that the new 'active-input' EyeToy device resulted in substantial increases in movement, muscle activity, energy expenditure and heart rate. Observed energy expenditure and heart rate with EyeToy were similar to those reported for fast walking, jogging, soccer and basketball. These results suggest active-input virtual electronic games may be useful in increasing overall physical activity and thus developing stronger musculoskeletal systems and better gross motor skills, and reducing obesity. This would indicate that there would be considerable advantages in ergonomists recommending to parents to encourage their children to switch from the traditional sedentary games to new active games. Children in affluent communities are immersed in electronic games and these data suggest exciting possibilities that children might be able play electronic games whilst gaining positive health benefits.

References

Abbott, R.A., Harkness, M. and Davies, P.S. 2002, Physical activity in children and its relation to flow mediated dilation of the brachial artery, *Atherosclerosis*, **160**, 233–239

Australian Bureau of Statistics, 2006, Children's participation in cultural and leisure activities, Australia. Australian Bureau of Statistics, Canberra: ACT

Briggs, A., Straker, L. and Greig, A. 2004, Upper quadrant postural changes of school children in response to interaction with different information technologies, *Ergonomics*, **47**, 790–819

Greig, A., Straker, L. and Briggs, A. 2005, Cervical erector spinae and upper trapezius muscle activity in children using different information technologies, *Physiotherapy*, **91**, 119–126

Harris, C. and Straker, L. 2000, Survey of physical ergonomics issues associated with school children's use of laptop computers, *International Journal of Industrial Ergonomics*, **26**, 337–346

Harrison, C., Comber, C., Fisher, T., Haw, K., Lewin, C., Lunzer, E., McFarlane, A., Mavers, D., Scrimshaw, P., Somekh, B. and Watlin, R., 2002, ImpaCT2: The impact of information and communication technologies on pupil learning and attainment. A report to the Department for Education and Skills. ICT in school research and evaluation series – No. 9. British Educational Communications and Technology Agency, London

Heft, T.M. and Swaminathan, S. 2002, The effects of computers on the social behavior of preschoolers, *Journal of Research in Childhood Education*, **16**, 162–174

Lanningham-Foster, L., Jensen, T.B., Foster, R.C., Redmond, A.B., Walker, B.A., Heinz, D. and Levine, J.A. 2006, Energy expenditure of sedentary screen time compared with active screen time for children, *Pediatrics*, **118**, 1831–1835

Maddison, R., Mhurchul, C.N., Jull, A., Jiang, Y., Prapavessis, H. and Rodgers, A. 2007, Energy expended playing video console games: An opportunity to increase children's physical activity?, *Pediatric Exercise Science*, **19**, 1–24

Marshall, S.J., Biddle, S.J.H., Gorely, T., Cameron, N. and Murdey, I. 2004, Relationships between media use, body fatness and physical activity in children and youth: a meta-analysis, *International Journal of Obesity,* **28**, 1238–1246

Marshall, S.J., Gorely, T. and Biddle, S.J.H. 2006, A descriptive epidemiology of screen-based media use in youth: A review and critique, *Journal of Adolescence,* **29**, 333–349

Moseley, D., Mearns, N. and Tse, H. 2001, Using computer at home and in the primary school: Where is the value added?, *Educational and Child Psychology,* **18**, 31–46

Roberts, D.L., Foehr, U.G. and Rideout, V., 2005, Generation M: Media in the lives of 8–18 year olds. The Henry J. Kaiser Family Foundation, Menlo Park, CA

Straker, L., Briggs, A. and Greig, A. 2002, The effect of individually adjusted workstations on upper quadrant posture and muscle activity in school children, *Work.,* **18**, 239–248

Straker, L. and Pollock, C. 2005, Optimising the interaction of children with information and communication technologies, *Ergonomics,* **48**, 506–521

Straker, L.M., Pollock, C.M., Zubrick, S.R. and Kurinczuk, J.J. 2006, The association between information and communication technology exposure and physical activity, musculoskeletal and visual symptoms and socio-economic status in 5-year-olds, *Child: Care, Health and Development,* **32**, 343–351

Vandewater, E.A., Shim, M.S. and Caplovitz, A.G. 2004, Linking obesity and activity level with children's television and video game use, *Journal of Adolescence,* **27**, 71–85

Villani, S.S. 2001, Impact of media on children and adolescents: a 10-year review of the research, *Journal of the American Academy of Child & Adolescent Psychiatry,* **40**, 392–401

Wang, X. and Perry, A.C. 2006, Metabolic and physiologic responses to video game play in 7- to 10-year-old boys, *Archives of Pediatrics & Adolescent Medicine,* **160**, 411–415

Yuji, H. 1996, Computer games and information-processing skills, *Perceptual and Motor Skills,* **83**, 643–647

MILITARY LOAD CARRIAGE INJURIES AND DISCOMFORT: A QUESTIONNAIRE SURVEY

Stewart A. Birrell[1] & Roger A. Haslam[2]

[1]*School of Engineering and Design, Brunel University, Uxbridge, Middlesex, UB8 3PH*
[2]*Department of Human Sciences, Loughborough University, Leicestershire, LE11 3TU*

Injuries are recognised as the leading health problem for the military services. However, research into the effect of load carriage on injury rates and discomfort seems to be lacking in the published literature. Work conducted for this study distributed a questionnaire to 90 students at a defence 6th form college. The questionnaire collected injury and discomfort data from students with particular focus on two weeks of military style exercises conducted prior to completing the questionnaire. Results from the study showed that the lower back was rated as the most uncomfortable region of the body following a typical period of load carriage, with the shoulders second. Approximately 20% of participants reported at least one injury, with the lower back accounting for 53% of the injuries sustained. In addition persistent discomfort was experienced by 57% of participants, with the lower back accounting for just under half of these.

Introduction

Injuries in the military have been termed a hidden epidemic and are now recognised as the leading health problem for the military services (Jones *et al.*, 2000). Research specifically focusing on the effect of load carriage on injury rates in the military seems to be lacking in the published literature. This may be because load carriage is often viewed as being a non-modifiable, extrinsic risk factor for injury. In fact many of the most detailed literature reviews and epidemiological studies fail to include any load carriage variables as risk factors for injuries. To the present author's knowledge no study has attempted to evaluate the role that load carriage has to play in the development of injuries resulting from military exercises. Nor, assessed the incidence and prevalence of load carriage related injuries in the days following a period of load carriage. Studies which have been conducted focus on marching itself and do not distinguish between load carriage and marching related injuries (Knapik *et al.*, 1992; Reynolds *et al.*, 1999). A previous study conducted by Birrell and Hooper (2007) did evaluate load carriage discomfort following a 2-hour treadmill march, but not injury. This research showed that early development of shoulder pain as a result of load carriage may be a risk factor for severe pain or non-completion of a period of prolonged load carriage.

Table 1. Mean subjective responses to load carriage discomfort.

		Combined	Males	Females
Upper Limb	Shoulders	2.60 (0.8)	2.53 (0.8)	2.73 (0.8)
	Neck	2.36 (1.1)	2.20 (1.0)	2.67 (1.1)
	Hands/Arms	1.79 (1.0)	1.83 (1.0)	1.70 (0.9)
Back	Upper Back	2.37 (1.1)	2.23 (0.8)	2.63 (1.0)
	Lower Back	2.98 (0.9)	2.78 (1.0)	3.37 (1.0)
Lower Limb	Feet	2.17 (1.1)	2.10 (1.1)	2.30 (1.1)
	Ankles	1.70 (1.0)	1.63 (1.0)	1.83 (1.0)
	Knees	2.02 (1.1)	1.85 (1.0)	2.36 (1.2)
	Hips	1.98 (1.1)	1.87 (1.0)	2.20 (1.3)
Whole Body	–	2.30 (1.3)	2.10 (1.0)	2.43 (1.1)

Method

A load carriage questionnaire was utilised for the data collection by this study, it consisted of 30 questions, split up into six categories; general, upper limb, back, lower limb, blisters and other. This was completed by 90 student participants from Welbeck Defence 6th Form College in Loughborough, Leicestershire. Of the 90 students, 60 were males and 30 female. All students were members of the upper 6th, aged either 17 or 18 years old. The students who completed the questionnaire were not full-time soldiers, but all had good experience with military load carriage. Many of the students also had some affiliation with the military outside of college.

Results

Table 1 shows the mean comfort rating given for each zone of the body questioned (standard deviation in parenthesis), and an estimation of whole body discomfort. Ratings represent participant responses using a 5-point comfort rating scale, with 1 being comfortable and 5 being extremely uncomfortable. Significant differences were observed with the Wilcoxon signed-rank test showing that the shoulders and lower back were rated the most uncomfortable zones of their respective regions, also as a region the back was rated the most uncomfortable. The Kruskal-Wallis test revealed significant differences between the genders, with females reporting more discomfort in the back region, neck and knee compared to the males.

Table 2 shows total of 67 persistent discomforts were reported by 51 participants. The back region accounted for almost half of these with 31 complaints. A total of 17 injuries were reported, by separate participants, on the questionnaire. Again the lower back was the most common site for injury occurring, accounting for 59% of total injuries sustained.

Finally, table 3 shows responses given by participants when asked to rank, from least effect (1) to most effect (10), which aspects of load carriage they considered would most increase upper limb discomfort. Participants rated both speed of march and distance hauled as increasing discomfort the least compared to weight of

Table 2. Number of persistent discomfort and injuries reported by students.

	U Limb	Back	L Limb	Total	Participants
Discomfort	17	31	19	67	51
Injury	2	10	5	17	17

Table 3. Aspect of load carriage which most affects discomfort to the upper limb.

	Weight	Time	Distance	Speed	Gradient
Combined	7.13 (1.8)	6.96 (2.0)	5.69 (2.1)	5.07 (2.1)	6.61 (2.2)
Male	6.97 (1.8)	7.22 (1.9)	5.73 (2.0)	4.75 (2.0)	6.45 (2.1)
Female	7.47 (1.6)	6.43 (2.2)	5.60 (2.4)	5.70 (2.2)	6.93 (2.3)

load, time carried and gradient/terrain. Male participants reported speed as having significantly less detrimental effect on upper limb discomfort compared to females.

Discussion

Upper limb

The shoulders were rated as significantly more uncomfortable than other regions of the upper limb, and second most uncomfortable overall. Injury to the upper limb as caused by load carriage is not frequently reported; however, discomfort is very apparent (Birrell and Hooper, 2007). Direct pressure from the backpack straps can cause the restriction of blood flow and sensory loss to the arm, thus affecting a soldier's ability to aim or shoot the rifle. A long-term effect of this may be Rucksack Palsy. Despite shoulder discomfort as a result of load carriage being a significant issue, only 17% of participants stated they were concerned about the long-term implications of load carriage on the upper limb.

The questionnaire also asked students to rate out of 10 which aspect of load carriage most increased the upper limb discomfort they typically experience during load carriage. Table 3 shows that as one might expect weight of load had the effect of most increasing discomfort in the upper limb, with speed having the least effect. Results also show that weight, time and terrain were all rated as significantly increasing discomfort compared to both distance and speed. In response to these findings this research suggests that only two of the three aspects that most significantly increase upper limb discomfort should be used in conjunction during a load carriage training exercise. For example, a march could involve carrying heavy loads over challenging terrain for shorter periods of time, or a long march over challenging terrain with relatively low loads carried. This approach may go some way to reduce the inevitable discomfort and potential injury to the upper limb as a

result of load carriage. The author recognises that all the aspects of load carriage rated are essential for complete training programmes, and that during operations it may not be possible to adhere to these guidelines due to operational needs. However, during training exercises this theory could be employed. Further research will highlight which aspects are best used in combination to reduce, or minimise, upper limb discomfort during load carriage.

Back and lower limb

As stated previously the lower back was rated as significantly the most uncomfortable region of the body, with a mean comfort rating of uncomfortable, or 3 out of 5. Participants were also asked if they were concerned about the long-term implications carrying loads has on the back, 34% of participants stated they were. These results are of interest as the student participants questioned were only 17 to 18 years of age and as yet not full time military personnel. However, these results show that a third of student participants are already concerned about their backs. This is with good reason as Songer and LaPorte (2000) state that low back pain and knee injuries are the leading cause for lifetime compensation within the US military.

The region of the lower limb that showed the most typical load carriage discomfort when questioned were the feet, at 2.2 out of 5 (using the 5 point comfort scale). Following the feet were the knee and hip at around 2 out of 5, or slightly uncomfortable. The ankles were rated as the significantly most comfortable region of the lower limb, and body as a whole.

Male and female responses

Table 3 shows the general trend for females to rate all aspects of load carriage as having a greater effect on upper limb discomfort compared to males. This difference is only significant for the speed of marching. A reason for this may be the fact that marching pace is usually set by the males or commanding officers. This speed may be faster than females would typically self-select. To maintain walking speed females will either have to increase stride length and/or stride frequency. Research suggests that these increases put females at a greater risk of injury, in particular pelvis stress fractures (Pope. 1999), and will also increase general fatigue.

Table 1 also shows that in general females rated all regions of the body as more uncomfortable than the male participants questioned, with the exception of the hands and arms. These differences were significant at the upper and lower back, knee and neck. There are numerous potential reasons for this including the female physiology (reduced muscle mass and strength) and biomechanical (Q-angle and kinematic differences during load carriage).

Load carriage injury

There were a total of 17 injuries reported by different students, this accounts for approximately 20% of the participants who completed the questionnaire (table 2).

Of these 17 injuries two were to the upper limb, five to the lower limb and 10 to the back. The lower back was the most common site for an injury to occur with 9 injuries being reported, accounting for over half of the total injuries sustained. Eleven of the injuries required medical attention, from here a course of treatment was prescribed. Treatment ranged from rest to a hospital visit. Six injuries required further medical attention; this was either in the form of physiotherapy (5) or on one occasion a hospital visit. The injuries sustained lasted for a total of 231 days. In addition to the injuries sustained there were 67 cases of persistent discomfort, experienced by 51, or 57% of participants (table 2). Persistent discomfort to the back accounted for 46% of the cases, with the remainder split relatively evenly between the upper and lower limb.

Relating the findings from this study to those in the literature can be achieved, again it is worth noting that this group of participants were not full-time soldiers. Reynolds *et al.* (1999) found that of 36% of soldiers completing five consecutive days of 20 km road marching with load suffered one or more injuries, with 8% of soldiers unable to complete the exercise due to injury. Knapik *et al.* (1992) recorded injury rates following a 20 km strenuous road march. A total of 24% of soldiers sustained one or more injury, with half of these requiring medical attention. The second most common injury, behind blisters, were back complaints totaling 20% of injuries; 50% of soldiers who did not complete the march attributed this to back pain.

This study suggests that for young part-time military trainees load carriage related injuries affects 20% of participants. With injury to the lower back accounting for over half of the injuries sustained. Although an injury rate of 20% is lower than in the literature above, these studies counted blisters in with the number of injuries. Persistent discomfort as a result of load carriage affected 57% of participants, with the back accounting for just under half of these.

Other

Participants were asked whether they felt load carriage reduced their ability to perform either a physical task (e.g. obstacle course) or mental task (e.g. map reading) at the end of a march. Sixty-eight percent of participants thought that load carriage would restrict their ability to perform a physical task, this dropped to 18% when a mental task was considered. Similar results were also observed by Birrell and Hooper (2007). This again highlights that ergonomic research to improve load carriage system design aimed at making the carrying soldier more comfortable will have a positive effect on improving military task performance.

Almost two-thirds of participants felt that carrying load increased the risk of them tripping or falling compared to no load, with 44% saying it would increase the severity of any fall.

Conclusions

The lower back was rated as the most uncomfortable region of the body following a typical period of load carriage, the shoulders were second. Females reported both

the upper and lower back as significantly more uncomfortable during typical load carriage than the males. The same was also true for the knee and the neck. This study suggests that in order to minimise discomfort as a result of load carriage a training exercise should not encompass the following three conditions; heavy loads, long durations or difficult gradient and terrain. Of interest to this study were the incidence and prevalence of load carriage related injuries. Approximately 20% of student participants reported at least one injury of any sort. The most common site for a load carriage related injury was the lower back, accounting for 53% of the injuries sustained. There were also 67 cases of persistent discomfort, experienced by 57% of participants, accounting for just under half of these discomforts was the lower back.

Acknowledgements

I would like to thank the students and staff at Welbeck College who took time to distribute and complete the questionnaire. Also thanks to Dr Robin Hooper for his previous supervision.

References

Birrell, S.A. and Hooper, R.H. 2007, Initial load carriage subjective injury data collected by interview and questionnaire, *Military Medicine,* **172**, 306–311

Jones, B.H., Perrotta, D.M., Canham-Chervak, M.L., Nee, M.A. and Brundage, J.F. 2000, Injuries in the military: A review and commentary focused on prevention, *American Journal of Preventive Medicine,* **18**, 71–84

Knapik, J., Reynolds, K., Staab, J., Vogel, J. and Jones, B. 1992, Injuries associated with strenuous road marching, *Military Medicine,* **157**, 64–67

Pope, R. 1999, Prevention of pelvic stress fractures in female army recruits. *Military Medicine,* **164**, 370–373

Reynolds, K., White, J., Knapik, J., Witt, C. and Amoroso, P. 1999, Injuries and Risk Factors in a 100-Mile (161-km) Infantry Road March, *Preventive Medicine,* **28**, 167

Songer, T.J. and LaPorte, R.E. 2000, Disabilities due to injury in the military, *American Journal of Preventive Medicine,* **18**, 33–40

RAIL ERGONOMICS

IMPROVING RAILWAY SAFETY CRITICAL COMMUNICATIONS

Claire Dickinson

Office of Rail Regulation (ORR), Grove House, Manchester, M16 0RB

Each year ORR investigates a number of railway incidents where failure of communication has played a contributory role. This paper will outline some of the key reasons for this and describe some of the tools and techniques being used for addressing this topic. The railway industry has made a concerted effort to improve the reliability of verbal communications when using telephones and radios to control train movements. The presentation will describe some of the initiatives underway within both the mainline and metro sectors of the industry.

Introduction

Whilst there is considerable automation involved in railway operations, everyday people working on the railway ensure the safe movement of trains by using effective verbal communications. Railway Safety Publications 3 and 4 (ORR, 2007) list the numerous safety critical tasks and activities that are carried out and nearly all involve verbal communication. On the mainline railway some key groups (e.g. Electrical Controllers, Signallers) have a designated role of 'lead responsibility' to ensure that before a conversation is finished, both parties involved understand the information that has been exchanged and are clear on any decisions or proposed activity. This is supported by the requirements of the industry's Rule Book (Section G1, Section 11 RSSB) that format the beginning of a communication, standard phrases, use of the phonetic alphabet, lead responsibility and the closure of a communication. However, the arrangements in place for controlling the risks from communication failure vary across the industry.

With 20,000 train movements each day there is a heavy reliance on telephone and radio systems. It is creditable to the railway industry that such frequent activity is undertaken without incident but occasionally failures of communication do occur. Initial estimates from reviews of rail incidents suggested that about 90% of incidents have communication as a contributory factor (e.g. rail industry communications website www.railscc.co.uk). Shanahan et al (2007) reported that about one third of all incidents on the railways are caused – at least in part – by miscommunications. He noted that track workers were particularly vulnerable, with more than half of all incidents involving a miscommunication component.

It was revealing that a failure to communicate at all and communication of incorrect information are the most common types of error. It is people "being human" that they forget to say something or do not carry out the required action, especially when under pressure. Like the instance of a driver that confirmed with his control the action he was to take, walked through a busy, crowded train to its end then failed to carry out the action. Instead he re-established contact with control and discussed the next steps to be taken. It is also a human response to demonstrate frustrations resulting from equipment failures or be motivated to misuse equipment. An example being train drivers turning the radio volume down to inaudible levels because of high frequency squeaks or whistles. People mishear, especially when not using the phonetic alphabet or when in noisy environments. They also use jargon that is specific to one group and not understood by another so that they interpret it wrongly. Occasionally wrong assumptions and expectation can play a role in communication failure. In the instance of trespassers applying graffiti the use of the phrase "being hit" being taken to mean the train was being sprayed with paint rather than something far more serious and consequential.

Recognising different types of communication failure though is not enough to initiate a company to devote resource to addressing it. What is more persuasive is knowing the impact on costs. Lowe and Nock (2007) make the point very well with the following example. An incident arising from a driver passing a signal set at danger (a SPAD) caused by miscommunication costs £22,000 and the industry has 40 SPADs a year is far more likely to engage senior management and enable resources to be allocated than quoting various percentiles for miscommunication involvement. But what really drives an industry into action is leadership and senior management commitment. It is indicative that the industry is able to learn from its failures that two incidents in particular gained the attention of Network Rail's senior management and led to a commitment to improve a situation. The result was a SAF6 programme from Network Rail that included government funding (via ORR) for the equipping of signal boxes with voice recording equipment. This has enabled NR to implement a communication risk strategic programme that significantly has also involved train, freight operators and contractors from the hundreds of companies involved in renewing the railway infrastructure. As a result standards have been rising within Network Rail and gradually we are able to "hear" improvements in communication performance of the drivers, track workers and others that communicate with the Signallers.

Industry approach for managing safety critical communications

Recognition of the role of communication failure in incidents in the late 1990's has led to highly effective activity within the rail industry to addressing such risks. Lowe and Nock (2007) summarise and highlight some of these activities: leadership from the establishment of the industry's Communication Working Group, strategic plans being implemented with railway companies, clarification across the industry of what constitutes good practice, revision of standards documents, rule book changes to make explicit areas of uncertainty, research activity. As already stated, Network

Rail's SAF6 programme has been particularly influential in raising standards across the industry. It is worth considering the strands within this programme:

- Raising awareness – sharing of good practice across the industry by website,
- Selection – development of tests for recruits on communication skills
- Competence – development and provision of free training across the industry
- Assurance – working with other companies to evaluate voice recordings using a standardised system and providing constructive feedback to those involved
- Leadership – everyone setting an example
- Rules and protocols – areas of deficiency being identified and action being taken
- Equipment – installation of voice recording equipment in signal boxes
- Investigation – continuous learning from incidents of communication.

ORR approach for regulating the management of safety critical communications

Both industry specific and general health and safety legislation are appropriately in place for ensuring that safety is not compromised by communication failure. But the regulator has a role to ensure the industry is adhering to such principles and regulatory requirements. This is carried out by ORR inspectors doing preventative inspection activity and on occasions being involved in incident investigation and enforcement. The ORR inspectors have adopted an educational role, raising awareness within the train, freight and metro community of the emergent good practice. They have required companies to demonstrate a higher standard of compliance in communication performance than was previously the case. Supporting NR's programme in encouraging the train and freight companies to participate in joint monitoring of voice recordings or training courses has been beneficial in driving compliance standards forward.

ORR's method for inspection of safety critical communication

Audits of voice communication are carried out using a standardised toolkit prepared by the Health and Safety Laboratory in collaboration with ORR's railway inspectors in 2005. This includes an Inspectors Guide, Summary Sheet and importantly, an Audit Checklist that covers a series of points to assess the adequacy of the management system concerning communication systems, communication culture, transfer of information, assessment of a degraded situation and equipment, plus evidence sheets for use when evaluating recorded conversations. This works by interviews being arranged with a "vertical slice" of people within the organisation from the operational people to the people in the management chain, and support functions such as the company training department. Interviewees are asked to discuss a recent scenario they have dealt with recently. There are guide questions provided but inspectors are encouraged to ask questions in their own way that establish evidence. This is then compared to the provided examples of satisfactory or

unsatisfactory evidence (see below for an example of one section). From the series of interviews recurrent themes and patterns emerge of the strengths and weaknesses of safety critical communications within that location. The following is an extract from the Equipment Section guide questions and evidence criteria:

Is equipment adequately maintained ? Is it appropriate for the task and fit for purpose ?
Are systems in place for ensuring the suitability and sufficiency of equipment e.g. battery life ?
Are complaints concerning equipment adequately addressed ?

Examples of Satisfactory Evidence	Examples of Unsatisfactory Evidence
☺ There is appropriate equipment available.	☹ Equipment fails to work when called upon.
☺ All staff are equipped with the means to communicate.	☹ No systems exist for the management of batteries.
☺ The equipment is available and fit for purpose.	☹ Sub standard equipment is accepted as the norm.

ORR's approach for investigation of communication failure

A different approach is taken during investigations of incidents. During investigations voice recordings are obtained and an analysis made of their compliance with good practice. So, if there was an emergency did the voice recording begin with "This is an emergency call. Am I speaking to the Signaller ?" There is a need for a structured and systematic approach to analyse the recording.
The stages currently in use for such a system are as follows:

1. Consider what systems were available at that location for communicating, e.g. NRN, CSR.
2. Are voice recordings available ?
3. Set down the content of the conversations in a tabular/sequential format,
   - Identify the time and date of each chunk of the conversation.
   - Identify which party said each recorded message and who had lead responsibility
4. Identify in the table the potential for human failure from the sequence of conversation
   - Indications of the right plan but a failure to execute the actions correctly e.g. someone who regularly identified a signal correctly using the phonetic alphabet demonstrating a slip of the tongue and instead of saying "Uniform" in the phonetic alphabet says "Umbrella"
   - Indications of the right plan but skilled operators forgetting to do something, e.g. a failure to repeat back or confirm that task, decision or action have been carried out

- Indications of a poor plan – key information may have been missing or insufficient but a decision and proposed action communicated anyway
- Indications of a second plan being developed and confusion on what to do
- Principles of lead responsibility not being heeded.
- Communication but no response elicited – mistaken understanding
- A decision or a task may not have been fully completed
- A decision or a task may be completed at the wrong time
- A decision or task may have been done in the wrong sequence
- Communication of an unworkable, complex or unsuitable plan of action
- Knowingly demonstrates violation behaviour

5. Evaluate compliance with industry good practice in applying voice communications, use of the protocols, phonetic alphabet, use of short sentences, action words, radio technique
6. Consider the potential reasons for failures in decision-making, a failure to carry out the activity or communication failure. E.g. a signallers giving permission to cross at a level crossing may have experienced a sharp rise in their workload or a driver may be using a radio that appeared fully charged at the start of the shift but the battery-life may be insufficient to last the entire shift.
7. Consider how to prevent a recurrence of the situation ie what controls should be present or could be used to mitigate the situation.

What is emerging from safety critical communication inspections and investigations

The audit inspection toolkit has been applied in seven inspections and in each case a large number of deficiencies have been identified but companies are receptive to raising their standard. In one case this is not being delivered fast enough or in a sufficiently comprehensive way such that formal enforcement action has been proposed. However, generally the companies involved have responded positively.

From the inspection and investigations undertaken to date a number of key themes are emerging of where further attention needs to be directed within the railway industry. These are as follows:

- Level of commitment – there is a need for more commitment from senior managers.
- Promoting good practice in the management of communication risks and training to enable greater application of raised standards within some pockets of the industry. Voice recordings are demonstrating that the opening of conversations to have been substantially improved over time but that it is the closure of conversations where weaknesses remain.
- Identification and practising communications used in infrequent encountered situations where there is the potential for higher consequence events. At present the focus is on communications used during normal situations and when making an emergency call.

- More assessment and monitoring of actual voice recordings to inform constructive feedback and where necessary, coaching and individual action plans. In depots this can be done by listening in on a third radio but infrequently do supervisors intervene with the shunter and driver to encourage more effective voice communications.
- Issues with equipment, batteries and black-spots need attention.
- Establishing voice protocols and setting down what is good practice within the company's competence management system. In particular, inclusion of communication and radio skills in the context of the many safety training courses.
- Establishment of better senior management performance indicators.
- Testing the performance of communication skills in safety audits of key procedures that include an important or significant verbal communication task.
- Greater employee involvement and action from railway trade unions.

Conclusion

There has been significant progress within the railway industry on raising the standard of verbal safety critical communications. Key to this is the commitment of senior railway managers emphasising its importance and implementing an organisation and systems to address the risks on a number of fronts. There is still substantial work to be done but it is heartening that the signs are that the industry is learning from previous failures and taking action to prevent recurrence.

References

Lowe, E. and Nock, P. 2007, Changing safety critical communications behaviour, In: *People and Rail Systems, J. R. Wilson, B. Norris, T. Clarke and A. Mills (Eds)*, (Ashgate : Aldershot), pp. 399–407.

Shanahan, P., Gregory, D., Shannon, M. and Gibson, H. 2007, The role of communication errors in railway incident causation, In: *People and Rail Systems, J. R. Wilson, B. Norris, T. Clarke and A. Mills (Eds)*, (Ashgate : Aldershot), pp. 427–435.

ORR, 2007 Safe movement of trains, Railway Safety Publication 3, www.rail-reg.gov.uk/upload/pdf/rsp3-smot-210507.pdf

ORR, 2007 Safety critical tasks – Clarification of ROGS Regulations Requirements. Railway Safety Publication 4, www.rail-reg.gov.uk/upload/pdf/rsp004-rogs-crtcl_tasks.pdf

RSSB Rule Book Section G1, Section 11 available at www.rssb.co.uk

The views expressed are those of the author and not necessarily those of the ORR.

SYSTEMS ENGINEERING MEETS USER CENTRED DESIGN: IN CAB SIGNALLING DISPLAYS

Mark Lowten & Glyn Roberts

Tube Lines Ltd, Canary Wharf, London, E14 4HD

This paper will show how using the combined principles of the systems engineering 'V' model and the user-centred design process, Tube Lines produced an in-cab signalling interface that enables an acceptable manual driving performance, is optimal to the user (taking into account size and location constraints), is integrated into the train cab environment and wider control system . This paper examines the work activities carried out at each of the main stages of the system 'V' model to ensure that the design is as usable as possible. This work started with de-risking activities and has been carried through to on-train pre-commissioning assessments which will be used as assurance evidence to show that the design and interfaces are acceptable for use on the railway.

Background

The traditional 'fixed block' signalling system that controls train movements on the Jubilee, Northern and Piccadilly London Underground lines is being upgraded to a modern 'moving block' signalling system to meet journey time capability requirements stipulated in the Public Private Partnership contract between London Underground and Tube Lines. The impact to the train operator (T/Op) is that the current line-side signals will be removed and replaced with an in-cab signalling display, called the train operators' display (TOD). As required by London Underground's category one standard, any new assets must be designed, installed and commissioned in accordance with the assurance gates as defined in the systems engineering 'V' model, and where appropriate should demonstrate best human factors (HF) practices. This ensures that a product is delivered that meets the requirements of the user(s) and the London Underground system as a whole.

To assist in the integration of HF activities and considerations in the various projects that Tube Lines is carrying out, an internal HF integration procedure has been written. This procedure details the work packages that should be carried out at each stage of the design life cycle to ensure that all HF issues are addressed. Within this procedure there are four assurance stages that the HF work packages are aligned with. These are:

- Conceptual design
- Detailed design

- Pre-commissioning
- Post-commissioning

The remainder of this document will describe the work packages that were carried out during each of these stages to ensure that the user was at the centre of the design process.

'V'Model stages

Conceptual design

The Initially HF work packages focused on capturing the requirements (including system functionality) that will affect how the end system is used. However, before any work packages can be undertaken the project needs to be scoped and the work planned. This scoping and planning is contained in the project Human Factors Integration Plan which also details how the HF issues will be managed, closed out, assured and how HF will be integrated into the wider project team and their activities. Following the completion of this document the work packages can be started.

Initially three work packages were carried out to capture the HF requirements for the TOD. These were:

- Train operator task analysis
- Day in the life/night in the life workshops
- Hardware evaluation

The task analysis was carried out to capture the specific tasks of the train operator on both the Jubilee and Northern line. This analysis looked at current and future tasks that will be performed by T/Ops under the migration stages of the project and when the full solution is finally implemented. Enough was known about the new system (as it is in operation throughout the world) to be able to predict how it would function on the London Underground. Therefore, via this comparison between the old and new systems it was possible to illicit operational requirements and raise issues that would need to be addressed throughout the product development.

Following the completion of the task analysis a number of 'day in the life/night in the life' workshops were arranged and carried out. These were system-wide activities and involved a number of the differing types of end users of the system, system designers, operational specialists, trainers, safety assessors and independent witnesses. These sessions looked at the operation of the railway as one system. From this, the interaction between the train operator and the service controller was established, the role of the maintainer within the system was understood and an initial impression of the new procedures that would be needed were realised. These workshops drew out new requirements that had not yet been realised. They also added additional information and understanding to current requirements so that a clearer vision of what was required of the interface was obtained.

In conjunction with these two activities, the hardware requirements about the display unit were being captured. Some of these requirements were taken from the

above activities and some were captured following a review of how the new system is currently operated across the rest of the world. A limiting factor with this upgrade is the fact that any new interface needs to be retro fitted to a current train. This means that there is limited space available for any new hardware whilst there is also the imperative to ensure that the new interface is in a usable location. Once the space envelope for the new hardware was established and the operational requirements of the new interface were captured, a technology type could be decided. Three choices were presented: Light Emitting Diodes, Vacuum Fluorescent technology or Liquid Crystal Display (LCD). In the end, LCD display was chosen as it provided a good quality image across all the required viewing angles, was much more flexible in terms varying the information displayed in specific areas and also as the decision had to be made very early in the design process it allowed for an increase in the flexibility of interface design as any changes would only be to software rather than hardware.

Once the requirements were captured initial ideas about how the interface should look were developed. Ideas were taken from current signalling interfaces, current speedometers, design guidance in reference material, requirements from standards and also domestically used devices. These ideas were developed and compared with the other types of information that could be presented to the train operator whilst controlling the train. At the same time a cognitive model was established to help the design team understand the mental processes that a T/Op would be carrying out whilst operating a train. This model was then used as a reference point to evaluate any ideas against to ensure that the interface design remained useable.

As the final deliverable for this stage of work, a design rationale was written which documented all the relevant design requirements, summarised the work and evidence to date and presented the concept design along with the logic behind the HMI. This design concept was then passed to the contractor as guidance so that they could develop it further in line with their system development.

Detailed design

Once the initial design and layout of the interface was established it need to be developed against the operating scenarios under which it would be used. It was decided to storyboard all the likely scenarios and establish what would be displayed on the TOD at the key stages throughout that scenario. These scenarios covered normal, abnormal, degraded and emergency events under which the TOD would be used. These scenarios were first developed using an Excel spreadsheet and once the appropriate tools were available these were transferred on to PowerPoint slides so that the visual implications could more easily be understood. By the end of this activity 62 operating scenarios were written. These scenarios were presented to both Tube Lines engineers and assurance agents as well as to London Underground engineers and operating specialists for comment and consultation.

The remaining work packages within the detailed design phase of the project were:

- Rapid prototype usability assessment
- Simulator usability assessment

- TOD glare assessment
- Programmable Audio Unit (PAU) usability assessment

Each of these activities were carried out to de-risk the interface and hardware design so that when it was fitted to the train there would be minimal chance of any adverse issues arising.

The rapid prototype usability assessment and simulator usability assessment were carried out to determine how usable the TOD HMI was from the end user's point of view. The rapid prototype usability assessment used an Excel macro to represent what would be displayed on the TOD under a number of operating scenarios. Subject matter experts (SME) watched the scenarios. They were asked questions about the scenarios (to establish their level of situational awareness (SA)), provided feedback on the icons used and completed a questionnaire about the usability of the interface. The results and feedback from this exercise were very positive with a number of minor design changes being suggested and consequently incorporated. Generally the interface was thought to be intuitive and usable, and this was confirmed by a very high proportion of SA questions being answered correctly (96%).

The simulator usability assessment was a higher fidelity assessment and used the newly upgraded training train cab simulator as a platform. This assessment involved SME's operating the train under a number of realistic operating scenarios (development in conjuction with London Underground). As this assessment was interactive rather than just observational it was possible to measure the SME's level of workload in addition to SA, general usability and acceptance of the symbology and the interface. Operator workload was measured successfully using the Instantaneous Self Assessment (ISA) workload measuring tool and SA was measured using SAGAT. The results from this assessment confirmed the findings from the first usability assessment and also showed that the operator had acceptable levels of workload even when operating under potentially very stressful conditions. Obviously the limitations of the assessment were understood but as a de-risking exercise this proved to be very encouraging.

The TOD glare assessment was carried out to ensure that the introduction of a new technology to the train cab would not affect the operability of the train as a result of any natural or artificial glare on the display. A trial was arranged to de-risk this issue. As the effect of glare is associated with the brightness of the light source and the angle of incident/ reflection the worse case scenario was established. A script of random TOD symbology was run so that all the icons, colours and messages were displayed. The subjects were then asked to state what had changed on the display as the script was run. The accuracy of the subjects' response and the speed of the response were recorded to determine how legible the interface was under worst-case conditions. The results from this assessment showed that under worse case lighting condition the T/Op can still see and read the signalling information on the TOD within the time allocated to undertake their tasks efficiently and effectively.

The final de-risking assessment involved the PAU. This provides an audible warning to the T/Op if there is a change in circumstances, for example, if the T/Op needs to slow for a speed restriction or if a text message has arrived. To ensure that these sounds are acceptable for use in the train two assessments needed to be carried

out: firstly, on the train to ensure that the sounds are audible against the ambient noise level and distinguishable from other train borne alarms. This assessment was carried out using a portable rapid prototype of the final alarm system. The alarms within this prototype were configurable so that the optimum sound type, frequency and level could be defined. The second assessment addressed how memorable the sounds were. The sounds were designed to be as intuitive as possible and were designed so that a mental model about what the sound means is easily learnt by the T/Ops. Both of these assessments produced positive results so that the system could go forward to on train testing.

Pre-commissioning

Following the completion of the detailed design phase the on train/pre-commissioning work packages could begin. The aim of this work is to provide additional evidence that the interface and the operating system are fit for use in passenger service on the operational railway. At the time of writing this paper, this work package has not yet been completed. However, the process and the measurements to be taken are very similar to those used during the simulator usability assessment. This is for two reasons: firstly, because they worked during the simulator usability assessment. The results gathered during that assessment were very constructive, both the subjective and object results agreed with each other and when discussed with the end user, they were happy that they represented their true feelings. Secondly, the scenarios tested during the pre-commissioning assessment are very similar to those used during the simulator assessment and it would be useful to be able to compare back and validate those initial results.

Post-commissioning

The post-commissioning assessment have not been started, however, the purpose of them is to:

- Ensure that the design is being used as expected
- Ensure that the design is fit for purpose i.e. it enables the operators to perform their tasks safely, effectively and efficiently, under normal, abnormal, degraded and emergency conditions (Validation)
- Ensure that the training and procedures developed can work under real operating scenarios

This work will be carried out once the system has been in operation for a reasonable period of time.

Project organisational tools

Throughout this paper, many references have been made to capturing issues, planning work and ensuring that there are not outstanding issues prior to operating of the system. There are two tools that are vital to ensure that this occurs. The first is the project issues log. This is a HF specific log where the HF issues are logged, tracked, actioned and closed with evidence referenced. This log is a living document that is

used at all stages of the systems 'V' model. It is actively managed by the HF team who are constantly working to close all the issues. At each stage of the 'V' model this log is used as a deliverable to show that there are no open issues that should have been closed for that stage of the project.

The second tool is the Human Factors Integration Plan. It was mentioned that this was first written at the start of the project, but what was not mentioned was that it is also expected that this will be updated as more information becomes available about the work packages required and the scope of the project. Therefore, this plan will probably end up being updated four or five times throughout the project lifetime (depending on the size of the project). These two documents are critical to the successful integration of HF with a project.

Conclusion

The process and work packages described within this paper have produced a design that is usable by those who will be using it. Following this process has meant the end users were involved in the design process and as a result all stakeholders are happy with and accept it. In addition, carrying out these work packages has meant that a suitable level of assurance evidence is available to show that the interface is usable and fit for purpose.

COMPARISON OF DIFFERENT TYPES OF REPORTS IN THE INVESTIGATION OF SIGNALS PASSED AT DANGER (SPADs)

Jessica Hutchings[1], Brendan Ryan[1], Michelle Nolan-McSweeney[2], Emma Lowe[2] & Wallace Weatherill[3]

[1]*Institute for Occupational Ergonomics, University of Nottingham, University Park, Nottingham, NG7 2RD, UK*
[2]*Network Rail, Melton Street, London, UK*
[3]*Participating Train Operating Company*

Evidence collected in the investigation of SPADs includes a range of technical (e.g. train speed, signal operation) and self reported information. This study aimed to provide an insight into the reliability of the evidence collected by self report methods during the investigation process. A qualitative approach informed the research methodology. Witness reports relating to six SPAD incidents were analysed in depth, with each incident providing four self report data types: tape recordings between the driver and signaller, driver statements, initial interview (between the driver and driver manager) and investigative interview transcripts (between the driver and incident investigator). It was concluded that drivers' reports, whilst valuable and necessary to understanding an event, may be inaccurate, unreliable and incomplete in some circumstances, due to the limitations of human memory and the approach to questioning. Further steps can be taken across the industry to inform those who are involved in collecting reports or in carrying out investigations of the factors impacting on memory and reporting.

Introduction

Signals passed at danger (SPADs) involve a train passing a stop signal without authority and have the potential to cause catastrophic accidents such as train collisions and derailments (Lowe and Turner, 2005b). Despite a fall in the numbers of incidents and aided by improvements in technology to reduce the likelihood of incidents, SPADs remain a concern for Network Rail and Train Operating Companies (TOCs).

Concerns were raised following the tragic accident at Ladbroke Grove (1999) about the effectiveness of investigations in obtaining the root causes of such incidents. Much has been done in the intervening years to strengthen the process of investigation and provide tools for investigators (e.g. the SPAD Hazard Checklist, Lowe and Turner, 2005). Information is available from various sources in the aftermath of an incident, such as data from train recorders, signalling equipment and track circuit technology. Another form of information which can be crucial in

understanding events leading up to an incident is that which can be provided by people involved, that is self reported information. This type of information can be affected by a number of factors, including the memory of the individual and their recollection of the events (Eysenck and Keane, 2000; Loftus, 1979), their motivation for reporting, the time since the incident (Eysenck and Keane, 2000) and even the approach to questioning by an investigator (Loftus, 1979). This paper presents findings from the analysis of four types of self reports that are collected routinely following a SPAD incident. These reports were:

- The tape recording of the initial conversation between the driver and signaller after the incident
- The written report or statement produced by the driver
- Transcripts from the initial interview of the driver by the driver's manager
- Transcripts (or audio tapes in two cases) from an interview of the driver by an incident investigator.

The first three of these reports are collected either immediately or very soon after an incident, whilst the fourth could be some time after the event. The analysis aims to give insight into the reliability of the information collected in the different types of report and identifies factors affecting the content of each of the different types of report. Preliminary recommendations are presented for development of the process of investigation.

Method

The researcher worked closely with representatives from Network Rail and a Train Operating Company to develop an understanding of the investigation process within the industry. This was achieved through discussions and preliminary interviews with senior managers involved in developing the investigation process and with key staff involved in conducting investigations on a day-to-day basis. This preliminary stage of the research also included the review of sample documentation (e.g. witness statements and interview transcripts) to develop an understanding of the likely content of the reports and develop a list of provisional themes, which could be used and developed in subsequent analyses.

For the main part of the research, the investigator at the TOC provided the researcher with self reports (in each of the four categories listed above) collected from six SPAD incidents. These incidents were selected by the investigator and were judged (within previous investigations) to represent a range of different causes of SPADs. A qualitative approach was used in the analysis of the self reported data, breaking down and coding different words and phrases within the reports according to the themes which had been identified in the preliminary stage of the research (described above) or other themes or sub-themes which were emerging during this part of the analysis. Qualitative analysis is an iterative and time consuming process (Miles and Huberman, 1994) and each of the reports were reviewed a number of times until it became clear that no further themes or sub-themes were emerging. Comparisons were conducted across the different types of report using tables to identify whether the same themes were common to each report. This allowed the

researcher to establish a framework of what information is reported by drivers at each stage of the investigation and to note whether the evidence changes in any way.

Results and discussion

Ten primary themes were identified and were decomposed into 71 sub-themes and these were each used in the comparison of content in each of the forms of self report. The primary themes are listed in Table 1. The analysis of the themes in the tape recordings of the initial conversation between the driver and signaller demonstrates that this conversation provides valuable first hand information, in particular, on technical details relating to SPADs (e.g. the signal indication). This part of the investigation process also gives an insight into the emotions of the driver (e.g. through evaluation of the driver's tone of voice and the communication style between the signaller and driver), information that may not be available from other sources of information.

This conversation is largely structured around the content of the form which has to be completed by the signaller at this time (the RT3189 form) and thereby influences what information is collected. However, it was clear that in some cases the signallers' questions could lead to the recollection of additional information. One interesting finding related to the attitude of some of the signallers to the use of the form and the way they asked the questions could have an effect of introducing bias to the responses (Memon *et al.*, 2003). Examples of the some of the signallers' comments are given in Table 2.

Table 1. Primary themes within the reports.

Primary theme
• Fatigue
• Driving practice/procedures
• Environmental conditions
• Emotions
• Attention, memory and distractions
• Technical factors
• Time
• Personal/lifestyle factors
• Communication
• Work organisation
• Job satisfaction

Table 2. Signallers' comments whilst speaking to the drivers.

"These questions are going to be a bit silly, but I got to ask them"
"(*Sighs*) . . . and now onto the other side"
"I know you don't but just for the sake of the form"

Preliminary interviews with the investigators at Network Rail and the TOC have suggested that the information in this conversation is important because it gives an early indication of the driver's state of mind and understanding of the event. However, it should be borne in mind that this communication between the signaller and driver was introduced with a different purpose in mind (to collect signalling related information) and was not originally intended for the purpose of investigation. Furthermore, the signaller has other objectives to re-instate service as soon as possible and continue regulating other trains in his/her area of control. Whilst there is some scope for collecting valuable information, opportunities to develop this communication between signaller and driver further for purposes of investigation could be limited. In this study the analysis has concentrated on the reports of the driver, though it is important to remember that the signaller, one of the parties in this conversation, could have a role to play in the events leading up to the incident and could introduce bias (even unwittingly) in the collection of information on the event.

The driver's written statement generally adopts a format of free recall of information, though there is some variation in the types of forms used within the industry. Free recall, especially when used soon after an incident, should help to limit bias in the self report information that is collected (Memon and Bull, 2000). The reports were often limited in length, usually containing information on the route travelled, recollections of signalling details (e.g. signal aspect or TPWS activation) and details of outcomes such as the distance which the driver went past the signal. Information was also given on weather and rail conditions, though this was possibly prompted by some preliminary information included on the form. Limited information was offered on likely causes of the SPAD, though occasionally, some information was provided which could be used in attempts to determine causes. For example a statement referring to how the driver saw workers on the line could be used as a subsequent line of enquiry to determine whether distraction could be an important factor in the incident. There is clearly a balance here, which needs to be considered in the collection of this type of report. There are definite benefits of free recall, but in this type of written report there are likely to be limitations in the amount of detail that will be produced by the witness. Additional information can be obtained with the use of suitable prompts, but this could result in the collection of some information, perhaps at the expense of other more relevant information. Further work would be needed to evaluate the effectiveness of the forms that are used currently for the collection of reliable and relevant information.

The interviews elicited reports with the most themes, in particular the investigatory interview, but this is perhaps not surprising as it is at these stages of the investigation that the driver's manager or the investigator begins to ask more specific questions about the incident, covering almost all of the themes that have been identified in Table 1. However, it is important to consider the reliability of new information that is provided. Information on the incident could be altered with the passage of time since any recollections of the event could be influenced by new information coming to light since the incident, by conversations with others, by the content of earlier accounts of statements of the event, or in the efforts of the witness to rationalise and understand the circumstances that lead up to the incident

(Haber and Haber, 2000; Loftus, 1979). Furthermore, the types of questions used by interviewers can affect the type of information that is reported (Haber and Haber, 2000; Loftus, 1979; Memon *et al.*, 2003). In many of the reports, investigators used leading questions, for example "The signal was showing a red aspect . . . " which may prompt the driver to give an answer that may not be accurate. This affects the reliability and accuracy of the information collected in an investigation and could inhibit the identification of the true root causes of the incident.

It was interesting to note that drivers commonly made statements such as "I can't recall", "I can't remember" and "I don't know", in each of the different types of report. These comments illustrate how a driver may, under some circumstances, have no reliable knowledge of information, such as the signal aspect at the particular point of a journey. A direct question on an issue such as this may encourage a respondent to provide an answer (based on what they think the signal aspect should have been). However, this would give little insight into the driver's perceptions or decision making at the time of the incident. Driving is a skilled task and many actions are executed automatically, therefore comments such as "I can't recall" are not unreasonable. However, this does present a challenge to the interviewer who may need to disentangle what a witness cannot remember and what a witness may, under some circumstances, wish to not disclose (Haber and Haber, 2000).

For each of these reasons it is important to consider the approach to collecting reliable information at this interview stage of the process. Where possible it can be helpful to begin interviews using open questions to encourage free recall, moving later in the interview to more direct questions to clarify specific details (Memon *et al.*, 2003). This type of approach was largely achieved in the investigatory interviews, but only in two of the six interviews by the drivers' managers. It is important to realise that any bias that is introduced within an earlier report or interview (e.g. through a leading question) can be encoded in memory and reported in subsequent reports or interviews with a greater prominence than would otherwise have been the case (Haber and Haber, 2000).

Conclusion

The findings provide Network Rail and the participating TOC with an insight into the effectiveness of the current methods of self reporting which are used in the investigation of SPADs. Further work is in progress to develop a greater understanding of the themes and sub-themes from the reports, and analysis of the types of questions used by investigators to elicit information from witnesses. The study reported within this paper has focused on the reports from drivers, but could be extended in due course to take into account reports from those in other roles, and reports relating to other types of incidents. There is some scope for work to support and enhance the collection of self reported information at different stages of the investigation process. This could be achieved in part through the development of tools (e.g. e-learning tools), and these could be made accessible to those involved in the collection of self report information and other investigatory activities across

the industry. This would help to encourage a common approach to investigation and greater exchange of relevant and reliable information.

References

Eysenck, M.W. & Keane, M.T. 2000, Cognitive *Psychology: A Students Handbook*, Fourth Edition (Psychology Press, Hove, UK)

Haber, R.N. & Haber, L. 2000, Experiencing, remembering and reporting events, *Psychology, Public Policy and Law*, **6**, 1057–1097

Loftus, E.F. 1979, *Eyewitness Testimony*, (Harvard University Press, Massachusetts)

Lowe, E. & Turner, C. 2005, A Human Factors SPAD Checklist. In Wilson, J.R., Norris, B., Clarke, T. & Mills, A. (eds.), *Rail Human Factors: Supporting the Integrated Railway*, (Ashgate Publishing Limited, England), 182–188

Memon, A. & Bull, R. 2000, *Handbook of the Psychology of Interviewing*, (John Wiley & Sons Ltd, New York)

Memon, A., Vrij, A. & Bull, R. 2003, *Psychology and Law*, (John Wiley & Sons Ltd, Chichester, England)

Miles, M.B. & Huberman, A.M. 1994, *Qualitative Data Analysis*, Second Edition (Sage Publications, London)

Wilson, J.R. & Norris, B.J. 2005, Rail Human Factors: Past, Present and Future. In Wilson, J.R., Norris, B., Clarke, T. & Mills, A. (eds.), *Rail Human Factors: Supporting the Integrated Railway*, (Ashgate Publishing Limited, England), 3–12

HUMAN SYSTEMS INTEGRATION IN RAIL: DEVELOPING THE INTERFACE FOR INTEGRATION

Jennifer McGovern Narkevicius

Jenius LLC, California, MD 20619

Human System Integration (HSI) is the management and technical approach to incorporating human considerations into the design and implementation of systems. HSI can contribute to improved system performance, safety, and reduced cost in Rail applications. While primarily used in military acquisition, application of HSI in transportation organizations such as railroads to achieve the safety and operational effectiveness goals will translate to long term reduced operational costs. Work to define the application of HSI to rail transportation is ongoing.

Introduction

It is self-evident that systems, especially complex ones, are successful if they work. That is if a system meets expectations for use, it is deemed successful. Usually the emphasis in systems focuses on the development and implementation of new technologies; however their success is highly dependant on the performance of the personnel in and around the system. Military establishments have been using a more user centred approach to improve the implementation of new technologies.

These same approaches can be applied to non-military systems such as transportation infrastructure elements. Railroads are essential infrastructure in many countries and railroad companies and properties need to make good capital investment choices that maintain service and market competitiveness. The application of HSI to non-military applications can be challenging but still successful. As in the military, transportation and rail in particular, are beginning to comprehend the tremendous potential Human Systems Integration (HSI) can contribute to both safety and efficiency.

Systems Engineering

Systems Engineering calls for a broad view of a system not only as a structure but also including the organization and environment in which the system must operate. One outcome of this approach is systematic requirements definition identifying the

requirements and interfaces for the system and its elements: are hardware, software, and human operators, maintainers and other users. The systems engineering process detailed in a number of references (International Council on Systesm Engineering, 2007; Kossiakoff & Sweet, 2003; Martin, 1997 and many others).

HSI

Integrating the HSI domains into systems engineering is more than just forcing human factors into system design. It includes awareness of the system performance, the effect of each of the components of that system, and the effects of allocating specific functions to each element of the system. Effective integration of the domains generates better, more balanced compromises with other specialty engineering disciplines.

The identified human domains focus on Manpower, Personnel, Training, Human Factors, Safety, Habitability, Survivability, and Environmental Occupational Safety and Health. While the names and distribution may vary across entities, the underlying technical data and analyses subtend the same areas. These human domains affect system design and decision making by clarifying the human cognitive, physical, organizational, and personal roles especially in automation and other complex technology solutions. Figure 1 illustrates HSI in the Systems Engineering process. By identifying work to be performed, the target audience, economical training; and information architecture design HSI can explore, design, and iterate solutions

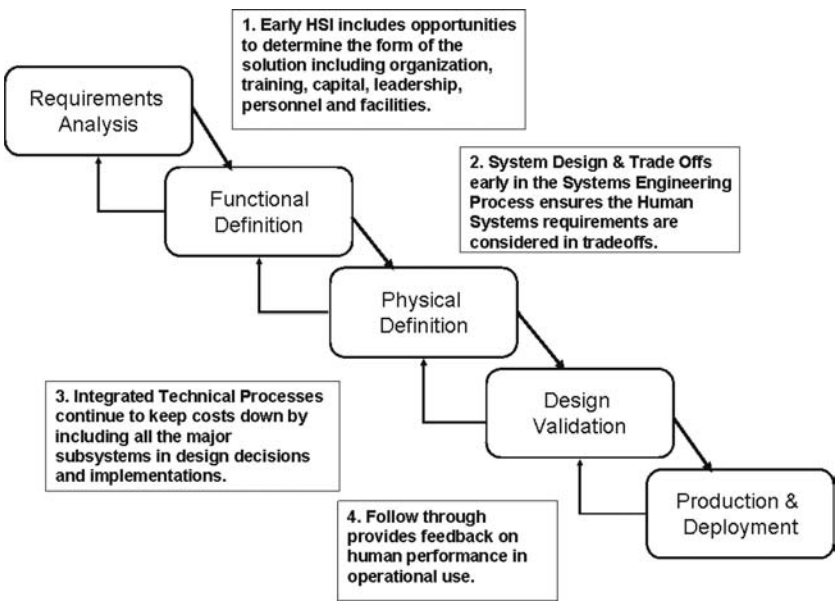


Figure 1. HSI elements embedded in Systems Engineering Process.

while requirements and function allocations are still being established. Tools and processes deliver insight to the potential outcomes of requirements from the earliest phases of development. The resulting systems have most favorable mixes of carefully selected and trained operators and maintainers working in a system architected to perform the required tasks.

The interaction of systems engineering and human factors results in defining and understanding the needs and requirements of the humans using the system in the operational environment. To define the requirements of humans as a major system component, it is essential to understand the inherent "capacity" of user populations and the operational environment in which they work. Few authors have detailed HSI and the place of the human disciplines in systems engineering (e.g. Booher, 2003, Chapanis, 1996, Pharmer, 2007). HSI is fundamentally more than anthropometry or cognition of the average member of the user population. It requires a detailed description of the targeted audience of operators, users and maintainers as well as an understanding of the Knowledge, Skills, and Abilities (KSAs) as well as other attributes of the people that will be operating (and maintaining) the system that may impact total system performance. It is essential to understand the work to be performed. All these diverse data must be included in the systems engineering process to ensure that the system will perform as desired in the operational environment. It is also necessary to address organizational issues. Organizational changes can affect the work to be performed and must be considered in global design decisions. Changes must be reflected in the information architecture of the system as well, especially if there are automation or decision making support elements. It is critical to understand the work humans perform and the context in which it will occur when designing any new system. That definition must be utilized locally in the human factors aspects of the design and globally in the overall systems design. In addition, it is important to socialize that definition of the work to be performed in the organization and among the claimants of that work. These business processes, organizational structures, and occupational work must be factored into the systems thinking and design.

Challenges in Implementing HSI for Railroad

Increasingly complex systems are being introduced into rail applications requiring better integration the humans that are central to the railroad (Wilson, et al, 2007). Military organizations have been using the HSI approach in include human requirements in complex systems for over 25 years. For example, the US Army used HSI processes in systems engineering in the LHX program to develop an engine that could be maintained with six tools. Predecessor systems required over 130 tools. This design change reduced manpower, personnel and training requirements while applying good human factors and enhancing safety. This solution was not uniquely military – there is no reason not to transition these processes and approaches to rail applications. There are similarities and differences between military and rail purposes.

The similarities include the acquisition and use of complex technology solutions in complicated organizations. These systems and their use can be costly. The application of technology to make processes more efficient (save time and reduce costs) is essential. Technology that improves safety is necessary (military needs to protect organizational assets for defense or rail protecting assets (locomotives and other equipment) for transportation requirements). For both military and rail purposes, procurement strategies may include phased implementation due to costs of the technology or logistical complexities.

The differences are equally striking. The user, maintainer, and constituency populations are very different for rail applications than for military. Rail work is unlike military work in form and structure. Rail work is profit oriented, the military is not. The concepts of employing technology (or concept of operations) and the relationships of the various parts of the (global) organization are dissimilar in rail than in military.

For both military and rail domains, the "tenets" of applying HSI are the same. These are initiate HSI early, identify issues and plan analyses, document HSI requirements, make HSI a factor in source selection, execute an integrated technical process, conduct proactive trade-offs, and conduct HSI assessments at appropriate milestones.

The primary challenge for applying HSI to non military applications is the organizational structure differences and the constituency differences. Military acquisition is governed by regulations and legal statute. Defense organizations have complex, documented processes to procure advanced technology solutions to perceived capabilities gaps. In addition, military organizations support research and development of future advanced technologies (e.g. Department of Defense Instruction 5000.2, 2003).

In rail, the integrated processes for acquisition of material and for recapitalization are not present (at least not to the same extent as in the military domain) and the organization and environment do not support the development of "universal" processes. Further, the organization, structure and environment are quite different. Additionally, the relationship of the constituents is also different. The regulatory relationship to system developers is different as is the acquisition structure. This works against integrating the human domains in the same way as in military organizations because enforcement methods for systems integration to include human operators, users and maintainers are weak, at best.

For example, the introduction and continued development of automated train control highlights the increased need for HSI in the railroad. If we are to realize a return on investment, then use of HSI as in military applications will result in decreased acquisition cost and lowered cost over the life of the system. The return on investment for the railroad properties acquiring rolling stock and other systems will include not only lower development and implementation cost but also lower overall lifecycle cost for maintenance and operation.

Well designed automated train systems that include the human user and maintainer capabilities as part of the design space will result in decreased accident rates. Reduced accident rates will not only reduce cost but will also enhance public perception of the railroad industry. Including HSI in the design process will also

improve potential use and acceptance of new technologies by the user population constituency (unions, management, and government). Safer use of advanced technology systems might also reduce the need for encroaching regulation.

Conclusion

There are exciting times ahead integrating HSI and Systems Engineering to develop systems that will work as advertised. While the highly structured environment of the defense industry is not available for railroad properties, the enhancements and efficiencies available from HSI can be incorporated to improve performance, safety, and cost. Work continues to define the best ways to incorporate HSI and Systems Engineering practice into the railroads.

References

Booher, H. R. (2003). Handbook of human systems integration. Hoboken, NJ: John Wiley & Sons, Inc.

Chapanis, A. (1996). Human factors in systems engineering. Hoboken, NJ: John Wiley & Sons, Inc.

Department of Defense (2003). Department of Defense Instruction 5000.2, Operation of the Defense Acquisition System.

International Council on Systems Engineering (2007). Handbook of Systems Engineering. From INCOSE website: www.incose.org

Kossiakoff, A. & Sweet, W. N. (2003). Systems engineering principles and practice. Hoboken, HJ: John Wiley & Sons, Inc.

Kubeck, L. C. (1995). Techniques for business process redesign: Tying it all together. NY: John Wiley & Sons, Inc.

Martin, J. N. (1997). Systems engineering guidebook: A process for developing systems and products. Boca Raton, FL: CRC Press.

Pharmer, J. A. (2007). The Challenges and Opportunities of Implementing Human Systems Integration into the Navy Acquisitions Process. Defence A R Journal, February 1, 2007.

Wilson, J. R., Farrington-Darby, T., Cox, G. & Hockey, G. R. J. (2007). The Railway as a Socio-Technical System: Human Factors at the Heart of Successful Rail Engineering. Proceedings of the Institution of Mechanical Engineers Part F: Rail and Rapid Transit.

THE INFLUENCES ON SAFETY BEHAVIOURS OF RAIL MAINTENANCE STAFF

Philippa Murphy

Network Rail, Ergonomics National Specialist Team,
London, NW1 2EE

As a result of an incident in April '05 and as part of an industry shift towards a more behavioural style of safety management, Network Rail (NR) set about better understanding their maintenance workforce. To do this, internal influences on behaviour were identified and researched and these included attitudes, motivations, social networks and cognitive capabilities. The research utilised both qualitative and quantitative methodologies for the first three workstreams and bespoke psychometric tests for the latter.

The research presented a number of results relating to attitudes including those towards rules, management and methodologies for change. Motivations to both work and safe working were examined and discussed the need for recognition and non-financial rewards as a motivating factor to work and perceptions of the rules and over regulation as a de-motivator to work. In terms of social networks, issues such as trust in management communications and the effect of these on safe behaviour were noted as were the greater effect of team communications and management on track behaviours. The cognitive capabilities work found the average reading age of the staff to be higher than that often debated and also examined other cognitive and personality traits and their relationship to safe working. The findings have subsequently been used as input into an organisational change program for Maintenance improvement.

Introduction

As a result of a fatality of a contractor's employee, working on Network Rail Infrastructure, a piece of research was commissioned by the then Deputy Chief Executive of Network Rail, to understand better the 'front line' employees in the Maintenance function to improve the management of their safety. This form of safety management is a reflection of the more mature organisation, where safety is managed proactively as well as reacting to events.

As around 90% of all accidents have been related to human error in the casual chain (McKenna, 1983), the aim was to create a 'psychological profile' of staff to both understand them better, and to baseline and thereby measure improvements from in the future. The railways are a potentially hazardous environment and instances where staff may depart from standardised procedures or may not be

aware of the correct procedures to use. Therefore, both the internal and external influences on staff and their individual differences need to be understood.

The concept of safety behaviours was broken down into four parts. The first three are internal factors, those being; understanding staff's motivations and attitudes to work and safety and an assessment of their cognitive capabilities. The fourth assesses the external influences on behaviours and decision making and is the assessment of the social network staff work in and the levels of trust and communication within the network.

Context of the research

Attitudes and motivation were selected as they have a major implication on both safety and productivity, and are the key drivers to decision-making processes and therefore behaviour (Flavell et al, 1993). Motivation influences safety behaviour in number of ways. For example organisationally preferred behaviours may be less optimal if choices are time-consuming, are disapproved of in certain group circumstances, or if people are encouraged to behave less safely given rewards the behaviour could bring (eg finishing job earlier, conforming to group norms). This is similar to the 'Control' and 'Behavioural' beliefs (Ajzen & Fishbein, 1980). By understanding the key motivators and de-motivators of a role, an assessment of work design, processes and practices can be made. Attitudes are important as they are shaped by a reaction to external factors and internalised. People may then adjust their attitudes to fit with group norms, when attitudes are examined in an industrial context (Katz & Lazer, 2003).

Cognitive capabilities were selected for study as these understanding the cognitive capabilities and differences within a group highlight either sections of staff or organisational designs where errors are more likely. Although specific correlations to accident rates is difficult (McKenna, 1983), it is an interesting area of study, both in terms of the results themselves and in terms of a gap analysis between the results and the cognitive demands of the roles.

The tests were selected due to their impacts on safety behaviours and these were: Personality: to examine, in particular, the concept of conscientiousness (Arthur and Graziano (1996); Workplace Cognitive Failures (Wallace and Chen, 2005); Information Processing: to examine speed of visual information processing in particular (Sternberg, 2006); Working Memory: to understand the capacity of the workforce and the potential for this to be involved in incidents (Kolich and Wong-Reiger, 1999), Literacy and Numeracy: to understand the reading age of the sample and to match against standards, procedures etc and briefing materials provided.

Finally social networks were examined as teams are the external influences on behaviours and can shape the attitudes and motivations of team members. This can be compared to the concept of normative beliefs, as they shape the subjective norms of the group. The element of trust in teams is important as levels of trust in management has been linked to the acceptance of organisational decisions and improved behaviour with regard to compliance and adherence to procedures (Mitchell, 2006). Within the team it builds and facilitates the ability to both question decision and

report near misses etc, and provide support to teams and fosters a more just culture (Reason, 1997).

The research used both qualitative and quantitative methods for data collection. For the motivations, attitudes and social networks research, a sample of 48 Maintenance staff were interviewed using a semi-structured interview to both give the Researchers exposure to railway jargon and norms and to understand certain areas in more detail. Surveys were then created and completed by 514 Staff. With regard to the cognitive capabilities work, a battery of tests were created using standardised tests, specific to the aims of the research.

Discussion of findings

Motivations

With regard to specifics to the role of track maintenance staff, the main sources of de-motivation factors for work were: ineffective organisational and safety communications, especially those relating to weekend engineering work. This affects motivations as conflicting instructions or incorrect details can be interpreted at changing goal posts or can impact on the provision and organisation of plant and manpower. The rule book was perceived as not being 'user friendly' nor completely clear in it's content and description of the sequencing of events. Again affecting motivation as there is a lack of clarity or in come cases conflicting advice on how duties should be carried out which creates conflict.

In terms of external sources of de-motivators, these included issues such as a perceived lack of respect from management. The effect of this on staff includes an increasing division between front line staff and management which, in turn, reduces the level of trust and quality of communications between the groups, and also increases the cohesion of the team and polarises them from management. It also displays a lack of empathy and support from management and could impact on improving communications and becoming a learning organisation as the propensity to report near misses will be reduced. Line managers were also perceived as having too high a workload, in particular the completion of paperwork, which impacts on the time the Supervisors can spend with their teams, again impacting on the communications and support.

Attitudes

The findings from the Attitudes work identified a number of issues which shape attitudes towards safety behaviours. The first focus of the research (Graham 2006) was the effect of team cohesion on safety. The attitude (and the bond within the group) was a major moderator of attitudes to safety and safety behaviours. The findings are similar to the work by (Turner and Fleck, 2005). Therefore, by improving group cohesion and monopolising on positive motivational factors, the management of and practise of safety behaviours should improve.

An additional influence of team cohesion is the ability of the group to moderate themselves and to manage workforce issues within the group. This has the benefit

of the group forming a common goal for behaviour and strengthening the bonds between them. However, it is then ultimately important that the group moderate themselves towards positive, proactively safety behaviours to ensure safe working.

A number of factors were identified as influencing cohesion and these include banter, time in service, mutual goals, mutual appreciation of the task in hand and consistency of the team (low turnover). In terms of wider influences on attitudes, the division between management and staff was noted. This basis of this attitude was identified as shaping behaviour by the social network research also in terms of levels of trust and the quality of the communication.

Social networks

The social networks research looked at the teams in which the Maintenance staff work and their team within the wider management network; and then both the levels of trust and quality of communications within these links. The aim of this was to understand how the teams functioned within themselves and the wider context. A number of interesting and interrelated findings emerged. The first is related to the above 'Attitudes' research in that management communications are moderated by trust. Therefore unless the maintenance staff trust in management communication, the effectiveness of this is minimised. From the attitudes research it showed that there are currently mid range levels of trust between the groups. The concept of trust has been proposed to have three component parts – ability, integrity and benevolence (Mayer et al, 1995). Management could be selected for and should be cognisant of this to both monopolise on this during communications and management and to improve the effectiveness for front line staff.

The respondents rated the quality of communication between and within the groups and the responses reflect the above. The respondents were asked to rate the quality of communications for their team, their line management and other management. Although this is a fairly crude breakdown of groups it serves as an indicative result. The mean rankings (out of 5) were team (3.7), Line Management (3.0) and Other management (2.8). This corroborates the results from the attitudes work and on both the perceived distance between management and staff and the strength of the bonds within teams. Whilst this has implications for management communications, it also highlights the importance of safety management and initiatives being formed by and managed within the group, given this is the strongest and most influential link. This gives credence to behavioural based safety programmes, backed by process improvements in the rest of the organisation and management chain.

A demonstration of the distrust in management communications is manifested in levels of near miss reporting. The surveys included a section on self-report questions which assessed participation in or observation of unsafe behaviours and near misses. The reported number was seventy, which if factored up to the population as a whole would account for 1620 near misses per year. This is in excess of the number reported last year. The reasons for underreporting were also ranked and the two most highly rated were the fear of being blamed and a feeling that nothing would be done as a result even if reports were made. The first is a not un-common reason in industries where a blame culture (Reason, 1997) may still exist. The second is

an interesting factor and may point to a feeling of 'learned helplessness', as it is perceived that nothing would be done by management anyway, which may breed the attitude of 'what is the point' leading to a feeling that they have minimal control. Both of these factors can be managed and are a key component of a learning culture and a more 'mature' organisation.

Cognitive profiling

The final piece of research on cognitive profiling found that the average reading age was above that for the norm group, with a reading age of twelve reported, which dispelled some organisational urban myths. However, the tests did show that 5% of the sample had levels of literacy which could prove to be an issue given the reading ages which some engineering standards and procedures are written for.

In terms of numeracy, the sample scored slightly below that of the norm group which does pose have any immediate concerns as their tasks do not involved much mathematical abilities but it may have an implication on how metrics or statistics are presented to staff. In terms of working memory capacity, the backwards digit span test revealed that against an average score of seven plus or minus two bits of information, the average for this sample was five bits of information with 17.5% of the sample falling within the range of three bits of information, This has implications for the way in which briefings etc are designed and for job aids which may be required to assist with recall or retention for more complex tasks.

Conclusion

The four pieces of research produced both a cohesive and consistent set of findings regarding both a baseline measure of the constituent parts of safety behaviours and also some useful, industrial based solutions for making improvements to current and future work systems. The aim of this research was to provide a basis for a more holistic and user – centred view of safety management, from which we can as an organisation better understand and more supportively manage front line staff. This aim was achieved as exemplified by the steps which have been taken by the organisations to implement some of the findings and to make improvements.

Recommendations from the final report (Clegg, 2006) suggested improvements to be made in the following areas: Pride and professionalism, Near miss reporting, Safety compliance, Safety ownership, Team working and team leadership, Communications and trust.

The component parts of the understanding of safety behaviours leads to both a better understanding of the workforce and better management of them. From taking this baseline measure of attitudes, similar surveys can be conducted in the future to objectively measure the changes and improvements from this improved management. From taking this view of process safety, as well as understanding issues which affect personal safety, the organisation is taking great steps to improve accident rates and behaviours of safety critical staff.

References

Ajzen, I. & Fishbein, M. (1980). '*Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Arthur, W. & Graziano, W. G. (1996). '*The Five-Factor Model, conscientiousness, and driving accident involvement*' Journal of Personality 64(3), 593–618.

Clegg, C. C. (2006). '*We are Trackmen*' Internal Network Rail report.

Flavell, J.H., Miller, H.P. & Miller, S. (1993). '*Cognitive Development*' Prentice Hill: New Jersey.

Graham, P. (2006). '*Social Safety: Group cohesion as a predictor and moderator of safety behaviour'* MSc thesis.

Katz, N. & Lazer, D. (2003). '*Building Effective Intra-Organisational Networks. Centre for Public Leadership*'. Working Papers.

Kolich, M. & Wong-Reiger, D. (1999). '*Emotional stress and information processing ability in the context of accident causation*'. International Journal of Industrial Ergonomics, 24, 591–602.

Mayer, R. C., Davis, J. H. & Schoorman, F. D. (1995). '*An integrative model of organizational trust*'. Academy of Management Review, 20(3), 709–734.

Mitchell, J. (2006). '*The impact of trust and communication on safety – Report to Network Rail*' (Internal Network Rail report).

McKenna, F. P. (1983). '*Accident proneness: A conceptual analysis*'. Accident Analysis & Prevention, 15, 65–71.

Reason, J. (1997). '*Managing the risks of organizational accidents*' Ashgate.

Sternberg, R. J. (2006). *Cognitive Psycholog*y (4th ed.). UK: Thomson Wadsworth.

Turner, D. & Fleck, S. J. (2005). '*Taking care: Relational motivation and proactive safety behaviour*'. (Ph.D research).

Wallace, J. C. & Chen, G. (2005). '*Development and validation of a work-specific measure of cognitive failure: Implications for occupational safety*'. Journal of Occupational and Organizational Psychology, 78, 615–632.

FATIGUE MANAGEMENT IN RAIL: THE ROLE OF ERGONOMICS

Fiona Kenvyn

Network Rail, Ergonomics National Specialist Team, London, NW1 2EE

Fatigue, and its impact on safety and performance, is a concern given the risks posed by not managing it sufficiently. This paper will look at the issue of fatigue as it impacts upon the rail industry and on Network Rail in particular. It will examine the challenges inherent in implementing a robust fatigue management strategy, and the role of the Ergonomics National Specialist team in integrating processes for scheduling, controlling and assessing working hours into the day to day functioning of a complex, distributed, round the clock delivery organisation. Ergonomics has played a significant role in ensuring fatigue management is at the forefront of Network Rail's safety management processes through the application of a holistic systems perspective to the consideration of fatigue within the rail system. This paper aims to add to the existing discussion on fatigue management as it relates to the rail industry, providing a case study that describes the reality of trying to integrate good practice into a dynamic and in many ways unique organisation.

Introduction

As the controller and maintainer of the rail infrastructure across England, Scotland and Wales, Network Rail holds great responsibility not only for the travelling public and its stakeholders, but for the growing workforce tasked with the operation and maintenance of a 24-hour railway network. Fatigue management forms a significant part of the company's health and safety management system given the requirements for shift work and night working by the many safety critical employees charged with signalling rail traffic, patrolling and maintaining the rail infrastructure, monitoring electrification and managing the information flow across all aspects of the rail network – activities that occur on a continuous basis in order to keep the rail system working safely, efficiently and productively.

Network Rail's Ergonomics National Specialist Team (NST) has taken the lead on fatigue management for the organisation, coordinating a number of interventions across the company's various functions. Recognising the need to ensure compliance with current legislation on managing working hours particularly with staff in roles that are safety critical, as on the railway, the team has taken a proactive approach to enhancing fatigue management practice across rail operations and maintenance to

ensure Network Rail's employees are able to work effectively within a system that is reliant on human intervention. Using a holistic systems perspective, the Ergonomics NST has tackled fatigue management from a number of different perspectives, all with the goal of ensuring the issue is on the forefront of the company's health and safety agenda.

The Ergonomics NST – taking the lead on fatigue

Network Rail's Ergonomics NST has led a number of initiatives to tackle the issue of fatigue for the company's safety critical employees. Interventions have been undertaken in four key areas: Rostering and risk assessment, education and training, monitoring and managing exceedances of working time, and workload and fatigue.

Rostering and risk assessment

The development of rules on working hours is one specific area of safety risk management that is pertinent to the rail industry today. Shift workers, in particular, have been identified as a vulnerable population given reported effects on health and well-being, performance and productivity, and attention and concentration (e.g. Rosa and Bonnett, 1993; Williamson and Feyer, 1995).

Research has also identified a number of specific components of shift working that contribute to an increase in the risk of fatigue, including with respect to night working (Scott, 1994), the availability of meal breaks (Tucker, 2003) and the length and arrangement of breaks between shifts (Barton and Folkard, 1993). Despite this, shiftwork remains a viable option for many industries including rail that require 24-hour operations given it allows the provision of round-the-clock services, including allowing opportunities for maintenance, with the aim of expanding productivity, maximising the benefits of technology and increasing profitability (Smith and Barton, 1994).

Within the rail industry, arrangements have been put in place for safety critical workers (that is, employees who are required to work on or with a direct interface to the railway line) to provide a degree of control over how working time is scheduled. The Rail and Other Guided Transport Systems (Safety) Regulations 2006 were recently updated, and Regulation 25 requires organisations to monitor and control the working hours of safety critical employees as part of the wider management of risk due to factors such as fatigue. The standard reflects current industry guidelines on working time limits that arose following the Clapham rail accident in 1988; the investigation into the incident identified fatigue resulting from excessive hours of working as a contributing factor (Hidden, 1989), and the following limits were put forward as an industry benchmark:

i. A maximum of 12 hours to be worked per shift
ii. A minimum of 12 hours rest between shifts
iii. A maximum of 13 consecutive shifts to be worked
iv. A maximum of 72 hours to be worked per week.

These limits, referred to within the rail industry as the "Hidden limits", have acted as the minimum requirements for working hours for safety critical employees since this time.

The Ergonomics NST has taken a key role in integrating the risk assessment process required by the Regulations into Network Rail's rostering practices. The Health and Safety Executive's Fatigue and Risk Index Calculator (Spencer, Robertson and Folkard, 2006) has been adopted as a useful tool to contribute to this process; the index enables the risk assessment of any rotating shift roster with respect to how an identified pattern of working hours might impact on a person's experience of fatigue and their likelihood of error. The index was used as part of a company-wide change to the 35-hour working week to assess the proposed rosters for safety critical shiftwork employees for potential fatigue and error risk. The aim is to integrate the use of the index as a form of risk assessment into day to day practice across the organisation.

Education and training

An effective method of ensuring Network Rail meets it's obligations with respect to the legislation on managing fatigue is to ensure those at risk of fatigue at work are appropriately informed about those risks, and of the methods they can apply to mitigate against them. The Ergonomics NST has taken the lead on fatigue education for Network Rail, integrating the principles of fatigue management into a number of the company's training courses, particularly those aimed at safety critical staff and shiftworkers. To date, this process has focused predominantly on operational staff including signallers and Mobile Operations Managers, though the module on managing shiftwork and fatigue is currently being adapted for briefing across Network Rail's maintenance organisation as well. Education has focused on a number of key areas, including:

- the effects of shiftwork, and particularly night working, on individuals
- a review of the relevant legislation and guidance aimed at managing fatigue
- the links between fatigue and performance, attention and safety
- the identification of strategies for managing fatigue
- a review of the key individual and organisational responsibilities to manage fatigue.

An education programme is also in development for managers of shift workers who develop rostering schedules, providing guidance to ensure rosters are designed and risk assessed in accordance with fatigue management good practice.

Monitoring and managing exceedances of working time

In theory, having clearly defined limits on working hours should make the process of controlling those hours an easier task –where legislation sets rules about working time, so those rules should be followed and mitigation against fatigue should be prioritised. Yet industry, and the rail industry more so than ever, is dynamic and constantly evolving; where safety is of the utmost importance, it is often competing

with productivity and efficiency to meet the demands of stakeholders, passengers, parliament and an over-scrutinising media (Knight, 2002). Safety critical employees may be rostered according to working time requirements, but when an incident occurs or delays in work accrue, there may be times when deviations from scheduled working time will occur.

The Ergonomics NST has designed and now owns Network Rail's company standard on the control and monitoring of working time exceedances for safety critical staff. Though rosters are devised in line with the requirements of the identified working time limits, the standard allows for exceptional circumstances where individuals may be asked to work beyond those, such as might occur in the case of a critical incident, equipment failure, extreme weather conditions or an unforeseeable shortage of staff that would cause significant operational disruption. In these instances, a process has been developed to capture the details of each exceedance and to ensure a formal risk assessment is undertaken and appropriate fatigue mitigation measures are put into place before any exceedance of working time is authorised. This information is then captured and reported as a corporate KPI, the aim being to monitor the level of exceedances across the network, to flag up locations where exceedances appear to be a persistent issue, and to then enable active management of causation factors and any blocks to the provision of effective mitigation strategies. The Ergonomics NST has been working closely with Network Rail's safety and compliance team to ensure the KPI remains targeted, that exceedances continue to be measured, and that appropriate action is taken where exceedances are identified.

As custodians of the standard on the control of working time exceedances, the Ergonomics NST regularly reviews the extent to which the standard and the processes it outlines are integrated into day to day practice across the organisation's large population of safety critical staff. As an example, an exploratory study on the standard's integration into practice was undertaken in early 2007 using a combination of Critical Decision Method (Klein, Calderwood and MacGregor, 1989) and semi-structured interviews with a sample of frontline staff and their managers in a small sample of Network Rail's maintenance depots (Kenvyn, 2007). The aims of the review included the following:

- To understand the extent to which fatigue management and the control of working hours were integrated into current practice within the maintenance depots sampled
- To understand the extent to which the process for monitoring and authorising working time exceedances were being carried out at a depot level
- To understand how decisions about authorising exceedances were being made, including the extent to which fatigue and risk were considered.

While the review involved only a small number of sites, the findings highlighted issues relating to real-time completion of exceedance forms, the undertaking of a formal risk assessment at the time of authorisation, follow-up by the wider organisation on reported exceedances, and perceived conflict between the need to complete work tasks to maintain the safety of the line and the management of fatigue risk in

staff who have been asked to work beyond their scheduled hours. The findings will inform a wider company review of compliance with the standard and its planned update during 2008/09.

Workload and fatigue

The study of workload forms a significant component of the work undertaken by the Ergonomics NST at Network Rail. A workload toolkit has been developed specifically for the railway operational environment by the Institute of Occupational Ergonomics at the University of Nottingham (see Pickup, 2006) and is currently applied in signalling environments. Assessments of workload are often required where there are increases in timetabled rail traffic, changes to an area of operational infrastructure (such as where a signalbox is closed and the managed infrastructure is transferred to another control centre) or where a signalling system is upgraded or enhanced in a particular location, as might occur where a signalling lever frame is upgraded to a VDU signalling workstation, for example.

With respect to fatigue, the assessment of workload in signalling environments enables the Ergonomics NST to inform decisions about issues including effective manning levels and appropriate arrangements for rostering and meal breaks, and to influence the arrangement of work tasks and the design of workstations and signalling systems to ensure that signallers are not at risk of fatigue as a result of their work system. The Ergonomics NST is also leading on the more specific issue of meal breaks through consultation with the Operations and Customer Services function within Network Rail. The goal is to develop a clear assessment process for establishing whether sufficient opportunity exists for adequate breaks from work activity among signalling staff, ensuring compliance with the ROGS requirements. This will also involve training local managers to undertake their own initial assessment to identify signalling locations that might be "at risk" due to a combination of both workload and the current rostering arrangements for that location.

Conclusion

Fatigue management is an essential component of any organisation's safety management system. Network Rail is one organisation that has recognised this, and has subsequently made great progress on integrating fatigue management into its day to day practice. Network Rail's Ergonomics National Specialist Team has played an integral role in placing fatigue management on the agenda as a key health and safety responsibility, identifying the potential challenges and taking the lead in identifying the way forward for the organisation.

References

Barton, J. and Folkard, S. (1993). Advancing versus delaying shift systems. *Ergonomics; 36* (1–3): 59–64.

Hidden, A. (Q.C.) (1989) *Investigation into the Clapham Junction Railway Accident*. Presented to Parliament by the Secretary of State for Transport, November 1989. London: HMSO.

Kenvyn, F. (2007) *Influences on real-time decision making about extending working hours within the rail industry (Masters Dissertation)*. Birkbeck College, University of London, London.

Klein, G.A., Calderwood, R. and MacGregor, D. (1989). Critical decision method for eliciting knowledge. Special issue: Perspectives in knowledge engineering. *IEEE Transactions on Systems, Man and Cybernetics*; *19*: 462–472.

Pickup, L. (2006). *An Overview of the Development, Testing and Application of the Workload Toolkit (IOE/RAIL/06/02D)*. IOE University of Nottingham, Nottingham.

Rosa, R.R. and Bonnet, M.H. (1993). Performance and alertness on 8 h and 12 h rotating shifts at a natural gas utility. *Ergonomics; 36*: 1177–1193.

Scott, A.J. (1994). Chronobiological considerations in shiftworker sleep and performance and sleep scheduling. *Human Performance; 7* (3): 207–233.

Smith, L. and Barton, J. (1994). Shiftwork and personal control: Towards a conceptual framework. *European Work and Organizational Psychologist; 4* (2): 101–120.

Spencer, M.B., Robertson, K.A. and Folkard, S. (2006). The development of a fatigue/risk index for shiftworkers. Health and Safety Executive Report No. 446. London: HSE Books.

Tucker, P. (2003). The impact of rest breaks upon accident risk, fatigue and performance: a review. *Work and Stress; 17* (2): 123–137.

Williamson, A.M. and Feyer, A.M. (1995). Causes of accidents and the time of day. *Work and Stress; 9* (2–3): 158–164.

APPLICATIONS OF HANDHELD COMPUTERS FOR THE RAIL INDUSTRY

Yasamin Dadashi[1], John R. Wilson[1,3], Sarah Sharples[1] & Theresa Clarke[2]

[1]*Human Factors Research Group, University of Nottingham, University Park, Nottingham, UK*
[2]*Ergonomics National Specialist Team, Network Rail, London, UK*
[3]*School of Safety Science, University of New South Wales, Sydney, Australia*

Advances in mobile computing technology and software applications have led to an expansion in potential uses for handheld computers. Network Rail has been progressively developing and applying handheld computers to field-based maintenance and inspection operations, with the aims of improving productivity, quality and personal and system safety. However, it is clear that these aims so far have been achieved with varying degrees of success. This research has been established to examine current applications of handheld computers in the management of infrastructure. As a first step, evaluative case studies have been conducted, with the objective of identifying the interaction issues that exist within current systems. This paper draws from the early stages of the research for Network Rail, explaining how handhelds are used for track engineering work. The results of analysing two handheld computer systems currently used in Network Rail are presented. The results obtained from evaluating these handheld computer systems have been summarised in terms of five principles which should be considered when designing user interfaces for handheld computers.

Introduction

The rail infrastructure in the UK is maintained by Network Rail. Handheld computers have been introduced to maintenance and inspection operations for over ten years and the company is trying to introduce and implement better and more comprehensive systems to enhance the productivity and safety. This research studies applications of handheld computers in the UK rail industry. The main objective of the research is to provide context specific guidelines for design of mobile user interfaces as well as methods and tools to support these guidelines.

A series of evaluative case studies were conducted with the aim of identifying interaction and usability issues of handheld computers. Two handheld computer systems used within Network Rail were examined. The first system has been implemented to assist Signalling and Telecommunication (S&T) inspectors with

their task. This system is used by S&T inspectors for recording the completion of inspection tasks which they perform in any shift. The handheld computer presents a list of all the locations that need to be inspected in one week. The users are required to determine whether the equipments in that location have been inspected by answering a few questions on the handheld computer. They should also register any detected defects on a separate form.

The second handheld computer system, still under trail, has been introduced to aid Level Crossing (LX) inspectors. This device is to supersede the paper-based inspection forms. LX inspectors use the handheld computer to fill in the inspection forms which record the condition of different assets at the level crossing. The information gathered by both handheld systems is used to update an asset management database which schedules maintenance operations and cyclic inspections.

The software for both systems have been designed by the same company and different features of the user interfaces such as the layout of data entry fields, colour schemes, text sizes, and structure of the menus are identical. Furthermore, the hardware used for both systems is a Motorola MC 9002 handheld device.

In this paper, the results of the usability questionnaires and interviews with participants are reported and the two systems are compared in order to determine the reason for the differences between the two systems.

Methodology

The body of research regarding the appropriate choice of method, data collection and analysis for handheld computer evaluation is not as strong as it is for conventional systems (Kjeldskov and Skov, 2003). It was important to ensure that the selected methods provide a good understanding of the context of use and the task. For these reasons it was decided to evaluate the LX and S&T handheld computer systems using different methods including heuristic evaluation, field visits, and interviews with users as well as a usability questionnaire which has been specifically developed for handheld computers within Network Rail.

The objective of the heuristic evaluation was to provide the evaluator with a better understanding of how the systems operate. In addition to heuristic evaluation, a series of field observations were conducted which provided valuable information about the task and context of use. Because of the nature of the rail industry and environmental constraints, it was not possible for the researcher to study users' interaction with the device on site. Therefore, the main focus of the field visits was on gathering data about the task and context of use. In order to observe the way users interact with the interface, they were asked to perform a typical task on the handheld computer once they were back at the depot.

A semi structured interview was also designed in order to enable the researcher to understand users' experience with the device as well as the way they feel about the system. The information gathered by the interviews was complemented by the data which were collected by the usability questionnaire. The usability questionnaire for handheld computers contained 12 usability factors which were measured by 42 statements. The questionnaire was designed based on an adaptation of a method

Table 1. Average overall ratings of usability factors for the S&T and LX handheld computers

Usability factors	Ease of use	User interface	Portability	Consistency & task relevancy	Feedback	Productivity	Adaptability	Affective design	Technology	Workload	Errors	Help
S&T handheld computer	2.70	2.84	2.79	2.92	2.35	1.94	1.88	1.69	1.50	1.94	2.58	2.79
LX handheld computer	4.36	4.38	4.44	4.53	3.83	4.28	3.67	4.00	3.67	4.33	4.54	3.32

suggested by Ryu and Smith-Jackson (2005). The participants were asked to rate the statements based on a five point likert scale where one represented "strongly disagree" and five "strongly agree" with three being the neutral point on the scale. Since the questionnaire has been designed for evaluating any handheld computer system within Network Rail, the scale also includes a Not Applicable option.

Results

In this section the results of the usability questionnaire and interviews with the users will be presented. Sixteen S&T inspectors and supervisors from four depots and six LX inspectors from three depots were interviewed and asked to fill in the usability questionnaire. The initial results of the studies showed that despite the similarities in the user interfaces of both systems, the S&T handheld computer is considered by the end users as a time consuming device that has just added to their problems. On the contrary, the level crossing inspectors believe that the handheld computer has enhanced the way they perform their tasks.

Table 1 summarises the overall average ratings of the usability factors for each of the handheld computer systems. It was decided that a score of three or higher represents respondents' agreement with the statement and therefore, an overall average of three or higher for any of the factors should in theory indicate that the system has been successful in supporting the measured usability factor. The results presented in this table clearly demonstrate the differences between the perceived usability of the two handheld computer systems.

The results obtained from the questionnaire show that none of the factors measuring the usability of the S&T handheld computer system have received a rating of three. This suggests that the S&T inspectors do not consider the handheld computer as a usable system. In fact, the data gathered from the interviews revealed that the S&T inspectors do not take the handheld computers on site. The main reason for users' reluctance in using the handheld computers is that the information provided by the system is not sufficient for performing the inspection task and at times it is

Table 2. Average rating of common statements for S&T and LX handheld Computer Systems

Statements	S&T	LX
2.3- It is easy to input text and information into the handheld computer.	2.88	4.5
3.3- The handheld computer is usable in all weather conditions.	3.06	4.83
3.4- The handheld computer is usable in all light conditions.	3.44	3.33
3.6- The handheld computer size is convenient for transportation.	3.13	4.5
3.7- The handheld computer is tough and would not break easily.	3.75	4.5
5.4- Highlighting the selected menu options on the handheld screen is useful.	2.53	4.33
9.2- The handheld computer is fast enough.	1.69	3.67

very difficult and time consuming to access the required information. S&T inspectors believe that their productivity has decreased considerably because of the time they spend filling in the forms on the handheld computers. The importance of this issue becomes clearer when the main objective of introducing handheld computers for S&T inspection is considered which is increasing productivity by reducing the amount of paperwork. Considering the factors that have received ratings closer to three, which was the minimum acceptable score, it becomes apparent that the interface is relatively easy to use and learn, the device is sufficiently portable and the terminology used within the system is consistent and relevant to task. The unsuccessful deployment of the device is likely due to the fact that the applications on the handheld computer do not comply with the way inspectors perform their task and it does not match their information requirements. The data gathered from semi structured interviews show that the only information the inspectors require to perform the inspection task is: 1- What the job is; 2- Where it is; and 3- What service it requires. The handheld computer contains this information, but in order to access this information, users have to navigate through several interfaces and search for the data which has made the task very time consuming.

As it can be seen in Table 1, the results obtained from evaluating the LX handheld computer are significantly different from the results of the first system. All of the 12 usability factors in the questionnaire have received ratings higher than three. In other words, LX inspectors regard the handheld computer as a usable system. The application on the LX handheld computer has been designed to substitute the paper-based inspection forms. These forms are the most important item of information LX inspectors require to perform their daily task. Presenting these forms on a handheld computer has reduced the amount of paperwork and increased the speed of performing the task without compromising the quality of the collected information. Therefore, despite a few usability issues which were mentioned during the interviews, the users are very satisfied with the system and the way it has changed their work pattern.

Table 2 presents the average ratings of seven individual statements. These statements are independent of the task and are only concerned with features of the interface which are not related to the tasks performed, so similarity of ratings would be expected. The data presented in this table show that S&T inspectors have given

lower ratings to these statements compared to the scores given to the same statements for the LX handheld computer. These data show that a mismatch between the applications on the handheld computer and the task performed by end users affects their view about other aspects of the interface as well.

Discussion

The following principles for developing a user interface for handheld computers have been derived from the results of the evaluative case studies reported in this paper:

1. Match between the applications on the handheld computers and information and tasks: Although this principle is important for any computer system, it is particularly vital for mobile computers, because while working on site, the information presented on the handheld computer would be the only point of reference.
2. Reliability of the technology: perhaps the most important characteristic of a handheld computer is its portability. This means that the handheld computer will be used for mobile tasks where users do not have access to other sources of information and rely totally on the handheld computer.
3. Suitability of the hardware: The results of the interviews with S&T inspectors show that many of the participants believe that a handheld computer is not a suitable hardware for the current application and unless the system is further developed to offer more task-related applications and information, a desktop computer is probably a more appropriate hardware. Therefore, it is important to ensure that the selected hardware matches the task requirements.
4. Consistency with other systems and between different platforms: It is important that any new computer system is consistent with other systems that users are familiar with and use frequently in their daily life. It is also essential to ensure that information is presented consistently across various platforms within the organization, e.g., the central database.
5. Feedback, training and support: Designing a task and context specific application is not sufficient for successful deployment of a system. Comprehensive training and continuous support are vital to successful implementation of a handheld system. Several users commented about the lack of any help documents within the system.

The findings of this study confirm the fact that an “easy to use” system is not necessarily “usable”. According to ISO 9241 – 11 (1998) usability is defined as “The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” S&T inspectors do not have any difficulties with operating the handheld computer. The system is not being used because it fails to meet the informational requirements of performing the inspection task.

Conducting research on applications of handheld computers in the rail industry was impossible without obtaining an understanding of how the current systems

operate. These evaluative case studies have provided important information about typical context of use of a handheld computer in the rail industry. Also, the information gathered from the field visits together with a literature review on human computer interaction models provided the necessary grounds for developing a theoretical framework for understanding users' interaction with a mobile computer. Furthermore, the findings of this study led to development of an applications matrix which can potentially be used as a reference to provide data about the information requirements of mobile workers and potential handheld computer applications for designers and system developers.

References

Borland, A., 2005, *S&T National Rollout Feasibility Support Study, 2005,* Technical Report, AMT Sybex Ltd.

ISO 9241 – 11, 1998, *Ergonomics Requirements for office work with visual display terminals (VDTS), Part 11: Guidance on Usability.*

Kjeldskov, J. & Skov, M., 2003, Evaluating the Usability of Mobile Systems: Exploring Different Laboratory Approaches. *Proceedings of the HCI International 2003, Crete, Greece* Vol. 2, p. 122–127.

Ryu, Y.S. & Smith-Jackson, T.L., 2005, Usability Questionnaire Items for Mobile Products and Content Validity. *In Proceedings of HCI International 2005, Las Vegas.*

STRUCTURED OBSERVATIONS OF AUTOMATION USE

Nora Balfe[1], John R. Wilson[1,3], Sarah Sharples[1] & Theresa Clarke[2]

[1] *Human Factors Research Group, School of Mechanical, Materials and Manufacturing Engineering, University of Nottingham, Nottingham, UK*
[2] *Ergonomics National Specialist Team, Network Rail, London, UK*
[3] *School of Safety Science, University of New South Wales, Sydney*

This paper concerns automation in rail signalling. A series of observations of rail signalling staff using an advanced automation system were undertaken at a number of locations in order to gather information on the way in which signallers' behave when using the system. Signallers' activities were coded into five categories – monitoring, intervention, planning, communications and quiet time, and signaller activity was coded at 5 second intervals. Analysis of the data gathered revealed insights into the strategies adopted by signallers in using and controlling the automated system and revealed differences between individual signallers' strategies of use. Major findings of the study are presented, including differences in monitoring and intervention levels between signallers and the regulation of workload by varying levels of monitoring. The implications of this research with respect to future rail signalling automation systems, and future research plans are discussed.

Introduction

Some signalling centres in the UK use an advanced form of automation called ARS (Automatic Routing System) to assist signalling staff in routing trains. The automation routes trains according to their timetable but signallers are still required in normal operations to monitor the route setting and to manually set routes for trains which do not have a timetable available to ARS. As with many other forms of automation (Bainbridge, 1983), ARS is not capable of dealing with non-routine events such as train or signal failures, and the signaller is required to take over in circumstances like these.

Observational studies of ARS at three Integrated Electronics Control Centres (IECC) were conducted in order to establish how the signallers currently use this automation and the factors that influence automation usage. There are many anecdotal reports and opinions of how signallers use or misuse this automation but there is a lack of empirical data on system usage. This study set out to document the activities signallers in IECCs routinely engage in and to determine the magnitude of the differences in these activities between individual signallers and between signal boxes (IECCs).

The aim of the study was to identify whether differences exist in individual signallers' use of automation. The objective of the study was to observe signallers' current use of ARS and determine what proportion of their time is spent monitoring, planning, controlling, communicating or not actively involved (quiet time) during normal operations.

Methodology

A method of systematic observation was adopted for this study. "Field observations support a discovery process, they serve to draw attention to significant phenomena and suggest new ideas, the validity and generality of which can then be evaluated through additional studies" (Vicente, Roth, & Mumaw, 2001, p. 835). Signallers were observed at their workstation and a coding system was used at time intervals of 5 seconds to capture their activities.

Three signallers were observed on two workstations in each IECC. The IECCs chosen for the study were picked on the basis of the reported quality of ARS in each. Perceived quality is reportedly high in Liverpool Street, low in Ashford, and mixed in. The choice of workstation within each IECC was based on attempting to match the relative complexity of the workstations as closely as possible. The Operational Demand Evaluation Checklist (ODEC) was used to evaluate each workstation's complexity (Pickup & Wilson, 2005) to aim to ensure matching complexity. Each observation was carried out at the same time of day so that as far as possible the same traffic levels and patterns were experienced by the signaller. Operational constraints meant that participant selection was by opportunity sampling only and participants were chosen on the basis of availability and willingness to take part in the study.

The activity of the signaller was coded every five seconds over a 90 minute period using the following coding structure:

- Monitoring – this occurred when the signaller was watching the signalling screens. It was not possible to tell how much attention he was paying.
- Intervention – Interventions were coded when the signaller was using the trackerball and keyboard to become involved in the signalling process, either to instruct the automation or to control a train manually.
- Planning – this referred to specific planning activities the signaller engaged in, such as looking up train timetables using a number of devices available to him. It was not possible to tell how much of the time coded as monitoring was spent developing a plan of action.
- Communications – signallers frequently need to communicate with a variety of other operational personnel, including drivers, other signallers and controllers.
- Quiet Time – this was any activity the signaller undertook which was not directly related to the signalling task. Common examples of this included making a cup of tea, having a chat with other staff, and checking email.

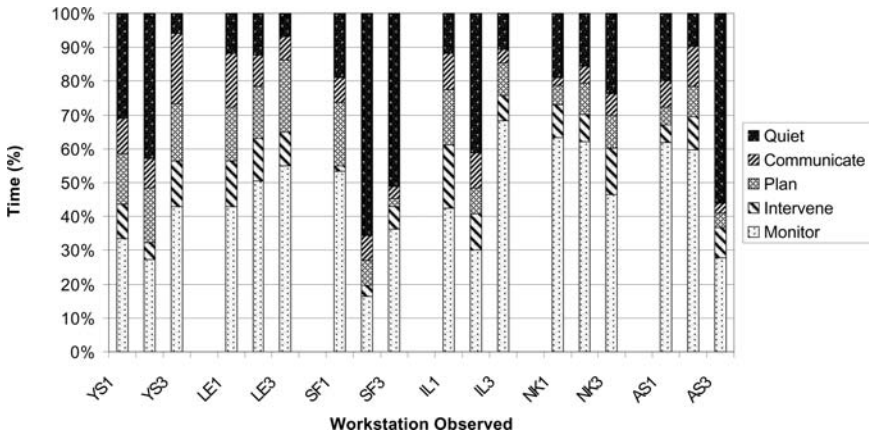


Figure 1. Signaller activity.

Results

Figure 1 shows the results of the observations in terms of the 5 coded activities. Each bar on the graph describes the distribution of activity for one signaller – workstations YS and LE in York IECC, SF and IL in Liverpool Street IECC, and NK and AS in Ashford IECC. Three signallers were observed at each workstation.

The graph shows clear differences in activity between signallers. The same signaller was observed on LE1 and YS3, and these two graphs are remarkably similar. SF2 and IL2 also show the same signaller, but these graphs are different, there was disruption on one of IL2 which contributed to the observable difference.

Discussion

From the graph it is clear that signaller activity can vary greatly despite experiencing very similar situations over the observation period. This is an indication that rather than workload being driven purely by the circumstances on the workstation, signallers drive their own involvement in the workstation to a certain extent. The similarity of the graphs of the same signaller on LE1 and YS3 provide further evidence that activity is signaller led rather than workstation driven.

Monitoring

During the course of the study it became clear that signallers do not appear to maintain consistent levels of monitoring, but engage in two distinct types of monitoring. Active monitoring was defined as the signaller sitting up while monitoring and was often associated with his hand on the trackerball and scanning behaviour. Passive monitoring was defined as the signaller sitting back while looking at the screens, although it was impossible to tell to what degree he was focusing on the screens. Passive monitoring was typically carried out for longer periods of time than active

monitoring, and was often interspersed with periods of quiet time. The data gathered in this study was insufficient to contribute to an understanding of the factors involved in changing between active and passive monitoring, but a later study is planned which will gather data on the triggers to different types of monitoring.

It is clear from the graph that the workstation does not drive the monitoring level, as very few of the graphs for each workstation are similar. Therefore that monitoring levels appear to be driven by the individual. It could be that although individual monitoring levels varied, overall they were highest in Ashford, lowest in Liverpool Street and medium in York. This would seem to relate to the anecdotal accounts of how well the automation runs in each IECC. Therefore, there is a possibility that monitoring levels are at least influenced by the quality of the automation but there is not sufficient evidence to conclude that this is certainly the case.

High monitoring (either active or passive) is associated with low levels of quiet time, and low monitoring is associated with high levels of quiet time. One possible explanation is that monitoring and quiet time are interchangeable and signallers who can find a 'distraction', e.g. someone to talk to, will use time otherwise used for monitoring.

Intervention

Two types of intervention were coded, use of the trackerball and use of the keyboard. The trackerball was used far more often than the keyboard. Use of the keyboard was highest in York IECC but as this study was interested only in intervention levels rather than the purpose of the interventions, the reason for this difference is not known. Further research could investigate common interventions, how they are achieved and the length of time required.

Following the observations, signallers were asked for the most common reasons for interventions during the observation period. Setting manual routes and applying reminders to prevent ARS from setting a route were the most common forms of intervention. Some signallers also reported that they intervened to set manual routes because they did not trust ARS to do it correctly. Interventions were occasionally necessary to manually route trains around a blockage or over a faulty track circuit. However problems like these on a workstation did not necessarily result in higher intervention rates, for example, the signaller during the YS1 observation was coping with severe disruption following earlier power failures but YS3, which did not have any disruption, had higher rates of invention. Similarly, there were incidents on both SF2 and IL2 during the observations but again these were not the highest observed levels of intervention. Therefore, intervention appears to be driven by the individual rather than the level of activity on the workstation.

Unfortunately it is not known what influences higher intervention levels. It is clear that it is not totally due to particular circumstances on the workstation but the difference in the times may be due to some signallers intervening in more circumstances than others, or some signallers using more efficient intervention strategies. Perhaps some signallers intervene in advance and are thus able to sort out potential problems quite quickly.

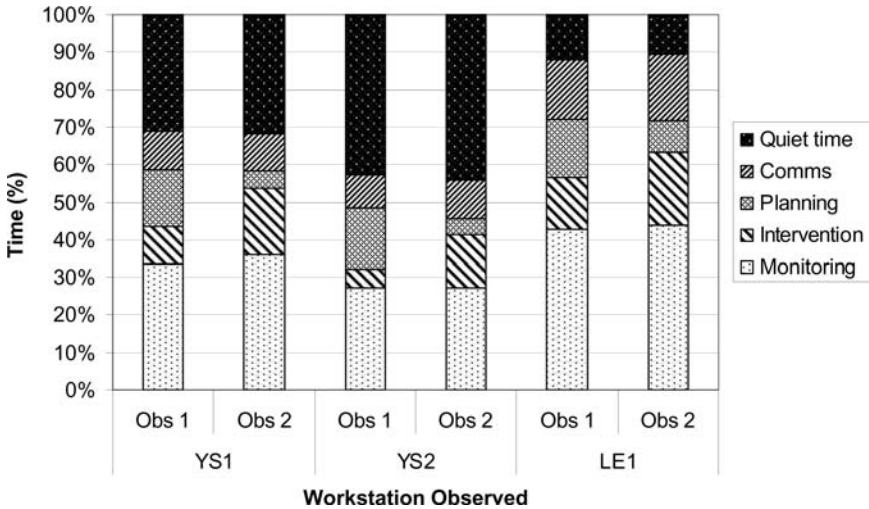


Figure 2. Inter-rater comparison.

Inter-rater reliability

Three of the observations were undertaken by two different observers to determine the inter-rater reliability for the method. There was a crucial difference between the observers as one had considerably more knowledge of the signalling task than the other. Figure 2 shows the comparison of the results for the two observers.

The graphs for each workstation are very similar in terms of monitoring, communications and quiet time, but there is a discrepancy in terms of intervention and planning. This is due to the difference in knowledge of the two observers as the more experienced observer was aware of which input devices were attached to the signalling system and which were for interacting with the planning systems while the less experienced observer coded all use of input devices as intervention. Cohen's Kappa (Robson, 2002) for the 3 observations YS1, YS2 and LE1 were 0.55, 0.52, and 0.52 respectively. This is not as high as might be expected given how closely the graphs mirror each other, however, as the data were coded at intervals of only 5 seconds, a small time lag between the observers would have had a large effect on the calculation of Cohen's Kappa while having almost no effect on the graphs produced.

Conclusion

The observations show a variation in activity between different signallers on the same workstation. This clearly indicates differences in the approach taken to monitoring and controlling the automation. There was an indication that monitoring levels may be related to automation quality; however intervention levels appear to be entirely driven by the individual. Two different types of monitoring activity were

identified during the observations and further research is required to examine the differences in signaller activity and workstation circumstances during active and passive monitoring. Finally, the inter-rater reliability of this method was found to be fair in terms of Cohen's Kappa.

References

Bainbridge, L. (1983). Ironies of automation. In J. Rasmussen, K. Duncan & J. Leplat (Eds.), *New Technology and Human Error* (pp. 271–283). Chichester: Wiley.

Pickup, L., & Wilson, J. R. (2005). *Mental workload assessment and the development of the operational demand evaluation checklist.* Paper presented at the Second European Conference on Rail Human Factors, London.

Robson, C. (2002). *Real World Research* (2nd ed.). Oxford: Blackwell.

Vicente, K. J., Roth, E. M., & Mumaw, R. J. (2001). How do operators monitor a complex, dynamic work domain? The impact of control room technology. *International Journal of Human-Computer Studies*, *54*(6), 831–856.

NETWORK RAIL SIGNALLER'S WORKLOAD TOOLKIT

Emma Lowe[1] & Laura Pickup[2]

[1]*Ergonomics National Specialist Team, Network Rail, London, UK*
[2]*Human Factors Research Group, School of Mechanical, Materials and Manufacturing Engineering, University of Nottingham, Nottingham, UK*

Within the operational rail industry there has been a need to understand the concept of workload in relation to signallers. Research to develop a concept of workload in signalling suggested that workload is a multi-dimensional concept and therefore requires a number of different techniques in order to assess it in an operational context. To that end the Network Rail workload toolkit is made up of a number of different workload tools designed to measures different dimensions of workload including subjective perceptions, actual activities and elements of the work system. A combination of tools will always be used and the results are integrated in such a way as to present a 'workload profile' of an area of control. This paper outlines the work that has been undertaken to develop the toolkit and provides an overview of its application in field based assessments.

Introduction

One of the most widely used and debated concepts in ergonomics is that of workload. The notion has found widespread acceptance as of value in assessing the impact of working arrangements, new tasks, and team sizes, in comparing the effects of different or job interface designs and in understanding the consequences of different levels of automation. However, workload is one of those seductive concepts that are so apparently meaningful to specialist and lay person alike, and which can be adapted to fit many contexts. This has resulted in a large number of (sometimes contradictory) viewpoints and theories on what workload is and how it should be measured. This was certainly the case in signalling in the rail industry where these multiple meanings have been the cause of great confusion. It highlighted the need for a clear understanding of what is embraced by the concept of workload before the most effective way of measuring it could be identified. A five phase project was undertaken involving: a literature review of the theory and existing tools; the development of a workload framework; development of a number of workload assessment tools; application of the workload tools in the field and workload tool evaluation. The outcome of this work was the Network Rail Signaller Workload Toolkit which consists of a number of different workload tools designed to measures different dimensions of workload including subjective perceptions, actual activities

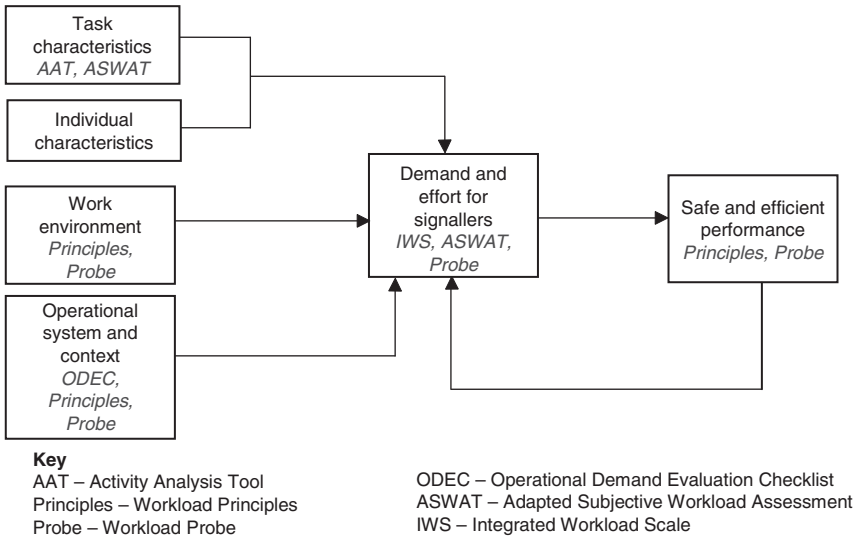


Figure 1. Workload framework showing different assessment tools.

and elements of the work system. A combination of tools will always be used and the results are integrated in such a way as to present a 'workload profile' of an area of control.

Workload framework

The literature review revealed that it is well established that workload is a multidimensional concept that can be considered as a combination of factors concerned with:

- the task – the number and combinations of tasks they have to complete
- the context – how and where they have to complete them and the urgency or accuracy necessary to ensure safety and organisational performance targets are met
- the individual – a signaller's own skill, experience and perception of their work

Given that workload is a multi-dimensional concept it follows that a number of different techniques are required in order to assess it in an operational context.

A conceptual framework was developed (Pickup et al 2003) to help direct where assessment tools are required to assess the different dimensions of workload. A simplified version of this framework is shown in Figure 1 above. This is not necessarily an operational model but has been proposed in order to develop and position the toolkit of methods to understand and assess workload. Thus it is explanatory of routes to measurement rather than of the mechanisms by which workload is caused.

The Network Rail Signaller Workload Toolkit

The Workload Toolkit is currently made up of six tools with an additional one, the Time and Pressure Evaluation tool (TAPE), still under development. Below is a description of the tools including details about the purpose of the tool and how it is used.

Workload Principles Tool

The Workload Principles Tool provides an assessment of the work system in relation to the degree to which it meets a number of ergonomics principles. The rationale is that if a work system has been designed to meet generally accepted ergonomics principles then any mismatches between people's capabilities and their work are avoided and consequently workload is not increased. Therefore the tools works by assessing the extent to which the principles are met is an indication as to the impact workload is having on the signaller's ability to perform safely and efficiently.

Using the tool simply involves considering all the information obtained through observation or discussion with the signaller, or line manager, and identifying whether each principle is met. Through a paired comparison exercise some of the principles were identified as primary principles and essential for safety and performance. If they are not met then remedial action is required. The remaining (secondary) principles are desirable.

These principles have been found to be useful, meaningful and generally consistent with other observations made during a number of workload assessments of signal boxes. However, they do rely on the judgment of the investigator to state whether each principle is fulfilled or not. This judgment is only as good as the skill of the investigator and the quality of the information gained from signallers and their manager.

The Integrated Workload Assessment Tool (IWS)

The IWS collects real time perceptions of signaller workload based on a nine-point scale. This tool can be used to identify peaks and troughs in the effort and demand experienced by signallers when responding to dynamically changing work conditions. If it is used in conjunction with video recording, subject matter expert commentary or the Activity Analysis Tool (AAT), it will assist in highlighting which combinations of tasks or situations are considered to produce high and low levels of effort and demand (workload).

The strength of the IWS lies in its ability to quickly and effectively provide data, which can be compared from minute to minute if necessary, from situation to situation or even between individuals. However, it does not differentiate between the different dimensions that are the essence of the multi-factorial concept of workload. One risk is that inappropriate interpretations are made from the data. This includes calculations of a 'mean IWS score'. This may not represent the dynamic nature or multiple dimensions of workload and misleading conclusions could potentially be drawn.

Task Activity Analysis tool

It is useful to be able to relate workload to activity. This can assist in assessing which activities or scenarios in signalling are more or less demanding than others. The Task Activity Analysis (TAA) tool involves observing and recording signaller's activities at certain times. It has proved to be very useful in providing estimates of time occupancy at key times. It is most effective when combined with workload scores from the Integrated Workload Scale (IWS), as described above, and/or with subject matter expert (SME) commentary, which can provide rich information regarding the nature of the work and the effectiveness and efficiency of the signaller including any compensatory strategies or deterioration in performance.

The main issue with the tool is that it cannot account for unobservable events such as mental processing, which may vary independent of the observable actions and events within the job.

The Adapted Subjective Workload Assessment Technique (ASWAT)

SWAT was developed in other industries but seemed to be generic in terms of workload dimensions. It includes dimensions that signallers suggested as representing their interpretation of the term workload. The tool provides a relatively quick and easy general comparison scale for signallers to assess three dimensions of workload retrospectively. It allows a comparison of signaller workload between two situations (e.g., change in timetable) or different times in the day over a work period of time. It also allows some degree of diagnosis about where the signaller's greatest demands or effort might be.

The original SWAT has been adapted in 2 main ways to facilitate use within the signalling context. Firstly the original SWAT has three dimensions – time load, mental effort and psychological stress load. Testing with signallers suggests that the 'stress' is inappropriate. The culture of signallers appears to view stress as a weakness; the term pressure was more frequently associated with workload and was suggested by signallers as a suitable alternative. Hence the Adapted SWAT refers to pressure rather than psychological stress load. Secondly, the original SWAT normally involves 2 stage scale development. However, using the SWAT scale by considering the dimensions as continuous and having equal weighting avoids the need for the first phase of scale development. This makes it easier to use in the field.

Operational Demand Evaluation Checklist (ODEC)

A signalling system or signalling workplace can be described in terms of a number of entities, which are constant (static characteristics, such as the number of signals or level crossings) or variable (dynamic characteristics, such as the number of unplanned or emergency possessions). All of these entities can influence the workload of a signaller. ODEC provides a systematic process to evaluate these entities within any one particular workstation, in order to represent the influence the overall system has on the signallers' workload.

ODEC can be used before employing any of the other workload tools, to understand and give context to the work of a signaller and it is increasingly being used in

early predictions of potential workload within new or proposed signalling systems. It is applied to one work area at a time and requires objective, empirical data from a number of sources which are then categorised as either low, medium or high. This does not refer to workload per se, but to the extent that different entities are found in the target system or workstation as compared to other systems. That is, the scores are relative.

There are two main issues associated with this tool. Firstly, the implications of Automatic Route Setting (ARS) to signaller workload have yet to be understood and are not currently factored in. Secondly, when ODEC was first developed a final score could be obtained which was purely intended to represent the overall demand integrated across the ODEC ratings. However, field use suggested that there was a temptation to impose 'redline limits' based on such scores, but without sufficient information on context or data on comparisons across sites. It was felt that such comparisons could be meaningless and even dangerous. The single score was removed, leaving the more descriptive data – the percentage of high, medium and low categories – as the final ODEC output.

The Workload Probe

The Workload Probe is an analytical interview based tool that explores workload issues considered to exist within the signaller's working environment. It is intended to identify how and where a mismatch exists in the signaller achieving their goals in the time available and the context of their workplace. This tool does not intend to provide a redline limit to judge a signallers workload against. An interview is completed by a human factors specialist either with an individual or a group of signallers and aims to elicit information on positive and negative experiences that influence the signaller's workload. The interview involves general questions about their workload and then systematically asks the signaller to consider a number of loading factors that have been previously recognised as influencing signaller workload. This aims to understand how certain factors such as. signalling equipment and the timetable may come together to create a workload issue. A fishbone diagram with an exploration table is provided to facilitate the documentation of why signallers believe each loading factor influences their workload.

Developing a workload profile

The workload profile is an overall view of the findings obtained from the tools. The choice of tools to build the profile will be dependent upon the questions being addressed by the workload assessment. For example, one common reason for a workload assessment is to assess the acceptability of an area of control for one signaller to manage. In such circumstances a combination of all six tools is usually most effective. However, assessing the impact on workload of an additional level crossing to the area of control, for example, may simply require a combination of ODEC, the TAA and the IWS to build up a profile. Understanding if the workload profile is acceptable is not a cut and dried decision but using a combination of tools

to evaluate how compatible the working environment or context is in accommodating the signaller to achieve their work does offers a pragmatic approach to making that judgement.

Future work

The Workload Toolkit continues to evolve as experience is gained with its use. Specific improvements are being progressed for the ODEC tool so that it can capture information on the impact of automation and arrangements are being developed to ensure the results from the workload assessments are captured centrally to facilitate validation of the tools and so more reliable information and guidance can be given in respect of making judgements about workload acceptability.

References

Pickup, L., Nichols, S. and Wilson, S., 2003, *Fundamental Examinations of Mental Workload in the Rail Industry – theoretical review*, IOE/RAIL/03/06/R. The Institute for Occupational Ergonomics, The University of Nottingham.

Pickup, L. 2006, *An Overview of the Development, Testing and Application of the Workload Toolkit* IOE/ RAIL/06/02D. The Institute for Occupational Ergonomics, The University of Nottingham.

USE OF VISUAL SCENARIOS TO UNDERSTAND AND IMPROVE ENGINEERING WORK ON THE UK RAILWAY

Alex Schock[1], Brendan Ryan[1], John R Wilson[1, 3], Theresa Clarke[2] & Patricia Dingwall[4]

[1]*Human Factors Research Group, School of Mechanical, Materials and Manufacturing Engineering, University of Nottingham, Nottingham, UK*
[2]*Ergonomics National Specialist Team, Network Rail, London, UK*
[3]*School of Safety Science, University of New South Wales, Sydney*
[4]*Programme Manager Future Engineering Project, Network Rail, London, UK*

Engineering work on the UK railway, from simple maintenance procedures to large-scale renewals and enhancements, is performed in a dynamic and potentially hazardous environment. Human factors support has been provided to a high profile industry project examining ways to improve the safety, efficiency, and effectiveness of engineering work. The human factors work has included evaluation and recommendations for continued development of proposed new rules and procedures for protection of engineering work, working closely with subject matter experts and frontline operational staff. A scenario-based approach has been used in different stages of the project to identify important human factors issues. This paper outlines the development of visual representations of engineering scenarios and comments on the usefulness of the approach so far and for future work.

Introduction

In early 2006 the Chief Engineer at Network Rail formed an industry wide working group to examine ways to improve the efficiency and effectiveness of rail engineering work, whilst improving or at least maintaining current safety levels. Consideration was given to the safe and efficient access to track and a simplification of the rules and procedures for protecting work. A number of high level goals were identified and a programme was drawn up to shape the future for engineering work. A benchmarking exercise involving visits to France and Canada identified a potential new method for protecting workers based on the Canadian Track Occupancy Permit (TOP) process. This process uses the fixed signalling system to provide protection and open channel communications for managing and controlling activities.

Earlier work had already been carried out to understand current approaches to protecting engineering work in the UK, including the investigation of key roles and their associated activities (Ryan et al. 2007). The human factors contribution to the wider Future Engineering Project was continued with the Network Rail

Ergonomics National Specialist Team working with the Rail Human Factors Group at the University of Nottingham. The work has included analyses of critical functions of engineering work. This has lead to the identification of human factors issues to be considered further. This paper aims to provide an overview of the development and use of scenarios in the project, and in particular the development of an approach to produce visual representations of relevant rail engineering scenarios.

Use of scenarios in research

Scenarios describe stories about people and the activities they carry out (Potts 1995; Carroll 2000). Research into the use of scenarios has been carried out mainly in the field of Human Computer Interaction (HCI). They can be used as a method for describing interactions between users and systems to aid in the development of new or changing systems. Potts (1995) comments on their ability to illustrate specific procedures and demonstrate system behaviour. He further highlights how they can avoid misinterpretation of the system state by helping people understand complex, abstract issues that are present in a system.

Using scenarios helps to clarify thinking and encourage designers to come up with robust systems that offer the flexibility that is required. They can be used in a number of ways: firstly to walkthrough activities for the comparison of competing proposals for a system, through a systematic walkthrough of activities; secondly to compare various alternatives for achieving goals; and finally in indentifying potential barriers to goals, their consequences, and possible ways to overcome them (Potts 1995).

Scenarios can be effective in task analysis as way of comprehensively identifying tasks that are meaningful to users of the system (Carroll and Rosson 1992). Initially, tailoring scenarios at the "basic task" level i.e. the level that can be interpreted easily by users, can be beneficial as they represent real user experiences, yet they can be applied at different levels if required dependent on the target audience. Previous studies within the railway domain (Pickup et al. 2005), have used simplified rail scenarios to assist in the development of Signaller workload measurement tools. The scenarios developed for use during the future engineering project are multifunctional, dynamic and offer wider flexibility.

Use of a scenario approach in the future engineering project

Preliminary approach

Scenarios have been used in a variety of ways in the future engineering project. Initially, they were applied informally during one to one interviews and structured workshops with frontline operational staff to inform the breakdown of activities, with staff verbally describing and helping to visually represent situations to understand how work is managed, controlled and protected. The participants produced rough drawings and even moved objects around the table in attempts to simulate engineering processes such as train movements and protection arrangements. As

work progressed it became clear that there was a need for a more standardised approach to the visual representation of these engineering situations. It was anticipated that these could be used in two ways: firstly within Network Rail and the wider rail industry to gain support and user input into the development of the proposed system; secondly, as a human factors tool for work going forwards in the analysis and systematic identification of critical issues. A number of review groups were formed and workshops held to gain an understanding of the implications of change and to obtain user input into system design. These sessions included the use of scenarios to demonstrate how revised arrangements for protecting workers could be applied to a range of work situations in the UK. Feedback from workshops demonstrated the usefulness of the scenario based approach as a method for demonstrating procedures and complex processes and has provided the thrust for further development.

Development of visual representations of engineering scenarios

A number of scenarios were designed in consultation with the project management team and industry experts, considering project requirements and availability of access to industrial contacts. Five scenarios were chosen to cover a range of methods used to carry out large-scale rail renewals jobs and related activities. The development of the scenarios involved 4 stages.

Stage 1 – Scoping

The scope of work to be covered by the scenarios was defined. This involved working closely with industrial contacts to identify recent renewals work which would be suitable for further investigation. In addition, the methods for carrying out the work were carefully considered to cover a range of techniques and cater for different locations. These early scenarios were based on very recent work where resources such as job plans were readily available and information could be provided by staff who could recall onsite experiences of the work.

Stage 2 – Systematic walkthrough and data collection

Workshops were organised to facilitate the collection of information and work details. These were attended by the human factors researcher, a member of the project team and industry experts (e.g. from the renewals contractors, Network Rail staff). Industry experts provided job plans, work scheduling sheets, route maps and diagrams and conducted a systematic walkthrough of the entire job, using their operational and planning experience. This provided the researcher with a rich understanding of the sequence and timings of events, the resources involved and important factors associated with particular geographic locations.

Details were recorded during the walkthrough and the researcher coded information accordingly on diagrams. The human factors researcher led the walkthrough, requesting information whenever there was a change in the phase of work: this included any change in the protection arrangements, on-site management, movements of people or machines/equipment and the introduction of any new resources.

In addition, any internal or external changes to working conditions were recorded. Critical timings for activities were recorded.

Stage 3 – Tabulation of data

During analysis of the information gathered at the workshops, relevant details were recorded in a tabular format, in preparation for the construction of visual representations. Each table contained the relevant information for all the activities occurring within a specific time frame for different phases of work (e.g. protection arrangements, resources required and a description of work activities and the timings for each phase).

Stage 4 – Construction of visual representation

A visual representation of the activities involved for each phase of work was constructed using Microsoft Visio. Figure 1 is an example of a single phase from one of the scenarios. This involved drawing a visual layout of the track, based on line diagrams which were used in the workshops. Other features such as the length of the track, mileages, structures, signals and points, and other specific identifiable locations were included. Icons were overlaid onto the track layout, positioned appropriately to represent the location of people and machines, based on the information that was recorded during the workshops. A colour key has been included to differentiate between roles. The protected area of railway has also been indicated using labelled icon. Where it was felt that there were alternative methods for providing protection, additional representations were constructed for use in comparing the different methods of working. Additional diagrams have been produced to represent each phase of the work.

Conclusions and work going forwards

The visual representations of rail engineering work have been effective in helping to provide project teams with a better understanding of the proposed system change. They have been valuable during the design stage, in the continued refinement of the proposed working arrangements, as a means for evaluating the different processes and their likely effects. In addition, the visual representations of the scenarios have been a useful means of gaining a richer understanding of some of the complex processes involved during engineering work. These have been used effectively during walkthroughs of the engineering process to highlight potential human factors issues for further analysis.

Work going forwards on the project will include the use of scenarios in the systematic identification of human contributions to system risk. Currently the scenarios offer static visual representations for a given phase of work. Future work will include exploring the value of adding extra dimensions such as animated sequences and 3-D effects if feasible to enhance the functionality and possibly explore activities in greater depth. The work will also investigate potential uses of the approach for further human factors analyses and other project requirements, such as the

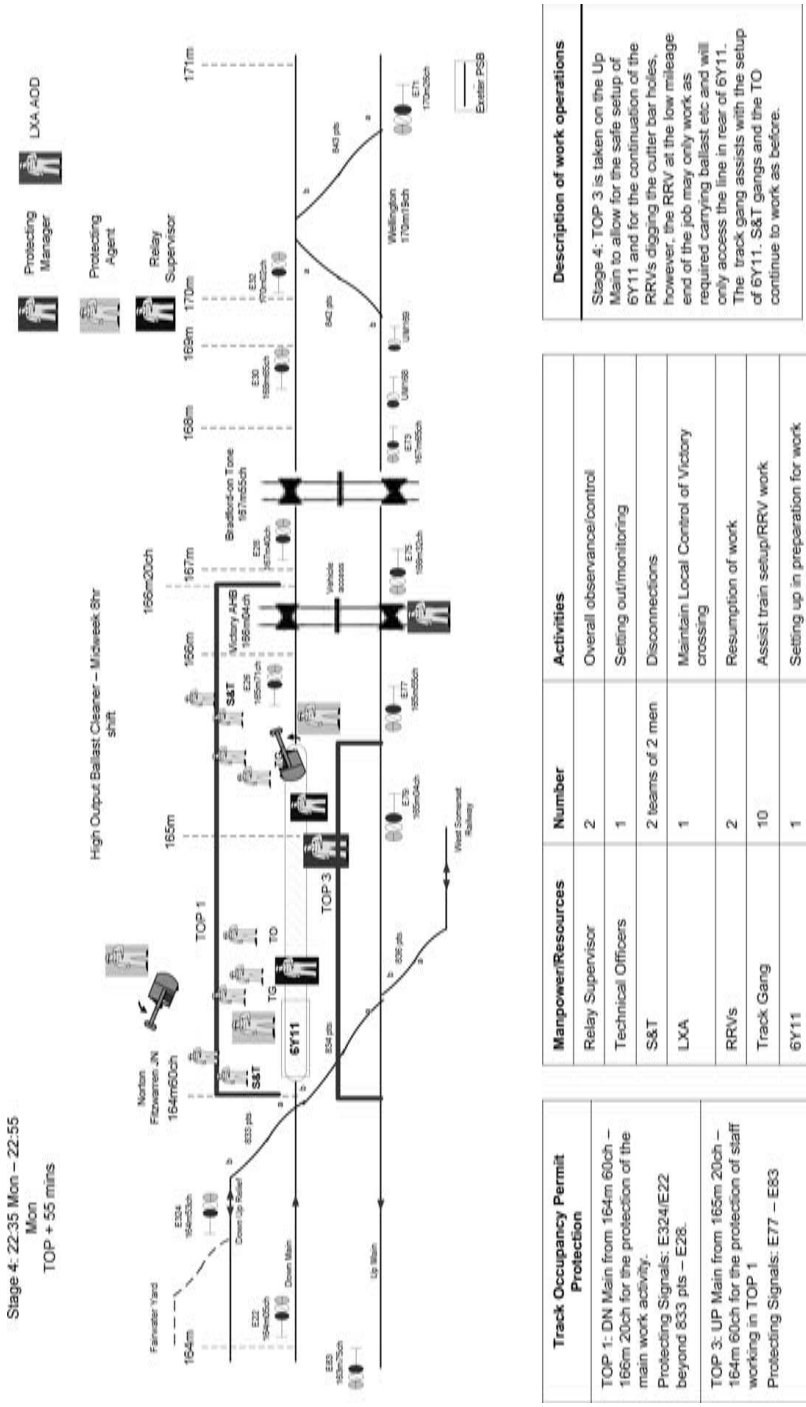


Figure 1. An example phase of work from a visual engineering scenario.

development of training packages where carefully designed scenarios can be used for training of those in new roles on some of the more complex issues.

References

Carroll, J. M. (2000). Five Reasons for Scenario-based design. *Interacting with Computers* **13**: 43–60.

Carroll, J. M. and Rosson, M. B. (1992). Getting Around the Task-Artifact Cycle: How to make claims and design by scenario. *ACM Transactions on Information Systems* **10**(2): 181–212.

Pickup, L., Wilson, J. R., Norris, B. J., Mitchell, L. and Morisroe, G. (2005). The Integrated Workload Scale (IWS): A new self-report tool to assess railway signaller workload. *Applied Ergonomics* **36**: 681–693.

Potts, C. (1995). *Using schematic scenarios to understand user needs*. Symposium on Designing Interactive Systems Ann Arbor, Michigan, United States ACM New York, NY, USA.

Ryan, B., Wilson, J. R., Schock, A. B., Lowe, E. and Kenvyn, F. (2007). Human Factors in the Management of Engineering Possessions: The Roles of the Engineering Supervisor and PICOP. People and Rail Systems: Human Factors at the Heart of the Railway, (Ashgate Publishing Ltd), 331–341.

USE OF OFF-WORKSTATION DISPLAYS IN RAIL CONTROL APPLICATIONS

Michael Carey

Network Rail, Ergonomics National Specialist Team,
London, NW1 2EE

The value and use of off-workstation displays with VDU-based control systems are currently under discussion within UK mainline rail. To inform this debate, an initial review has been carried out of related research literature and case studies carried out of the use of off-workstation displays at three signalling centres. The research literature suggests that there are a number of possible benefits for individual users and for team co-ordination. Questionnaire surveys at all three signalling centres indicated that the signallers have found the displays to be highly useful in sharing information with colleagues and in providing information for visitors/managers without encroaching on their workstation space. At two of the signalling centres, the displays are also used at shift handover and in enabling the signallers to adopt more varied working positions. Implications for future railway control systems are now being considered.

Introduction

As part of the renewal and expansion of the UK mainline railway network, VDU-based control systems are progressively replacing older technologies such as control panels and lever frames. At the same time, advances in display technology are offering the opportunity to employ large size off-workstation overview display. Within the United Kingdom, the mainline railway system has operated VDU-based control for signalling for almost 20 years without the need for such off-workstation displays. Power control systems have also been installed using desk displays alone. Given the cost and space demands of large screen displays, the value and use of widespread adoption is under active discussion within Network Rail.

To inform the debate, a review has been carried out of relevant research findings on the utility of such displays. In addition, experience in the use of such displays has been reviewed across three signalling centres which have each implemented the technology in slightly different ways. (note: electrical control centres within Network Rail have not yet adopted this technology and hence were not available for study). The conclusions of this review are intended to feed into policy on the adoption of such technology for current and future generations of control systems.

Research on large screen displays

Individual use

Various potential benefits have been identified for individual user in adding a large off-workstation display to those displays already provided on a desk.

In terms of accessing displayed information, the additional display space can reduce the need to change displays in order to access information. For example, the off-workstation display could be used for overview information, releasing desk displays for detailed information and input actions. The problems with access to information in dynamic control applications are already well established (Bainbridge 1991). The value of increased display area has also been established empirically for office applications (Czerwinski *et al*., 2003; Ball and North 2005).

The off-workstation display can also provide a constantly available, whole system overview in a fixed physical location. For applications that involve the control of a complex network, the continuous visual presentation of the network may assist in planning and responding to faults (refer to den Buurman, 1991 for a case study of where the removal of a system overview display had major consequences). However, as pointed out by de Groot and Pikaar (2006), such displays should focus on the provision of useful information, minimising the use display space for redundant information on system structure or non-dynamic elements.

Other physiological and practical benefits have also been identified for providing large-screen off-workstation displays (Carey and Barnes 2001). For example, an off-workstation display can offer the user with greater flexibility in posture in operating the system from a seated desk position. Depending on the size and location, the off-workstation display may allow a user to adopt a more reclined, relaxed seated posture when monitoring system status than desk mounted displays that require a more upright posture to enable them to be read.

The size of the display may also allow the information to be read at a distance away from the seated desk position. This can allow the individual user able to move away from the desk in the control room and still retain an overview of the system under control. It also can allow supervisors and visitors to review the status of the system without encroaching on the seated desk position of the users.

Research studies have also hinted at other possible benefits to the individual user that may result simply from providing a larger display. For example, investigations into navigation tasks in 3D virtual environments has highlighted that there can be performance benefits from providing a physically larger display screen (Tan *et al*., 2004). The larger monitor was found to generate a more immersive experience, which in turn aided the artificial navigation task. Whilst this form of task has little relevance to rail control tasks, this does suggests that there may be some subtle psychological benefits from simply providing larger displays for information.

Team use

Where control and communications tasks are carried out by an interacting team within a control room, there is clear potential for a large shared display to assist with

co-ordination and integration of team activities. This is particularly relevant where team members are working on parallel, interacting aspects of a specific problem domain and require a co-ordinated picture of the various on-going activities. The shared display can inform about inter-related activities and become a bulletin board for key developments. Dudfield *et al*. (2001) describe such a potential application in air combat operations.

Stubler and O'Hara (1995) provide a useful review of all the various ways in which 'group-view' displays may be used in supporting teamwork. Shared knowledge of team activities may enhance team performance and cohesion. This type of team interaction is self-evident in many of the previous generation signal boxes that utilise large multi-position panels. There can be a high level of shared team awareness through overheard conversations and observed actions.

A shared display may also assist in non-verbal communication of information between team members. Heath and Luff (1992) give examples of information being shared through glances at a shared track display in a London Underground control room. Such shared information, combined with overheard conversations, may be used to maintain an awareness of the activities being undertaken by adjacent and interacting team members. It is also possible to point to a shared display to highlight information to a colleague or as a support for briefings when taking over a workstation from a colleague.

Explicit evidence that a large shared display delivers measurable benefits is harder to obtain. Emery *et al*. (2001) report on an extensive simulation exercise that tested the usefulness of a large screen display to support a military command team. The results provided no evidence that the shared display had provided any performance benefit, though there was strong subjective support for the display from the participants.

Case studies

Approach

As a first step in exploring what benefits off-workstation displays provide for railway control operations in use, a questionnaire survey was applied in three signalling control centres that have been provided with off-workstation displays (in addition to desk-mounted displays):

- Manchester South Signalling Control Centre has a single, large screen track display. The display is intended to show a combined diagram of an area controlled from a number of individual signalling workstations. To date only one signalling workstation has been commissioned and, as a result, the large screen display currently is an exact replica of the detailed track display shown on the signaller's desk monitors.
- Rugby Signalling Control Centre will eventually control a significant proportion of the West Coast Main Line from six side-by-side signalling workstations. The display wall shows a continuous overview diagram of the whole route. The signallers have both overview diagrams and overview displays available on their workstations.

Figure 1. Manchester South.

Figure 2. Rugby.

Figure 3. Marylebone.

- Marylebone Signalling Centre controls a significant proportion of the Chiltern route between Birmingham and London Marylebone Station. The signallers control their section of the route from detailed track displays on their workstation. Each workstation is provided with an overview display of the entire route on a series of floor-mounted large flat screen monitors. If necessary, the entire route can be controlled from a single workstation. The off-workstation monitors then provide a whole route overview, though the signaller is required to swap desk displays as necessary to carry out control.

Findings

In each signalling centre, the large screen display repeats information provided on the desk displays, which are also the displays used for implementing control

actions. Hence, it would have not been surprising to find the large screens reported to be relatively infrequently used. However, at both Manchester South and Rugby SCCs, the large screens were reported to be frequently used by staff for train service regulation purposes. The screens at Rugby are designed to assist with regulation of the intensively used main line and the survey results suggest that the screens are frequently used for that purpose. At Manchester South, the large screens are a complete replica of the desk displays and yet staff still reported the large screen to be used frequently across a range of signalling tasks. In contrast, the screens at Marylebone Signalling Centre were reported to be used much less frequently. The signallers indicated that they only tended to use the large screens to observe trains approaching from the adjacent workstation. All other signalling information was predominantly sought from the desk displays.

The signallers at all three locations rated the screens as being useful in sharing information with their colleagues and in providing a means for others who enter the control room (including managers) to obtain information without encroaching on their console. In addition, the signallers at Manchester South and Rugby rated the large screen display as being useful in reviewing traffic conditions when taking over a shift and also enabling them to adopt a more comfortable seated position. In comparison, the Marylebone signallers did not see any advantage either during shift handover or in varying seating position.

In summary, at two of the three case study sites, the signallers had found a number of positive benefits in having the large screen displays provided. These were both individual (practical/physiological) and team-related (communications) benefits. It is possible that such attributes of the displays could contribute to improved team coherence in the long-term (especially as additional signalling areas are added at each centre). Some of the team-related benefits were also evident at Marylebone, though not to the same extent. The physical location and size of the displays at Marylebone is not as optimal as at the two other sites, which may have had an affect on the frequency of use of the off-workstation displays.

Conclusions

There are a number of theoretical individual and team-related benefits to be gained from the provision of large off-workstation displays in a control centre. These benefits are likely to be strongest where operators have interacting and overlapping roles related to the system under control and have to co-ordinate and prioritise their actions. In signalling and electrical control centres, the workstations divide the railway into discreet geographical areas of control and there tends to be limited overlap between operating positions. Nonetheless, the feedback from use to date suggest that there still can be some individual and team-related benefits from the provision of large screen displays, though it is not yet clear whether the benefits are substantial (i.e. measurable in terms of control performance). There are possible long-term benefits for team working in terms of facilitating incidental communications and shared responsibility.

The next generation of signalling and electrical control systems are likely to utilise a higher level of automation and involve greater centralisation of control into fewer, larger control centres. It is likely that workstation areas will be capable of being split under fault conditions and combined under low demand conditions. The area controlled by individuals will therefore be more dynamic. The flexible nature of workstation control areas will limit the potential to utilise large shared display screens, though there are possible benefits in having some form of shared display to assist with co-ordinating responses to failures and performance issues. The next planned step is to analyse likely cognitive functions in both future signalling and electrical control centres to inform the selection of workstation designs and information displays.

References

Bainbridge, L. 1991, "Multiplexed VDT Systems: A framework for Good Practice". In, *Human-Computer Interaction and Complex Systems*. pp. 189–210. (London: Academic Press).

Ball, R. and North, C. 2005, "An Analysis of User Behaviour on High-Resolution Tiled Displays". In, *IFIP INTERACT'05: Human-Computer Interaction*. pp. 350–363. (New York: Springer).

Carey, M.S. and Barnes, M. 2001, "Integration of Large Screen Projected Displays into Control Rooms". In, *Safe and Reliable Control Room Operations*, London, December 2001. (London: IBC).

Czerwinski, M., Smith, G., Regan, T., Meyers, B., Robertson, G. and Starkweather, G. 2003, "Toward Characterizing the Productivity Benefits of Very Large Displays" *IFIP INTERACT'03: International Conference on Human-Computer Interaction.* pp. 9–16. (Amsterdam: IOS Press).

De Groot, N. and Pikaar, R.N. 2006. "Videowall Information Design: Useless and useful applications". In, R.N. Pikaar, E.A.P. Koningsveld & P.J.M. Settels (Eds.), *Proceedings IEA 2006; Meeting Diversity in Ergonomics*, 10th–14th July, 2006, Maastricht.

den Buurman, R. 1991, "Ergonomic Design for Monitoring Displays for Control Rooms". In, *Ergonomics and Environment: Proceedings of the 27th Annual Conference of the Ergonomics Society of Australia*. pp. 75–80. (Downer ACT, Australia: Ergonomics Society of Australia).

Dudfield, H.J., Macklin, C., Fearnley, R., Simpson, A. and Hall, P. 2001, "Big is Better? Human Factors Issues of Large Screen Displays with Military Command Teams". In, *People in Control: The Second International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres*, pp. 304–309 (London: IEE).

Emery, L., Catchpole, K., Macklin, C., Dudfield, H. and Myers, E. 2001, "Big is Better? Empirical Results of an Assessment of Command Teams with Large Screen Displays". In, *People in Control: The Second International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres*, pp. 86–91. (London: IEE).

Heath, C. and Luff, P. 1992, "Collaborative and technological design: Task coordination in London Underground control rooms". *Computer Supported Cooperative Work (CSCW)*, Vol. 1, No. 1–2, pp. 69–94.

Stubler, W.F. and O'Hara, J.M. 1995, "Group-View Displays for Enhancing Team Performance". In, *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, pp 1199–1203. (Santa Monica, CA: HF&ES).

Tan, D.S., Gergle, D., Scupeli, P.G. and Pausch, R. 2004, "Physically Large Displays Improve Path Integration in 3D Virtual Navigation Tasks". In, *ACM CHI 2004 Conference on Human Factors in Computing Systems*. pp. 439–446. (New York, NY: ACM).

IMPROVING PRODUCTIVITY IN RAIL MAINTENANCE

Trudi Farrington-Darby, George King & Theresa Clarke

Network Rail, Ergonomics National Specialist Team,
London, NW1 2EE

This paper describes a sixteen week intervention in a rail maintenance delivery unit as part of a programme of Depot Improvement Projects (DIP) in the UK rail industry. Nine deliverables were set as objectives for the project with the main focus on a quantified productivity improvement. The project is an example of the positive impact on the well being of people and an organisation – the twin aims of ergonomics (Wilson, 2005)- when a team of technical specialists work alongside colleagues with expertise in human performance and behaviour. The nine deliverables were met through several phases of a project framework. The first phase involved assessment of the delivery unit through "Day in the life of" (DILO). Data from these observations was analysed to identify opportunities for improvement through any aspect of work and organisational functioning. The project then focussed on working with staff of the delivery unit using participatory methods and offering support in the form of time and expertise to meet the deliverables. Key activities included "Blocks and Barriers" sessions, development and implementation of individual development plans and training matrices, management and provision of equipment including safety equipment in the form of PPE, management team restructuring, job role profiling, the implementation of processes specifically for planning, setting up opportunities for communication between different levels of the organisational hierarchy and involving staff at different levels to become involved in planning the work they would then perform. Within the paper the project outline, deliverables and intervention approach including specific techniques and approaches are described. The value to organisation and individual employees is evaluated. The proposed route for future organisational and individual development that will benefit the company and the people that enable it to function is discussed.

Introduction and background

Railway infrastructure maintenance is a function of Network Rail's business that is geographically dispersed nationally. There are variations in practice that have arisen due to maintenance previously being provided by external companies called infrastructure maintenance companies (IMCs) and from cultural differences across geographical areas which have emerged over time. Whilst individual differences

were recognised as a positive contribution by experienced staff the need for some standardisation of best practice is in the interest of both safe and productive railway maintenance was also appreciated. Any attempt to improve productivity in an industry that makes use of the expertise of its staff and needed to recognise that human behaviour and its change would be at the core of such a programme. This programme relied on the input of people with expertise in human behaviour and performance.

The Depot Improvement Project (DIP) was designed to take a firsthand look at how maintenance was performed within a delivery unit (DU) and then offer support through a focused DIP team in enabling the unit to identify best practice and to implement it through a change process. The first of these projects was run in a high density traffic area, with a high staff turnover within the delivery unit. The second pilot was run in a lower density traffic area but where the workforce were more stable and where experience was rich. In this paper we discuss the third DIP which took place in a location where there was moderate traffic and a mixed workforce and where the lessons from the implementation of the previous project had been consolidated into the modus operandi of the project.

The DIP team

The DIP team comprised of a change leader (project leader), three 'behaviouralists' (a term inherited from the original consultants involved in the pilots and now under review), a data analyst, four technical change agents (specialists in railway maintenance), and one LEAN expert. To facilitate delivery unit involvement and to begin the leadership that would be required for sustainability once the DIP team had left the delivery unit, the maintenance delivery unit manager (MDUM) was also part of the core DIP team. The rest of the delivery unit management were considered to be 'associate' DIP team members to promote a sense of the whole team being responsible for the improvement of productivity within the delivery unit. The DIP facilitated and enabled a change, the desire for which is steadily built into the psyche of the delivery unit. Progressively they buy into thinking about how they do things how they could do things.

The DIP deliverables

The project deliverables, of which there were nine, were the top level aims for the project. The focus was on an increase of 50% in productivity and getting section managers from the delivery unit and their supervisors out on track for 80% of their time.

The DIP programme

The programme ran for sixteen weeks and the DIP team worked from the Delivery unit everyday for that period. The programme was flexible with an emphasis on

responding to what were considered the opportunities for that delivery unit and the sections within it. This did however occur within a framework divided into DILOs to gauge DU status, Action Planning and DILOs for review of actions. Finally there was sustainability planning where the DU was required to draw up a plan for sustaining the improvement and gain a further improvement of an additional 50%.

Step 1: Go look see

"Go look see" was a phrase coined from the LEAN (Page, 2004) approach to work analysis used by Network Rail's Six Sigma specialists. Direct observation and in particular naturalistic observation were tools that had already been used by the Ergonomics National Specialist Team (ENST) in various other projects and research programmes (Farrington-Darby et al, 2007) and yield real data for real problems. The first four weeks involved capturing data for the various sections at the delivery unit through Day In the Life Of (DILO) sessions. The change agent (behavioural or technical) would spend a full shift with a team or gang observing. The DILO was designed to specifically seek out opportunities to reduce or eliminate what LEAN specialists refer to as 'waste'. Some of these categories included travel, inventory, waiting, external influence, over processing etc. The report sheet also included sections about factors that would support staff in working productively and also support their well-being through job satisfaction and clarity of task and role. Examples of which include; clearly understood team objectives, gangs having support from section management, having the resources on site to perform work safely and effectively etc.

Step 2: Quick wins

Each day the DILO findings, concerns and queries were discussed by the DIP change agents using a conference call facility. Each week the findings from the DILO sheets were discussed firstly with the relevant section managers and then shared with the whole core and outer DIP team at a weekly meeting. The strategy was to implement some solutions that would provide "quick wins" and show that the DIP project was listening to the needs and concerns of the workforce and would take action. Also, would provide encouragement that it was possible to change and improve the way work was done. Then as the project progressed more complicated and long term issues were addressed.

Step 3: Listening to concerns

The DILO sessions provided the change agents with chances to informally interview delivery unit staff about what they thought the problems were and where they thought opportunities for improved working lay. Added to these interviews, the change agents ran group Blocks and Barriers sessions, where the same questions were posed at both gang and section management levels. The different groups received feedback from the other groups' sessions and subsequently provided a

basis for achieving a mutual understanding of the problems and opportunities which could then be taken forward.

Step 4: Action planning

After four weeks of observation the DIP team change agents worked with the Section managers and their assistants to develop plans to deal with the issues that had arisen. The Section Manager took ownership of this plan and from this point in the project the takes increasing responsibility for improving productivity.

Feedback, support and growth

The project relies on two main themes to bring about change and continuous improvement; feedback and support. The relationship between visiting DIP and the rest of the delivery unit was what made this project different to so many of the projects and initiatives that had gone before. Whilst initially sceptical and cautious (sometimes even hostile) the sections came to know and trust the visiting DIP team because they worked to "hold a mirror" to the delivery unit but not to embarrass or threaten but to support the delivery unit but in a way that enabled them to achieve more of their own potential. This drive to make the most of people is the third theme; one of personal and professional growth at all levels.

The changes

There were a large number of issues that were brought up and addressed through the project many of them particular to the section. Some examples of issues are shown in table 1 along with the approaches to addressing them.

The repeated DILOs saw a decrease in waste, and increase in time on track and of that time and increase in the productive work done when gangs were on track. Alongside these more quantifiable measures there were observable changes in behaviour and attitude of staff, in particular section managers who were the main role models for other staff levels in the organisation.

Future plans

The project intended to specifically look at identifying waste and opportunities as a means of increasing productivity. What emerged was a more holistic approach to a broader problem. Large inconsistencies existed not only in procedures and work practices with gangs out on the track but also with the job roles and associated behaviours amongst staff of all levels across sections and geographical areas. This product of an emergent job culture combined with the effect of a fragmented maintenance function only recently brought together has refocused the programme. Specific is now being place on job roles and their associated profiles. Here methods like work and task analysis are being used to define not only

Table 1. Examples of issues and actions.

Issue	Action
• Wasted time signals and telecommunication (S&T) technicians waited on site for access to track • Gang and team leaders feeling undervalued And • Poor work daily work planning	• Signallers and S&T technicians scheduled time to join each other on shift to understand challenges and opportunities in working together • Weekly Plan Due Review (PDR) meetings with section management to plan and feedback on work making the best of everyone's local and general maintenance knowledge And • Daily PDR reporting with supervisors
• Gaps in gang and team competence • General demotivation of staff at gang and team level no personal future direction at work • Section Manager overloaded and office bound	• Training courses arranged and training matrices drawn up • Individual personal reviews with broader scope than there had been in the past. Equipping managers with ideas and skills to elicit individual motivation in an innovative way • Redefine job roles and support in allocation of tasks according to team skills And • On site activity coaching
• Staff shortages due to high sickness and absence	• Support in reviewing and implementing sickness and absence policy through liaison with Human Resources, interview skill coaching and record keeping coaching

task steps but the behaviours and attitudes associate with a post. With the right people in the right posts with the right support the work practices generated by processes such as LEAN will be more resilient, as will the maintenance function as a whole.

Conclusion

This project shows how powerful an approach that considers the well-being of staff and the needs of the organisation can be. In a short amount of time some fundamental shifts as well as some superficial quick wins took place within this delivery unit. Productivity improved by 50% and section managers wanted to be the best they could. They were not out on track for 80% of their time but a more important qualitative result was that they understood why they needed to be, why they personally were not out there and what they would do when they got out with their men. The contribution from the change agents specialising in human performance, psychology, participatory, field data collection and analysis methods made a big difference to the traditional approach of either developing people through yet another training course or improving work practice by imposing change.

References

Farrington-Darby, T.J., Wilson, J.R., Norris, B.J., Clarke, T. 2006. A naturalistic study of railway controllers. *Ergonomics*, **49** (12–13), 1370–1394

Page, J. 2004. Implementing LEAN manufacturing techniques: Making your system lean and living with it (Henser and Gardner, Cincinatti)

Wilson, J.R. 2005. Methods in the understanding of human factors. In: J. R. Wilson and N. Corlett (eds), *Evaluation of human work* Third Edition, (Taylor and Francis, London), 1–31

FINGERS ON YOUR BUZZERS PLEASE

Anita Scott

Rail Safety and Standards Board (RSSB), Evergreen House, London, NW1 2DX

The automatic warning system (AWS) has existed on the GB railways since the 1950s. The AWS provides indications to the train driver about the aspect of the next signal. If the next signal is restrictive, a warning horn will sound until the driver acknowledges the horn by pressing a reset button. If the warning horn is not acknowledged within a specified time, an automatic brake application occurs and the train is brought to a stand. A train operating company sought assistance from the RSSB Human Factors Team to investigate the potential reasons for an increase in the number of 'late to cancel' AWS events, which had resulted in service delays. This paper details the findings of the study, in which a number of underlying issues were identified, namely: workload, equipment ergonomics and reliability, and the reporting and recording of events.

Introduction

The automatic warning system

The automatic warning system (AWS) was introduced to GB railways in the 1950s to help train drivers to observe and respond to signals. The automatic warning system provides indications to the train driver about the aspect of the next signal. These warnings are in the form of an audible alarm and a visual indicator in the train cab (the 'sunflower'). If a signal being approached is displaying a clear signal (green), the audible indication will be a 'bell' and the sunflower will remain all black. If the signal being approached is restrictive (red, single yellow or double yellow) a warning horn will sound until the driver acknowledges the horn by pressing the reset button on the drivers' desk. At this point the sunflower changes to a yellow and black display which acts as a reminder to the driver that the last signal passed was at caution. If the warning horn is not acknowledged within 2 seconds (high speed trains) or 2.7 seconds (lower speed trains), an automatic emergency brake application occurs, and the train is brought to a stand (See Table 1).

The automatic warning system is also used to warn the driver of approaching permanent speed restrictions (NB, AWS is not fitted at all permanent speed restrictions) and all temporary and emergency speed restrictions.

The problem

A train operating company approached the Human Factors team at RSSB following an apparent increase in the number of 'late to cancel' AWS events in which the

Table 1. Summary of the AWS in-cab indications.

Signal aspect	Audible alarm	Acknowledgement action	Sunflower display
Green	Bell	N/A	All black
Single yellow, Double yellow, Red	Horn	Press AWS reset button within cancellation time (If unsuccessful, automatic brake application occurs)	Yellow and black

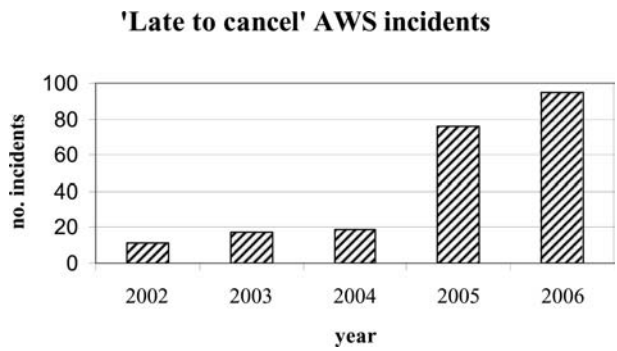


Figure 1. Number of 'late to cancel' AWS incidents over five years.

AWS horn was not acknowledged in time by the driver, and the emergency brake was applied (Figure 1). The company wanted to understand the underlying reasons behind these events because the increased number of emergency brake applications was having an impact on train delays (average of 27 minutes lost per incident). In addition, the relatively vigorous emergency brake application had the potential to increase incidents for passengers and staff, such as falls and scalds from hot drinks. Concerns were also raised that if the occurrence of 'late to cancel' events were to further increase, and drivers were aware of this increase, drivers could become preoccupied in acknowledging the AWS and potentially be distracted from observing the operational environment.

The company wanted to investigate whether it was appropriate to increase the acknowledgment period of the warning horn from 2 seconds to 2.7 seconds (only two settings are available to be used) as the company believed that this should decrease the number of 'late to cancel' events because the driver would have more time to respond to the audible alarm.

Method

A number of methodologies were used to investigate this issue:

- A literature review was undertaken to look into the reaction times of humans in relation to task type.
- Incident data was reviewed.

- Interviews were conducted with train drivers of the company to look into the issue of 'late to cancel' events.
- A train cab ride was undertaken to observe the AWS during operation.
- A meeting was held with an engineer of the train company to understand the mechanical workings of the reset button.
- Investigations were also carried out with four other high speed train operating companies to establish whether they had an awareness of, or had experienced the problem of 'late to cancel'. Interviews were conducted with train drivers (n = 21) and managerial roles with knowledge of the problem across these different companies.
- Liaison with equipment suppliers was undertaken to look into issues of equipment reliability.
- Finally, work was conducted to consider the potential risks of introducing a changed response time for the AWS.

Results and discussion

A number of interesting findings were revealed over the course of the study:

Prevalence of 'late to cancel' events

Interviews conducted with train drivers both within the target company and other high speed train operators suggested that drivers, on average, experienced a 'late to cancel' AWS event a few times over their career. A previous driver error study revealed that 37% of drivers experienced one or more brake applications due to errors in their operation of AWS in a twelve month period (Crick et al., 2004). These figures demonstrated that these 'late to cancel' events are a relatively common experience among drivers but occur rarely at an individual level. Given that drivers successfully acknowledge the AWS horn many times per journey, the task actually has a relatively good rate of execution from a human performance perspective.

Reaction times

One possible reason for the occurrence of the 'late to cancel' events is that the response time is not long enough. A review of literature revealed that human reaction times, which include limited decision making with some physical response range from 0.39 to 0.77 seconds (Boff and Lincoln, 1988). In two studies concerning the AWS, the average reaction times of drivers to the AWS horn was measured to be 0.6 to 0.9 seconds (Crick et al., 2004 and McLeod et al., 2003). In most cases therefore, the acknowledgment period to the AWS appeared to be adequate (2 or 2.7 seconds).

When the Train Protection and Warning System (TPWS) was introduced onto GB trains, the reaction time to the AWS was effectively reduced from several seconds to 2 or 2.7 seconds. This is because the original AWS was a pneumatic

system which gave drivers more time to respond to the warning horn compared to the AWS when it was integrated within the electronic TPWS. In addition to this change in response time, the actual action of cancellation changed slightly, such that drivers were required to press *and* release the AWS reset button to cancel the audible alarm. Previously, only a press action was required to acknowledge the alarm. It was therefore possible that some of the reported events could reflect the shorter time period available to cancel the AWS horn and the slight change to the acknowledgement action.

Workload

Another potential reason for 'late to cancel' events was that drivers had other tasks to do which competed with the task of acknowledging the AWS. Driver interviews revealed that 'late to cancel' events commonly occurred when the driver was distracted from the cancellation task, and doing something else. These other tasks included:

- responding to the vigilance device (requires the driver to provide a positive input to the train system on a regular basis, otherwise the brakes are applied)
- responding to other train systems with alarms
- sounding the warning horn
- using the cab radio
- speaking with the train guard
- taking a drink
- using the screen wipers
- pulling down the window blind.

Distractions external to the train cab were also commented upon such as warning track workers by the line and objects striking the train.

The company at the centre of the study reported a significant number of brake applications in high workload conditions. Such events included: drivers responding to a different safety system with a high urgency tone, or where drivers were taken by surprise at speed restrictions where the AWS horn also sounds.

Equipment ergonomics

The ergonomics of the AWS equipment was another source of interest. Interviews with drivers revealed that 'late to cancel' events sometimes occurred due to variation in the positioning of the AWS reset button between different train types. On occasion a driver might move his/her hand towards the AWS reset button, only to remember that the button is in a different position to what they were expecting. This could result in the driver being late to acknowledge the warning horn and the brakes applying.

Some drivers reported the surface of the reset buttons to be slippery such that the fingers could slip off, resulting in failure to acknowledge the AWS. For one button design, drivers reported that if it is not struck 'square on' the acknowledgment could fail.

Equipment reliability

A number of 'late to cancel' events occurred when the driver had attempted to acknowledge using the reset button, but the first attempt had failed (e.g., due to not hitting the button 'square on', fingers slipping off, or the button simply not working) and the second attempt to cancel being too late. Although the engineering design was reported to be very simple and highly reliable, the occurrence of first attempt acknowledgement failures appeared to be relatively common.

Reporting of events

The driver interviews revealed that there were inconsistencies between drivers as to whether they believed it was necessary to report 'late to cancel' AWS events to the signaller. Following a review of the GB Master Rule Book, it was found that there is a requirement to report failed cancellations of the AWS, however the rules could be subject to interpretation. The reason for drivers not reporting 'late to cancel' AWS events seemed to be that they believed it was an issue of performance rather than of safety (the train has been brought to a stand, therefore the hazard has been controlled). Where drivers said that they would report 'late to cancel' events to the signaller, a common view was that this was out of courtesy to inform the signaller of any delays which might be incurred.

Recording of events

When other train operating companies were contacted to investigate this topic of 'late to cancel', there was very little awareness of the issue or of it being problematic. It was established that for a company to become aware of a 'late to cancel' event, a performance delay would need to be incurred and the reason for this delay attributed. As such, when one train operator looked into their delay attribution, they found very few incidents associated with 'late to cancel' AWS. One reason for the low occurrence of 'late to cancel' events was that the drivers might be able to 'make up' the lost time and therefore be less likely to report the event to the signaller.

Conclusion

This study was particularly interesting because the answer to the question of 'Is the cancellation time enough?', rapidly became more complex than initially anticipated. Issues were identified at the individual level (distraction), task level (workload, equipment ergonomics and equipment reliability) and organisational level (reporting and recording of incidents). This 'systems approach' therefore enabled a number of key issues to be identified to the industry, of which a number were taken forward. For example, the issues identified with the equipment design were referred to the equipment supplier and the inconsistencies surrounding the reporting and recording of events were referred to an industry group for further consideration. The company at the centre of the study decided not to pursue a change to the cancellation time

specified in the Rail Group Standard. It was deemed that little safety benefit would be derived from progressing the issue further.

References

Boff, K. R. & Lincoln (1988) Engineering data compendium. New York: John Wiley and Sons.

Crick, J., Gibson, H., Lear, A. & Page, H. Driver error data collection project: error analysis. Rail Safety and Standards Board research project T148.

McLeod, R. W., Walker, G. H., Moray, N. & love, G. 2003. Extended use of the Automatic Warning System. Rail Safety and Standards Board research project T021.

INTEGRATION IS MORE THAN A HUMAN FACTORS ISSUE

Fiona Bird

Rail Safety and Standards Board, London, NW1 2DX

Human Factors professionals often discuss the integration of human factors within projects and the development of new systems. This paper looks at the wider issue of multi-disciplinary integration such that it ensures the inclusion of human factors in system development products. It identifies methods by which multi-disciplinary integration can be achieved on systems development projects and the actions that Human Factors professionals can take to ensure integration from an early stage. These methods will be backed up by a real world case study from a safety systems development project, where multi-disciplinary outputs have been achieved through requirements processes. The paper will finish with an approach that human factors specialists can promote within multi-discipline projects.

Introduction

Human Factors professionals have been working relentlessly over the last 60 years to successfully integrate their discipline with that of engineers and designers. Over the last 10 years, the profile of human factors in the rail industry in Great Britain (GB) has increased significantly, in particular in system development projects.

The objective of human factors integration (HFI) is to ensure that human factors methods and principles are applied appropriately and consistently during systems development in order to achieve a safe and effective design for end users (Elder et al, 2001). This leads us to question "**how** [do we] make practical, efficient, contributions to a large project, across disciplines, in an integrated fashion"? (Fuld and Harmon, 1992)

The aim of this paper is to offer further experience of human factors integration to the human factors and ergonomics community and contribute to the continual developments of methods to ensure sound integration of human factors knowledge with that of other disciplines.

Barriers to integration

The variety of disciplines that can be found on any project are generally working toward the same goal, i.e. develop a system that meets the needs of the industry, is

reliable and meets its performance requirements. Despite this common objective, integration is not always successful. Here are a few of the barriers:

- Engineers, designers and operational specialists do not necessarily understand the data that will be produced (or the data to expect) and what they should be doing with it.
- Human factors outputs may not necessarily be easy to understand by other disciplines, in terms of their meaning and how they can be applied in a practical sense.
- Many disciplines adopt different approaches to provide solutions to the same problem and this leads to different outputs, e.g. human factors may involve some empirical research.
- A review by MW Kellogg identified that existing human factors procedures, relating to engineering work, did not tie-in effectively with the Company's engineering process (Elder et al, 2001).
- The definition of system scope can vary, e.g. the human factors definition of a system would include the human elements, whereas the systems engineers may simply focus on the hardware and software and not the operational elements that are needed to support them.

The role of Integration Plans

The Human Factors Integration Plan (HFIP) is often promoted as the best method for enabling integration. However on multi-disciplinary projects there is no requirement placed on other disciplines to produce such a plan, and the HFIP is written by the human factors team from a human factors perspective. There is no guarantee that the other disciplines will buy into the HFIP; Garner (2003) challenges us "to recognise that a formal HFIP does not mean that **our work** will be fully integrated". This suggests that whilst the human factors team may be integrated, their outputs may still be treated as a separate entity.

Davis (2001) describes how Railtrack's Thameslink 2000 programme used "several strategy documents to define the processes by which the programme objectives will be met", which included a Human Factors Integration Strategy. The general approach adopted by the Human Factors Operability Group was to conduct benchmarking exercises and analyses of existing operating processes; then through consultation with end users operational strategies and functional requirements were produced to inform the design scheme. From the author's experience of the project, the missing element in the Thameslink 2000 programme was a single repository for collating the requirements produced by the various disciplines. Requirements resided in separate documents and relied on programme-wide awareness of all reports if the requirements were to be implemented.

Integration Plans have a place and it is important to recognise that integration needs to occur between all disciplines and that this would be most effective by a Project Integration Plan, that is agreed by all disciplines at the beginning of a project.

Twenty four carat human factors

We have found that, for various reasons, some industry changes have to be made – with or without human factors input. Garner (2003) goes on to say that "we need to recognise that it is not always possible to complete things by the book…sometimes partial compromise is the only way forward". So sometimes, by necessity, we have to find alternative solutions to human factors problems, e.g. training, rather than risk creating unrealistic requirements that have negative implications in terms of performance and cost, which will be ignored. We have to strike a balance between human factors good practice, safety, cost and performance.

Requirements Management

What is Requirements Management?

Requirements Management can be defined as "the science and art of gathering and managing user, business, technical, functional and process requirements within a product development project" (Wikipedia, 2007). Requirements Engineers and Requirements Management Processes are becoming a more regular fixture in systems development projects and this has proved to be an effective method for enabling integration of human factors.

This process enables the different disciplines to work separately using their own technical methods, e.g. applying human factors tools and techniques. The conclusions of any research are produced as requirements and the research report is available as part of the audit trail identifying how that particular requirement was identified.

Producing requirements

The advantage of a requirements management approach is that the human factors requirements will be placed in a requirements database, rather than rely on an engineer or designer to extract the relevant human factors information from a technical report which can lead to misinterpretation. Nevertheless it is still important to remember that the users of human factors requirements are not those who receive the final product; our users are usually those who have to be able to apply our findings and requirements to the design of a product.

For example, the Rail Safety and Standards Board (RSSB) and their supplier are currently developing an information system for train drivers; the RSSB human factors team is providing support and human factors is an integrated part of the development. A requirements management process is in place to support the design and development; the user of the requirements will be the systems manufacturer so the human factors requirements will have to be clear enough that they can be practically and accurately applied to a physical product, both during production and the acceptance process. This process was also applied on a RSSB signalling system project and feedback from the engineers at then end of the initial development stage was that human factors integration had worked well.

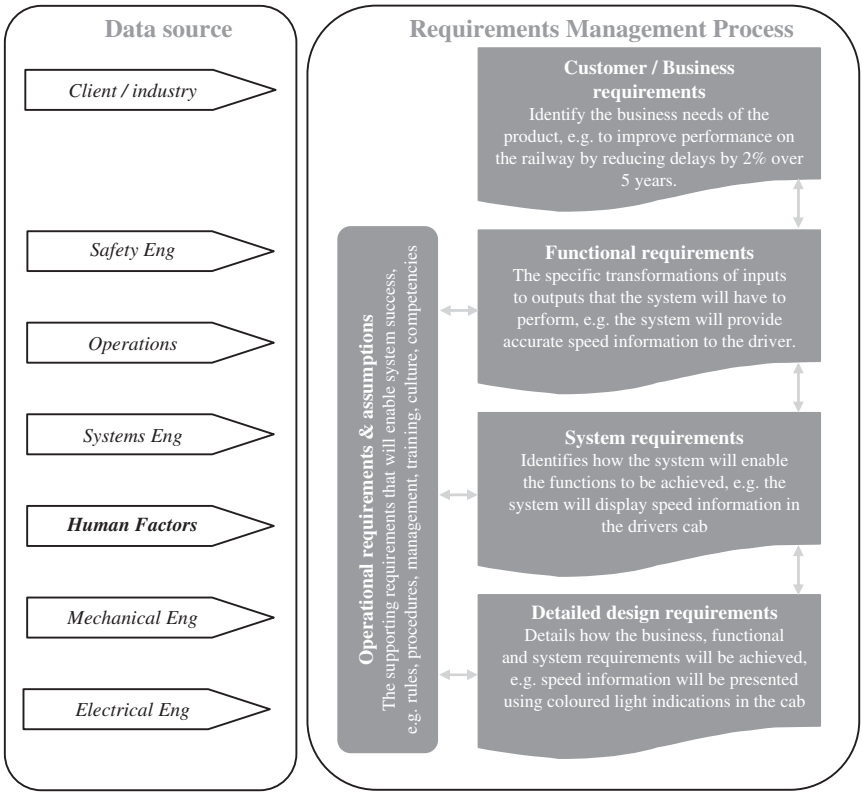


Figure 1. Requirements Management Process.

Checking for conflict across the disciplines

The requirements management process will usually include a checking process to identify conflicts between requirements produced by different disciplines. Therefore where a human factors requirement may conflict with that of another discipline, a technical review would be held to enable the conflict to be discussed and resolved in a manner that is acceptable to both disciplines. This enables the majority of conflict to be resolved before anything is physically produced, thus reducing cost in the latter stages of the project, when rectifying problems could be costly.

Final output

The final output is a set of requirements that has been produced and agreed by all disciplines, the programme and the stakeholder group, that can be used by the supplier to produce the system.

Figure 1, above, describes the Requirements Management Process

Integration throughout the programme lifecycle

Integration needs to occur across the programme throughout the lifecycle and the following proposes a method for achieving it.

1. ***Project start-up*:** at the start of the project, identify all the disciplines who will be involved e.g. systems engineer, operations, human factors, safety, then co-ordinate the team (even if practical involvement is at a later stage). An integration plan should be produced identifying the activities for all disciplines throughout the programme and an agreed process for integrating the outputs of their activities. The scope of the project should also be defined, identifying the scope of the system and where the human elements, including procedures, rules, training, fit in.
2. ***Requirements management process*:** development of a requirements management process that is clearly defined and enables the requirements produced by all disciplines to be recorded. Human factors specialists can conduct their research as necessary, with their final outputs being requirements that form part of the design and reports that contain evidence/support for the requirements.
3. ***Usability testing*:** during the procurement, development and pre-commissioning stages, the requirements are used to assess the system, i.e. does it meet the requirements? Multi-disciplinary testing, including usability testing, can also occur at this stage and the requirements amended as necessary to reflect the outputs of the testing.
4. ***System monitoring*:** during and after commissioning, a process should be in place to capture feedback from end users on usability, training, conflicts etc. These can then form requirements for future upgrades of the system.

Conclusion

The use of a Project Integration Plan and a single Requirements Management Process, will enable human factors to be integrated into multi-disciplinary projects. The advantages are:

- One Integration Plan for the project, avoiding conflict between the procedures, strategies or plans that are produced by separate disciplines.
- One set of multi-disciplinary requirements, including human factors, that are clear, practical, and usable by those who will be responsible for their application.
- All disciplines are able to use their own methods of working and their own outputs; these are then translated into requirements that form the design.

References

Davis, 2001, Human Factors Integration for Thameslink 2000, *People in Control: An International Conference on Human Interfaces in Control Rooms*, Cockpits and Command Centres 19–21 June 2001

Elder, Hamilton & Martin, 2001, Human Factors Integration in the Petro-Chemical Industry, *Contemporary Ergonomics 2001*, Edited by M.A. Hanson. Taylor & Francis, London, 2001, pp. 561–566

Fuld, & Harmon, 1992, Human Factors Integration for large scale facility designs, *Fifth Conference on Human Factors and Power Plants, Monterey, California, June 7–11 1992*, pp. 16–19

Garner, 2003, Human Factors Integration – A Company Case Study, *Proceedings of the First European Conference on Rail Human Factors, York, 13–15 October 2003*, pp. 508–514

SAFETY AT SEA

TOO TIRED TO START? TRAVEL TO THE VESSEL AND SUBSEQUENT FATIGUE AT WORK AMONG SEAFARERS

Emma J.K. Wadsworth, Paul H. Allen & Andy P. Smith

Centre for Occupational and Health Psychology, Cardiff University, Cardiff, CF10 3AS

Seafarers working in the UK offshore oil support, short-sea and deep-sea shipping industries completed questionnaires and daily diaries about their work and fatigue levels. Two-thirds (66%) reported having no opportunity to sleep between travelling to the vessel and starting their first shift. Almost half of this group (47%) had travelled for six or more hours, and 19% for 12 or more hours. Having no opportunity to sleep was strongly related to fatigue, and being unprepared for work at the start of a tour was strongly related to fatigue prior to work the following day. A significant proportion of seafarers, therefore, begin their tour of duty and crucial first shift very or extremely tired, and this has knock-on effects into the next day. Lack of opportunity to sleep has the potential to impact on personal and vessel safety both immediately and in the longer-term as fatigue builds during the tour.

Introduction

Work-related fatigue and tiredness at work are associated with poorer performance (Beurskens *et al*, 2000; Charlton and Baas, 2001) and may increase the risk of work-place accidents and injuries (Bonnet and Arand, 1995; Hamelin, 1987). Seafaring is a global, safety critical industry in which the potential consequences of fatigue at work are huge. Indeed, fatigue has been identified as a cause or contributory factor in accidents (Holland-Martin, 1696; MAIB, 2002; Majumdar *et al*, 2004; Raby and McCallum, 1997; Schilling, 1971) and impaired collision risk awareness (Wellens *et al*, 2005) at sea over many years, and it is clear that the relationships between fatigue and human error, and consequent safety and health, are crucial in seafaring (Jensen *et al*, 2004).

Seafarers' preparedness for their first work shift at the start of a tour of duty is particularly important for two reasons. First, efficient and effective hand-over is crucial to both safety and performance; and second fatigue during a tour of duty increases, primarily during the first week (Wadsworth *et al*, 2006), so beginning a tour of duty with a greater level of fatigue will impact on subsequent fatigue during the rest of the tour.

Method

As part of a larger programme of research focusing on seafarers' fatigue (Smith *et al*, 2006), 549 seafarers from the UK offshore oil support (N=96), short sea and coastal (N=133) and deep sea (N=320) sectors of the industry completed questionnaires. In addition, 203 seafarers (77 offshore oil support, 94 short sea and coastal and 32 deep sea) completed daily diaries. Both the questionnaire and daily diary included questions about seafarers' travel to the vessel to begin a tour of duty, and are described in more detail elsewhere (Smith *et al*, 2006). In the questionnaire, respondents were asked about how long they had had spent travelling to the vessel, whether they had the opportunity to sleep before starting their first shift, and how they would class their state when starting their first shift (coded on a 5-point likert scale from "about normal/not fatigued" to "extremely tired/exhausted"). In the diaries, respondents were asked about how prepared they felt for work when joining the vessel in terms of tiredness (coded on a 5-point likert scale from "well prepared" to "very unprepared". In addition, respondents rated how tired they were on a visual analogue scale by marking with a cross the place on a 10 cm line (with "not at all tired" and "extremely tired" at each end) which best corresponded to how they felt on waking prior to work each day for a tour of duty.

Travel to the vessel

Among the questionnaire respondents, mean time spent travelling to the vessel since their last main sleep period was 8.63 hours (sd=8.02, median = 6 hours, min. = 0, max. = 52). Two-thirds (66%, N = 356) had no opportunity to sleep before their first work shift. Among this group, mean travel time was 7.22 hours (sd = 7.39, median = 5 hours, min. = 0, max. = 47), and almost half (47%, N = 164) had travelled for 6 hours or more, one third (33%, 114) for 8 hours or more, and 19% (66) for 12 hours or more. This suggests that seafarers are travelling for a significant amount of time to reach their vessel, and that most have no opportunity to sleep between arriving at the vessel and starting their first shift.

First shift

Ten percent (N = 56) of all the questionnaire respondents described their state when starting their first shift as very tired, extremely tired or exhausted. Those who had no opportunity to sleep before starting their first shift were significantly more likely to report starting their first shift very or extremely tired or exhausted ($p = 0.008$) (Figure 1).

Those who reported being very or extremely tired or exhausted on starting their first shift reported travelling for longer (17.04 hours (se = 1.67) compared to 7.64 (0.31), $F(1, 524) = 77.45$, $p < 0.0001$). Repeating this analysis among only those with no opportunity to sleep before starting their first shift showed the same significant difference (15.52 hours (se = 1.73) compared to 5.95 (0.32), $F(1, 343) = 81.01$, $p < 0.0001$). This suggests a strong association between seafarers' state starting their

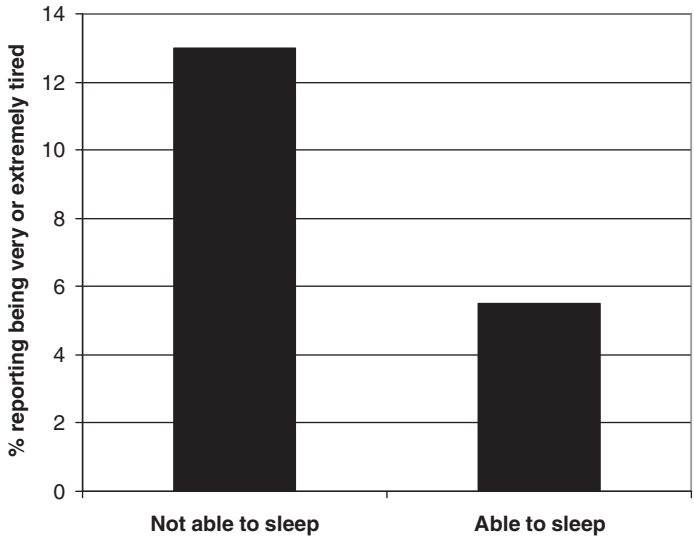


Figure 1. Proportion of seafarers starting work very or extremely tired or exhausted by opportunity to sleep before first shift.

first shift and both having the opportunity to sleep and time spent travelling to the vessel.

Subsequent fatigue

Among the 203 seafarers who completed daily diaries, 10% (N = 20) described themselves as quite or very unprepared for their first day's work because of tiredness and 18% (35) as moderately unprepared. Those who described themselves as quite or very unprepared for work on their first day were significantly more tired prior to work the following day ($p < 0.0001$) (Figure 2).

This suggests that preparedness for work the first day onboard the vessel is strongly associated with subsequent fatigue.

Discussion

These results suggest that many seafarers travel for significant amounts of time to join their vessel for a tour of duty. The majority have no opportunity to sleep following their journey before starting their first shift. Both travel time and the lack of opportunity for sleep were associated with tiredness during this first shift. In addition, not being prepared for work on the first day of a tour because of tiredness was strongly associated with level of fatigue prior to starting work on the second day of tour. This is important because this study has shown elsewhere (Wadsworth *et al*, 2006) that fatigue on waking is a more sensitive measure of cumulative occupational fatigue than fatigue on retiring, and that fatigue on waking increases

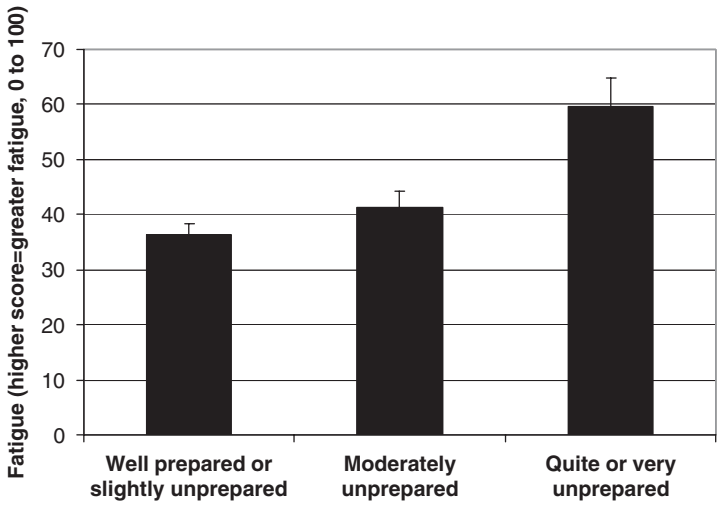


Figure 2. Level of fatigue prior to starting work on day 2 of a tour of duty by preparedness for work on day 1.

between the start and end of a tour, primarily during the first week at sea. Those who start with an increased level of fatigue because of their journey to the vessel and subsequent lack of opportunity to sleep, therefore, will continue with progressively higher fatigue levels as the tour progresses.

Seafaring is a global, safety critical industry and seafarers work to increasingly tight schedules and turn-around margins. This makes the hand-over to crew members starting their tour of duty both pressured and crucial for safety and performance efficiency. Research has shown that working at sea has the potential to be fatiguing (Brown, 1989). This is clearly exacerbated when seafarers are not given to opportunity to recover from sometimes very lengthy travel to the vessel. Being too tired to start work has the potential to impact on personal and vessel safety not only immediately but also in the longer-term as fatigue builds during the subsequent tour of duty.

Acknowledgements

This research is supported by the Maritime and Coastguard Agency, the Health and Safety Executive and NUMAST. We would like to acknowledge the contribution made by those who participated in the research. The views expressed here reflect those of the authors and should not be taken to represent the policy of the sponsors.

References

Beurskens, A.J.H.M., Bultmann, U., Kant, I.J., Vercoulen, J.H.M.M., Bleijenberg, G. and Swaen, G.M.H. 2000, Fatigue among working

people: validity of a questionnaire measure. *Occupational and Environmental Medicine,* **57**, 353–357.

Bonnet, M.H. and Arand, D.L. 1995, We are chronically sleep deprived. *Sleep,* **18**, 908–911.

Brown, I.D. 1989, *Study into hours of work, fatigue and safety at sea*. (Medical Research Council, Cambridge).

Charlton, S.G. and Baas, P.H. 2001, Fatigue, work-rest cycles, and psychomotor performance of New Zealand truck drivers. *New Zealand Journal of Psychology,* **30**, 32–39.

Hamelin, P. 1987, Lorry drivers' time habits in work and their involvement in traffic accidents. *Ergonomics,* **30**, 1323–1333.

Holland-Martin, D. 1696. *Trawler safety: final report of the committee of inquiry into trawler safety.* London, UK: HMSO.

Jensen, O.C., Sorensen, J.F.L., Kaerlev, L., Canals, M.L., Nikolic, N. and Saarni, H. 2004, Self-reported injuries among seafarers. Questionnaire validity and results from an international study. *Accident Analysis and Prevention,* **36**, 405–413.

MAIB, 2002, *Report on the analysis of fishing vessel accident data, 1992 to 2000.* Marine Accident Investigation Branch.

Majumdar, B., Browne, G., Roberts, J. and Carpio, B. 2004, Effects of cultural sensitivity training on health provider attitudes and patient outcomes. *J Nurs Scholarsh,* **36**, 161–166.

Raby, M. and McCallum, M.C. 1997, *Procedures for investigating and reporting fatigue contributions to marine casualties.* Paper presented at the Human Factors and Ergonomics Society 41st Annual Meeting.

Schilling, R.S.F. 1971, Hazards of deep-sea fishing. *British Journal of Industrial Medicine,* **28**, 27–35.

Smith, A., Allen, P. and Wadsworth, E. 2006, *Seafarers' fatigue: the Cardiff research programme*. Southampton: MCA.

Wadsworth, E., Allen, P., Wellens, B., McNamara, R. and Smith, A. 2006, Patterns of fatigue among seafarers during a tour of duty. *American Journal of Industrial Medicine,* **49**, 836–844.

Wellens, B.T., McNamara, R.L., Allen, P.H. and Smith, A.P. 2005, Collisions and collision risk awareness at sea: data from a variety of seafarers. In P. D. Bust & P. T. McCabe (Eds.), *Contemporary Ergonomics,* 573–577.

A DIARY STUDY OF FATIGUE OFFSHORE

Andy Smith

Centre for Occupational and Health Psychology,
Cardiff University, Cardiff, CF10 3AS

The present paper describes a diary study comparing offshore workers (seafarers and installation workers) and onshore workers. Results from day workers showed that those offshore slept for a shorter time, woke more often during the night and had greater difficulty falling asleep. Comparisons between installation workers and seafarers showed that the installation workers felt less alert at the start of their working day. Night workers reported lower levels of alertness at the end of their working day. Sleep problems were reduced for day workers over the course of their tour of duty. Indeed, the first night offshore was associated with reduced sleep, especially for installation workers doing night shifts. Participants also completed diaries when they were on leave and the results showed that sleep and alertness were still abnormal at the start of the leave period. Overall, the present results confirm findings from surveys and onboard testing. In conclusion, the present approach has proved very useful in helping to increase our knowledge of offshore fatigue.

Introduction

Recent research has examined offshore fatigue and studied both seafarers (see Smith, Allen and Wadsworth, 2006; Smith, 2007; Allen, Wadsworth and Smith, in press, a, b) and installation workers (see Smith, 2006). A variety of different methodologies have been used to investigate the topic. Results from several surveys demonstrate that fatigue is frequently reported offshore (Wadsworth et al., in press; McNamara and Smith, 2002). These findings have been confirmed by measuring alertness and performance onboard ship (Smith, Allen and Wadsworth, 2006) and on installations (Smith, 2006). Another approach has involved the use of daily diaries. Burke et al. (2005) and Wadsworth et al. (2006) have used this methodology and it has shown some interesting features of seafarers' fatigue.

In the larger of the two studies (Wadsworth et al., 2006) participants from the offshore support, short sea, coastal, and deep sea sectors of the UK shipping industry completed daily diaries. Information about sleep and fatigue was collected each day on waking and retiring. The results showed that fatigue on waking was a more sensitive measure of cumulative occupational fatigue than fatigue on retiring. Fatigue on waking increased between the start and end of tour. This increase took

place in particular during the first week at sea, and was most apparent amongst those on shorter tours of duty. Such findings suggest that day to day changes in fatigue are measurable, and may vary among particular sub-groups of seafarers. Used among different groups, such a diary method could help identify times and activities associated with the greatest risk. The aim of the present study was to compare two types of offshore worker (seafarers and installation workers) and also include an onshore group. Comparisons were made between day shift workers and night shift workers. The use of the diary methodology meant that it was possible to examine how fatigue indicators changed over the tour of duty. Indeed, the study also provided an opportunity to investigate anecdotal reports that fatigue induced by work offshore often continues into the first week of leave.

Method

Diaries

The diaries were completed for 7 days and assessed quality and duration of sleep periods and alertness (measured using a 100 mm visual analogue scale from "Very sleepy" to "Very Alert" at the start and end of the working day. An example is shown below:

Day 1 **Date: _____**

To be completed just before starting work:

- Time you went to bed:
- Time you went to sleep:
- Time you woke up:
- Time you got up:
- Sleep duration:
- Number of awakenings during a sleep period

Rate your sleep:

	Least (1)				Most (5)
Ease of falling asleep	1	2	3	4	5
Ease of arising	1	2	3	4	5
Was this sleep period sufficient?	1	2	3	4	5
How deep was your sleep?	1	2	3	4	5
Did you wake earlier than intended?	1	2	3	4	5

How alert do you feel now?
Put a mark on the line to indicate how you feel right now.

Very Sleepy Very Alert

_____

0 100

Study 1: A comparison of onshore and offshore workers

Logbooks were completed by 58 onshore day workers and 42 offshore workers working day shifts. The results showed that two groups differed significantly on a number of sleep variables. Offshore workers slept for a shorter time (mean sleep duration: offshore = 6.38 hours, onshore = 6.97 hours), woke up more often during the night (offshore mean number of awakenings = 1.8; onshore mean = 1.2), had greater difficulty falling asleep, and were less likely to consider that they had a deep sleep or enough sleep. Although these differences were statistically significant, the magnitude of the effects were small.

Study 2: A comparison of different offshore groups

These analyses compared 31 installation workers and 29 seafarers working on support ships in the offshore oil industry. Forty two were day workers and 18 night workers. Twenty five were in the first week of their tour of duty and 35 were in either their second or third week offshore.The results showed that installation workers felt less alert at the start of the day (installation workers mean alertness = 45.5, s.e. = 2.2; seafarers mean alertness = 54.1, s.e. = 2.0). Those working nights reported lower alertness at the end of their shift (day workers mean alertness = 47.8, s.e. = 2.2; night workers mean alertness = 32.9, s.e. = 2.6) even though they perceived their job to be less physically demanding. Day workers starting their tour of duty awoke more frequently (mean number of awakenings = 2.4) than those in their second or third week of the tour (mean number of awakenings = 1.4). The reverse was true for night workers (mean number of awakenings at start of tour = 1.4; mean number of awakenings in 2nd/3rd week = 2.0). Sleep duration was reduced for the first night offshore, especially for installation workers doing night shifts. The alertness levels at the end of the first day were lower for the seafarers than installation workers, with the reverse pattern being present on days 6 and 7.

Study 3: Diaries completed during leave

Forty-three volunteers completed the weekly log while they were on leave. Twenty-two were installation workers and 21 seafarers. Thirty-four had worked day shifts before leave and 9 had worked nights. Of these 43 participants 22 had just started their leave and 21 were on their second week of leave. The results showed clear evidence that alertness was reduced at the start of leave and took about 4 days to return to normal levels at the start of the day and a week at the end of the day (see Tables 1 and 2).

Seafarers starting leave generally slept for at least an hour a day longer than in the second week of leave. This was not found for the installation workers who generally slept for a shorter time at the start of their leave period.

Conclusion

Overall, the diary data show that offshore workers report sleep problems and reduced alertness (greater fatigue). These data also confirm the view that shift

Table 1. Means and standard errors for start of day alertness (high scores = greater alertness) for the first and second week of leave.

Start alertness		Day 1	Day 2	Day 3	Day 4
Starting Leave (N = 22)	Mean	46.61	49.49	53.60	57.69
	S.E	4.79	4.99	4.69	4.49
2nd Week Leave (N = 21)	Mean	53.31	58.06	56.94	56.33
	S.E	4.83	5.09	5.28	4.60

Table 2. Means and standard errors for end of day alertness (high scores = greater alertness) for the first and second week of leave.

End Alertness		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Starting Leave (N = 22)	Mean	38.41	45.25	55.99	59.96	62.79	62.56	67.02
	S.E	5.80	4.76	4.90	5.00	5.26	5.46	5.24
2nd Week Leave (N = 21)	Mean	63.84	64.78	68.78	76.42	74.63	68.84	69.00
	S.E	5.67	6.36	5.61	4.83	4.95	6.13	5.51

patterns may have an important impact on fatigue and that effects may change over the course of the tour of duty. Our surveys have shown that offshore workers feel that their jobs impact on their quality of life when on leave (see Smith, Allen and Wadsworth, 2006) and this was also confirmed by the diary data.

Acknowledgements

This research was supported by the Maritime and Coastguard Agency, the Health and Safety Executive and NUMAST. We would like to acknowledge the contribution made by those who participated in the research. The views expressed here reflect those of the author and should not be taken to represent the policy of the sponsors.

References

Allen, P.H., Wadsworh, E.J. and Smith, A.P. In press, a, Seafarers' fatigue: A review of the recent literature. *International Maritime Health.*

Allen, P.H., Wadsworh, E.J. and Smith, A.P. In press, b, The prevention and management of seafarers' fatigue: A review. *International Maritime Health.*

Burke, A., Ellis, N. and Allen, P. 2003, The impact of work patterns on stress and fatigue among offshore worker populations. In P.McCabe (Ed.), *Contemporary Ergonomics 2003* (Taylor and Francis, London), 131–136.

McNamara, R. and Smith, A.P. 2002, The combined effects of fatigue indicators on health and well-being in the offshore oil industry. *Archives of Complex Environmental Studies*, 14 (3–4).

Smith, A.P. 2006, Shiftwork on oil installations. In P. D. Bust (Ed.), *Contemporary Ergonomics 2006,* (Taylor and Francis, London), 565–569.

Smith, A.P. 2007, *Adequate crewing and seafarers' fatigue: The International Perspective.* (International Transport Federation, London)

Smith, A., Allen, P. and Wadsworth, E. 2006, *Seafarers' fatigue: The Cardiff Research Programme*. (MCA, Southampton)

Wadsworth, E., Allen, P., McNamara, R. and Smith, A. In press, Fatigue and Health in a seafaring population. *Occupational Medicine*.

Wadsworth, E., Allen, P., Wellens, B., McNamara, R. and Smith, A. 2006, Patterns of fatigue among seafarers during a tour of duty. *American Journal of Industrial Medicine*, **49**, 836–844.

SEAFARERS' FATIGUE: CONCLUSIONS AND THE WAY FORWARD

Andy P. Smith, Paul H. Allen & Emma J.K. Wadsworth

Centre for Occupational and Health Psychology,
Cardiff University, Cardiff, CF10 3AS

The present paper will present some of the main conclusions from our recent programme on seafarers' fatigue. Different sectors of the industry experience fatigue for different reasons. The crew of fishing vessels often report excessive fatigue due to long working hours and the physical conditions associated with the job. A second cause of fatigue is the rapid port turnarounds that have to be made by ships transporting dry cargo between ports in Europe. Similar problems occur in other transport sectors and another of our surveys showed that drivers who made the most deliveries reported the greatest fatigue in the road haulage sector. Other causes of fatigue may also occur (e.g. travelling to the ship). Recommendations are given about possible ways to prevent and manage fatigue. Given the variety of causes and consequences of fatigue, it is unlikely that a "one size fits all" approach will be successful and it will be necessary to address fatigue using several different approaches.

A Review of seafarers' fatigue

The potential for fatigue at sea

Global concern with the extent of seafarer fatigue is widely evident everywhere in the shipping industry. Maritime regulators, ship owners, trade unions and P & I clubs are all alert to the fact that in some ship types, a combination of minimal manning, sequences of rapid turnarounds and short sea passages, adverse weather and traffic conditions, may find seafarers working long hours and with insufficient recuperative rest (Smith, Allen and Wadsworth, 2006). A holistic view is needed of the effects of stress and health factors associated with long periods away from home, limited communication and consistently high work loads on seafarers. In these circumstances fatigue and reduced performance may lead to environmental damage, ill-health and reduced life-span among highly skilled seafarers who are in short supply (Smith, 2007). A long history of research into working hours and conditions and their performance effects in process industries, road transport and civil aviation, where safety is a primary concern, could be usefully compared to the situation in commercial shipping (Allen, Wadsworth and Smith, in press, a).

The fatigue process

Fatigue is best viewed as a process and one should assess factors that induce fatigue, perceptions of fatigue and the outcomes that are associated with fatigue. These outcomes may influence both health and safety and it is crucial that occupational fatigue is viewed as a major health and safety issue. Indeed, there has been considerable research on fatigue at work and onshore studies show that as many as 20% of the working population experience symptoms that would fall into the category of extreme fatigue (Smith, 2007). Many of the established risk factors for fatigue are clearly relevant for seafarers: lack of sleep, poor quality sleep, long working hours, working at times of low alertness (e.g. the early hours of the morning), prolonged work, insufficient rest between work periods, excessive workload, noise and vibration, motion, medical conditions and acute illnesses. Many of these problems reflect organisational factors such as manning levels or the use of fatigue-inducing shift systems. It is often the combination of risk factors that leads to impaired performance and reduced well-being and few would deny that seafarers are exposed to such high risk combinations (Wadsworth et al., in press).

Effects of fatigue on health and safety

Seafaring is a global, safety critical industry in which the potential consequences of fatigue at work are huge. Indeed, fatigue has been identified as a cause or contributory factor in accidents (Raby and McCallum, 1997) and impaired collision risk awareness (Wellens et al., 2005) at sea over many years, and it is clear that the relationships between fatigue and human error, and consequent safety and health, are crucial in seafaring. The health risks associated with fatigue are well established in onshore populations and there is no reason to believe that such associations do not occur in seafarers, although information on this topic is limited (e.g. Wadsworth et al., in press) and further research is required.

Risk factors for fatigue and the prevalence of fatigue

Surveys have shown that risk factors for fatigue at sea are frequently present, as are reports of perceived fatigue. An ITF report (1997), based on responses from 2,500 seafarers of 60 nationalities, serving under 63 flags, demonstrates the extent of excessive hours and fatigue within the industry. Almost two-thirds of the respondents stated that their average working hours were more than 60 hours per week and 25% reporting working more than 80 hours a week (42% of masters). In addition, 36% of the sample were unable to regularly obtain 10 hours rest in every 24, and 18% regularly unable to obtain a minimum of 6 hours uninterrupted rest. Long periods of continuous watch keeping were also reported, with 17% stating that their watch regularly exceeded 12 hours. Over half the sample (55%) considered that their working hours presented a danger to health and safety. Respondents also provided a wide range of examples of incidents that they considered to be a direct result of fatigue. The early hours of the morning were the most difficult in

terms of feeling the effects of fatigue and it is important that safe manning assessments, watch systems and procedures reflect the potential decline in individual performance at these times. More than 80% of the sample reported that fatigue increased with the length of the tour of duty. Long tours of duty were also common (30% reporting usual tour lengths of 26 weeks or above). This cumulative fatigue may also reflect the reduction in opportunities for rest and relaxation ashore, due to the reduced port turn-around times now required.

Our recent survey (Wadsworth et al., in press) has shown that fatigue is still a major issue offshore and that about 30% of seafarers report that they are very fatigued. Fatigue may be present during work, after work and may even extend into the person's leave. Fatigue-related symptoms such as loss of concentration are widespread and these have implications for safety. Indeed, about 25% of respondents reported fatigue while on watch, many reported that they had fallen asleep while on watch, and 50% of the sample reported that fatigue leads to reduced collision awareness. The results have confirmed that there are a number of risk factors for fatigue: tour length, sleep quality, environmental factors and job demands were associated with all of the measures of fatigue. Hours of work, nature of shift, and port frequency/turnaround times were associated with fatigue at work. The likelihood of reporting fatigue and impaired health increases as a function of the number of risk factors a person is exposed to (e.g. 1–2 factors doubles the risk of being highly fatigued but 7 or 8 factors increases the risk 30 times). Diary data confirm results from the survey, as does research from New Zealand (Gander, 2005).

Fatigue reported by the crew of fishing vessels

The crew of fishing vessels often report excessive fatigue due to long working hours and the physical conditions associated with the job. In a small survey of 81 fishermen we have confirmed that fatigue is a major problem in the fishing industry. For example, nearly a third (n = 25, 31%) considered their working hours a danger to their own health and safety, and a quarter (n = 20, 26%) considered their working hours a danger to safe operations onboard. Most of the fishermen (n = 61, 81%) felt that the effects of fatigue increased the longer they were at sea, and 60% (n = 48) said their personal safety had been at risk because of fatigue at work. Thirteen (16%) had been involved in a fatigue related accident, 36 (44%) said they had worked to the point of collapse, 33 (41%) had fallen asleep at the wheel, and 34 (43%) had been so tired they had slept on deck or in the gangway (Smith, Allen and Wadsworth, 2006).

Fatigue induced by rapid port turnarounds

A case study presented from research onboard a mini-bulker illustrated a number of factors which may be critical when trying to account for fatigue on this class of vessel. Whilst the case study highlights that working hours are justifiably considered a key area of concern when addressing fatigue across all ship types, on certain

vessels it may be more profitable to consider the challenging issue of crewing in terms of strategically covering for the disparate demands of port and sea work within a competitive market. Evidence from this case study suggests that ship owners may be inclined (and authorised) to favour leaner 'open-sailing' crewing arrangements which inevitably struggle to cope when faced with demanding port turnarounds. A comparison of seafarers and a small sample of road haulage drivers showed some interesting similarities in terms of risk factors for fatigue. Among the seafarers number of port turnarounds was related to fatigue and a similar trend was seen for the drivers, where those who made the most deliveries were more fatigued. This suggests that lorry drivers and seafarers show parallel trends in terms of fatigue and that fatigue can be observed in contexts which are to some extent operationally comparable (Smith, Allen and Wadsworth, 2006).

Prevention and management of fatigue

Given the diversity of activities undertaken in the maritime sector, and the different profiles of fatigue risk factors in different work groups, it is clear that a range of strategies will be needed to deal with fatigue (see Allen, Wadsworth and Smith, in press, b). The International Maritime Organisation has issued guidance material for fatigue mitigation and management but voted against making fatigue education mandatory. Convention 180 of the International Labour Organisation requires that States fix maximum limits for hours of work or minimum rest periods on ships flying their flags. There is a high degree of agreement among prescriptive regimes with regard to minimum rest requirements. They are generally consistent with current scientific understanding about the sleep required, on average, for people to continue to function at a reasonable level.

Problems with existing legislation and guidance

Two studies from the Cardiff research programme suggest that the legislation aimed at preventing fatigue at sea is not particularly effective. The first examined the impact of the Working Time Directive and the second evaluated the IMO fatigue guidelines. With regard to the Working Time Directive, it was clear from the survey data that excessive working hours and inadequate periods of rest are still problematic onboard a range of vessels. Furthermore, hours are likely to be under-recorded, either by management, or by individual seafarers wary of jeopardising their employment by bringing their company under legislative scrutiny. Another paper from the Cardiff programme evaluated the IMO guidelines on fatigue. It concluded that lengthy, all inclusive guidelines are no substitute for specific and implementable recommendations. Other reports have also evaluated possible methods of preventing and managing fatigue at sea. Houtman et al. (2005) found that the measures that were considered most necessary and effective in reducing fatigue were: proper implementation of the ISM-Code; optimising the organisation of work on board vessels; lengthening of the rest period; and reducing administrative tasks on board vessels.

Conclusions

The evidence for fatigue at sea

The first conclusion from this review is that the potential for seafarers' fatigue is high. An evaluation of the fatigue process shows that seafarers are exposed to many risk factors for fatigue, often report extreme fatigue and may have impaired performance, well-being and health due to fatigue. This statement is supported by a number of studies from different countries, using different samples and methods to evaluate the topic.

Current legislation and guidance is not working

The second conclusion is that current legislation and guidance on fatigue does not appear to have had the desired effect across the industry. One approach to improving the situation is to keep but improve on the current approaches to fatigue (e.g. improve guidance; better enforcement of working time directives).

The way forward

What is obvious from the present review is that fatigue is a health and safety issue. It should be tackled using standard approaches (e.g. appropriate training given; audits etc.) and any increased risk dealt with in a similar way to other breaches of health and safety. Industry wide, cultural change is needed to address fatigue. There will always be the debate as to whether legislation or codes of practice are most appropriate for this area. One approach would be that fatigue policies must be globally in place and that these would be penalised if minimum standards were not met and rewarded if additional desirable features were included. The aim, therefore, would be to have an evolving move towards "best practice" in fatigue prevention and management. The complementary approach is to remove the "worst case" scenarios and "at risk" vessels, such as the mini-bulkers, have been identified as those which need to change manning levels and shift patterns. Fatigue is a complex issue and its prevention and management may require a number of different approaches. The first stage of dealing with fatigue is to get the relevant people to acknowledge that there is a problem to address. The evidence base for this view is strong and has been developed by multi-disciplinary research studying a wide variety of ships in different countries. A second step is to discuss fatigue and how it is influenced by organisational practices and individual factors. Fatigue awareness training and the development of measures to identify fatigue and counter it are becoming common place in other transport sectors and may be a useful part in any package developed to prevent and manage fatigue at sea. However, their efficacy needs to be evaluated. Future research should, therefore, not be restricted to demonstrating that fatigue exists but be concerned with evaluation of methods of preventing and managing seafarers' fatigue.

Acknowledgements

This research was supported by the Maritime and Coastguard Agency, the Health and Safety Executive and NUMAST. We would like to acknowledge the contribution made by those who participated in the research. The views expressed here reflect those of the authors and should not be taken to represent the policy of the sponsors.

References

Allen, P.H., Wadsworth, E.J. and Smith, A.P. In press, a, Seafarers' fatigue: a review of the recent literature. *International Maritime Health.*

Allen, P.H., Wadsworth, E.J. and Smith, A.P. In press, b, The prevention and management of seafarers' fatigue: A review. *International Maritime Health.*

Gander, P. (2005). *A review of fatigue management in the maritime sector*: Massey University Sleep/Wake research centre.

Houtman, I., Miedema, M., Jettinghoff, K., Starren, A., Heinrich, J., Gort, J., et al. (2005). *Fatigue in the shipping industry*: TNO report 20834/11353.

Raby, M. and McCallum, M.C. 1997, *Procedures for investigating and reporting fatigue contributions to marine casualties.* Paper presented at the Human Factors and Ergonomics Society 41st Annual Meeting.

Smith, A.P. 2007, *Adequate crewing and seafarers' fatigue: The International Perspective.* (International Transport Federation, London).

Smith, A., Allen, P. and Wadsworth, E. 2006, *Seafarers' fatigue: The Cardiff Research Programme*. (MCA, Southampton).

Wadsworth, E., Allen, P., McNamara, R. and Smith, A. In press, Fatigue and Health in a seafaring population. *Occupational Medicine*.

Wellens, B.T., McNamara, R.L., Allen, P.H. and Smith, A.P., 2005, Collisions and collision risk awareness at sea: Data from a variety of seafarers. In P.D. Bust and P.T. McCabe (Eds), *Contemporary Ergonomics 2005*, (Taylor and Francis, London), 573–577.

DESIGNING HYPERBARIC DECOMPRESSION CHAMBERS

Elaine Mackie, Andree Woodcock & Chris Johnson

Coventry School of Art and Design, Industrial Design Department, Coventry University, CV1 5FB

Deep sea divers live and work under extreme conditions. It is not untypical for a diver to spend over 20 days living and working underwater. In such instances they share a decompression chamber with other divers (between 2 and 5) who may be working on different shifts. Given the need to redesign and update this accommodation, a questionnaire survey was undertaken of divers to determine what the main issues were with the accommodation. Issues related to not just the quality of communication with the outside world. The recommendations were fed into a series of concept designs and a prototype build which is currently undergoing client approval.

Introduction

The objective of the work was to improve the current conditions within hyperbaric decompression diving chambers based upon an understanding of the range of different diver requirements. Current chambers are sealable pressure vessels with entry and exit hatches for divers that are used after the end of a dive in order to safeguard against decompression sickness (DCS) more commonly known as the bends. DCS is the name given to the variety of symptoms that are suffered by divers exposed to either a decrease or more often an increase in the pressure around their bodies.

Putting the user at the centre of the design of the chamber involves consideration of the divers living environment in terms of the restricted physical space available for them to be able to conduct everyday tasks as well reducing the cases of potential infection caused by the need to control the atmospheric conditions within the chamber. The divers' cognitive requirements are also necessary to take into account when they are faced with remaining in the chambers for unbroken time periods of over a month. The application of human factors can therefore lead to improvements in health and safety, ease of use, comfort and quality of working life and leisure.

Research

The initial stage of the project was to review how the current chambers were being used by the divers. Due to the unpredictable availability of the diving chambers it was not possible to conduct direct observation, interviewing and task analysis. Instead divers' experiences and the understanding of their living conditions were

achieved using a questionnaire approach; the most appropriate method based on published guidance (Boynton & Greenhalgh, 2004; Kirakoski, 2000).

Additionally the physical workplace was assessed by the client and a report compiled based on divers' reviews of the prototype build and compliance with the general guidelines pertaining to diving chamber space and furniture, namely:

- Sintef 'Design Requirements for Saturation Diving Systems: The living chamber'
- Offshore Standard DNV-OS-E402

These Norwegian guidelines give very prescribed recommendations on the internal components for the diving chamber e.g. bunks, seats, tables etc. and are based upon the size and construction of diving chambers that the design aims to satisfy as opposed to British Standards which are not as specific and open to interpretation. Panero & Zelnick (1979) was used as a reference for the design details not mentioned by Sintef in terms of determining seat incline, heights of cabinets and mirrors.

Questionnaire development and design

The questionnaire was designed by third year intern students and the author following a series of meetings with the project manager and client. The meetings flagged up assumptions that might have occurred about the use of the chamber without initial client feedback. Its principal purpose was to provide detailed information about what kind of activities divers would like to do in the chambers, the amount of private space that they have and their working shift patterns that could provide the designers with an insight into the life of the divers in the diving chamber and the problems they experienced. From this a set of user requirements could be devised.

Questionnaire distribution

The questionnaire was distributed in mid July 2006 to the client company and divers currently engaged in deep sea diving who would form the end user population of the chamber. They had a three fold purpose:

1. To enable the Coventry design team to understand more about the nature of the work and leisure time of deep sea divers
2. To provide initial input into early concept designs
3. To understand communication with the outside environment.

Results

Personal details

The results are based on 28 (sometimes partially) completed questionnaires. The age of the divers varied from 30+ to under 60 years of age, with both the modal and arithmetic mean age lying in the 41–50 year age range. The majority of divers in the sample have over 20 years diving experience. This trend indicates that diving has been their main occupation, and they are likely to be highly skilled, used to

diving conditions and the decompression chamber. Most of the participants' longest period of time in a decompression chamber was between 26–35 days.

Nature of work

36% of the respondents reported muscular problems, 57% stated they had no problems. Much of the work undertaken was of a highly demanding, physical nature. There was also a general reporting of arm, neck, shoulder and lower back pain, sore wrists and fingers.

Physical fitness – away from work

All respondents rated themselves as having either average (10 respondents) or above average fitness (17 respondents), with one respondent rating himself as having a high level of fitness. When not at work, the hours per week spent exercising varied from under 4 to up to 16 hours, with the average time lying in the 4–8 hour range.

The types of exercises undertaken were varied with most of the respondents undertaking more than one form of exercise. The most popular form of exercise being that related to leg strength; which could be explained by the fact that divers do not use their legs extensively when they are diving and therefore need to make sure that sufficient circulation is maintained in these limbs. The majority of activities undertaken were non competitive. Some of the gym based activities, such as resistance bands were also used in the decompression chamber.

Physical fitness – during work

The amount and type of exercise undertaken changed when the divers were in the decompression chamber, with the majority of divers never exercising inside. However the confined space of the chamber and the conditions of the space where exercise could be carried out (i.e. the shower/wet room) made this difficult to achieve. For some, exercise started when the decompression started. Exercises undertaken required little (e.g. resistance bands, bungee tubes, stepper, dumb bells and light weights) or no equipment (sit ups, press ups, lunges, stretch, lateral raises and bicep curls, crunches). Sit ups and steps were the most frequently mentioned (three occurrences each).

Shift work

It was very clear that having two different shifts operating from one chamber caused disruption to sleep with 87% of the divers commenting upon the affect of noise intrusion from such items as the operation of locks and hatches and teams passing through the chamber when others were asleep. There was also a change in circadian rhythms, with a drift towards shorter days.

Communication patterns

Communication with colleagues outside the chamber was mainly related to instructions about the dive and took place when needed via radio, email via flashcard, fax

and intercom. This communication would vary from 5–10 minutes 3 to 4 times a day, once or twice a shift from a few minutes to up to 2 hours.

Everyone communicated with their family and friends with the amount of time, method and frequency of the communication varying. Some of the divers were in contact with their families every day, in some cases 2 or 3 days a week.

Living environment

In the chamber there are essentially three different areas; namely the shower/toilet, living/eating quarter and sleeping space.

Ingress and egress from the chamber is currently through a circular hatch located within the shower/toilet area. Another circular hatch is located from the shower/toilet area into the main living/eating quarters. Approximately 75% of the divers questioned had not experienced problems in this respect.

Only two respondents had no issues about the interior. For the rest of the divers comments were associated with:

- Shower/toilet: privacy within this area.
- Living/eating quarter: seats – no support for back and generally a lack of seating with much time being spent on bunks. Moreover the space inside the chamber did not allow for a table. Neither were there any communication facilities such as the internet or e-mail or wi-fi. The lack of integral entertainment facilities meant that many divers had to take their own personal items in with them ranging from the majority taking books, through to personal electronic equipment, food and spare clothes.
- Sleeping space: bunks – considered too small and narrow and mattress uncomfortable, and not enough privacy provided by the curtains.

In terms of responses covering all areas of the chamber, both general and specific recommendations can be made regarding the use and design of space, changes in work pattern down to the design of shaving mirrors and mattresses. The most important areas would seem to relate to:

- Provision of separate chambers for different shifts
- Better toilet, sleeping and sitting areas – spaces, fixtures and fittings
- Better communication facilities
- Noise reduction
- State of the art entertainment facilities, and built in leisure facilities
- Provision of an exercise area.

Questionnaire interpretation into design

The detailed report was used to direct design in terms of converting concept sketches into computer generated images generated by Rhino Photoshop and Illustrator, examples given in Figure 1. Original dimensioned information came from the client as Solidworks files, e-drawings, pdf drawings and jpgs. Rhino work was translated into iges files for 5 axis machining to create some of the components to

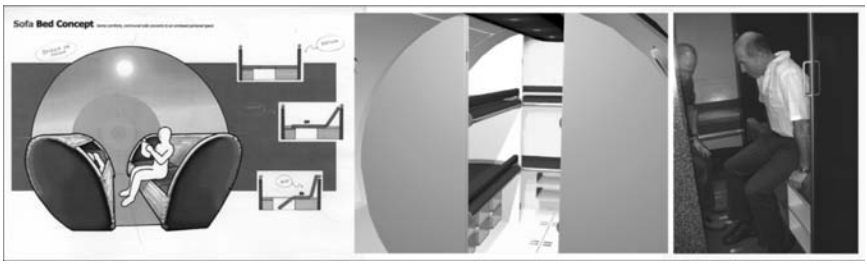


Figure 1. Sketch concept translation into computer file and mockup evaluation.

be incorporated in a full sized mockup built by a professional team supported by eight third year intern students.

This mockup has been used to facilitate the next development of the chamber design by having eight divers who had all completed the questionnaires and 4 medical experts from Norway further comment upon the suitability of the interior and compliance of relevant standards. These reviews consisted of allowing the divers to test the seating and bunks inside the mockup in terms of sturdiness.

A further visit was made by a retired diver and a Dive System Coordinator who together with a member of the ergonomics teaching staff informing evaluated the design against the requirements that emerged from the questionnaire. This approach provided a more in depth understanding of the scenario of usage associated with the furniture and equipment inside the chamber.

Conclusions

Not all the issues raised by respondents could be addressed directly as some issues related to personnel, management and financial concerns.

However, what emerged from the questionnaire approach was that divers strongly expressed a need to be included and have ownership of the design. The final design needs to be refined by divers' first hand knowledge of the problems they face within the chambers as the design team have no first hand knowledge or experience of the working conditions, the work routine or characteristics of the end user population.

To date the reviews of the prototype build have been favourable and the designers have a better awareness of divers' requirements.

Recommendations

It is recommended that further site visits be conducted with a more structured approach in order for the designers to better understand 'a day in the life of' divers by videoing divers highlighting the current issues with the use of the interior. These would not only provide invaluable information and insight into the design requirements, thereby leading to a better design, but also ensure that the divers are active stakeholders in the design of their new environments.

Acknowledgements

The authors would like to acknowledge Divex Limited, Aberdeen, the divers and designers for their participation and contribution towards the research and assessment of the chamber mockup.

References

Boynton, P.M., & Greenhalgh, T. 2004, Hands-on guide to questionnaire research: Selecting, designing, and developing your questionnaire, *British Medical Journal*, Volume **328**, 1312–1315

Kirakoski, J. 2000, *Questionnaires in Usability Engineering: A list of frequently asked questions* (3rd Edition), Human Factors Research Group, Cork, Ireland

Panero, J., & Zelnick, M. 1979, *Human Dimensions and Interior Space: A Source book of Design reference standards*', Watson-Guptill

Toal, J. 2007, *Comparison of Standards for Living Environments*, Divex Limited, Westhill, Aberdeen, UK

Sintef 'Design Requirements for Saturation Diving Systems: The living chamber'

Offshore Standard DNV-OS-E402

SLIPS, TRIPS AND FALLS SYMPOSIUM

AN EVALUATION OF THE DIFFERENT STYLES OF FOOTWEAR WORN AT THE TIME OF A STAIR FALL

Victoria Halford & H. Harvey Cohen

Error Analysis Inc., La Mesa, CA, USA

This study aims to address some of the gaps in current falls research; namely the lack of studies focusing on occupational stair falls, and the effects of different types of footwear worn at the time of a fall. Forty-nine subjects, all of whom were employed as care workers, and who had reported a stair fall, were interviewed. The footwear worn at the time of the fall was assessed using the standardised Footwear Assessment Form. The Chi-square test was used to analyse the relationships between footwear types and type and severity of stair fall. Results revealed that the majority of subjects had worn sub-optimal footwear, and indicated an association between flat slip-on shoes and a number of different types of fall.

Introduction

Most falls occur in the domestic environment, with 61% of fatalities from falls occurring on stairs (Cayless, 2001). However, it has been estimated that each year in the UK there occur over 100,000 injuries on non-domestic stairs, and around 100 fatalities (Scott, 2005). In spite of these figures, there exists relatively little available literature about falls on stairs in the workplace, compared with those in the domestic setting (Cohen *et al.*, 1985, Templer *et al.*, 1985). Much of the literature has focused on gait disruption, caused either by poor or inconsistent design, or by other environmental factors (Chang, 1999, Cohen, 2000, Cohen and Cohen, 2004, Roys, 2006).

Many such footwear research studies have tended to focus on the material properties of the sole of the shoe, and how those properties might be implicated in a fall; fewer studies have considered the style of footwear. A few researchers have considered footwear styles in terms of balance and these have tended to be amongst older subjects (Lord *et al.*, 1999, Sherrington and Menz, 2003). Most of the studies which have considered footwear as a risk factor have tended to use older users as subjects. This study considers younger subjects, all employed in the care home environment (care assistants, managers, nurses). The aim of the study was to analyse the footwear worn at the time of a stair fall, in order to consider a link between footwear features and likelihood of falling amongst care workers.

Method

Six care and nursing homes were approached for the study. Forty-nine subjects were eventually included (all female, average age 38.6 years), who had all reported a stair fall incident at work between the dates February 2005 and June 2006. Subjects needed to be able to recall the shoes worn at the time of the fall. The location and type of fall were recorded for each subject. Footwear was assessed using the validated standardised Footwear Assessment Form formulated by Menz and Sherringham (2000). Assessment was based upon condition of the shoes as well as design features. The stairs were assessed according to UK Building Regulations (BS5395-1:2000), to ensure that dimensions were consistent and within the acceptable limits. All stairs were checked for handrails.

Results from the survey were coded according to category. Data was analysed using SPSS Release 10 for Windows (SPSS Inc, Chicago, USA). The Chi-square test was used to analyse the relationships between footwear types and characteristics, location of fall on stairs, type and location of stairs (interior/exterior, straight/dogleg in design) and severity of fall. The relationship between footwear characteristics of subjects who fell and floor covering was also analysed. Type of fall (whether slip, trip, loss of balance or fall) was recorded, based upon subjects' own description.

Results

The most common type of footwear worn at the time of a stair fall was flat slip-on shoes (42%), followed by Court shoes (28%) and backless sandals (21%). The least common type was trainers (9%). The majority of subjects wore shoes with one or more sub-optimal characteristic, such as poor or absent fixation (89%), worn 'polished' tread pattern (66%) and heel height of >5 cm (28%). Most stairs (61%) were carpeted, 8% had a linoleum cover, and 31% were concrete. The vast majority of subjects (85%) did not need any medical assistance, although (93%) did experience bruising. Two subjects (4%) suffered a fracture, requiring hospital treatment.

Association between footwear types and characteristics

Flat slip-on shoes were more likely to feature sub-optimal or hazardous features such as no fixation (Chi-square = 10.51, df = 1, $P < 0.001$), 'polished' soles (Chi-square = 9.88, df = 1, $P < 0.001$) and excessively stiff or soft soles (Chi-square = 11.29, df = 1, $P < 0.001$). Court shoes were the most likely type of shoe to have a sub-optimal heel height (Chi-square = 11.05, df = 1, $P < 0.001$).

Association between footwear type and type of fall

Subjects who reported wearing heeled shoes of >3 cm (such as Court shoes) were more likely to have reported a loss of balance (Chi-square = 10.56, df = 1, $P < 0.001$). There was also an association between wearing flat slip-on shoes and the reporting of slipping (Chi-square = 12.09, df = 1, $P < 0.001$). Trips were associated

more with Court shoes (Chi-square = 12.28, df = 1, $P < 0.001$). There was no association between footwear type and severity of fall.

Association between footwear type and location of fall on stairs

There was an association between subjects who wore shoes with no fixation and subjects who reported a stair fall on the last three steps whilst descending (Chi-square = 11.87, df = 1, $P < 0.001$).

Certain types of footwear were more likely to have sub-optimal characteristics: flat slip-on shoes appeared to be more worn than other styles of shoes (Chi-square = 10.51, df = 1, $P < 0.001$) and more likely to have excessively flexible heel counters (Chi-square = 16.22, df = 1, $P < 0.001$).

Discussion

These results indicate that the majority of care home staff surveyed wears inappropriate and sub-optimal footwear. Furthermore such footwear choices may have been implicated in the stair falls that were analysed in this study. A number of subjects commented that they felt they had made good footwear choices because they had selected flat shoes over high-heeled. Although high heels have indeed been implicated in falls (Lord and Bashford, 1996) due to loss of balance, subjects were not aware of the full range of sub-optimal risk factors associated with their footwear choices. This means there is a contradiction between stair-user perceptions of appropriate footwear and research findings. Interestingly, few falls recorded in this study were reported to be as a result of loss of balance; while many were reported as being a result of trips and slips. There was, however, an association between the wearing of Court shoes with a heel >3 cm and a reported loss of balance on stairs. This is in keeping with previous findings (Lord & Bashford, 1996).

Footwear was evaluated as sub-optimal from the point of view of design characteristics and wear-and-tear, although this distinction was not recorded. In a number of cases, subjects' shoes were well-designed (with appropriate fixation etc.), but in poor condition. The presence of, for instance, a worn-out sole, may well have contributed to a subject's stair fall, in much the same way as an under-used sole would, in terms of loss of friction (Manning *et al.*, 1985, Gronqvist 1995). Some footwear characteristics (such as lack of fixation) may have been a result of sub-optimal design; others, such as excessively flexible heel counters may implicate both inadequate design and wear-and-tear. A more in-depth study would analyse the significance of footwear maintenance as well as design factors in occupational stair falls. This may well have implications in terms of user-awareness of the risks and hazards of stair-use.

It might also extend to management attitudes to staff footwear: a number of subjects commented that workplace regulations restricted the type of shoes they were able to wear. Some managers included in the study commented that trainers were not deemed 'appropriate' footwear for staff, as they did not look 'professional' or presentable. This accounts for why less than 10% of subjects reported wearing trainers at the time of their fall. Interestingly, the trainers inspected for the study

were significantly less worn than other shoes, and had a number of optimal design features: good fixation, appropriate sole stiffness, heel height of <3 cm, and so on. On paper at least, such footwear would appear to be ideally suited to care-home work. Unfortunately the study did not investigate user perceptions of comfort, and the comments made by managers suggest that comfort is not a factor for consideration in staff footwear guidelines.

This study backs up Templer's (1992) research, which found that most stair falls occur at the first and last 3 steps, suggesting that such falls occur when the body has to adapt to a change in environment (from a flat walking surface to stairs, and vice versa). This study has, in fact revealed an association between wearing shoes without fixation and falling at descent of the last three steps of a staircase. This is possibly because as well as navigating a change in walking surface, descending stairs has the additional gravitational force acting upon a subject. Footwear without fixation may not be able to withstand the forces acting upon it during descent, resulting in an increased likelihood of falling.

A particular drawback of this study design is the lack of consideration of fit. Subjects did not have the size of their shoes checked against their actual foot size. Earlier research (Finlay, 1986) has found that less than 50% of subjects wore shoes of the correct size. Ideally appropriateness of fit would be included in the footwear evaluation in any future study. It was noted that accident logs were not consistent between care homes: this may suggest a larger issue of unawareness of the importance of recording accurate data, as well as the possible institutional need for standards in footwear choice. Furthermore, it should be noted that the literature has shown that missteps are far more common on stair treads than slips due to loss of traction (Bakken *et al.*, 2007) although laypersons commonly report such incidents as "slips." Further research will look at whether this result is an artefact of self-reporting, or whether slip incidents on stairs are actually more prevalent due to the types of footwear worn.

Conclusions

While it is not possible to ascertain the exact role that footwear plays in stair-falls, it is clear that many care workers in this study had suffered a stair-fall whilst wearing sub-optimal footwear. Some footwear was evaluated as being of sub-optimal design, some poorly maintained/worn out. Awareness amongst subjects of optimal shoe design, though not formally measured, appeared low. There appears to be a complex dynamic at work between footwear, stair surface and stair design. More rigorous studies, beyond the scope of this report, are clearly required to interrogate the exact nature of this relationship.

References

Bakken, G.M., Cohen, H.H., Abele, J.R., Hyde, A.S. and LaRue, C.A. 2007, *Slips, Trips, Missteps and Their Consequences*, Second Edition, (Lawyers and Judges Publishing Co., Inc., Tucson, Az)

Cayless, S.M. 2001, Slip, trip and fall accidents: Relationship to building features and use of coroners' reports in ascribing cause, *Applied Ergonomics*, **32**, 155–162

Chang, W.R. 1999, The effect of surface roughness on the measurements of slip resistance, *International Journal of Industrial Ergonomics*, **24**, 299–313

Cohen, H.H. 2000, A field study of stair descent: Why and how people fall down stairs. *Ergonomics in Design*, **8**(2), 11–15

Cohen, H.H., Templer, J. and Archea, J. 1985, An analysis of occupational stair accident patterns, *Journal of Safety Research*, **16**(4), 171–181

Cohen, J. and Cohen, H.H. 2001, Hold on: An observational study of staircase handrail use. Proceedings of the 45th Annual Meeting, Human Factors and Ergonomics Society, Minneapolis, MN, USA, October 2001

Finlay, A.E. 1986, Footwear management in the elderly care programme, *Physiotherapy*, **72**, 172–8

Gronqvist, R. 1995, Mechanisms of friction and assessment of slip resistance of new and used footwear soles on contaminated floors. *Ergonomics*, **28**, 224–241

Lord, S.R. and Bashford, G. 1996, Shoe characteristics and balance in older women, *Journal of the American Geriatric Society*, **44**, 429–33

Lord, S.R., Bashford, G., Howland, A. and Munro, B. 1999, Effect of shoe collar height and sole hardness on balance in older women, *Journal of the American Geriatric Society*, **47**, 1–4

Manning, D.P., Jones, C. and Bruce, M. 1985, Boots for oily surfaces. *Ergonomics*, **28**, 1011–1019

Menz, H.B. and Sherringham, C. 2000, The Footwear Assessment Form: A reliable clinical tool to assess footwear characteristics of relevance to postural stability in older adults, *Clinical Rehabilitation*, **14**, 657–64

Roys, M. 2006, Steps and stairs. In R. Haslam and D. Stubbs (eds) *Understanding and Preventing Falls*, (Taylor & Francis, London), 51–68

Scott, A. 2005, *Falls on Stairways – Literature Review.* (HSL Human Factors Science Group Report number HSL/2005/10)

Sherrington, C. and Menz, H.B. 2003, An evaluation of footwear worn at the time of fall-related hip fracture, *Age and Ageing*, **32**(3), 310–314

Templer, J.A., Archea, J. and Cohen, H.H. 1985, Study of factors associated with risk of work related stairway falls, *Journal of Safety Research*, **16**, 183–196

Templer, J. 1992, *The Staircase: Studies of Hazards, Falls, and Safer Design*. (MIT Press, Cambridge)

REFLECTIVITY AS AN INDICATOR OF SLIP RISK

Paul Lehane & Rob Clark

Environmental Health & Trading Standards
London Borough of Bromley, Civic Centre, Stockwell Close
Bromley. BR1 3UH England

Traditional methods of assessing slip resistance have utilized various devices based on measuring friction. It is recognized that a dry clean floor offers the best slip resistance and that this is reduced by the presence of contaminants such as water, oils or powders.

Quirion et al 2007 used light reflection in laboratory assessments of cleaning efficiency on floor tiles noting that the reflection of light from the test tile increased as oil was added reaching a plateau value when the tile was saturated with oil. The more efficient the cleaning method that was employed the more oil contamination that was removed the lower the reflectivity measurement.

This paper examines the potential of reflectivity as a method of assessing slip risk by comparing the reflectivity measurements obtained from 18 floor surfaces in restaurants in the London Borough of Bromley England with the ratings obtained using existing assessment techniques – SlipAlert and the Slip Assessment Tool (SAT) developed by the HSE.

Introduction

Slip accidents continue to be a significant source of occupational injury and lost time in England and Wales. Reducing their incidence is a priority for the Health and Safety Executive (HSE), Local Authorities and businesses in working towards achieving a 10% reduction in the rate of fatal and major injuries by 2010 as set out in the 'Revitalising Strategy' for health and safety (DETR 2000).

Slips occur when a pedestrian's frictional requirement exceeds those of the floor surface on which they are walking. There has been a significant amount of research concerning the measurement of friction and applying it to the complex process of human locomotion, yet this knowledge largely remains within a small specialist community.

To achieve a reduction in slip accidents this specialist knowledge needs to be incorporated in to simple and workable solutions that can be understood and applied by non-specialists and duty holders. This is beginning to happen with the publication of HSE guidance and the emergence of a number of tools designed to simplify the assessment of the slip risk presented by floors. The HSE's preferred method for assessing slip resistance is the pendulum (HSE 2007) but this requires specialist

training to use and is very cumbersome. 'SlipAlert' is a commercially available friction based assessment tool which has been tested by the Health and Safety Laboratory (HSL 2006) who reported that it should be regarded as providing a good indication of available friction, lending itself to risk assessment, monitoring of floor surfaces and evaluating & monitoring cleaning regimes. The reliability of some other friction measurement devices has been questioned and there is still a need to develop further simple and reliable on-site testing methods.

If duty holders are to be able to make objective assessments of slip risks they need to be able to 'measure' or quantify the risk and the effect of any interventions designed to reduce it. In the main 'measurements' are restricted to frictional properties (coefficients of friction or surface roughness) by devices such as the Pendulum (HSE 2007), SlipAlert (SlipAlert 2007), or Tortus (Tortus 2007) or by using HSE's Slip Assessment Tool (HSE 2007a) which makes an overall assessment of a floor based on a range of variables including surface roughness, contamination, cleaning and usage.

Ergonomists and safety practitioners need to be open to the potential offered by alternative methods of assessing slip risks to support duty holders fulfil their responsibilities and this paper examines how reflectivity might offer another way to assess slip risk.

Quirion et al (2007) developed a technique for assessing the oiliness of flooring using reflected light. The method is based on an observation that the value of R (reflected light) increased as the amount of oil on a tile increased until its value reached a plateau when the tile was saturated. Francois Quirion constructed a 'reflectivity' meter, which was made available to the researchers to undertake field trials. The meter uses a 1 cm diameter red LED as the light source. The light is directed at an angle of 45° to the surface to be measured. Light reflected from the surface strikes a photoelectric cell also mounted at 45° to the surface. The greater the amount of light that is reflected and strikes the photoelectric cell the higher the digital reading on the meter.

The amount of oil on a tile has been shown to affect its slipperiness and therefore its slip risk (Underwood 1992) and this field trial of the reflectivity meter was designed to allow an initial assessment of its operation and performance and compare this with two existing methods of assessing slip risk – SlipAlert (SlipAlert 2007) and HSE's Slip Assessment Tool (HSE 2007a).

Methodology

A variety of different types of catering establishments in the London Borough of Bromley were approached by the authors who explained the purpose of the survey and sought their voluntary participation. Arrangements were made to undertake the measurements at a time suitable to the businesses. This was usually during the mid/late afternoon when they were less busy.

The premises were surveyed and suitable flooring and positions for testing were identified. Measurements were initially undertaken in the food preparation areas and then from serving areas and public areas where they had a different type of

flooring and it was practical and safe to undertake the measurement procedure. Each type of flooring took between 15 and 20 minutes to test using three different measurement methods – Reflectivity, SlipAlert and SAT.

Reflectivity testing

The reflectivity measurements of the flooring were taken under three conditions using a meter and method devised by Quirion et al 2007. Ten Reflectivity readings for each condition were taken in a small area of the flooring, in the following sequence 1) as found, 2) clean and 3) saturated with oil. The meter was calibrated before and after each set of readings against a standard tile provided with the meter. The 'battery' reading was also noted at the beginning and end of each set of readings and divided by .184 to give the BAT value. Mean values from each set of readings (floor condition (TILE), reference (REF) and battery (BAT) were used in calculating the reflectivity value (Fig 1).

After measuring the floor in the 'as found' condition it was cleaned using a 1 in 4 mix of a degreaser and water. An area of floor was cleaned by spraying the diluted degreaser on to the floor wiping it dry with a paper towel. This was repeated three times. 'Clean' reflectivity readings were then taken. A small area of the floor was then covered with extra virgin olive oil spread evenly over the test area (approximately 150 mm × 150 mm). Oily (saturated) reflectivity measurements were taken and the floor was then thoroughly cleaned and dried.

SlipAlert

SlipAlert was used as supplied and operated according to the instructions except that the mean of 10 readings was used rather than 3. Testing was undertaken on the floor as found and wetted using clean tap water applied by a hand held sprayer. The floor was dried after testing. SlipAlert is released from a ramp and the distance travelled across the floor is recorded from the digital display and converted to a COF and a 'risk' rating from the graph provided. (Low, medium and high risk of slipping)

Slip assessment tool (SAT)

SAT is a computer based assessment and prediction tool developed by HSE and freely available in the HSE Website. On-site surface roughness measurements (Rz) were taken using a Surtronic Duo roughness meter and a pro-forma was used to collect other information. SAT ratings were calculated off site. The SAT reading is a numerical indicator of the slip risk presented by the floor. The higher the number the greater the risk. (Low, medium, significant and high risk of slipping)

$$R = 100 \times \frac{(BAT - REF) \times TILE}{(BAT - TILE) \times REF}$$

Figure 1. Calculation of reflectivity (R).

Table 1. Mean results by floor type.

Flooring type	Mean Rz	Mean SAT Rating	Mean SlipAlert as found	Mean SlipAlert wet	Mean Reflectivity as found	Mean Reflectivity clean	Mean Reflectivity oily
Safety	35	18 low	125 low	160 medium	33	33	82
Quarry	20	23 medium	123 low	211 high	51	51	86
Ceramic	11	39 significant	149 medium	236 high	65	57	94
Vinyl	12	14 low	125 low	183 high	67	75	99
Wood	10	13 low	112 low	171 medium	57	63	90
All floors types	18	22 medium	128 low	195 high	53	54	109

Results

18 floors were tested using the three test methods. (4 safety floors, 6 quarry tile floors, 4 vinyl/PVC floors, 3 ceramic tiled floors and 1 wooden floor). The mean results for the flooring types are presented in table 1.

Reflectivity readings were able to distinguish the 'oily' floor from the 'as found' and 'clean' floors. Paired samples T tests showed significant differences between the 'oily' floor and the 'as found' and 'clean' flooring measurements. (Oily/As found $t = -12.036$, 17df, $p = .000$ and oily/clean $t = -10.185$ 17df, $p = .000$), however no significant difference was identified between the 'as found' and 'clean' readings.

The readings across all floor types for the three measurement techniques were tested for correlations using two tailed Spearman's Rho as the sample size was small. No correlations were found between the Reflectivity measurements obtained and those from either SAT or SlipAlert, neither was any correlation found between the values from the SAT or SlipAlert.

Significant positive correlations (Spearman's Rho two tailed) were found between the three reflectivity readings – as found, clean and oily. (as found/clean .924 $p = .000$, as found/saturated .682 $p = .002$ and clean/saturated .701 $p = .001$).

In addition a significant negative correlation was found between reflectivity (as found and clean) and Rz roughness. (Spearman's Rho two tailed $-.783$ $p = 0.01$ for Rz and R as found and $-.684$ $p = 0.01$ for Rz and R clean).

Discussion

Reflectivity measurements clearly differentiated between artificially saturated oily floors and those conditions found on site or following cleaning. There are two possible explanations for the lack of difference between the 'as found' and 'cleaned' reflectivity measurement.

The floors tested 'as found' were already clean from the routine cleaning procedures employed in the restaurants or the test method of cleaning was ineffective,

however as Quirion et al (2007) reported that degreasers are the most effective type of detergent to use on oily floors, it is suggested that the floors were probably already clean.

The reflectivity measurements were expected to follow the order, clean, as found and then oily, with increasing oiliness resulting in higher readings. This was only observed in half of the floors tested. In the other half of the floors tested the 'clean' and 'as found' readings were reversed. This is possibly due to the presence of dirt on the floor in addition to oils. Dirt has the effect of dulling the surface and reducing the reflectivity. Removing the dirt allows the floors natural reflectivity to be measured. However, as reported earlier, the difference between the 'as found' and 'clean' floor measurements was not statistically significant.

The reflectivity meter was designed for laboratory use but field trails have indicated that it is robust enough for on site use. This small-scale study has not found any correlation between reflectivity and other test methods (SAT and SlipAlert) but neither were any correlations identified between SAT and SlipAlert. We had hypothesized that there would have been a relationship between the existing tools assessing the risk of slipping, if only to the extent of the risk categories (low, medium & high slip risk). This effect may be due to the small sample size.

There is a certain amount of electronic drift in the reflectivity meter readings in that the calibration setting did not remain static and the adjuster screw is very small and difficult to adjust. To obtain the oily reflectivity it was necessary to put oil on the floor. This is messy and relied on a visual assessment to ensure proper coverage. The oil had to be thoroughly cleaned from the surface after testing to ensure the floor was left in a safe condition.

The meter uses a red LED light source and this may influence the reflectivity from different coloured floorings. Further testing to determine the effect using a white light or another coloured light source might be considered beneficial.

As reflectivity readings (as found, clean and oily) were strongly correlated between themselves and as the 'as found' and 'clean' were strongly correlated to surface roughness (Rz), it has the potential to offer a method of assessing slip risk. Just as there are a number of roughness parameters available, the relationship between aspects of reflectivity and slip risk need further research to establish its relevance and reliability.

The support of François Quirion for the loan of the Reflectivity Meter is gratefully acknowledged.

References

DETR 2000, *Revitalising Health and Safety: Strategy Statement*. London, Department of the Environment, Transport and the Regions. www.hse.gov.uk/revitalising/strategy.pdf

HSL 2006, Evaluation of the Kirchberg Rolling Slider and SlipAlert Slip Resistance Meters PED/06/09 Health and Safety Laboratory Buxton UK

HSE 2007, Assessing the slip resistance of flooring (A technical information sheet) www.hse.gov.uk/pubns/web/slips01.pdf

HSE 2007a, Slip Assessment Tool. HSE, www.hsesat.info/ accessed 4 December 2007

Quirion, F., Poirier, P. and Lehane, P. 2007 Optimal Cleaning to prevent slippery floors in restaurants *JEHR Volume **6** Issue 1*

SlipAlert 2007, http://www.slipalert.com/ accessed 4 December 2007

Tortus 2007, http://www.tortus.co.uk/ accessed 4 December 2007

Underwood, D.C. 1992, Effect of floor soil on coefficient of friction in food service operations, *Ceram. Eng. Sci. Proc*. **13**, 78–85

ACCIDENTS ON ENGLISH DWELLING STAIRS ARE DIRECTLY RELATED TO GOING SIZE

Michael Wright & Mike Roys

BRE, Garston, Watford, WD25 9XX, UK

A postal survey of dwelling stairs in England was carried out, asking occupiers to measure the rise and going of three successive steps. The questionnaire also asked if an accident had happened on the stairs in the last two years. After tidying the data there were 1,223 responses. The mean going of stairs of those people who recorded that they knew of an accident occurring on the stairs in the last two years, and those that did not, was compared and found to be significantly different, measured by the Mann-Whitney test ($p < 0.005$). There were no significant difference between for the mean rises ($p = 0.08$), or the number of steps in a flight ($p = 0.72$) between these two groups.

Across the range of stair geometries found in a random sample of English dwellings there is a very significant influence on accident rate of going size, with larger goings reducing the likelihood of a stair accident being recorded.

Introduction

Stairs in the home are involved in over 500 deaths and over 300,000 injuries in England and Wales each year. There may be a further 1.5 million minor falls on stairs in the home that are not serious enough for the victim to attend A&E (Roys and Wright, 2003). This equates to a minor fall every 20 seconds on average, and an injury every two minutes. Although previous research (Roys, 2001) indicated that step dimensions play a major part in the risk of an overstep, which is a major contributor to slips and falls down stairs, there was little hard evidence that smaller goings actually lead to an increased risk of an accident.

Experimental methodology

In 1999 a self completion postal survey of dwelling stairs in England was carried out. Ten thousand stair survey questionnaires were printed and distributed to a random selection of addresses in England. Of these 10,000 questionnaires, 1,939 were returned, recording information about stairs in homes. The survey collected a large amount of data about the main stair, including;

- the type of stair (straight, turning with winders, turning with landing, or spiral),
- the rooms at the top and bottom of the stair,

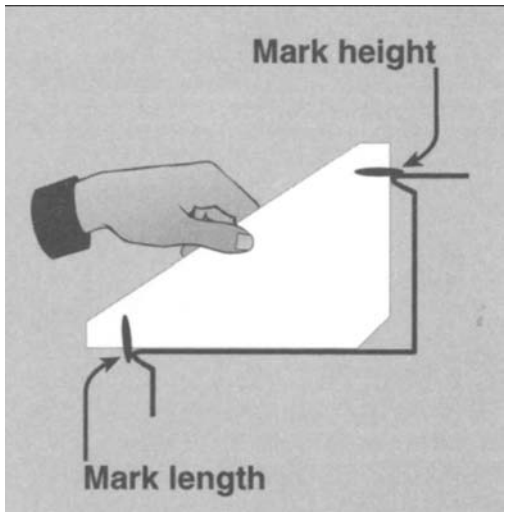


Figure 1. Diagram in questionnaire showing how to use the card templates.

- the number of steps in the stair,
- whether the rises were open or closed,
- the banister arrangement on left and right sides (wall, with handrail, without handrail, banister with vertical supports, with broad horizontal supports, intricate patterning, solid banister, none), and
- the stair covering (carpet, exposed wood, exposed metal or concrete, lino or vinyl, other).

The survey asked the occupiers to measure the straight part of the stairs using the following procedure and three card templates that were provided with the questionnaire;

1. Select three straight steps.
2. Place the card marked "TOP" on the highest of the three steps, as shown in the diagram, and mark the height and length of the step on the card.
3. Repeat this process on the next two steps down with the cards marked "MIDDLE" and "BOTTOM".

The questionnaire also asked the occupier if they, or someone they knew, had had an accident on the stairs in the last two years. It was felt that two years was a reasonable timescale to ask people to recall, long enough to provide a reasonable sampling frame, yet not so long that people's memories would be unreliable.

Data preparation

Of the 1,939 responses 168 did not return the card templates, leaving 1,771 returns with some physical measurements. These returns were processed to produce a robust

data set that was providing information about stairs which are likely to be the main stair in a dwelling. The data processing involved:

[i] Only including a response if the flight of stairs has between 9 and 16 steps (inclusive).
[ii] Deleting all responses where any of the rise and going measurements are missing.
[iii] Deleting all responses where the difference in any successive going values is 100 mm or greater. This was done to reduce cases where there was either a doubt about the measurement, or the stairs were so non-uniform to make them un-representative.
[iv] Deleting all responses where the difference in any successive rise values is 50 mm or greater. This was done to reduce cases where there was either a doubt about the measurement, or the stairs were so non-uniform to make them un-representative.
[v] Delete all cases where it is not known if there was an accident in the last two years.
[vi] Delete all cases where the average going is 100 mm or less, or is greater than the size of the template, 300 mm.

After these measures had been taken there were 1,223 cases: 147 where the respondent said an accident had occurred on the stairs in the last two years, and 1,076 where there they recorded no such accident.

Results

The above cases were analysed using SPSS. The three rise and three going measures were averaged to produce one rise and going pair for each response. These distributions were tested for normality, using the one-Sample Kolmogorov-Smirnov test. The going distribution was not normal (significance = 0.001) and hence only non-parametric tests were conducted.

To give an idea of the distribution of the stair geometries that were present in England in 1991 a scatter diagram of the average rise against the average going is presented in Figure 2. It is interesting to note that only 718 of the 1,223 cases have an average going of 220 mm or more. This means that 41.3% of the sample does not meet the current Building Regulation (DETR 1998) for this measure alone, ignoring other restrictions such as the pitch and rise values.

Comparisons were made between the cases with an accident reported in the last two years, and those without, for the average rise, the average going and the number of stairs in the flight. The Mann-Whitney U-test shows that only the going is significantly different between these two sets, at a significance of <0.005. The rise has a significance of 0.08 and the number of steps has a significance of 0.72. The distribution of the average goings for the two sets of stairs, those with a recent accident and those without, is presented in Figure 3.

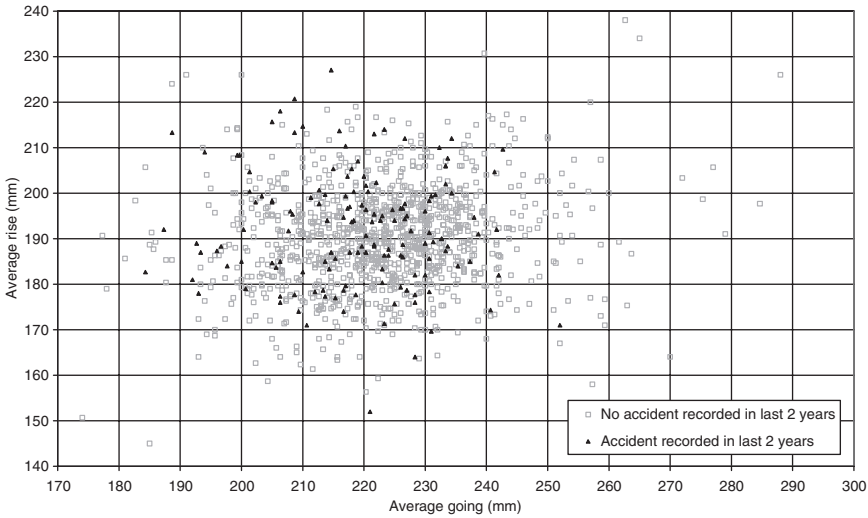


Figure 2. Scatter diagram of stairs in the sample, from English dwellings in 1991.

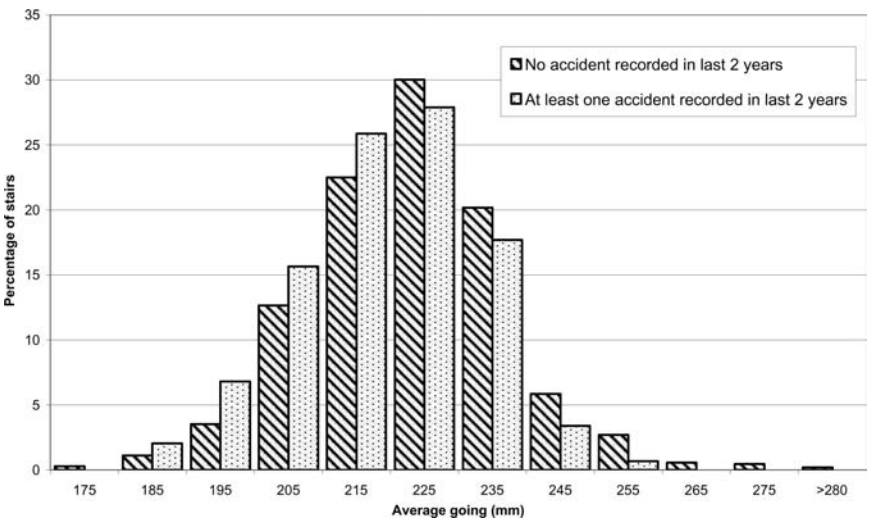


Figure 3. Percentage of stairs that have had at least one accident, or no accident, recorded as taking place on them in the last two years, as a function of average going.

A different way of looking at the data is to show what proportion of the stairs in a set range have had a recent accident recorded for them, this data is presented in Figure 4. Due to the extremes of the distribution having a small number of cases present, the graph is only drawn where the total number of stairs in any 10 mm bin is 20 or more.

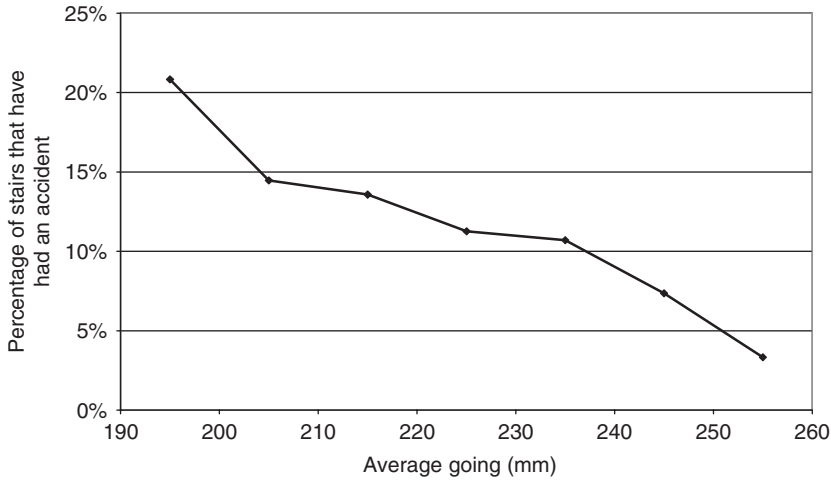


Figure 4. Percentage of stairs that have had at least one accident, recorded as taking place on them in the last two years, as a function of average going.

Discussion

It is clear from these results that having a larger going reduces the likelihood of an accident being reported. It is also suggested from Figure 4 that by increasing the going to at least 250 mm there will be a significantly reduced risk of an accident. It is to be recalled that these stairs are taken from the population of domestic stairs in England in 1991, and so do not fully measure the possible range of stair geometries that could be built. In particular, the number of stairs that were included in the sample with a going of more than 260 mm was 13 out of 1,223 (1.1%). Clearly this sample cannot provide useful data for stairs with a going greater than 260 mm.

These accidents are for domestic settings, and so the findings are likely to reflect typical domestic use, which means probably only one user on the stairs at a time, and widths narrower than stairs found in non-domestic settings. The suggestion that by increasing the minimum allowed going to 250 mm will significantly reduce the risk of an accident should not be applied to other situations where the stairs may be wider, or crowds may be using the stairs. It is also possible that if stairs are used in an emergency, for an evacuation, different ways of using the stair would be employed, so different going sizes may be needed to significantly reduce the risk of an accident.

Conclusion

If accidents on stairs were a random occurrence, with design having no effect, statistical tests would show no significant difference between a group of occupiers who reported a recent accident and those who reported no recent accident when

compared across different physical measures. Across the range of rise heights found in a random sample of English dwellings there was no influence on accident rate, nor was there any influence on the number of stairs in a flight on the likelihood of an accident. However, across the range of going sizes found in English dwellings, there is a very significant influence on accident rate, with larger stair goings reducing the likelihood of a stair accident being recorded.

References

DETR 1998, *The Building Regulations 1991, Approved Document K. Protection from falling, collision and impact*, (The Stationery Office)

Roys, M.S. 2001, Serious stair injuries can be prevented by improved stair design, *Applied Ergonomics*, **32**, 135–139

Roys, M. and Wright, M. 2003, *Proprietary nosings for non-domestic stairs*, BRE Information Paper, IP 15/03 ISBN 1 86081 6525

BEHAVIOURAL SAFETY – REDUCTION OF SLIPS AND TRIP ACCIDENTS IN A STATION ENVIRONMENT

Ian Rowe

Ove Arup & Partners, Arup Campus, Blythe Gate, Blythe Valley Park, Solihull, West Midlands

This case study paper contains details of a pilot programme run by Arup/GNER at York station to reduce the occurrence of Slips and Trips accidents to members of the public and Station staff. The two year programme used an holistic approach to the problem which takes into consideration how the people (users and staff) behave in the context of the station environment and how this behaviour could be effectively modified in order to reduce accidents. Socio-technical system and Trans-theoretical models were used to develop an intervention and working with GNER management and front line staff, consultants from Arup implemented a programme that has delivered a significant improvement in safety culture on the station and has resulted in a 55% reduction in accidents.

The approach taken is designed to be self sustaining and deliver continuous improvement in station safety.

Introduction

In recent years, UK railway safety has improved considerably due to enhanced technology and a focus on human factors. However, for train operators, more frequent hazards still occur at stations across the UK in the form of Slip and Trip accidents. These cost the industry millions of pounds every year and are now a number one priority in addressable safety issues for many train operating companies.

Unlike high profile incidents such as Signals Past at Danger (SPaDs), hundreds of personal injuries occur each year and are often unreported. As a result, collecting sufficient data is often difficult and analysis of accident statistics rarely allows for identification of root causes.

Slip and Trip accidents are statistically one of the highest safety risks for railway users in the UK. Whilst the severity of this type of accident is generally low, the frequency of occurrence is high. Normally, one or two actual fatalities occur every year on the railway directly due to Slips and Trips. These accidents affect staff and customers and are extremely difficult to tackle due to their apparent random nature. The tracking of accidents is also made more difficult by occurrences and near misses going unreported, largely due to personal embarrassment and societal attitude towards falling over. Reported Slip and Trip accidents therefore represent only ‘the tip of the iceberg’.

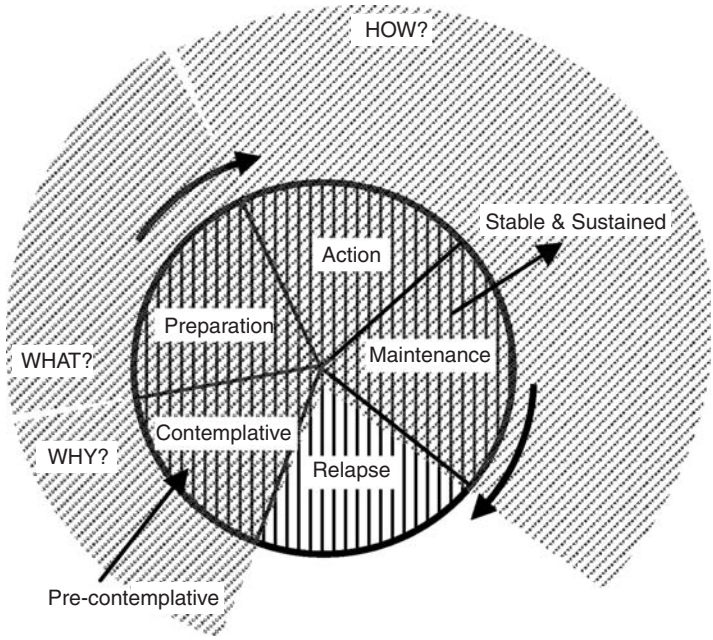


Figure 1.

According to the accident statistics registered on the centralised Safety Management Information System (SMIS), most of the reported Slip and Trip accidents involve elderly people, those with disabilities or the able bodied who are in some way encumbered. These users often sustain serious injuries and many require hospital treatment.

In March 2004, GNER, the train operating company responsible for 11 stations on the East Coast mainline, decided to review Slip and Trip accidents at its stations and propose a solution for improving passenger safety. The GNER safety team monitored the Slip and Trip accident statistics for several years and had noted that occurrences were increasing. Their analysis of the statistics showed that occurrences were highest in two of the major interchange stations; York and Newcastle. However, their analysis failed to establish the cause of these accidents, therefore making it difficult to prevent them. With trends of Slip and Trip accidents showing an increase, a diagnostic study was conducted and it was agreed that York station would be used as a pilot for a reduction programme.

Approach

The approach adopted for the intervention was based upon the Trans-theoretical (or state of change model) (Fig. 1) supported by Socio-Technical System Model (Fig. 2) and the intervention was based around acknowledged good practice in behavioural safety.

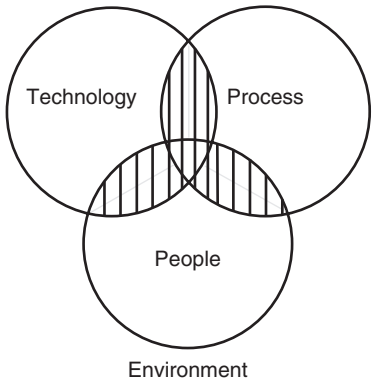


Figure 2.

Trans-theoretical model

This model is an adaptation of the model known as the 'Trans-theoretical' Model established by Prochaska et al, 1982 and originally used to provide a framework of understanding for the state of change within a health belief/behaviour context. This model has subsequently been reviewed for its use in other 'state of change' applications such as that of Human Factor interventions within organisations (J.H. Barrett et al, 2005).The model is useful in framing an understanding of the status of an organisation. It enables the design of interventions that progress an organisation through a programme to affect a reliable and sustainable STF reduction programme.

Understanding the model

The model shows a continuum that takes an organisation through a 'journey' from that of being pre-contemplative (they don't know there is an issue) to that of Maintenance (where they have acknowledged the issues, planned and designed interventions and have sustained to the point that they are maintained) and to a possibility of Relapse where there is a loss of focus on the issues and subsequent return to the contemplative state. The overlays on the inner circle show what type of messages and support is required for organisations at each of the states.

During pre-contemplative and through to contemplative states

The messages concerned with 'Why' are important at this stage. The primary requirement here is to educate and provide an understanding of why action should be taken, to demonstrate that there is an issue that requires attention and acknowledgement of the threats and opportunities relevant to the issue.

During preparation state

The messages pertaining to 'What' to do are important at this stage. In the case of STF interventions, the actions that can be taken can often include simple guidance

such as: 'Mop up spills', 'Remove trip hazards', 'Clear gangways' etc. These actions can be supported by simple tools such as checklists, audits, inspections etc.

During the action state

Once the 'What' to do message is understood the action state requires support in the form of 'How' to accomplish the required actions. This is far more complicated than the previous messages. It needs to address the education, knowledge and experience of stakeholders and also the organisational and behaviour issues around implementation of a programme.

During the maintenance state

The support of 'How' to continues to be the theme in this state. For a programme to be maintained the programme must contain an element of measure and progress, enforcement and re-enforcement of desired behaviours and a continued and demonstrated management commitment to the programme.

Socio-technical system model

This model takes into consideration the interaction of people, processes and technology within the specific environment and relies on a holistic approach to understand the dynamics associated with a particular problem. It is important to understand the relationship between all elements of the model when planning the intervention.

Understanding the situation at York station

York Station can be considered as a Socio-Technical System. It consists of People (passengers, staff, maintenance personnel, visitors, etc.), Processes (boarding/alighting, platform changes, safety procedures, emergency procedures, etc.) and Technology (information screens, signage, public address systems, lifts, etc.), which need to be considered. These are all within the station environment (temperature, humidity, lighting etc.). All of these can have some effect on the causation of Slip and Trip accidents, and normally, combinations of these are found to be responsible.

The intervention

An observational study was conducted to begin to understand the issues. This involved observing the day to day workings of the station, customer and staff behaviours and understanding stimuli and responses brought about by the station 'system'. The study identified a number of potentially contributory factors. These included Customer route choice, management of surfaces, maintenance of surfaces, station processes and staff attitudes towards Slip and Trip accidents.

The study also confirmed the significant likelihood of under reporting incidents. This under reporting was identified as a key issue when developing an approach to

address the problem. If data about unreported incidents could be collected, the resultant larger data set could be used to direct interventions and thus reduce accidents.

As with any behavioural safety intervention, front line staff were identified as key stakeholders and their engagement was considered critical.

The elements of the developed reduction programme are as follows:

- The development of a practical and pragmatic training programme for all station staff. This training programme explored the root causes of accidents and demonstrated the importance of staff vigilance and actions that could be taken. All station staff ranging from the station manager to cleaners were trained using this programme.
- Staff attitudes towards safety were measured using a simple customized questionnaire.
- A method of collecting more detailed data about the occurrences of near misses, hazards and dangerous behaviours was developed with the front line staff and implemented. This relied upon operational staff logging observed incidents and mapping these.
- Station managers were trained to analyse the reports and take appropriate actions.
- The staff appraisal system was also changed to focus station staff on the importance of their role with regard to safety and in particular actions pertaining to STFs.
- The behaviour of Senior management, union commitment was also an important contributor to this programme. Meetings and briefings were held for the safety executive group (board level managers) and the safety committee (union and senior management). During these meetings, progress reports and actions required by the stakeholders were communicated and agreed.

The implementation of the programme resulted in dramatic effects. The staff attitude questionnaire was administered at the start of the intervention and then again after nine months. This resulted in a significant improvement in staff attitudes towards Slips and Trips. The highlights being:

- A greater understanding of the issues associated with Slip and Trip accidents
- A greater ownership of the problem being taken by station staff
- The staff perception of senior management commitment increasing
- The proactive reporting of near misses etc. became habit for station staff and this information was then used to concentrate efforts in the areas with most accident potential.

Spillages and leaks on terrazzo floor tiles were identified as a focus area and the reporting and management of spillages were changed accordingly. This included:

- A redesign of the cordoning and warning process
- A complete review of cleaning methods and practices
- Introduction of new ‘Dri-mops’
- Alteration of signage to encourage alternative safer passenger route choice
- Improved response times for maintenance on identified hazards.

Challenges when addressing safety issues

Some of the actions identified required modifications and an upgrade of the station infrastructure. These require major financial investment and as such will take time to implement as they may need to be incorporated into the ongoing upgrade of stations by Network Rail.

The apparent lack of progress with these perceived major issues can be a de-motivator for station staff. However, the establishment of regular feedback via briefings, newsletters and informal talks are being used to minimise this potential down side and maintain enthusiasm.

Finally, in any change programme, there is normally some resistance. The main cause of this is usually due to the legacy left by previous change initiatives. In this case, there was some initial scepticism but this has been addressed by the persistence of the intervention and displayed commitment from GNER management. The project culture adopted was one of openness and honesty and this was applied at all levels including the feeding back of relevant weaknesses encountered in the senior management team.

Project outputs

The pilot Slip and Trip Reduction Programme at York was completed in August 2006 and has delivered significant and tangible results. These include:

- A 28% overall reduction in Slip and Trip accidents
- A 55% reduction of accidents with causes under direct control of station staff
- A significant number of logged reports by station staff of near misses, dangerous behaviours and hazards, enabling proactive prevention
- A significant shift in staff attitudes regarding Slip and Trip issues.

As well as these measurable indicators there are also less tangible, but vitally important outcomes. The station is visibly tidier and there is improved confidence amongst all stakeholders that there is a high degree of control. Staff are enthusiastic about their programme and are committed to continuing with the improvements in safety.

Implications

The study highlights the importance of acknowledging slips, trips and falls and their consequences amongst managers, station staff and customers. Interventions have the potential to prevent and reduce the number of reported and unreported accidents; therefore they warrant further attention and investment. Staff attitudes are also an important catalyst to preventing the number of accident occurrences.

References

Prochaska, J.O. and DiClemente, C.C., 1982 – Trans theoretical therapy: toward a more integrative model of change.

Barrett, J.H., Haslam, R.A., Lee, K.G. and Ellis, M.J., 2005 – Assessing attitudes and beliefs using the stage of change Paradigm-case study of health and safety appraisal within a manufacturing company.

KINEMATIC AND TRIBOLOGICAL ANALYSIS OF SLIP PHENOMENA DURING WALKING

Takeshi Yamaguchi, Shintaro Hatanaka & Kazuo Hokkirigawa

Graduate School of Engineering, Tohoku University,
Aoba-ku, Sendai, Miyagi 980-8579, Japan

In this investigation, the effect of step length, walking speed and subject on traction coefficient between a shoe sole and floor was clarified and desirable walking conditions for reducing the chance of slipping were also suggested. From the gait experiments, a walk with shorter step length and higher walking speed is effective to reduce the chance of slipping irrespective of subjects. Based on a kinematic analysis of COM and COP during walking, the peak value of traction coefficient at heel-strike period $|F_h/F_n|_h$ or that at toe-off period $|F_h/F_n|_t$ on dry walkway increases with an increase of $\tan|\theta_h|$ or $\tan|\theta_t|$. θ_h or θ_t is an angle between vertical line through COM and the line connecting COM and COP at the moment when traction coefficient takes peak value at heel-strike or toe-off period. These results indicate that a walk with smaller θ_h and θ_t is effective for prevention of slip during walking.

Introduction

Slip during walking will not occur theoretically as long as the traction coefficient between a shoe or foot sole and a walkway, also known as RCOF (Redfern *et al*., 2001), is smaller than the static friction coefficient. Thus, a walk with the smaller traction coefficient will be an effective gait to reduce the chance of slipping. It has been reported that gait and anthropometric characteristics may affect the traction coefficient based on the gait experiments and kinematic and kinetic gait analyses (e.g. Grieve, 1983, Lockhart *et al*., 2003, Yamaguchi and Hokkirigawa, 2008). However, a comprehensive understanding of their effects on the traction coefficient during walking has not been obtained. Therefore, an effective gait to reduce the chance of slipping has not been well understood.

The objective of this study was to clarify the effect of gait characteristics and anthropometric characteristics of pedestrians on slip, based on the gait experiments and kinematic analysis of the whole body center of mass (COM) and the center of pressure of a shoe sole (COP). The effective gait to reduce the chance of slipping was also suggested.

Methods

Four healthy male adults participated in this study. Age, stature and weight information is presented in table 1. Informed consent was obtained prior to participation. Figure 1 shows experimental set-up for gait experiments. Eight force plates (OR6-5, Advanced Mechanical Technology Inc.) were used to measure the three orthogonal components of the ground reaction force (F_x, F_y, F_n). The gait experiments were carried out under dry and glycerin-lubricated walkway conditions. The data acquisition system consisted of the force plates, infrared cameras, A/D converter and a personal computer was used. Commercially available men's shoes (0.27 m length) were used, and flat NBR (nitrile butadiene rubber, shore hardness HS70) sheet was attached to the shoe sole. Full body kinematics were recorded using a three dimensional motion capture system (vicon 612, Vicon Motion Systems Ltd.). Thirty-four infrared-reflective markers were attached to bilateral upper and lower extremities and torso as shown in figure 2. Each subject was tested with a step length of 0.55, 0.75 or 0.95 m, and a walking speed of 1.0, 1.4 or 1.9 m/s for each

Table 1. Subject information.

Subject No.	Age, years	Stature, m	Weight, kg
1	24	1.71	57
2	22	1.75	56
3	45	1.87	86
4	29	1.81	85

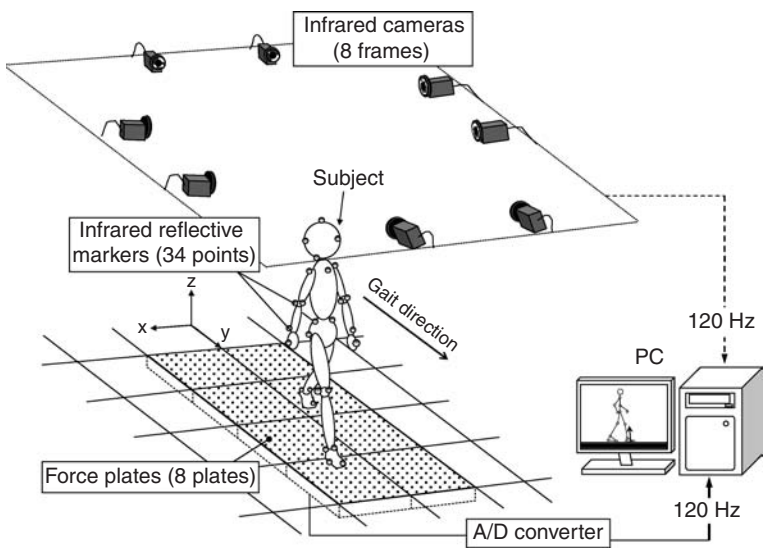


Figure 1. Experimental set-up.

step length. The whole body COM position was computed from the kinematic data using known Japanese segmental parameters (Ae *et al.*, 1992) in 14-segment configurations consisting the head, trunk, bilateral upper arm, bilateral forearm, bilateral hand, bilateral thigh, bilateral shank and bilateral foot.

Results and discussion

Figure 3 shows the normal (F_n) and horizontal components (F_h) of ground-reaction force (figure 3(a)), traction coefficient ($|F_h/F_n|$) (figure 3(b)) for subject No.3 walking on a dry walkway surfaces with step length of 0.75 m and walking speed of 1.4 m/s. It can be seen in figure 3(b) that there exists peak value of traction coefficient at heel-strike period $|F_h/F_n|_h$ or that at toe-off period $|F_h/F_n|_t$. In order to prevent slipping in stance phase, $|F_h/F_n|_h$ and $|F_h/F_n|_t$ should be smaller than the static friction coefficient between a shoe sole and floor. Therefore, smaller $|F_h/F_n|_h$ and $|F_h/F_n|_t$ will reduce the chance of slipping during walking.

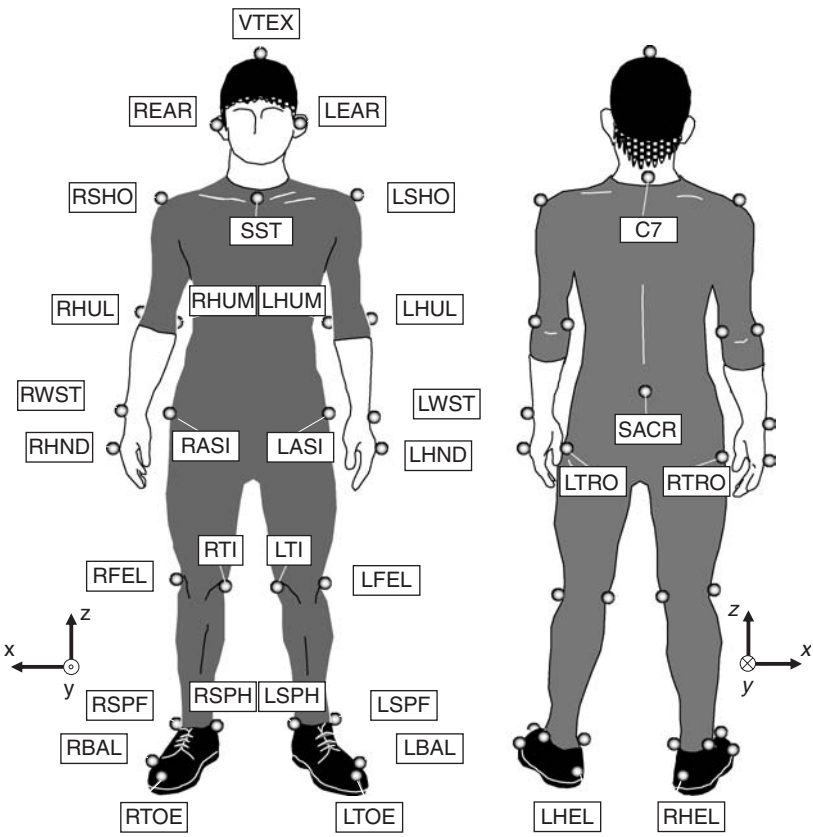
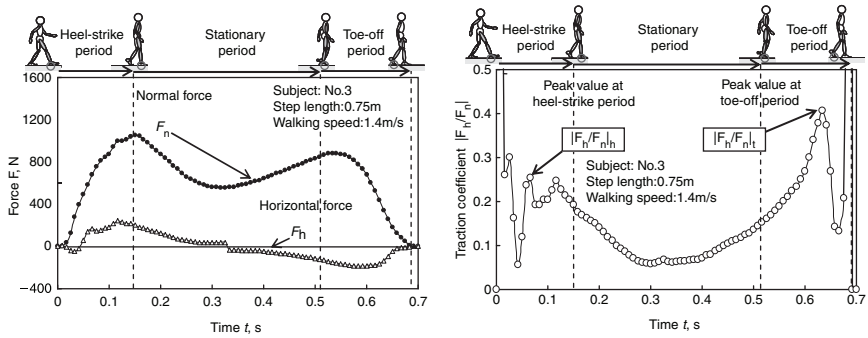
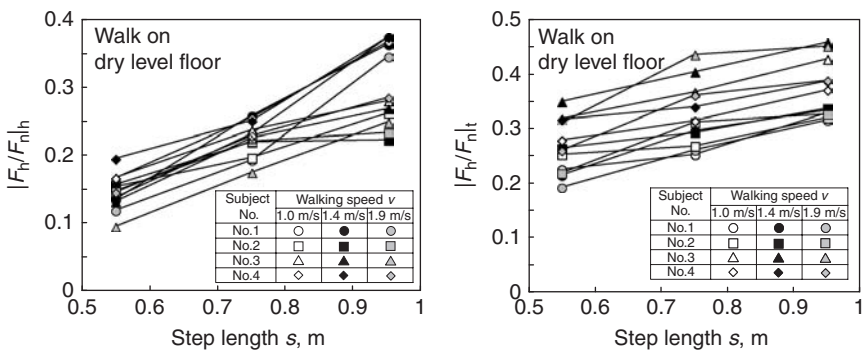


Figure 2. Marker placements.



(a) Ground reaction forces (b) Traction coefficient

Figure 3. Typical ground reaction forces and traction coefficient $|F_h/F_n|$.



(a) Heel-strike period (b) Toe-off period

Figure 4. The effect of step length on the peak values of traction coefficient.

Figure 4 shows the effect of step length on $|F_h/F_n|_h$ (figure 4(a)) and $|F_h/F_n|_t$ (figure 4(b)) for walking on dry walkway surfaces. It can be seen that both $|F_h/F_n|_h$ and $|F_h/F_n|_t$ increase with an increase of step length for all subjects. It can also be seen in figure 4 that $|F_h/F_n|_h$ and $|F_h/F_n|_t$ are subject dependent.

Figure 5 shows the effect of walking speed on $|F_h/F_n|_h$ (figure 5(a)) and $|F_h/F_n|_t$ (figure 5(b)) for walking on dry walkway surfaces. It can be found in figure 5(a) that $|F_h/F_n|_h$ slightly decreases with an increase of walking speed. On the other hand, $|F_h/F_n|_t$ slightly decreases with an increase of walking speed at step length of 0.55 m.

Figure 6 shows walking mode maps as a function of step lengths and static friction coefficients between a shoe sole and floor, which are constructed from the results of gait experiments on a dry walkway surface, based on slip/non-slip criteria (Yamaguchi and Hokkirigawa, 2008). These walking mode maps demonstrate three regimes of slip occurrence for each walking speed: no slip, slip may occur depending on subject and slip will occur. As shown in figure 6, the desirable static friction coefficient for walk without slip and the critical static friction coefficient for all

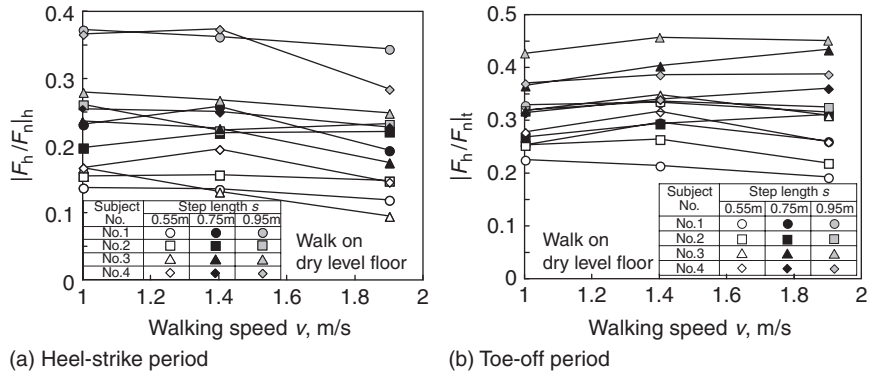


Figure 5. The effect of walking speed on the peak values of traction coefficient.

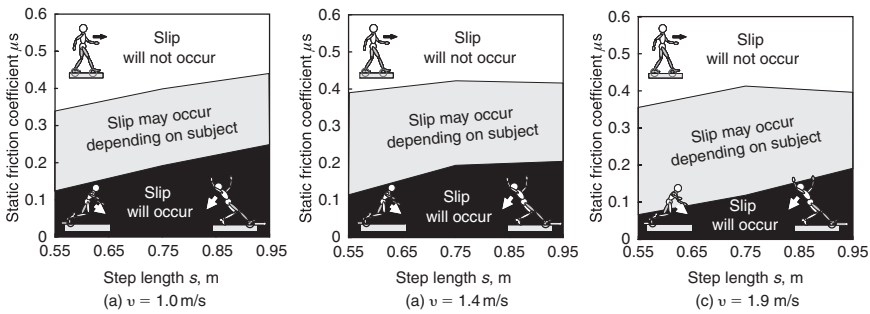


Figure 6. Walking mode maps in stance phase based on slip/non-slip criteria.

subjects slipping decrease with a decrease of step length for each walking speed. In addition, "slip" regime area gets smaller with higher walking speed. Thus, it can be expected that a walk with shorter step length and higher walking speed will be effective for reducing a chance of slipping. The effectiveness of such walking conditions on reduction of the chance of slipping was verified based on the gait experiments on glycerine-lubricated walkway surfaces.

The reason why smaller value of $|F_h/F_n|_h$ and $|F_h/F_n|_t$ can be obtained in such walking conditions is discussed from the viewpoint of the three dimensional geometric relationship between COM and COP which is modelled as shown in figure 7. In figure 7, $\tan|\theta|$ is expressed as following equation;

$$tan|\theta| = \left|\frac{l}{h}\right| = \frac{\sqrt{(x_{COP} - x_{COM})^2 + (y_{COP} - y_{COM})^2}}{z_{COM}} \tag{1}$$

Figure 8 shows $|F_h/F_n|$ and $\tan|\theta|$ in stance phase for subject No. 1 walking on a dry walkway surface at step length of 0.75 m and walking speed of 1.4 m/s. Figs. 9(a) and (b) show the relationship between $|F_h/F_n|_h$ or $|F_h/F_n|_t$ and $\tan|\theta_h|$ or $\tan|\theta_t|$, where θ_h or θ_t is the θ at the moment when $|F_h/F_n|$ takes peak value at heel-strike or toe-off period as shown in figure 8. As shown in figs. 9(a) and (b),

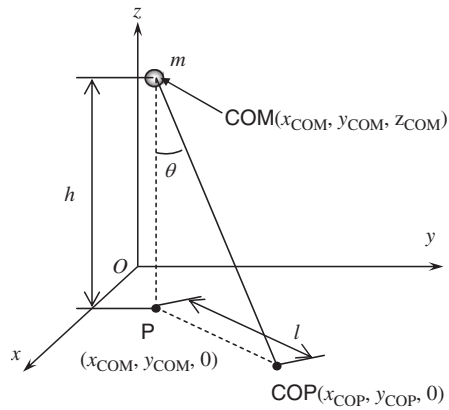


Figure 7. Model for kinematic analysis.

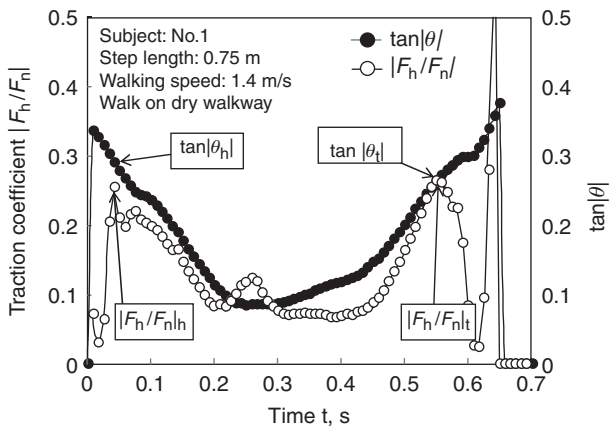


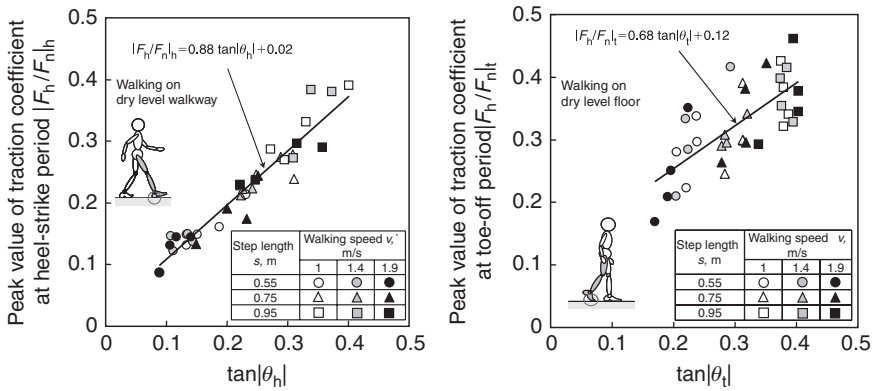
Figure 8. $|F_h/F_n|$ and $\tan|\theta|$ in stance phase.

$|F_h/F_n|_h$ or $|F_h/F_n|_t$ increases with an increase of $\tan|\theta_h|$ or $\tan|\theta_t|$. The regression curves of these relationships can be expressed as follows;

$$|F_h/F_n|_h = 0.88 \tan|\theta_h| + 0.02 \quad (2)$$

$$|F_h/F_n|_t = 0.68 \tan|\theta_t| + 0.12 \quad (3)$$

Thus, the peak value of traction coefficient at heel-strike or toe-off period can be reduced by reducing θ_h or θ_t irrespective of subjects. Furthermore, this finding indicates that the peak values of traction coefficients, i.e. the chance of slipping, can be controlled by the geometric relationship between COM and COP.



(a) Heel-strike period (b) Toe-off period

Figure 9. The relationship between $|F_h/F_n|_h$ or $|F_h/F_n|_t$ and $\tan|\theta_h|$ or $\tan|\theta_t|$.

Conclusions

(1) The peak values of traction coefficient at heel-strike and toe-off period increase with an increase of step length and with decrease of walking speed on dry walkway.
(2) The peak value of $|F_h/F_n|_h$ or $|F_h/F_n|_t$ increases with an increase of $\tan|\theta_h|$ or $\tan|\theta_t|$.
(3) Based on kinematic and tribological analyses, a walk with smaller θ_h and θ_t is effective for prevention of slipping during walking.

References

Ae, M., Tang, H-P., Yokoi, T. 1992, Estimation of inertia properties of the body segment in Japanese athlete, *Journal of the society of biomechanism* (in Japanese), **11**, 23–33

Grieve, D. W. 1983, Slipping due to manual exertion, *Ergonomics*, **26**, 61–72

Lockhart, T. E., Woldstad, J. C. and Smith, J. L. 2003, Effects of age-related gait changes on the biomechanics of slip and falls, *Ergonomics*, **46**, 12, 1136–1160

Redfern, M. S. *et al*. 2001, Biomechanics of slips, *Ergonomics*, **44**, 13, 1138–1166

Yamaguchi, T. and Hokkirigawa, K. 2008, 'Walking-Mode Maps' Based on Slip/Non- Slip Criteria, *Industrial Health*, **46**, 1 (in press)

IMPACT OF SPREADING FLOUR AND SALT ON FLOOR SLIPPERINESS

François Quirion & Patrice Poirier

QI Recherche et Développement Technologique Inc., Montréal, Québec, Canada, H1H 3R2

Spreading salt or flour on an area spoiled with food is often mentioned as an act of prevention against slips and falls in the workplace. We ran some laboratory tests in order to evaluate the impact of salt and flour particles on the slipperiness of different types of flooring (glazed ceramic, porcelain and slip resistant vinyl sheeting). In all cases, the presence of salt and flour particles reduces significantly the floor friction as determined with the Brungraber Mark II equipped with either a Neolite slider.

Introduction

While promoting floor cleaning as a preventive measure against slips and falls in the food industry, some restaurant workers and owners told us that they throw flour or salt on food spills or greasy areas in order to make these areas less slippery. This approach seems questionable because salt or flour will also be spread on uncontaminated areas and it is not certain that this will make flooring less slippery. As suggested by Kim (2007), the presence of small particles at the slider-flooring interface could act as a third body layer and reduce the friction. This investigation deals with the impact of these particles on floor slipperiness. Notice that we did not investigate the impact of adding salt or flour on the slipperiness of flooring initially contaminated with fats.

Methodology

Flooring and contaminants

Four types of flooring were tested in order to cover a range of surface roughness and slip resistance. Their characteristics are summarized in Table 1. A one cm^2 scan of the flooring is also shown in Figure 1.

The ceramic tiles are solid white. The porcelain tiles have a speckled surface but there is no particles standing out of the surface. The two Forbo vinyl sheeting have quartz and carborundum particles standing out of the surface. TS76 also has a textured surface with small 0.3 cm × 0.3 cm squares popping out of the surface. The texture made the measurement of the average roughness unreliable for the TS76

Table 1. Origin and average roughness of the flooring tested.

Name	Brand	Surface	Ra (μm)
Ceramic	Vitra	Smooth, Glazed	1.3 ± 0.1
Porcelain	Megagres	Homogeneous	2.8 ± 0.2
TS75	Forbo vinyl sheeting	Quartz and Carborundum	11 ± 3
TS76	As TS75	As TS75 +Texture	nd

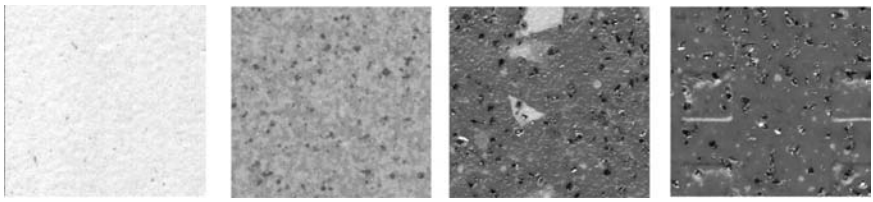


Figure 1. One centimetre square scans of the four flooring tested. From left to right: ceramic, porcelain, TS75 and TS76.

samples. All samples were 20 cm × 20 cm (the Forbo samples were cut from vinyl sheeting).

Salt, coarse salt and white flour were chosen as typical floor contaminants. The coarse salt particles have a mesh size between 18 and 40 while the table salt is in the range 40–60.

Brungraber Mark II friction

The friction is determined using a Brungraber Mark II equipped with either a Neolite or a Stainless steel slider. Both sliders are 4 cm × 6 cm. The Neolite is hand sanded with a 400 grit paper before each series of experiment while the stainless steel is sanded with an orbital sander and a 220 grit paper.

The initial position of the slider is different from the usual parallel position. The "heel" of the slider is tipped down so that it touches the test surface before the articulated strut mechanism is released. That position has two advantages: it simulates the angle of the heel with the floor at heel strike and it makes it much easier to reproduce the initial position of the slider.

The first measurement is that of the uncontaminated flooring. Then, a given amount of the contaminant is spread homogeneously on the flooring. Using the 4 cm × 6 cm sliders, it is possible to perform about 6 strikes on one sample as long as there is no slide. This method proved to be sufficient to identify the friction limit.

Once the friction has been determined, the contaminant is removed, the slider (Neolite or stainless steel) is cleaned and another measurement is made at another surface concentration of the contaminant. This procedure results in a friction vs. surface concentration data set. Data sets were obtained for the three contaminants,

Table 2. Uncontaminated friction and friction at 100 g/m^2 (150 g/m^2 for coarse salt) of contaminant between the Neolite slider and the flooring.

Neolite	Uncont.	Salt	Flour	Coarse salt
Ceramic	0.63 ± 0.05	0.18	0.21	0.28
Porcelain	0.57 ± 0.11	0.13	0.34	0.26
TS75	0.78 ± 0.09	0.17	0.43	0.30
TS76	0.92 ± 0.02	0.26	0.51	0.34

the four flooring and the two sliders. For coarse salt, the friction was measured only at one surface concentration.

Results and Discussion

This investigation deals with the impact of salt or flour on the slipperiness of four different flooring. These contaminants are likely to fall on a kitchen floor. They may also have been spread voluntarily on the floor. In that case, the worker usually wants to make an area contaminated with oil, grease or food less slippery. However, some of the salt or flour will also be spread on uncontaminated areas. How will that affect the floor slipperiness?

The friction data were determined as a function of the concentration of the contaminant on the flooring. Over the range of concentration investigated, 2 to 200 g/m^2, the friction does not change much. For comparison purposes, Table 2 and 3 reports the friction obtained at 100 g/m^2 (150 g/m^2 for the coarse salt) of contaminants. At that concentration, the tiles are almost completely covered with salt or flour.

Neolite friction

Table 2 compares the friction of the Neolite slider on the four flooring for the three contaminants with the friction of the uncontaminated flooring. The first observation is that the addition of salt, flour or coarse salt on the four flooring tested decreases significantly (by 1.7 to 4.7) their friction with Neolite. This strongly suggests that adding salt or flour to an uncontaminated area increases the risk of slips and falls.

Except for the glazed ceramic, the friction seems to increase with the roughness of the flooring. That is true for the three contaminants. So, the rougher flooring seem to offer a better slip resistance in the presence of contaminants such as salt, flour and coarse salt.

In most cases, the friction remains relatively constant with the surface concentration of the contaminant. Nevertheless, there are some trends that seem to repeat. First, as the amount of table salt increases, the friction between Neolite and the four flooring decreases before reaching a plateau value. On the other hand, when the contaminant is flour, the friction between the flooring and Neolite remains constant (Ceramic and Porcelain tiles), increases (TS75) or decreases (TS76). Notice

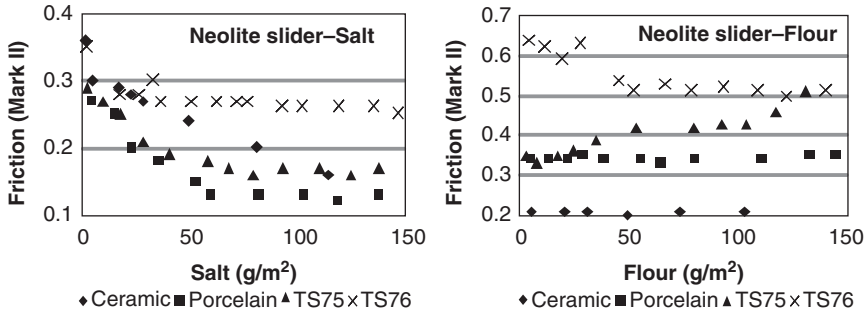


Figure 2. Evolution of the friction of the Neolite slider on the flooring covered with salt (left) or flour (right). Ceramic (♦), Porcelain (■), Forbo Tractionstep TS75 (▲), Forbo Tractionstep TS76 (X).

Table 3. Uncontaminated friction and friction at 100 g/m^2 (150 g/m^2 for coarse salt) of contaminant between the Stainless steel slider and the flooring.

Sta. steel	Uncont.	Salt	Flour	Coarse salt
Ceramic	0.32 ± 0.01	0.33	0.25	0.28
Porcelain	0.31 ± 0.02	0.36	0.29	0.34
TS75	0.39 ± 0.01	0.33	0.30	0.36
TS76	0.48 ± 0.03	0.35	0.37	0.32

however that the changes remain fairly small when compared to the difference between the uncontaminated and contaminated surfaces.

Stainless steel friction

The first question that comes to mind is "*Why measure the friction against a stainless steel slider if stainless steel is not representative of sole or heel material?*" When we first started to use stainless steel as a reference sole material (Quirion, 2004), it was because we were interested in the intrinsic contribution of the flooring to the friction. Moreover, stainless steel is easy to clean and it is not easily contaminated as most elastomers so that it can be used over and over with a good reproducibility.

But after measuring friction with stainless steel sliders under various conditions, there is another reason for using it as a heel or sole material: it seems that in many cases, stainless steel would indeed provide better friction than other elastomers. For instance, for the friction of stainless steel on oily surfaces (Quirion and Poirier, 2006).

Table 3 compares the friction of the stainless steel slider on the four flooring for the three contaminants with the friction of the uncontaminated flooring. This time, the friction is almost the same for the three contaminants (0.25 to 0.37) on the four flooring. These values are also close to the friction of the uncontaminated floorings suggesting that the friction between the stainless steel slider and the flooring is not much affected by the presence of salt, flour or coarse salt. The only significant loss

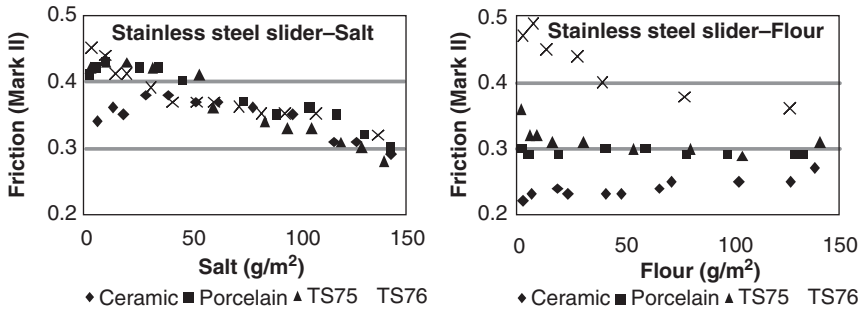


Figure 3. Evolution of the friction of the Stainless steel slider on the flooring covered with salt (left) or flour (right). Ceramic (♦), Porcelain (■), Forbo Tractionstep TS75 (▲), Forbo Tractionstep TS76 (X).

of friction occurs for the textured vinyl sheeting (TS76) and TS75 in the presence of salt and flour.

Figure 3 shows that for the hard flooring (ceramic and porcelain), the friction of the stainless steel slider first increases with the amount of salt and then decreases at high salt concentrations. For the softer and rougher vinyl sheeting, the friction decreases with the concentration of salt. The impact of the concentration of flour on the flooring differs from one flooring to another. This time, these effects are significant because there is not much difference between the frictions of the smooth stainless steel with contaminated or uncontaminated flooring.

The friction between the uncontaminated flooring and the stainless steel slider is always smaller than the friction with the Neolite slider with the same uncontaminated flooring. However, the situation changes in the presence of contaminants. For instance, the friction of stainless steel in the presence of salt or coarse salt is higher or similar to the friction of the Neolite slider under the same conditions. The only situations where the Neolite slider performed significantly better than the smooth stainless steel was for the two Forbo vinyl sheeting covered with flour.

This does not mean that shoe sole or heel should be made of stainless steel. But maybe the combination of stainless steel and rubbery materials could be beneficial in terms of slip resistance on contaminated surfaces. The approach is similar to that of Hokkirigawa and Yamaguchi (2007) who developed high grip rubber-ceramics composites as sole material.

Conclusion

The results obtained in the course of this investigation suggest that spreading salt, flour or coarse salt, intentionally or not, reduces significantly the friction between Neolite and four flooring ranging in hardness, roughness and texture. Interestingly, the impact is not as important for the friction of smooth stainless steel.

A possible explanation for the "success" of adding salt or flour to a slippery area is that it makes the area more visible and thus easier to go around. But it would be much better, and probably as fast, to remove the spill instead of looking for flour or salt to spread on it!

Acknowledgements

This investigation was supported by the *Institut de recherche Robert Sauvé en santé et sécurité du travail* (IRSST). We wish to thank *Forbo* for providing free samples of their vinyl sheeting and *Tarkett* for lending to us the Brungraber Mark II apparatus.

References

Kim, I.-J. (2007) *A new concept for the measurement of slip resistance properties*, In Proceedings of the International Conference on Slips, Trips and Falls-From Reasearch to Practice, August 23–24, Hopkinton, MA, IEA Press, pp. 168–172.

Hokkirigawa, K. and Yamaguchi, T. (2007) Development of high grip shoe soles by using rubber/RB ceramics composites, In Proceedings of the International Conference on Slips, Trips and Falls-From Reasearch to Practice, August 23–24, Hopkinton, MA, IEA Press, pp. 163–167.

Quirion, F., (2004) *Improving slip resistance with optimal floor cleaning: A preliminary field study*, In, Contemporary Ergonomics (2004), Ed. Paul T McCabe, pp. 48–52, CRC Press.

Quirion, F. and Poirier, P. (2006) *Correlation between floor slipperiness and surface saturation with oil*, In Proceedings IEA 2006 Congress, Maastricht, Netherland, Ed. R.N. Pikaar et al, Elsevier Ltd.

A PROCEDURE TO PREPARE FOULED AND WORN QUARRY TILES

F. Quirion, A. Massicotte, P. Poirier & N. Hadjev

QI recherché et Développement Technologique Inc., Montréal, Québec, Canada, H1H 3R2

Quarry tiles are common floorings found on commercial and institutional kitchen floors. When new, it presents a fair roughness and a relatively high porosity. With time, cooking oil and spilled fat become trapped within the pores and this fouling, combined to mechanical abrasion, leads to smooth and impermeable tiles that are quite slippery in the presence of a little water or oil. Hence, laboratory tests on new quarry tiles are not representative of the fouled and worn quarry tiles that are most likely to be found on site. We present a laboratory procedure to prepare fouled and worn quarry tiles by successive sanding, fat application and cooking of the tiles. The resulting fouled and worn tiles are quite similar to those encountered in the real life.

Introduction

Since 1997, we are interested in the optimisation of floor cleaning as a preventive measure to reduce slips and falls in the workplace and in particular in the restaurant and food industry. The optimisation of the cleaning efficiency was optimised for different types of flooring, including quarry tiles. These tiles have been very popular for the past fifteen years and they are quite common in restaurant and institutional kitchens. In the course of a field investigation (Quirion, 2004), it was noted that the physical characteristics of onsite quarry tiles were often quite different from the characteristics of the new tiles that were tested in the lab, to a point that onsite quarry tiles were often mistaken for smooth and glazed tiles while the new tiles are rather rough and matte.

It then became obvious that the physical characteristics of the quarry tiles change a lot with time and that these changes had a drastic impact on the original slip resistance of this type of flooring. So, in order to perform floor cleaning experiments that are relevant to the restaurant and food industry, it became necessary to test the same type of flooring as usually encountered in these premises (Quirion et al, 2007).

Underwood (1992) found that 90% of the dirt trapped in quarry tiles originated from fats and oils with the particularity that the fat reacts with air to polymerise into a hard film that is very difficult to remove. This effect combined to continuous abrasive wear due to traffic leads to smooth and glossy tiles. Underwood also presented a method to prepare a “model soiled tile” that combined sanding to simulate abrasive wear and immersion in a “dirt mixture” followed by cooking

to simulate fouling. That method was tested in our laboratory but we had difficulties in obtaining reproducible surfaces. For that reason, we developed our own procedure which is mostly inspired form the Underwood paper and presented in this article.

Methodology

This section describes briefly the experimental methods that were used to evaluate the friction, roughness, gloss (light reflectivity) and oil penetration of the quarry tiles.

Quarry tiles

Two types of quarry tiles were tested in the course of this investigation. They were originally 15 cm × 15 cm tiles, with a thickness of 7 mm. The first type is referred to as CIMONE quarry tiles while the second one is referred to as MONO quarry tiles. The CIMONE tiles are smoother ($Ra = 2.8 \pm 0.2\,\mu m$) than the MONO tiles ($Ra = 5.4 \pm 0.8\,\mu m$) and their porosity ($6.3 \pm 0.6\%$ v/v pores volume) is twice that of the MONO tiles ($2.9 \pm 0.1\%$ v/v pores volume).

Friction and roughness

In this investigation, all the friction measurements were performed on clean and dry tiles. At this point, it is important to note that fouling, even if it comes from fats, produces a hard solid film that is neither oily nor greasy. So the friction measured is measured on clean and dry surfaces.

Friction was determined by pulling a 126 g slider mounted on three hemispherical stainless steel caps at a speed around 2.5 cm/sec and measuring the pulling force with a force gauge (Shimpo, FG-1) as a function of time. A typical run is shown in Figure 1. The dynamic friction coefficient was obtained by averaging the friction measured while the slider was moving on the tiles. Each measurement is the average of 15 to 20 individual readings sampled at 10 Hz and the dynamic friction coefficients reported in this paper are the average of 4 to 6 individual measurements at different locations on a tile.

The roughness was evaluated using two different procedures. The first one follows from the measurement of the dynamic friction coefficient. It was observed that the standard deviation, $\boldsymbol{\sigma}_{\mu}$, of the mean dynamic friction coefficient of a single measurement seemed to be correlated with the roughness of the tile tested. Based on this observation, the roughness index, $\mathbf{R_I}$, was arbitrarily defined as the standard deviation of the individual friction measurements multiplied by one hundred, $\mathbf{R_I} = 100 \cdot \boldsymbol{\sigma}_{\mu}$. The second procedure uses the DekTak 3030 (at 5 mN and a scan length of 5 mm) to determine the average roughness, **Ra**. These results are always the average of at least five independent determinations on different locations of the tile tested.

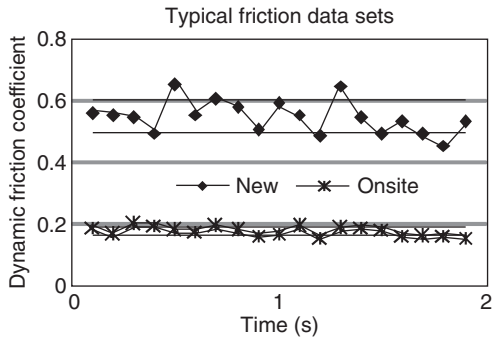


Figure 1. Typical data sets for the friction as a function of time on new and onsite quarry tiles. Horizontal lines indicate the $\pm\sigma$ of the mean value which is an indicator of the roughness.

Light reflectivity and oil coverage

Light reflectivity is a measurement of the intensity of the light reflected on a surface. Typically, it increases with the oil concentration at the surface, from the reflectivity without oil, $\mathbf{R_0}$, to a plateau value when the surface is saturated, $\mathbf{R_P}$. The reflectivity at any oil concentration, $\mathbf{R_C}$,

$$Coverage(\%) = 100 \cdot \frac{(R_C - R_0)}{(R_{100} - R_0)} \tag{1}$$

can be used to calculate the oil coverage on a surface (see equation 1).

The reflectivity, or the oil coverage, can be used to evaluate the rate of penetration of oil from the surface into a tile. To do so, a thin film of oil is spread over the tile and the light reflectivity is measured as a function of time. As the oil penetrates into the tile, the reflectivity decreases.

Results and discussion

Table 1 compares the physical characteristics (friction, roughness and reflectivity) of the quarry tiles treated in our laboratory with the characteristics of quarry tiles encountered during field investigations in twelve restaurants and four hospitals.

CIMONE and MONO quarry tiles both look like typical quarry tiles but it is impossible for us to determine if these tiles correspond exactly to the tiles tested during the field investigations. However, we feel that most quarry tiles will respond similarly to wear and fouling. The time it will take for the tiles to become fouled and worn may differ from one type to another but in the end, their physical characteristics should be similar.

New vs. onsite quarry tiles

The new CIMONE and MONO quarry tiles have the same characteristics within experimental uncertainty. When compared to the characteristics of the onsite quarry

Table 1. Physical characteristics of the different states of the quarry tiles tested.

State	R(%)	μ_K	R_I	Ra
ONSITE				
Fouled and worn	nd	0.188 ± 0.027	1.6 ± 0.4	nd
CIMONE quarry tiles				
New	46 ± 1	0.491 ± 0.019	6.1 ± 0.8	nd
Sanded grit 220 (-13 g/m^2)	45 ± 1	0.266 ± 0.023	2.4 ± 0.8	nd
After fouling and wear procedure	51 ± 2	0.206 ± 0.013	1.4 ± 0.2	nd
MONO Quarry tiles				
New	48 ± 2	0.467 ± 0.036	7.7 ± 0.7	5.8 ± 0.4
Cooked oil ($+34$ g/m^2)	42 ± 1	0.440 ± 0.035	6.7 ± 1	5.0 ± 0.5
After fouling and wear procedure	63 ± 4	0.165 ± 0.013	1.0 ± 0.3	1.1 ± 0.3

tiles, one finds that the aging process resulted in a drastic decrease of the dynamic friction ($\mu_{\mathbf{K}}$ decreases from ~0.48 to 0.19), and roughness index ($\mathbf{R_I}$ decreases from ~6.9 to 1.6). At the time of the field investigation, the reflectivity of the onsite tiles was not determined. However, their surface appeared to be smooth and shiny, as if it were glazed. The onsite tiles were also impermeable to water and oil, which is not the case for new quarry tiles into which water and oil penetrate quite easily.

One may wonder what happened to these tiles over the years. The answer resides in two phenomena: abrasive wear and fouling.

Impact of abrasive wear

The impact of abrasive wear is simulated by sanding the tiles. Table 1 reports the characteristics of the CIMONE tiles after 13 g/m^2 of the tile was removed through orbital sanding (grit 220). The abrasive wear by itself can account for a great part of the friction and roughness loss. However, it has little impact on the light reflectivity of the tiles. Results obtained with coarser grit size led to even smaller light reflectivity. Moreover, the sanded tiles are still permeable to water and oil. Thus, abrasive wear alone cannot account for all the characteristics of the aged quarry tiles (onsite).

Fouling: Oil penetration

As mentioned above, new quarry tiles are generally porous. For instance, the pore volume represents about 6% of the CIMONE quarry tiles and 3% of the MONO quarry tiles. When first installed, especially if they are not properly sealed, the high porosity absorbs water and oils as it falls on the floor. Since the contaminants are taken away from the surface, the floor does not become slippery, giving the impression that it is slip resistant. Unfortunately, we noted that once the fat was trapped into the pores of unsealed quarry tiles, it was almost impossible to remove even with aggressive floor cleaning so that with time, fats and oils accumulate in the tile, react with oxygen and eventually clog the pores leading to impermeable flooring.

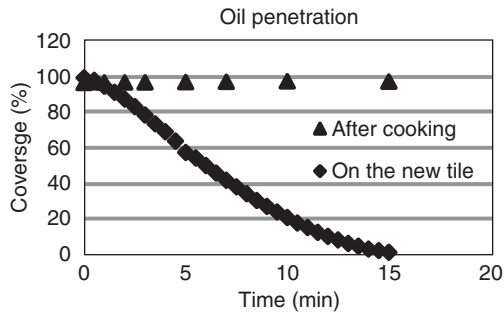


Figure 2. Penetration of vegetable oil at 17 g/m^2 on a new MONO quarry tile, ♦, and after cooking 38 g/m^2 of oil in the tile, ▲.

This is shown in Figure 2 where vegetable oil was applied on top of a new MONO tile at a surface concentration of 18 g/m^2. At the beginning, the oil is at the surface covering almost completely the tile. But within minutes, the oil penetrates within the pores. When oil is added on top of this tile, the penetration rates decreases and so on.

After 38 g/m^2 had penetrated in the pores, the tile was cooked in a convection oven to simulate accelerated fouling. As seen in Figure 2, the oil penetration rate on this fouled and cooked tile is now almost zero. Note that the pore volume of MONO tiles could accommodate about 200 g/m^2 of oil so that at 38 g/m^2, the tile is far from being saturated with oil.

Fouling is thus responsible for the permeability loss of the tiles. However, as seen in Table 1, it has little impact on the dynamic friction and roughness of the tiles. It even decreases the light reflectivity of the tiles because it makes them appear darker (wet look). So, in order to simulate both the permeability loss and the friction and roughness loss, the tiles must be submitted to both abrasive wear and fouling.

Fouling procedure

In the real life, wear and fouling occur progressively. In order to simulate that evolution, the fouling and wear procedure that we have developed is based on consecutive cycles of fat application, cooking and sanding of the tiles until the fouled and worn surfaces are similar to those observed onsite.

The two brands of quarry tiles were treated with two slightly different procedures but the steps are basically the same and they reported in Table 2. The tiles are first sanded with an orbital sander to remove about 13 g/m^2 of tile. Then, the fouling and wear cycles begins with the application of shortening (vegetable fat), cooking of the greasy tiles in a convection oven at 125°C for 17 to 20 hours, sanding of the cooked tiles and cleaning of the sanded surfaces with a degreaser. When the tiles are removed from the oven, the surface may be spotted with burned fat that did not penetrate completely. The sanding step removes these spots and contributes to increase to homogeneity of the surface. Ultimately, it increases the shine of the tiles as observed onsite.

Table 2. Summary of the fouling and wear procedures applied to new quarry tiles.

	CIMONE	MONO
Initial pore volume (% v/v)	6.3 ± 0.6	2.9 ± 0.1
Grit for the initial sanding	220	150
Type of fat used	Shortening	Shortening
Amount of fat applied per cycle (g/m^2)	From 35 to 15	From 44 to 18
Oven temperature (°C)	125°C	125°C
Cooking time (h)	17	20
Grit size for sanding the cooked surface	220	220
Surface cleaned with a degreaser	yes	yes
Number of cycles for fouled and worn surfaces	5–6	3–4
Average mass gained after the cycles (g/m^2)	87	68

The number of cycle required to get fouled and worn quarry tiles with a homogenous surface depends on the porosity of the tiles. As seen in Table 2, the more porous CIMONE tiles required up to 6 cycles while the MONO tiles were usually ready after three cycles. Assuming that sanding the cooked surfaces eliminates only cooked fat then the CIMONE tiles accumulate more fat (87 g/m^2) than the MONO tiles (68 g/m^2), in accordance with their higher initial porosity. These values are however much smaller than the maximum amount of fat that these tiles could accommodate (400 g/m^2 and 200 g/m^2 for the CIMONE and MONO tiles, respectively).

Conclusion

A simple method is presented to prepare fouled and worn quarry tiles that have friction and roughness similar to those typically found in restaurant and institutional kitchens. These tiles can be used a model tiles to investigate slip resistance and cleaning efficiency.

Acknowledgements

This investigation was supported by the *Institut de recherché Robert Sauvé en santé et en sécurité au travail* (IRSST).

References

Quirion, F., (2004), *Improving slip resistance with optimal floor cleaning: A preliminary field study*, In, Contemporary Ergonomics (2004), Ed. Paul T McCabe, pp. 48–52, CRC Press.

Quirion, F., Poirier, P. and Lehane, P. (2007), Optimal cleaning to prevent slippery floors in restaurants.

Journal of Environmental Health Research, 6(1), 25–33.

Underwood, D. C. (1992), Effect of floor soil on coefficient of friction in food service operations. *Ceram. Eng. Sci. Proc.* 13, 78–85.

PERCEPTION AND COGNITION DURING WALKING WHILE CONCURRENTLY USING A CELLULAR PHONE

M. Kuzel[1,3], M. Heller[2], R. Gray[3], S. Di Jorio[1] & S. Straughn[3]

[1]*Exponent, Inc., Phoenix, AZ 85027, USA*
[2]*Exponent, Inc., Suite 300, Philadelphia, PA 19104, USA*
[3]*Department of Applied Psychology, Arizona State University, Sutton 340J, Mesa, AZ 85212, USA*

This research investigated the effect of cellular phone (cell) use during walking on an individual's ability to visually perceive and gather detail about objects in their environment. Volunteers navigated a walking course in an office building where they encountered salient "out of place" objects (such as a cluster of balloons, various children's toys, cooking utensils, etc.). Each subject walked the course while exposed to three treatment conditions: no cell, cell with easy cognitive task (answering simple questions), and cell with hard cognitive task (verbally working through brain teasers). After each pass through the course, subjects completed a survey addressing what they visually perceived. The order of the conditions was counterbalanced and different objects were presented in each condition. The results indicate cell phone use significantly affects the ability to both detect objects and gather details about those objects.

Introduction

Cellular phones (cell) are now prolific in our society. As of mid-2007 there were more than 2.3 billion global cell subscribers, with an estimated 3.96 billion subscribers projected worldwide by the end of 2011 (Plunkett Research, Ltd.). This is equivalent to approximately 56% of the Earth's population having a cell. Thus, it is relatively unsurprising that seemingly everywhere one goes, commuters, construction workers at job sites, students on-campus, shoppers, and even joggers are seen (and heard) talking as they go about their daily activities.

The majority of research on the effects of cell use has been directed toward the driving environment (Horrey & Wickens, 2006). In general, these studies have sought to evaluate the effects of cell use on driving performance, assessing such variables as speed, reaction time, stopping time, stopping distance, following distance, increased driving errors and/or glances away. Within this environment, the research is even more limited with regard to the effect of cell use on object detection and recognition, and what has been done tends to address eye fixations, gaze directions and detection of signage.

To our knowledge, the effects of cell use while walking has not been investigated. The most closely related research was an observational study of pedestrian crossing behaviour at intersections (Hatfield and Murphy, 2007). It would seem likely that the distraction produced by cell phone conversations that has been found in driving studies may affect walking behavior and could be a contributing factor to roadway pedestrian and slip-trip accidents. If a person walking is distracted by using a cell phone they may be less likely to detect potential hazards in the walking environment (e.g., cracks, steps, cars, other people, etc.): the "happily connected" may actually be "haplessly connected."

The goal of the current study was to evaluate the effects of cell use on the ability to both visually detect objects and recognize details about those objects in a real world environment during walking. These results may be further generalized to assessing the ability to detect surface hazards while engaged in other cognitively distracting tasks, such as listening to headphones, or mentally going over a to-do list.

Methods

Subjects

Twenty-four subjects participated in the study. This included 16 women and 8 men, with an age range of 18–38. Each participant was compensated for involvement and participation was voluntary. Only subjects with very limited exposure to the test environment were recruited.

Setting and Materials

Testing was conducted at a private business, which offered a ground floor office area where subjects could walk a counter clockwise path around the approximately 67 m square hallway starting and ending at the same place. All testing was conducted on a weekend day while no other business activity was occurring in the building. The carpeted hallway was approximately 1.6 m wide, with offices lining both the exterior walls (to the right of the subjects) and interior walls. All office doors were closed and all photographs were removed from the walls to limit any extraneous visual information. Prior to testing, an Institutional Review Board evaluated and approved the proposed protocols.

Objects for the study were selected for their ability to standout in this environment (i.e., saliency). In total, 33 objects (11 per condition) were dispersed throughout the hallway in places considered to be most visually perceivable. Objects that were placed on the floor were approximately 18-inches away from the wall, such that a trip-slip hazard would not be created, but the object would still likely fall within foveal vision. As shown in Table 1, objects were classified as either being small (S) or large (L) and there were an equal number of objects on the right and left side of the halls. The placement and combination of objects was selected such that subjects would experience as dissimilar conditions as possible between the three object groupings (A, B, and C).

Table 1. Listing of objects placed in hallway.

Group A	Group B	Group C
Full size skeleton with blue bath robe and sunglasses (L)	Bicycle missing front tire and turned upside down (L)	Ladder w/ blue hardhat on top step & large yellow spider on hat brim (L)
$20 bill (S) 4-foot tall Halloween witch (L)	Stuffed teddy bear (S) Purple sheet on office door (L)	4-aces spiraled out (S) Truck tire on rim (L)
Stuffed turtle (S)	$20 in quarters on paper plate with large plastic scorpion (S)	Stuffed dog (S)
Shovel spade side up with cell phone opened over spade (L)	2 red and 1 pink helium balloons (L)	Cup & saucer overflowing with pennies (S)
Pink sheet on office door (L) 18″ frying pan with lid (S)	6-foot tall Leprechaun (L) Silver praying monk figurine holding a tooth brush (S)	Pink shirt on hanger (L) Princess piggy bank with hammer laid on it (S)
Baby doll (S)	Bottle of wine and 2 bottles of beer (S)	Cowboy boots – 1 black adult size, 1 white child size (S)
Knife and spoon crossed (S)	Soup ladle w/ salt & pepper shakers (S)	Crash test dummy on chair with child's tea set in lap (L)
Letter "A" printed in black on sheet of white paper (L) 3-foot King of Hearts card (L)	Letter "M" printed in black on sheet of white paper (L) Mallard duck decoys (S)	Letter "R" printed in black on sheet of white paper (L) Black mannequin head with white ball cap (L)

Procedure

After filling out the consent form, subjects were told: "*We are interested in how people's use of a cell phone changes while they are walking.*" Subjects were informed they would walk around the first floor of the building, complete a short survey, take a 20-minute break and watch reruns of a popular television show, and then complete another pass, repeating this procedure until they had walked the hallway in each of the three conditions. Subjects were informed that while talking on the cellular phone that they would be asked a series of questions, but that none of the answers would be recorded or graded.

Subjects were taken one at a time to the hallway where the experiment took place, fitted with a pedometer and read final instructions. The same script was read to each subject: "*Safe walking requires both controlled physical motion and mental attention. As you walk through the hallway, please pay attention to the environment around you. You will be asked questions about the objects you pass while walking.*" Subjects were also read one of three scripts specific to the condition they were about to perform, each of which provided the instruction to walk through the hallway at their normal pace, continuously, until returning to the starting area.

For each of the two cell conditions, they were then handed the phone, to which a female researcher on the line introduced herself, provided additional instructions about the conversation and then began asking questions. The script for the easy task (CE) described that they would be having a conversation similar to what one may have with a new acquaintance. The questions included: How old are you?, How tall are you?, What is your middle name?, etc. The script for the hard task (CH) described that they would be asked questions intended to cause them to have to think about the answer, and that if they needed to do so, it should be out loud. They were also told that if completely stumped they could say "pass". The questions for the difficult task included: Please recite the alphabet in reverse?, What is your age times 2?, How many words can you make from the letters a, e, k, l?, etc.

After completing the course, subjects were taken to an empty office where they were given forms related to the specific condition just completed. The first form (the questionnaire) asked them 15 questions about 10 of the 11 objects in the group they just viewed. The questions asked "Did you see a . . . ? (Yes / No)" and "If yes, was it/did it have a . . . . ?" In each case, there were five questions for objects that did not exist. For the ten objects they did see, questions either provided a multiple-choice answer set or required a one-word descriptive answer. None of the questions or answers was intended to confuse or mislead the subjects. In each of the two cell conditions, a second form asked subjects to write down as many questions as they could remember from the cell conversation.

A counterbalanced design was used, giving six orderings for the three conditions (NC – no cell, CE – cell easy, and CH – cell hard). Four subjects were randomly assigned to each ordering. The object groups were distributed among the treatment conditions such that each treatment was paired with each object group twice.

Results

The answers on the questionnaires from each treatment condition were scored as correct (+1) or incorrect (0) for both object detection (Did you see a . . . ?) and detail recognition (If, yes, was it/did it have . . . ?). The five "dummy" questions were scored for false detection (i.e., +1 incorrect "Yes" answer). The total number of objects detected, details recognized, false detections, time to complete pass, number of steps during pass and questions recalled was tallied for each subject. Table 2 provides summary statistics for these variables.

The complete data set was analyzed using a one-way within-subjects ANOVA. Pairwise comparisons of variable means were made with a Bonferroni correction for multiple comparisons. As shown in Table 3, the cell treatment condition produced highly significant effects on object detection, detail recognition, walking time, and number of steps. Pairwise comparison revealed that object detection and recognition was significantly more accurate in the NC condition than either CE or CH. In addition, walking time and number of steps was significantly shorter in the NC condition that in either the CE or CH conditions. It is also noteworthy that the number of false detections increased from the NC task to the cell tasks.

Table 2. Descriptive statistics for each treatment condition.

Treatment/Response	NC	CE	CH
Objects detected (Mean # (sd))	8.71 (1.49)	7.12 (2.56)	5.42 (1.99)
Details recognized (Mean # (sd))	7.79 (1.91)	4.79 (2.66)	3.12 (2.11)
False objects detected (Overall #)	3	12	8
Correct detection and detail (%)	89%	67%	58%
Time (Mean time in min:sec (sd))	1:16 (0:12)	1:22 (0:14)	1:31 (0:16)
Steps (Mean # taken (sd))	149.9 (22.54)	163.5 (25.68)	163.9 (20.68)
Questions recalled (Avg./Max/Min)		10.7/16/7	5.5/9/2

Table 3. Results of ANOVA Analysis and Pairwise Comparisons.

Treatment/ Response	Main effect F(2,22)	Pairwise comparisons (Mean difference)		
		NC–CE	NC–CH	CE–CH
Object detection	15.59 ($p < 0.001$)	1.58**	3.29**	1.71*
Detail recognitions	28.47 ($p < 0.001$)	3.00**	4.67**	1.67*
Time	16.48 ($p < 0.001$)	0:09.25**	0:16.19**	0:06.94
Steps	6.16 ($p < 0.001$)	15.12	15.82*	0.71

**significant at $p < 0.001$ level. *significant at $p < 0.05$.

Discussion

The data indicate that the use of a cell results in a significant negative change in the ability to detect and recognize objects that are salient within the environment. Furthermore, the magnitude of this deficit appears to increase as the demands of the cell phone conversation increase. The lack of detail recognition is particularly important because it may have a greater influence on the potential for the occurrence of a slip-trip fall event. For example, a person may very well be able to detect (i.e., "see") general information about the sidewalk, but if they fail to recognize the details of the expansion joints, for example, then the 1-inch heave in the slabs may still result in the trip-fall event.

The use of the cell also had a significant effect on time to complete the course and the number of steps taken, causing an increase in both. This finding is consistent with reports in driving studies that cell phone use leads to reduced driving speed (Horrey & Wickens, 2006) and suggests that subjects in our study may have been attempting to partially compensate for their impaired cognition in the cell phone conditions by slowing down. However, this did not appear to be an effective strategy: while the subjects took longer to complete the course, there were still significant differences in object detection and recognition. This would indicate that just because someone walks slower while on the phone, they are not necessarily walking safer.

The literature on cell use while driving often has used the term "inattentional blindness" to describe the cell user's inability to focus on the driving task while

attending to a conversation. To our knowledge, no previous study has sought to relate this potential effect to walking. The results of the current study show that the use of a cell may cause "inattentional blindness" while walking as well. This occurred even though the subjects received specific instructions (and reminders) about being tested on the out-of-place objects they saw in the hallway and were having an essentially meaningless conversation. Our study was unlike the typical real world in that it included these reminders, and not all potential trip-slip hazards are as salient as a stuffed teddy bear on an office hallway floor. Given these differences and the results observed, the cognitive load created by having a truly interesting conversation on a cell while walking may constitute a significant risk.

The results presented here indicate that the use of a cell results in a significant negative change in the ability to detect and recognize objects that are salient within the environment and the magnitude of this deficit appears to increase as the demands of the cell phone conversation increase. Although only cell use was studied, other electronic devices prevalent in society, such as portable music players, may also cause a similar cognitive load. Therefore, future work is planned to further explore the deficits created by these devices.

References

Hatfield, J. and Murphy, S. 2007, The effects of mobile phone use on pedestrian crossing behaviour at signalised and unsignalised intersections, *Accident Analysis and Prevention*, **39**, 197–205.

Horrey, W.J. and Wickens, C.D. 2006, Examining the impact of cell phone conversations on driving using meta-analytic techniques, *Human Factors* **48**, 196–205.

Plunkett Research, Ltd. 2007, Wireless, cellular and RFID industry trends. http://www.plunkettresearch.com.

EFFECT OF SHOE TYPE ON KINEMATICS OF STAIR NEGOTIATION IN WOMEN

M. Heller[1], S. Di Jorio[2], M. Kuzel[2], L. Ciccarelli[1] & M. Carhart[2]

[1]*Exponent, Inc., Suite 300, Philadelphia, PA 19104, USA*
[2]*Exponent, Inc., Phoenix, AZ 85027, USA*

The purpose of this study was to understand the effect of shoe type on adult female stair negotiation. Although previous literature has assessed related measures such as joint kinematics during stair ascent/descent and the effects of shoe type on fall risk for over-ground walking, none has addressed the effects of shoe type on stair negotiation. This research was conducted on a real set of stairs, and volunteers each wore multiple shoe types (sneaker, loafer, high heel, and flip-flop sandal) during the course of testing. Subjects were videotaped while ascending and descending a typical staircase, allowing for observation of several kinematic parameters, including foot placement, speed, and handrail use. The results of an observational analysis indicate that shoe type influences gait patterns during stair negotiation; women wearing high-heeled shoes demonstrate a more controlled gait pattern reflecting an apparent greater demand on foot placement and attention.

Introduction

Women in today's society typically wear a variety of shoe types. A casual glance around town will yield views of sneakers, flip-flop sandals, high heels, loafers, and many other shoe types on women's feet. Many women will admit noticeable differences in personal comfort and walking speed between shoe types. However, shoe selection is often based on fashion and not practicality. There is the potential that gait differences caused by shoe type may influence the likelihood of a slip-trip-fall event on stairs. Thus, it is important to understand women's gait in various shoe types including the kinematics of stair negotiation.

During the previous decade, an increased effort has been made among researchers to understand human kinematics on stairs, including studies by Mian *et al*. (2007), which assessed joint kinematics during stair negotiation, and by Christina and Cavanagh (2002), which examined the required coefficient of friction during stair decent. Though each of these studies explored human motion while negotiating stairs, neither addressed the effect of shoe type on gait kinematics. Studies that have addressed effects of shoe type have focused on fall risk (Koepsell *et al*, 2004; Tencer *et al*, 2004; Sherrington *et al*, 2003) or gait kinematics limited to level over-ground walking (Perry *et al*, 2007; Kerrigan *et al*, 1998), or a combination thereof (Gefen *et al*, 2002). Furthermore, these studies were conducted in a laboratory

environment instead of a real-world setting. To our knowledge, no studies have investigated stair negotiation kinematics on a real set of stairs with each subject wearing multiple shoe types.

The goal of the current study was to assess the effect of shoe type on gait patterns during stair negotiation in adult females. Data from the study will help us to better understand the effects of shoe type in this environment and provide insight as to whether shoe type may influence the likelihood that a women could experience a fall event on stairs.

Methods

Subjects

Sixteen subjects participated in the study. Subjects were required to be 18 years or older, and participation was limited to women who had no previous history of lower torso or back injuries, who were not knowingly pregnant at the time of the study, and who were feeling well and healthy on the day of the study. A summary of subject information is provided in Table 1. Each participant was compensated for her involvement and participation was voluntary. Subjects brought with them a pair of sneakers, high-heeled shoes, loafers, and flip-flop sandals that they wore regularly and were comfortable wearing.

Setting and materials

Testing was conducted at a private business, which offered a standard 10-step plus landing staircase on an exterior wall connecting the ground level to the second floor. The stairs were 103 cm with a riser height of 17.8 cm and tread depth of 30.2 cm. A handrail was available for use at the subject's discretion. Four video cameras were placed within the test environment: two looking down onto the staircase at different angles and two at different heights viewing the subject in the sagittal plane. All testing was conducted during daylight hours in dry weather. Prior to testing, an Institutional Review Board evaluated and approved the proposed protocols.

Procedure

After the volunteer signed the consent form, a researcher photographed the right shoe from each pair and the subject's right foot for subsequent measurements. During this time, the volunteer was fitted with circular, adhesive, cross-hair

Table 1. Subject data.

	Age [years]	Height [cm]	Weight [kg]	Shoe size [US]
Mean	30 ± 5	163 ± 7	66 ± 11	7.5
Min	21	152	51	5.0
Max	40	175	89	9.5

joint markers. The order in which shoe types were worn was randomized and pre-determined for the subjects. After donning their first pair of shoes, volunteers were led to the upper landing of the test environment and asked to go down the stairs at a comfortable pace, then follow the defined path on the ground level to a specified area, turn around and return up the stairs to the second floor landing, completing one trial. Each volunteer completed two trials for each shoe type. When finished with all trials, the volunteer filled out a survey regarding her comfort in various shoe types, her comfort walking on stairs in various shoe types, and frequency of wearing high-heels.

Video analysis was performed to allow observation of the parameters of interest during stair negotiation, including whether or not the volunteer chose to use the handrail, the occurrence of toe or heel overhang, and the speed of stair ascent and decent. These observations of stair negotiation were tabulated to evaluate differences associated with shoe type. Observations could not be made for every trial due to either an obscured view or a technical problem with the digitized video file, but efforts were made to make the data set as complete as possible.

Results and discussion

Although the volunteers performed the trials in four different types of shoes, the results of the loafer shoe type were not analyzed due to the wide variety of styles provided by the volunteers. The observational analysis described herein provides insight into the assessed parameters, without evaluating statistical significance. Table 2 provides a summary of several of the observed parameters in terms of percentage of occurrence.

One of the first noted and most obviously observable characteristic differences between the kinematics associated with each of the shoe types was a decrease in speed associated with wearing high heels. This was observed in both the ascent and descent of the stairs.

Table 2. Summary of observed occurrences of assessed parameters.

		Hangover of heel (ascend) or toe (descend)		First foot on step during first trial		First foot on step in second trial compared to first trial	
Shoe type	Ascend/ Descend	Yes (%)	No (%)	Right (%)	Left (%)	Different (%)	Same (%)
Sneaker	Ascend	93	7	64	36	57	43
	Descend	93	7	68	32	50	50
Flip-flop	Ascend	82	18	64	36	57	43
	Descend	86	14	57	43	57	43
High heel	Ascend	24	76	59	41	47	53
	Descend	41	59	52	48	73	27

The foot that the volunteer first used to step up or down was recorded for all trials. For all shoe types, the right foot was used first more often than the left during ascent. Similarly during descent, in all shoe types, the right foot was used first more often than the left, although the high heel condition was fairly evenly distributed, with only one more right foot first step observed. Interestingly, approximately half of the volunteers chose to begin their ascent with a different foot during their second trial for each shoe condition. This was also seen during descent, when approximately half of the volunteers chose to step down with a different foot during their second descent while wearing a given shoe. This phenomenon was even more prevalent during descent in the high heel shoe. Of the sixteen subjects, fifteen were right-handed. This would indicate that there did not seem to be a preference of footedness among the volunteers, and that there was no definitive correlation between handedness and footedness.

Observations were made regarding whether or not the volunteer's heel was hanging off the edge of the riser during ascent or their toe hanging off during descent. Although many of the volunteers did have the heel of their foot hang off the back edge of the tread while wearing flip-flops and sneakers, the women were more likely to keep their foot within the bounds of the tread while wearing high heels. The video analysis indicated that most of the women made an attempt to avoid allowing their heel to overhang the riser when wearing high heels, which allowed them to have a more even weight distribution by sharing the load between the ball and the heel of the foot. The women that allowed their heel to overhang the edge of the riser had adopted a gait pattern relying on balancing on the ball of the foot while ascending the stairs. Conversely, during descent the volunteers would allow the front of their foot to hang over the edge of the step in flip-flops and sneakers, but were less likely to have any part of their foot hanging over the edge while wearing high heels. One on the volunteers in her post-test survey specifically noted the personal need or requirement to place her heel and toe within the confines of the tread to be comfortable while walking in high heels on stairs. Based on the data analyzed to date, it is unclear if the differences between individuals' gait patterns are definitely related to comfort and experience in high heels or the kinematics of the specific individual.

No volunteers were observed using the handrail when ascending or descending the stairs in sneakers or flip-flops. However, subjects were more likely to use the handrail on either the right or left side while wearing high heels during both ascent and descent. Arm swing during ascent and descent was analyzed as small swing (barely braking plane of body), large swing (well beyond plane of body), or no swing (arm stayed within plane of body). The subjects were observed to be more likely to swing their arms during stair ascent than stair descent. Specifically, greater than 20 percent of the subjects had no arm swing during descent, while less than 10 percent had no arm swing during ascent. This may be indicative of either a greater comfort level during ascent than descent and/or due to a tendency toward energy conservation during stair climb.

An observation was made at the transition point between stairs and level ground. Specifically, the analysis involved comparison of the last step *on* level ground prior to the bottom riser during ascent and their first step *onto* level ground from

the bottom step during descent. A mark on the pavement equivalent to the tread depth was used as a measure of comparison. Volunteers were more likely to take a larger step immediately prior to stair ascent when wearing sneakers. That is, in the sneaker condition, there was a greater distance between the stance leg and the riser for the bottom step; while in the flip-flop and high heel conditions, they tended to position themselves closer to the riser before initiating the step-up. During descent, women were observed to be less likely to step beyond the threshold while wearing sneakers, almost as if not breaking the rhythm that they had created while coming down the stairs. In flip-flops and high heels the women were more likely to take a step further away from the last riser, indicating that they chose to alter their stride as they transitioned to walking on level ground.

The volunteers' survey responses indicated that they were least "comfortable" in high heel shoes (as reflected by the responses of 15 of the 16 volunteers, with the 16th representing flip-flops as least comfortable). Further, they indicated that in general, walking on stairs causes a decrease in their overall comfort level, with high heels being mentioned as a contributing factor to this discomfort. Based on their indicated frequency of wearing high heels, the lack of regular experience may be a factor in this perceived discomfort, however, the women may also be associating discomfort with the increased task demands that occur while negotiating stairs in high heels. The observational analysis is consistent with these responses, given the differences in behaviour and kinematics between the shoe types. In total, the observed changes women make when wearing high heels is highly consistent with the increased task demands on motor control and postural stability associated with wearing high heels. It is most likely that women perceive this increase in task demand as a decrease in personal comfort. This is further likely to cause a woman to alter her gait kinematics in an attempt to use more caution while wearing high heels on stairs.

Conclusions

This study presents preliminary observational analysis comparing female stair negotiation in different shoe types. The results illustrate that most women exert additional effort, caution, and attention when ascending and descending stairs in high heels, including utilizing the hand rail, keeping the foot within the confines of the stair tread, and walking more slowly. These observations reflect a greater demand on task kinematics during stair negotiation while wearing high-heeled shoes, resulting in behavioural and kinematic adaptations. Our survey responses indicate that women are generally less comfortable in high heels versus other shoes types. Thus, women appear to demonstrate and mentally note that gait in high heels on stairs is more challenging from a motor control and postural control standpoint, resulting in a diminished margin of error and an increase in the required level of attention to the task. In a case where attention is diminished, either due to becoming too comfortable or complacent about the task, the risk of fall increases. Future work will attempt to establish which differences are statistically significant and most probable to be a factor in a fall event.

References

Christina, K.A. and Cavanagh, P.R. 2002, Ground reaction forces and frictional demands during stair descent: effects of age and illumination, *Gait and Posture* **15**,153–158.

Gefen, A., Megido-Ravid, M., Itzchak, Y., and Arcan, M. 2002, Analysis of muscular fatigue and foot stability during high-heeled gait, *Gait and Posture*, **15**, 56–63.

Kerrigan, D.C., Todd, M.K., and O'Riley, P. 1998, Knee osteoarthritis and high-heeled shoes, *The Lancet*, **351,** 1399–1401.

Koepsell, T.D., Wolf, M.E., Buchner, D.M., Kukull, W.A., LaCroix, A.Z., Tencer, A.F., Frankenfeld, C.L., Tautvydas, M., and Larson, E.B. 2004, Footwear Style and Risk of Falls in Older Adults, *Journal of the American Geriatrics Society*, **52**, 1495–1501.

Mian, O.S., Thom, J.D., Narici, M.V., and Baltzopoulos, V. 2007, Kinematics of stair descent in young and older adults and the impact of exercise training, *Gait & Posture*, **25**, 9–17.

Perry, S.D., Radtke, A., and Goodwin, C.R. 2007, Influence of footwear mid-sole material hardness on dynamic balance control during unexpected gait termination, *Gait & Posture*, **25**, 94–98.

Sherrington, C. and Menz, H.B. 2003, An evaluation of footwear worn at the time of fall-related hip fracture, *Age and Ageing*, 32: 310–314.

Tencer, A.F., Koepsell, T.D., Wolf, M.E., Frankenfeld, C.L., Buchner, D.M., Kukull, W.A., LaCroix, A.Z., Larson, E.B., and Tautvydas, M. 2004, Biomechanical Properties of Shoes and Risk of Falls in Older Adults, *Journal of the American Geriatrics Society*, **52**, 1840–1846.

TO SLIP OR NOT TO SLIP: A COMPARISON OF MATCHED SLIP AND NON-SLIP TRIALS UNDER SIMILAR CONDITIONS

Raymond W. McGorry, Chien-Chi Chang & Angela DiDomenico

Liberty Mutual Research Institute for Safety,
Hopkinton, MA, USA

This paper evaluates the kinematics and kinetics during early stance phase for matched trials of slips and non-slips that occurred under the same experimental conditions. Data were collected as part of a larger study where participants walked over several floor surfaces presented with and without contaminant, during a protocol designed to elicit a wide range of small slip distances. A motion analysis system tracked displacement of two heel-mounted markers and a forceplate provided ground reaction force measurements. Of the 867 trials collected, 47 instances occurred where two replications were performed by the same participant under the same conditions yet produced disparate outcomes, non-slips and slips. Analyses were performed to compare and contrast the kinematic and kinetic features that were associated with the different outcomes produced during the matched trials. ANOVA found no significant differences between the non-slip and slip trial for heel velocity, foot angle or utilized coefficient of friction (UCOF) at or about heel strike. However, significant differences in heel velocity, and UCOF were observed at greater than 25 ms post-heel strike and around the time of foot flat.

Introduction

Falls resulting from slips are a major source of injury in the home and workplace. One of the scientific approaches to preventing slips and falls has been to develop an understanding of what constitutes a slippery condition. Gait studies have typically reported a required COF ranging from 0.17 to 0.20 for normal walking on level terrain (Redfern et al, 2001). Typically, when the available COF provided by the shoe-floor interface is less than the required COF, the probability of a slip increases (Cham and Redfern, 2002; Hanson et al, 1999).

As part of a study of an approach to detect small post-heel strike heel movements, a laboratory protocol was designed to create “marginally slippery” conditions, where small slips or “micro-slips” were likely to occur (McGorry et al, 2007). Full slips also occasionally occurred during the experiment. Upon review of the 867 trials collected during the larger study, 47 instances occurred where two replications performed by the same participant walking under the same conditions of floor, contaminant and walking speed produced disparate outcomes, in one case

resulting in a slip and no slip in another. This observation led us to examine how similarities and differences in the measures collected during early stance phase might explain difference in outcome (slip or non-slip) in matched trials.

Method

Participants

Participants were recruited, gave written informed consent and participated in a study conducted in our biomechanics laboratory. Data for each participant were examined and trials where full slips occurred were extracted. For each case the replicate trial of that experimental condition was examined, and if no slip occurred in the replicate, that trial was also extracted, and the two trials were treated as a matched pair. Details of the classification criteria are given below. Matched pairs were found for 47 trials from 22 participants. The means (standard deviations) of ages, heights and weights of the 11 male and 11 female participants were 43.1 (12.2) years, 169.8 (10.4) cm, and 78.6 (18.6) kg, respectively.

Experimental protocol

Participants wore shoes fitted with a hard plastic (Delrin™) heel when walking over randomly presented surfaces of Delrin™-dry and Teflon™-dry and Teflon™-with contaminant. The contaminant was an aerosol furniture polish (thickness < 0.5 mm). Available COF measurements of the floor surfaces and the contaminants, made with a Brungraber Mark II slipmeter, ranged from 0.12 to 0.21. Participants walked an 8 m walkway, while connected to an overhead fall arrest system at walking paces of 1.5, 1.8 and 2.1 m/sec. All test surfaces and the surrounding areas were white and were lit indirectly. To minimize visual cues as to floor condition participants were instructed to look straight ahead at a distant point. Experimental conditions were presented randomly. Two reflective markers placed on the heel of the right shoe (Chambers et al, 2003) were tracked by an infrared motion tracking system (Motion Analysis Corp.) for a two-second period beginning just prior to heel strike. Marker displacement was sampled at 400 Hz. A piezoelectric forceplate (Kistler # 9281CA) mounted near the end of an 8 m long walkway was used to record the ground reaction forces during stance phase, and were sampled at 1200 Hz. See McGorry et al, (2007) for further details on protocol.

Data processing and analysis

The forward-directed shear and vertical (normal) ground reaction forces were processed using a fourth order, zero lag Butterworth filter, with a 4 Hz cut-off frequency. A three-fold interpolation was performed on the heel displacement data to synchronize it with the ground reaction forces. Heel strike (HS) was identified and the subsequent analyses were limited to the 200 ms stance phase period following HS. The start of foot flat (FF) was also identified. Forward slip distance (FSD)

Table 1. Means and ANOVAs for non-slip and slip groups.

Variable	Units	Non-slip Mean (s.d)	Slip Mean (s.d)	p
Ins. heel velocity at HS	m/s	1.00 (0.51)	1.10 (0.74)	.47
Mean velocity before HS	m/s	1.69 (0.64)	1.77 (0.82)	.63
Foot angle at HS	deg	25.1 (4.48)	25.3 (5.43)	.85
Foot angular velocity	deg/s	342 (54.3)	331 (74.1)	.46
Time to foot flat	ms	74.3 (12.7)	78.1 (16.4)	.22

was defined as the horizontal distance travelled by the heel from heel strike until the heel stopped moving (or for some slips, until the end of the 200 ms window). Foot angle and instantaneous horizontal heel velocity were calculated at HS. Mean horizontal velocity was calculated for the 10 ms period immediately antecedent to HS. The time period from HS to FF was determined, and the mean angular velocity of the foot over that period was calculated. The UCOF, or horizontal shear force/normal force, was calculated for each data point during the post-heel strike period, with sign denoting directionality (forward, "+", and rearward, "−").

The FSD, UCOF, and instantaneous heel velocity (HV) for each group were examined at 5 ms increments for the 30 ms period following heel strike, and at 10 ms increments from 60 ms interval around the time of foot flat.

For purposes of analysis it was necessary to categorize trials by slip distance. A Slip was defined as horizontal displacement of the heel greater than 100 mm, and a Non-slip was defined as displacement less than 30 mm. These criteria were developed from review of the literature, details presented elsewhere (McGorry et al, 2007). Single factor analyses of variance (ANOVA) were conducted to test differences in group means. The criteria for statistical significance was set at $p < 0.05$. Slip distance, the factor used to dichotomize the data set, was significantly different between the two groups [9.38 (8.80) mm – Non-slip group, 359 (154) mm – Slip group, $p < 0.0001$].

Results

Table 1 presents the means (standard deviations) and ANOVA results for the Non-slip and Slip groups. No significant differences were found for the heel velocities at or prior to heel strike, or for the variables related to foot angular displacement.

Bar graphs representing each variable are presented in Figures 1a, b and c, respectively. Significant differences between groups are indicated by an asterisk.

Discussion

Respective of the fact that FSD was the variable used to *a priori* define the two groups, the magnitude of the between-group differences in slip distances was quite

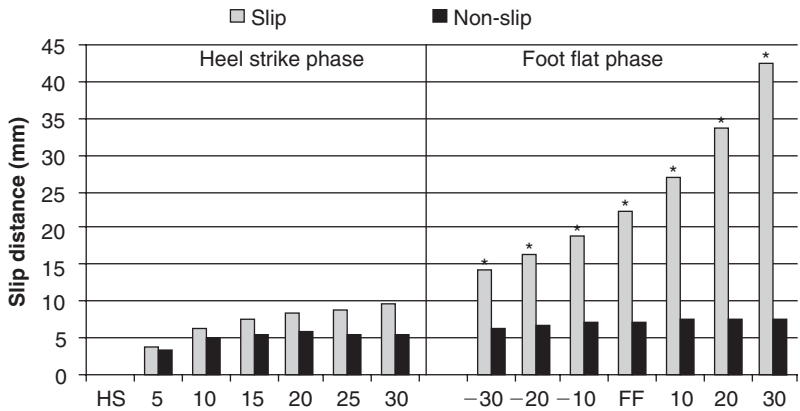


Figure 1a. Slip distance.

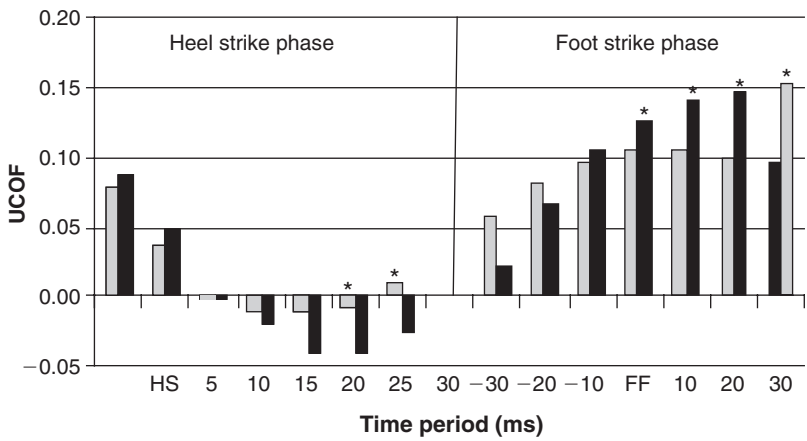


Figure 1b. UCOF.

large, with a mean for the Slip group almost 40 times that of the Non-slip group. Despite this large difference in outcomes, the differences between groups in the "initial conditions" were non-significant. Clearly, differences in the kinematics and kinetics of heel strike were not responsible for in the markedly different outcomes produced by participants under matching experimental conditions that were marginally slippery. In reviewing figure 1a, we see that slip distances are not significantly different, and remain relatively small throughout the 30 ms period following heel strike phase. However, the differences were significant for each instance during the 30 ms periods before and after foot flat, and the rate of increase was greater for the Slip group as the slip developed, but was relatively constant throughout foot flat for the Non-slip group.

Figure 1b shows the UCOF created by participants' ground reaction forces. Here both groups show a large downward trend immediately after heel strike, but the

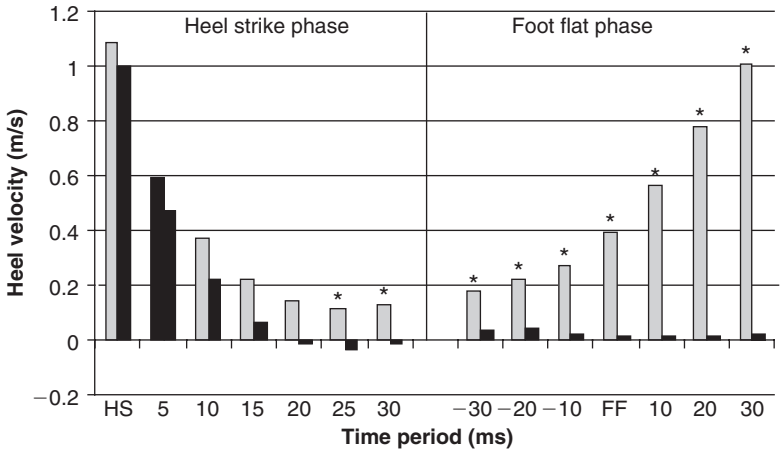


Figure 1c. Heel velocity.

groups differ significantly after 25 ms. At that time, the Non-slip trials have a significantly more negative (rearward directed) UCOF. This trend reverses as weight acceptance increases during the foot flat phase, and between group differences are significant from foot flat onward. During this period the Non-slip trials are able to develop an ever-increasing forward-directed COF, since static COF prevails once the heel stops moving. In contrast, during Slip trials UCOF initially rises and then flattens out around the time of foot flat. This occurs because the heel accelerates forward (see the rate change of slippage in figure 1a and the related increase in heel velocity in the foot flat phase of figure 1c) and thus only the smaller magnitude dynamic available COF is available to develop shear force to decelerate the heel. Figure 1c also shows that during Slip trials the heel never completely stops following heel strike, as it does in Non-slip trials where the heel has the advantage of a static state and thus the larger static available COF. It is also of interest to note that in the Non-slip trials the mean UCOF is actually directed rearward by 20 ms following heel strike (see figure 1b). It may be that the ability to slow (stop) the heel during this period is critical, allowing the opportunity for rearward directed shear forces to develop in these marginally slippery states.

The results suggest that under conditions where the available COF from shoe-floor interface closely approaches that of UCOF typical of normal gait, slip propagation may not depend on conditions just prior to or at the time of heel strike. Rather, differences in early stance phase, 25 ms or more following the heel strike event may be more critical under such conditions. During the Slip trials, the UCOF developed during that post-heel strike period was insufficient to bring heel movement to halt, allowing for marked acceleration in the slipping foot during the subsequent foot flat phase. This study investigated gait under surface conditions that were very near the typical range of UCOF. It is possible, if not likely, that the results would differ under much lower friction conditions, such as when a thick contaminant pool is present. Further analysis is needed to explore differences in

the slip propagation patterns relative to the presence or absence of a contaminant, or the effect of walking speed.

References

Cham, R., Redfern, M.S. 2002, Heel contact dynamics during slip events on level and inclined surfaces, *Safety Science*, **40**, 559–576.

Chambers, A.J., Margerum, S., Redfern, M.S., Cham, R. 2003, Kinematics of the foot during slips, *Occupational Ergonomics*, **3**, 225–234.

Hanson, J.P., Redfern, M.S., Mazumdar, M. 1999, Predicting slips and falls considering required and available friction, *Ergonomics*, **42**, 1619–1633.

McGorry, R.W., DiDomenico, A., Chang, C-C. 2007. The use of a heel-mounted accelerometer as an adjunct measure of slip distance, *Applied Ergonomics*, **38**, 369–76.

Redfern, M.S., Cham, R., Gielo-Perczak, K., Gronqvist, R., Hirvonen, M., Lanshammar, H., Marpet, M., Pai, C.Y., Powers, C. 2001, Biomechanics of slips, *Ergonomics*, **44**, 1138–1166.

THE MINIMUM COST OF NON-FATAL FALL ACCIDENTS TO THE NHS IN ENGLAND

Mike Roys & Michael Wright

BRE, Garston, Watford, WD25 9XX, UK

Reliable data on home injuries are difficult to find since the injury surveillance system was disbanded in 2002. However, some useful data can be obtained from Hospital Episode Statistics. There were 15 million episodes recorded in 2005–06, 1 million of which were due to external causes. Of this 1 million, 37% are related to falls. The Hospital Episode Statistics also record the number of bed days taken up by admissions. For admissions related to external causes 49% relate to falls. On average a person involved in a falls related episode is 63 years old, and will be admitted for between nine and ten days. Other admissions resulting in fractures may also relate to falls. The HES data can therefore be used to estimate the total number of bed days relating to fall injuries that are treated within the NHS in any one year. Estimates fall between 2.8 million and 4 million bed days, with a total cost to the NHS in excess of £27 Billion.

Introduction

For many years the UK had an injury surveillance system to provide an indication of the number of injuries occurring in homes and leisure locations. The last year of records was taken in 2002 (DTI, 2003). However, from these older data we can estimate that there are over 1.2 million falls in homes every year serious enough for the victim to visit A&E. There were 400,000+ falls on the level, 300,000+ falls on stairs, 35,000 falls from ladders or stepladders, 11,000 falls from buildings and a further 475,000 other falls, where sufficient details were not provided. Falls equated to 46% of all estimated accidents in the home. There were a further 1.4 million falls estimated to take place at leisure locations, 50% of all estimated leisure accidents.

Today we must rely on alternative data sources, the best of which is the Hospital Episode Statistics (HES). While this database does not provide any information about the activities that lead to the injury, the location or the types of injuries, they still provide an important indication of the cost to the NHS and hence the seriousness of falls. Permission from *The Information Centre*[1] has been obtained to illustrate the extent of fall related hospital episodes during 2005–06.

[1] "Copyright © 2007, Re-used with the permission of The Information Centre. All rights reserved." Allocated Permission Reference Number: 06090107.

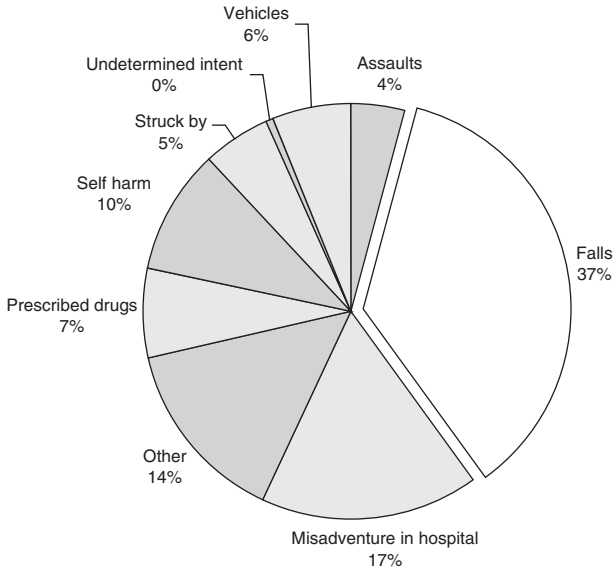


Figure 1. Distribution of external cause episodes.

Hospital episode statistics fall data

Every time an individual is treated within the NHS an Episode is recorded within the HES records. These records include information about the presenting individual's age, what cause has been allocated to their episode, how long they were treated for, and additional information about the location of the consultant supplying the treatment. This database forms the backbone of information on treatment requirements and costs for Primary Care Trusts across the country, but can also be used to give an indication of the extent of problems within a particular field, for example falls.

In 2005–06 there were over 15 million hospital episodes recorded. Of these one million were recorded as due to an external cause using the standard WHO ICD-10 classification system (WHO, 1992). Figure 1 shows how these external causes are distributed amongst the main categories of harm. Clearly the largest proportion (37%) of external cause episodes are falls related.

The, majority (90%) of the 373,000 fall episodes resulted in admissions, which is the norm for all external causes. However, the average fall incident results in a longer stay in hospital, such that falls nearly make up 50% of the 5.7 million bed days associated with external causes, see Figure 2. On average a person involved in a falls related episode is 63 years old, and will be admitted for between nine and ten days.

Five of the top ten external causes, as rated by number of bed days, are fall related, see Table 1.

These top five fall related ICD-10 codes represent 45% of the total bed days for external causes, with the remain 15 codes making up the other 4% recorded in Figure 2. It is also worth noting the large number of episodes recorded under X59,

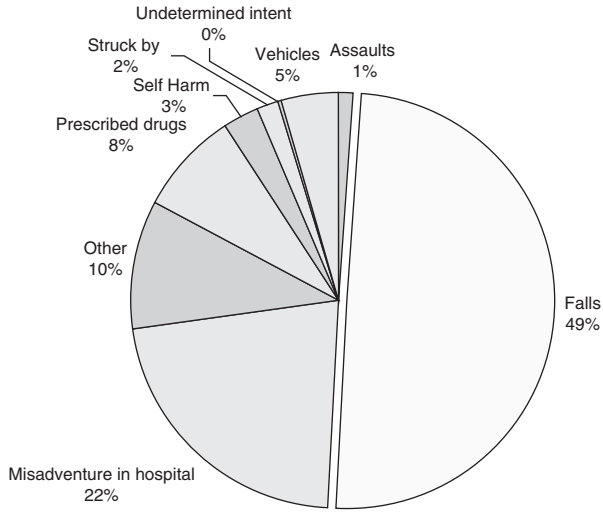


Figure 2. Distribution of bed-days for external cause episodes.

"exposure to unspecified factor". When the ICD recording system changed from ICD-9 to ICD-10 this code was looked at more closely to see how it would have been coded under ICD-9. It was reported that 65% of these would have been coded "Fracture of unspecified cause (E887)" under ICD-9 which was often a fractured neck of femur, and as a result of this reclassification, fall deaths fell by 30% for males and 50% for females (HSQ, 2002). It may be that this poor coding is still apparent today in the Hospital Episode Statistics suggesting that a further 2.5% of the total bed days from external causes are related to falls.

It is worth also noting that within the ICD-10 coding system there is a set of codes which relate to fractures, indicating an outcome of an event rather than the cause. Within this section there is a code S72 which is labelled fracture of femur. This code was used to record 97,000 episodes, resulting in a further 1.8 million bed days. It is difficult to know how many of these fractures were related to falls but it is highly likely that many of them are.

The HES data can therefore be used to estimate the total number of bed days relating to fall injuries that are treated within the NHS in any one year. Estimates reasonably fall between 2.8 million and 4 million bed days.

Calculating the cost of falls

The Housing Health and Safety Rating Scheme (ODPM, 2004), based on data from the Building Regulation Health and Safety Reports (BRE(a),(b), 1995) provides some indication of the severity of different harms resulting from hazards in dwellings. Each hazard has four levels of harm, labelled class I, which is a death, through to class IV relating to minor injuries which are still serious enough to require medical attention. For the purposes of this paper the cost to the NHS really

Table 1. Top ten external causes by number of bed days.

ICD-10	Description	Episodes (K)	Mean length of stay	Mean age	Bed days (K)	% of total bed days for all external causes
W19	Unspecified fall	148	12.2	73	1,371	23.7
Y83	Surgical operations and other surgical procedures as the cause of abnormal reaction of the patient, or of later complication, without mention of misadventure at the time of the procedure	131	9.9	56	991	17.1
W01	Fall on same level from slipping tripping and stumbling	85	8.5	65	639	11.1
W18	Other fall on same level	37	9.7	67	303	5.2
X59	Exposure to unspecified factor	50	5.1	44	225	3.9
W10	Fall on and from stairs and steps	29	6.1	54	161	2.8
W06	Fall involving bed	15	11.7	69	145	2.5
Y84	Other medical procedures as the cause of abnormal reaction of the patient, or of later complication, without mention of misadventure at the time of the procedure	22	8.9	60	142	2.5
Y95	Nosocomial condition	6	33.0	75	137	2.4
Y43	Primarily systemic agents	14	8.8	53	94	1.6

revolves around the middle two classes of harm, Class II and Class III. This scheme included fall related hazards, the main ones being falls on the level and falls on stairs. It is proposed that the Class II harms are similar in magnitude to a neck of femur fracture and the Class III harms are similar in magnitude to a wrist fracture. Typically there are twice as many Class III harms estimated for fall related hazards, compared to Class II.

The HES data covers injuries in all locations, not just in the home, but a large majority of these accidents are generated in the home. It is therefore reasonable to assume that a similar proportion applies to all fall accidents. We are therefore able to deduce that for the 1 Million neck of femur fractures recorded under HES, we expect a further 2 Million class III harms similar in magnitude to a wrist fracture.

Table 2. Cost of an individual hip fracture.

Category	Cost
Hospital care	£4,760
Ambulance	£171
Long stay residential care	£20,010
GP use	£164
Outpatient use	£319
Total cost	£25,424

The NHS cost associated with neck of femur injuries has been well documented. Table 2 is taken directly from the Economic Cost of Hip Fracture in the UK report (University of York, 2000) which highlights the cost associated with an individual hip fracture, estimated to be just over £25,000. When multiplied by the expected 1 million fractures associated with falls the total cost to the NHS for these injuries can be estimated to be in excess of £25.4 Billion. Approximately £5 Billion of this is the costs of providing beds (about £200 per day), since the mean length of stay for neck of femur fractures is 25 days.

The cost of wrist fracture type injuries are less well reported, but we do know that the average length of stay for falls is about 9 days. Using the cost for providing beds alone the cost is an additional £1.8 Billion.

The total cost of falls

Using these estimates, the total cost of falls to the NHS will be in excess of £27 Billion, with at least £7 Billion of this associated with caring for patients on the ward of NHS hospitals. It is therefore reasonable to assume that substantial savings could be made, both in the running costs of NHS hospitals, and in the bed spaces freed up, through intervention measures designed to reduce the potential for harm from falls.

The conclusions from the University of York report are still relevant. "The treatment of hip fractures . . . " [and other injuries generated through falls] " . . . places a heavy burden on the NHS and social care services. As the UK population ages, the cost of falls is expected to escalate. Therefore, measures to reduce the incidence of falls will generate significant savings to the NHS and society as a whole."

References

Dti, 2003, *24th (Final) Report of the Home and Leisure Accident Surveillance System: 2000, 2001 and 2002 Data*, Department of Trade and Industry: London

WHO, 1992, *International Statistical Classification of Diseases and Related Health Problems*, Tenth edition, World Health Organization: Geneva

Health Statistics Quarterly 19 (Autumn 2003)
ODPM, 2004, *Housing Health and Safety Rating System: Guidance*, Version 2, Office of the Deputy Prime Minister: London
BRE, 1995, *Building Regulation and Health*, G J Raw and R M Hamilton (eds), BRE Report BR 289, Construction Research Communications Ltd.
BRE, 1995, *Building regulation and safety*, S J Cox and E F O'Sullivan (eds), BRE Report BR 290, Construction Research Communications Ltd.
The University of York, 2000, *The Economic Cost of Hip Fracture in the UK*, Department of Trade and Industry: London

OPTIMUM FRICTION LEVEL FOR WALKING OVER A FLOOR

Tomás Zamora[1], Enrique Alcántara[1], Fran Payá[1], Javier Portolés[2] & Enrique Algora[2]

[1]*Instituto de Biomecánica de Valencia (IBV)*
[2]*TAU Cerámica*

The existence of an optimum range for floor friction level encompassing safety and comfort has been suggested before. The present study combined subjective and biomechanical analysis of 6 women walking over a sample of floors to assess the existence of an optimal coefficient of friction interval. Five ceramic tiles with friction coefficients (DCOF) ranging from 0.19 to 0.63 were tested. Pain at different body parts, comfort and fatigue were gathered after 15 minutes walking, whereas the Required Coefficient of Friction at landing and take-off were computed from ground reaction forces measured using a force plate.

A great DCOF related to pain under the metatarsal heads, toes and at the knee, whereas a low DCOF related to tights pain and perception of low friction- This supports the existence of this optimal range. The analysis of RCOF with respect to DCOF suggested that there are 3 walking strategies depending on DCOF. First, slow walk to avoid slipping (DCOF < 0.25); second, optimum walk for DCOF ranging between 0.25 and 0.55. Thus, the optimum friction coefficient should be around 0.4 to 0.55 to include safety and third, attenuated walk for a high DCOF (>0.55).

Introduction

Most research about floor friction has traditionally focused in establishing a minimum friction level to avoid slips and falls (Grönqvist et al, 2001; Redfern et al, 2001) and from that, developing reliable testing methods to measure the floor dynamic coefficient of friction (DCOF). Hanson et al (1999) related the required coefficient of friction (RCOF) of subjects with the measured DCOF to build a good estimator of the probability of suffering a slip and/or a fall.

In this context, it could seem that once that minimum is achieved, the more the friction the better. However, it is known that a too high friction can produce several problems as trips, fatigue, reduce comfort or even cause joint injuries. Thus, the existence of a maximum value for DCOF has been suggested in a way that friction should be kept in between an optimal interval. But, this has not been addressed before.Your introduction goes here. Only the first line of each paragraph is indented.

Table 1. Floors chosen for the experiment and their coefficient of friction.

Floor sample	DCOF
P1	[0.37]
P2	[0.63]
P3	[0.25]
P4	[0.51]
P5	[0.19]

This paper presents a research on the optimum range of friction to avoid slipping at the same time that keeping comfort and avoiding joints pain and fatigue.

Material and methods

Six healthy women without mobility or equilibrium diseases (35–45 years old) walked over a sample of floors in dry conditions. Each subject wore her own shoes on the five floors.

Five different floors (Table 1) from the same ceramic tile manufacturer (TAU CERAMICA) were selected to have a wide range of DCOF as measured with IBV dynamical friction testing device (Dura et al, 2005). All subject tested the floors in the same randomly chosen order, from P1 to P5. Every floor was installed in a surface of 7×2 meters, simulating a corridor.

Both, biomechanical data and subjective perceptions were collected during the experiments. Subjective data included perception of floor comfort and friction before and after 15 minutes walking. Also, pain at different body parts was collected. The variables were gathered in a five points scale.

Biomechanical measures consisted of the acquisition of Ground Reaction Forces using a force plate embedded at the walking corridor. The RCOF was obtained as the division of the horizontal forces by the normal forces (Hanson et al, 1999). The peak at landing (RCOFL) and at take-off (RCOFT) were computed as study parameters.

A Descriptive statistics of subjective variables was conducted to assess whether pain, fatigue and discomfort appeared followed by a Kruskal-Wallis Non Parametric Analysis of Variance to assess whether floor had a significant effect on perception. A SPEARMAN correlations analysis to assess the relationship of subjective perceptions with DCOF was also done.

An analysis of variance (ANOVA) for RCOFL and ROCFT parameters considering floor as factor was done to assess any differences due to them. Walking Speed was included as covariable. A Bonferroni post-hoc analysis was done to test the existence of homogenous groups.

A SPEARMAN correlations analysis was done to assess the relationship of subjective perceptions with RCOFL and RCOFT, as well as with the difference

Table 2. Post-hoc analysis results for the RCOFL.

			Group		
	Floor	N	1	2	3
Tukey B(a,b,c)	P5	27	1459		
	P2	28		1757	
	P1	29			1924
	P3	29			1938
	P4	29			1945

Table 3. Post-hoc analysis results for the RCOFT.

			Group		
	Floor	N	1	2	3
Tukey B(a,b,c)	P5	30	2887		
	P2	28	3000	3000	
	P4	30		3103	
	P3	28			3264
	P1	28			3321

between both. Regression analysis of RCOFL and RCOFT with DCOF was finally carried out to analyze the influence in gait of DCOF.

Results

The results of descriptive statistics of subjective data showed that, after 15 minutes walking, pain appeared in different body parts with a rather high frequency: low back (26%), thighs (29%), knees (13%), legs (23%), the heel pad (20%), metatarsal heads (46%) and toes (13%). However, perception of poor comfort occurred only in less than 10% of cases.

The results of Kruskal-Wallis Non Parametric Analysis of Variance showed that there were statistically significant differences only in friction perception before and after contacting the sample. Thus, floors caused not differences in either comfort or in pain after walking for 15 min.

As it could be expected, there was a significant correlation between dynamical coefficient of friction (DCOF) and friction perception (Spearman +0.64) in a way that the greater the DCOF the more friction was perceived by people. Also, it was found high negative relationship between DCOF and pain appearance in tights (Spearman −0.97) and high positive between friction and pain under the metatarsal heads and toes, as well as at the knees (Spearman 0.97).

Values measured for RCOF ranged between 0.13 and 0.33, with a Coefficient of Variation around 15%.

The ANOVA results issued statistically significant differences for RCOFL and ROCFT among floors. Tables 2 and 3 show the results of the post-hoc analysis for the

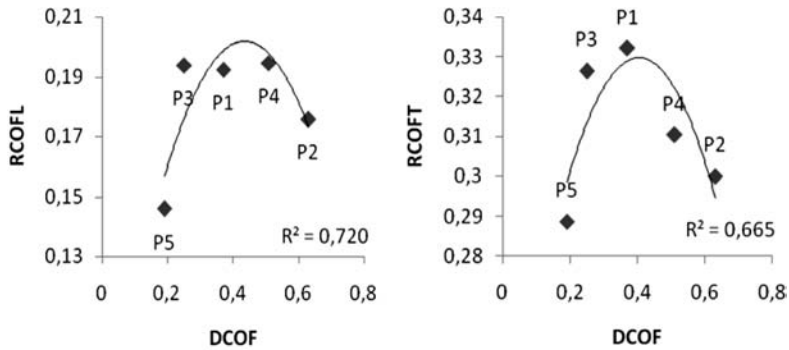


Figure 1. (a) RCOFL compared with the DCOF; (b) RCOFT compared with the DCOF.

landing (RCOFL) and take off (RCOFT) coefficient of friction respectively. There appear to be three homogeneous groups for RCOFL. Group 1 with low RCOFL grouped (P5), group 2 with medium RCOFL (P2) and group 3 with high RCOFL formed by floors P1, P3 and P4.

Results for RCOFT identified 3 groups. Group 1 with low RCOFT (P5 and P2). Group 2 with medium RCOFT (P2 and P4) and group 3 with high RCOFT formed by floors P3 and P1.

The correlation analysis between biomechanical and perception variables showed that perceived comfort had significant relationship with gait speed and RCOF. Bigger comfort was positively correlated with gait speed (Spearman 0.87) and negatively with RCOFT (Spearman −0.97) and RCOFL (Spearman −0.9). That means that the more comfortable people feels, greater walking speed and friction required. Also, bigger comfort was related with bigger differences between the RCOFL and RCOFT (Spearman 0.99).

Regression analysis issued a quadratic relationship of RCOFL and RCOFT with DCOF. For RCOFL (Figure 1), subjects increased the RCOFL as DCOF increased to stabilize around RCOFL 0.19 and then decrease after DCOF reached 0.5. A similar behavior was observed for RCOFT, but it decreased after DCOF reached 0.4.

Both RCOF showed also a clear quadratic relationship (Figure 2) depending on DCOF. When DCOF increases from 0.19 to 0.25, both increase but then, as DCOF increases, while RCOFL keeps constant, RCOFT changes and decreases, then a further increase in DCOF causes both to decrease. That can be seen as trying to keep slipping risk low (RCOFL below 0.2) whereas maintaining walking efficiency (RCOFT around 0.3).

Discussion and conclusion

In agreement with previous experiments, walking for 15 minutes caused pain at different body parts, especially under the metatarsal heads, but people reported a low

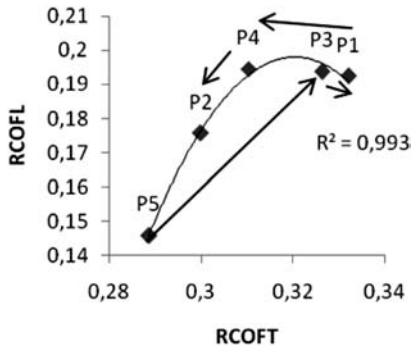


Figure 2. RCOFL versus RCOFT.

frequency of discomfort and fatigue. Moreover, not differences among floors were found. That could indicate that longer tests could be required to elicit discomfort and floor influence.

Results of correlation analysis supported the existence of an optimum DCOF range. The lower the DCOF the more frequent the pain at the tights and less friction perceived whereas the greater the DCOF the more frequent the pain at knees, under the metatarsal heads and toes.

In this sense, results showed differences in required coefficient of friction, both for landing and take-off, among floors. Also, bigger comfort appeared to be related to these parameters in a way that the more comfortable people feels greater friction is required. That can be seen in terms of walking strategies. RCOF is the tangent of the angle of the resultant of ground reaction forces with respect to the vertical. Thus, the greater the RCOF the greater the angle, what is probably due to greater walking speed or a longer step.

Hence, the observed relationship of RCOFL and RCOFT with DCOF may indicate that once a safety level is achieved (around DCOF 0.2), increasing available friction increases the RCOF, that is the angle or walking "intensity". For take-off, RCOF reaches a maximum around 0.34 for a DCOF of 0.4 and then decreases as DCOF increases, specially for DCOF 0.5, whereas RCOF at heel strike remains around 0.2. For DCOF 0.6 both decrease significantly. That can mean that as more friction is available a greater take-off angle is used to walk faster, but this increases pain at different body parts and it is reduced even more friction is available. Landing angle is kept constant till DCOF 0.6, when probably a greater walking strategy change is demanded. This should be contrasted by analyzing the walking pattern kinematics.

Concluding, people seems to adapt to available DCOF by mean of three strategies:

1. Low DCOF (<0.25): To avoid slipping, gait pattern is adjusted so that RCOF < DCOF.
2. Optimum DCOF range (0.25 to 0.55): maximum difference between the landing and the take off, there is an equilibrium between a smooth landing and an energetic take off. In any case, the safety DCOF demands a minimum of 0.4

(Durá et al, 2005). Thus, the optimum friction coefficient should be around 0.4 to 0.55 which includes safety and comfort.

3. High DCOF (>0.55): The strategy is to modify the gait to keep a smooth landing (short steps) by reducing take-off angle.

References

Durá, J. V., Alcántara, E., Zamora, T., Balaguer, E. and Rosa, D. 2005, Identification of floor friction safety level for public buildings considering mobility disabled people needs. *Safety science*, **43**, 407–423.

Grönqvist, R., Chang, W-R., Courtney T.K., Leamon, T.B., Redfern, M. S. and Strandberg, L. 2001, Measurement of slipperiness: fundamental concepts and definitions, *Ergonomics*, **44**, 1102–1117.

Hanson, J., Redfern, M. and Mazumdar, M. 1999, Predicting slips and falls considering required and available friction, *Ergonomics*, **42**, 1619–1633.

Redfern, M.S., Cham, R., Gielo-Perczak, K., Grönqvist, R., Hirvonen, M., Lanshammar, H., Marpet, M., Pai, Y.C. and Powers, C. 2001, Biomechanics of slips, *Ergonomics*, **44**, 1138–1166.

MULTIDISCIPLINARY RESEARCH TO PREVENT SLIP, TRIP, AND FALL (STF) INCIDENTS AMONG HOSPITAL WORKERS

James W. Collins[1], Jennifer L. Bell[1], Raoul Grönqvist[2], Theodore K. Courtney[3], David A. Lombardi[3], Gary S. Sorock[3,7], Wen-Ruey Chang[3], Laurie Wolf[4], Sharon Chiou[1], Bradley Evanoff[5], Helen M. Wellman[3], Mary Matz[6] & Audrey Nelson[6]

[1]*National Institute for Occupational Safety and Health, Morgantown, WV 26505 USA*
[2]*Finnish Institute of Occupational Health, 00250 Helsinki, Finland*
[3]*Liberty Mutual Research Institute for Safety, Hopkinton, MA 01748 USA*
[4]*BJC Health Care, Corporate Health Services, St. Louis, MO 63110 USA*
[5]*Washington University School of Medicine, St. Louis, MO 63110 USA*
[6]*Veteran's Health Administration, Tampa, FL 33612 USA*
[7]*Johns Hopkins University, Baltimore, MD 21205 USA*

Hospitals are diverse work environments with slip, trip, and fall (STF) incidence rates that are 67% higher than other U.S. private industry. **Objective:** To conduct lab and field research to identify risk factors for STF incidents and evaluate a 'best practices' STF prevention program in three hospitals. **Methods and Results:** 1) A descriptive analysis of six-years of workers' compensation data identified 316 STF claims, 2) A case follow-back field study identified contaminants and surface transitions as transient risk factors, 3) lab evaluations of flooring and footwear identified promising slip-resistant shoes and flooring, 4) assessments identified STF hazards, and 5) a field study found a 58% reduction in STF incidents after a 'best practices' program was implemented. **Conclusions:** Food service workers, housekeepers, nurses, EMS/Transport, and office staff are at highest risk of STF injury. This study applied multiple research methods to design a STF prevention program and demonstrate it's effectiveness in three hospitals.

Background

Falls are the second leading global cause of accidental death, after motor vehicle collisions. In the United States (US), STFs on the same level are a leading cause of occupational injury resulting in an average of 52 deaths and 220,000 non-fatal injuries annually and account for the largest proportion of lost time injuries to health care workers (21%) (Courtney *et al*, 2001). The health services sector is the largest employer in US private industry with an estimated 13 million workers. The large population of health care workers and the high frequency of STF incidents create a substantial risk for health care workers.

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

Study design and methods

Descriptive analyses

Six-years of STF workers' compensation injury data were analyzed to identify the circumstances and locations of STF incidents to target prevention efforts.

Case-crossover/case follow-back studies

Case-crossover (case follow-back) methodologies examined transient risk factors and described STF circumstances that could be targeted for prevention. Health care workers were interviewed by telephone with a structured questionnaire (Courtney *et al*., 2006).

Laboratory testing of shoe-floor slip resistance

Laboratory studies evaluated the slipperiness of shoes (most commonly worn and slip-resistant), and hospital flooring (existing and slip-resistant) tested with 'soapy' and 'oily' contaminants. The test apparatus was a slip simulator instrument which can closely reproduce the biomechanical parameters during heel strike in normal gait. In this study, the test parameters were: normal force 500 N, sliding velocity 0.4 m/s, and heel contact angle 5 degrees. The dynamic friction coefficient was computed during the time interval 100–150 ms from heel strike, which in level walking represents a critical moment for a slip and fall.

Seven shoe types were first pre-tested on a stainless steel surface (roughness, Rz 1.6 μm) as new (intact heel and sole) and after abrasion according to draft standard (European Standard EN 13287: 2004). Glycerol 85% wt was used to simulate the 'oily' condition.

Ten different floorings were tested using two significantly different performing shoe types (no. 2 and no. 3) with abraded heels and soles with glycerol 85% wt ('oily' condition) and natrium lauryl sulphate 0.5% wt in water ('soapy' condition) used as contaminants. Paired t-tests compared the differences between initial (intact) and abraded shoe heels and soles, and between the different floorings and contaminants. The following slip-resistance classification for dynamic coefficient of friction was used:

Dynamic friction coefficient (DCOF)	Slip resistance class
>0.30	Slip-resistant
≥ 0.20–0.30	Moderately slip-resistant
<0.20	Slippery

Hospital hazard assessments

On-site STF hazard assessments examined the condition of walkway surfaces, projecting objects or cords, lighting, handrails, and drains. Areas examined inside included entrances, stairs, ramps, operating rooms, the emergency room, scrub sink areas, nursing stations, pharmacy, histology lab, hallways, kitchen, dishwashing areas, cafeteria, patient rooms, bathrooms, surgical instrument decontamination, engineering and carpenter shops and the morgue. Outside areas examined included the parking garage, street, handicap ramps, and sidewalks. Photos of STF hazards and preventive recommendations were described in a written report that was provided to hospital management, safety staff, the housekeeping manager and the groundskeeper manager.

Field study-intervention trial

An intervention trial evaluated a 'best practices' STF prevention program in three acute care hospitals. The 'best practices' program was a multi-faceted effort based on findings from the descriptive analyses, case follow-back study, on-site hazard assessments, and lab tests. The field study compared STF injuries during a 4-year pre-intervention period with a 3-year post-intervention period in a cohort of approximately 22,000 full-time equivalent workers. Workers' compensation data and personnel records were used to compute injury rates.

Results and discussion

Descriptive analysis

During the six-year period from January 1, 1996 through December 31, 2001, 12,098 individuals worked a total person-work time of 44,488,639 hours. The analysis identified 316 workers' compensation claims for STF incidents (19% of all claims). Of the 316 fall incidents, 77% were STFs on the same level, 12% were slips and trips without a fall, and 11% were falls from elevation. The most common scenarios for STFs on the same level involved workers slipping on grease, ice/snow, spilled drinks, and food. The most common circumstances for trip and fall incidents involved employees tripping over phone or medical equipment cords, surface irregularities in flooring and carpeting, and holes in the parking lot. The highest injury rates for STF incidents occurred among food service workers (4.5 STF injury claims/100 food service workers), and the most frequent injuries ($n = 99$, 31%) occurred among nursing staff and office workers ($n = 64$, 20%). The most frequent injuries were sprains, strains, dislocations, and tears (47%) involving the lower extremities (45%).

Case-crossover/case follow-back studies

Through February 2005, 123 health care workers who reported a STF to the occupational health department were recruited and interviewed using a structured

Table 1. Shoes pre-tested on the reference stainless steel surface – 'oily' condition.

Shoe no. and type	Intact heel/sole DCOF (SD)	Abraded heel/sole DCOF (SD)	p-value
1. Nursing shoe with laces	0.159 (0.001)	0.198 (0.022)	<0.01
2. Slip-resistant shoe w/laces – a	0.328 (0.026)	0.375 (0.028)	<0.001
3. Shoe with laces – b	0.149 (0.015)	0.173 (0.013)	<0.001
4. Clog	0.073 (0.012)	0.141 (0.008)	<0.001
5. Shoe with open heel	0.084 (0.011)	0.142 (0.012)	<0.001
6. Shoe with laces – c	0.113 (0.009)	0.138 (0.019)	<0.01
7. Safety shoe with laces	0.140 (0.009)	0.153 (0.010)	<0.05

* Statistically significant difference between new and abraded soles (t-test paired 2-tailed).

telephone questionnaire. Participants were predominantly female (86%) with a mean age (range) of 46 (19–67). Nurses (33%), secretaries or clerks (13%), and health technicians (9%) were the most frequently reported occupations. One hundred and eight subjects (88%) fell: 53% after slipping, 32% after tripping. Liquid contaminants (e.g., water) were involved in 36% of the events. Sixty-four percent of the STFs occurred at a transitional area: dry/wet (32%), one type of floor to another (20%), or uneven surfaces (15%). The back, knees, ankles/feet were most frequently injured. Strains and sprains (29%), contusions (27%), and non-specific pain and soreness (22%), were typical. Other injuries included abrasions, fractures, edema and lacerations. Overall, 94% of subjects were injured. Interested readers should consult Courtney *et al*, (2006) which presents these results in detail.

Laboratory study results

The lab study identified slip-resistant shoes and flooring that performed optimally under 'soapy' and 'oily' conditions. The results for the shoes are presented in Table 1, and for the floorings in 'oily' and 'soapy' conditions in Table 2. Results confirmed previous results (Grönqvist, 1995) showing that heel and sole abrasion significantly improved slip-resistance.

Flooring testing and classification

Slip-resistance evaluations of ten hospital floorings were conducted with both good and poor performing shoes (nos. 2 and 3, Table 1). Shoe no. 2 was classified as 'slip-resistant' (DCOF > 0.30) and shoe no. 3 as 'slippery' (DCOF < 0.20) on stainless steel ('oily') test condition.

All DCOF differences between the two contaminant conditions 'oily' vs. 'soapy' were statistically significant ($p < 0.01$) for all floorings, except flooring no. 4 and no. 8 tested with shoe no. 2 (Table 2). All DCOF were significantly different (p-values < 0.01) between the two test shoes (no. 2 and no. 3) (Table 2). Quarry tile was the only tested flooring that was slip-resistant with both test shoes under all contaminant conditions (Table 2).

Table 2. Floorings tested in the 'oily' and 'soapy' conditions with two shoes.

Flooring no. and type	Shoe no. 2* 'oily' condition DCOF (SD)	Shoe no. 2* 'soapy' condition DCOF (SD)	Shoe no. 3** 'oily' condition DCOF (SD)	Shoe no. 3** 'soapy' condition DCOF (SD)
1. Waxed vinyl tile	0.356 (0.044)	0.263 (0.029)	0.133 (0.015)	0.163 (0.008)
2. Slip-resistant – a	0.325 (0.027)	0.452 (0.034)	0.129 (0.012)	0.254 (0.015)
3. Slip-resistant – b	0.369 (0.027)	0.353 (0.021)	0.155 (0.023)	0.283 (0.012)
4. Slip-resistant – c	0.367 (0.034)	0.378 (0.021)	0.144 (0.017)	0.290 (0.011)
5. Slip-resistant – d	0.335 (0.023)	0.277 (0.022)	0.131 (0.013)	0.202 (0.007)
6. Quarry tile	0.580 (0.026)	0.753 (0.021)	0.288 (0.021)	0.539 (0.018)
7. Safety – a	0.352 (0.023)	0.405 (0.016)	0.163 (0.008)	0.243 (0.034)
8. Safety – b	0.311 (0.018)	0.319 (0.012)	0.146 (0.008)	0.242 (0.023)
9. Safety – c	0.351 (0.022)	0.483 (0.019)	0.154 (0.011)	0.248 (0.032)
10. Safety – d	0.365 (0.020)	0.437 (0.014)	0.168 (0.008)	0.267 (0.025)

* Shoe no. 2 – slip-resistant shoe w/laces. ** Shoe no. 3 – common tennis shoe w/laces.

Hazard assessments

The hazard assessments identified a range of slip, trip and fall hazards. The major hazards found were: 1) flooring and walkway surface irregularities, 2) changes in walkway level, 3) grease, water and food on floors in food cooking and serving areas, 4) loose cords under computer workstations and in operating rooms, and 5) stair visibility and handrail problems.

Field study

Analyses of workers' compensation data indicated the comprehensive 'best practices' STF prevention program resulted in a 58% reduction in workers' compensation injury incidence rates attributed to STF incidents (Bell *et al.*, manuscript under review).

STF prevention strategies emerging from this study

1. Conduct hazard assessments to proactively identify STF hazards.
2. Offer slip-resistant shoes to employees whose work in wet or greasy floors.
3. Place a sufficient number of absorbent mats at hospital entrances so that water is not tracked on the floor on rainy or snowy days. Replace mats with curled edges.
4. Implement effective procedures to degrease floors in food preparation/serving areas.
5. Provide plastic disposable umbrella bags at hospital entrances.
6. When mopping, disinfecting, stripping or waxing floor surfaces, use highly visible wet floor signs with barricades to block pedestrians from walking onto wet floors.
7. Develop a system for promptly cleaning up spills.

8. Lids should be required for drinks.
9. Use clamps/cord holders to secure medical equipment cords and cords under desks.
10. Check parking areas and covered garages, to ensure that lighting is adequate.
11. Check drains on a regular basis to ensure that there are no clogs or standing water.
12. Slip-resistant flooring and shoes may reduce the risk of slipping and falling.

References

Bell, J.L., Collins, J.W., Wolf, L., Grönqvist, R.A., Chiou, S., Chang, W.R., Sorock, G.S., Courtney, T.K., Lombardi, D.A., Evanoff, B. 2008, Evaluation of a comprehensive STF prevention program for hospital employees. *In review, Ergonomics*

Courtney, T.K., Sorock, G.S., Manning, D.P., Collins, J.W., Holbein-Jenny, M.A. 2001, Occupational slip, trip, and fall-related injuries–can the contribution of slipperiness be isolated? Ergonomics **44**, 1118–1137

Courtney, T.K., Lombardi, D.A., Sorock, G.S., Wellman, H.M., Verma, S., Brennan, M.J., Collins, J.W., Bell, J.L., Chang, W.R., Grönqvist, R., Wolfe, L., DeMaster, E., Matz, M. 2006, Slips, trips and falls in US hospital workers-detailed investigation. IEA Triennial Congress Proceedings

European standard EN 13287, March 2004, Personal protective equipment – Footwear – Test method for slip resistance

Grönqvist, R. 1995, Mechanisms of friction and assessment of slip resistance of new and used footwear soles on contaminated floors. Ergonomics, **38**, 224–241

SLIPS AND/OR TRIPS AT THE MENOPAUSE

John C. Davies, Graham J. Kemp[1] & Chris Bowes[2]

Computing Services Department, The University of Liverpool, Liverpool, L69 7ZF
[1]School of Clinical Science, [2]SafetyAdviser's Office

We report the analysis of two large accident databases to investigate the occurrence of slips and trips on women near the menopause. After the age of 45, the incidence of trips shows a gradual and continuing increase with age. In partial contrast, the incidence of slips increases sharply near the age of the menopause, and then appears to decline with further increase in age.

Introduction

A number of authors (McNamee *et al*, 1997, Davies *et al*, 2001, Cherry *et al*, 2005) have reported an increase in the incidence of slip, trip and fall accidents (STFA) in women over 45 years. The underlying causes of this increase have not been established, in part because accident recording systems often do not preserve sufficient information to allow detailed analysis of causes. Fundamental questions such as the nature of the unexpected underfoot events are difficult to establish. Slips and trips are often grouped into a single accident classification, but their mechanisms are significantly different. We have analysed two accident databases which do preserve this distinction: the Home Accident Surveillance System (HASS) 1998 data, and the University of Liverpool Safety Office database. In both datasets we classified accidents to differentiate slips and trips in order to investigate their occurrence in women near the age of menopause.

The HASS system, operated by the Consumer Safety Unit of the UK Department of Trade and Industry (DTI) collects data at selected hospitals chosen to give a representative (national) view of accidental injuries in the home. The DTI provided us with data on approximately 150,000 accidents from their 1998 study. The University of Liverpool Safety Office database records all incidents on campus reported to the Safety Office since 1979. It has approximately 13,000 records and the coding for these accidents has separated slips, trips, falls on stairs and falls to a lower level as distinct classifications.

Methods

We view accidents as a set of events, starting with an unforeseen/unexpected event related to the immediate cause of the accident. Underfoot accidents are defined to have the following characteristics: *tripped, slipped, missed footing, lost balance,*

twisted/turned ankle/foot as the first unexpected event is a sequence that leads to injury. We concentrate on slips and trips within underfoot accidents, in part because of the limitations of the datasets, but also because our expectation is that these form major groupings.

Home accident surveillance system data

The data and methods used here have been described in an earlier paper (Davies *et al*, 2001). Briefly, the HASS data include age, gender the nature of the injury and a free text description of the circumstances of the accident. Accidents were classified as *underfoot* (true|false) based on the narrative text in the database. We have extended the analysis reported earlier by using the keywords *slip* and *trip* within the narrative text to differentiate slipping and tripping within underfoot accidents. Fractures were identified within the free text injury description.

University of liverpool safety office data

The University of Liverpool Safety Office (UoL) database has been in place since 1979. One of its purposes is to record incidents and accidents to inform policy to improve safety on campus. The range of injury recorded is from none (near miss or incident) to major injury. Incidents and accidents involving staff, students and members of the public are recorded separately. The occupation of staff members was also recorded. Accidents have been assigned one classification – the classifications of interest in this study are: [Fall on stairs, Fall to lower level, Slip (same level), Trip (same level after striking object), Sports (normal course of play)].

The demographic of members of staff is available from the University Payroll Database for the period 1997 to 2007 for male and female staff. In this period the age and gender profile of the population of staff has been stable, although the total number of staff has increased. The range of variation is indicated by error bars showing standard deviation. Two different groups within payroll data have distinct demographics, namely hourly-paid manual staff and all other staff, and these have been analysed separately. The demographic data covers the last 10 of approximately 30 years of data. We assume that the demographics profile has remained constant over the whole period. We use the demographic data to normalise the incidence of slips and trips by calculating appropriate scale factors to produce a flat demographic with respect to age.

For the purpose of this study we have excluded accidents to students and members of the public, so that we have known groups defined by the payroll database. Sports accidents and incidents with no injuries have also been excluded.

Results

Home accident surveillance system data

In the age range 20–79 years there were 74,002 accidents in the HASS database. 28,381 were classified as *underfoot* (true|false) and 7,049 of these were further

classified as either slips or trips. The distribution of accidents for men and women by *underfoot* by age is shown in Figure 1. The overall totals for men and women are comparable in the age range 20–44 years, but proportionately more women than men attend hospital because of underfoot accidents. The incidence of non-underfoot accidents has similar characteristics for men and women, both showing a peak at about 35 years followed by a falling trend with increasing age. The incidence of underfoot accidents for women over 45 years shows a relative increase compared with men.

Figure 2 shows the distribution of fractures within underfoot accidents. As discussed in an earlier publication (Davies *et al*, 2001), there is a discontinuous increase in the number of fractures suffered by women in the age band 50–54 years.

Figure 3 shows slipping and tripping accidents by women by fracture. In Figure 3(a) the incidence of slips shows a well defined peak in the age range 50–54 years, falling back to a level which apparently continues the trend from the 45–49 years band. The increase in fractures at the 50–54 years band corresponds with the peak in slipping and is sustained with increase in age. Tripping accidents, in Figure 3(b), have a different characteristic, showing a less well defined peak followed by an increase with age. The incidence of fractures associated with tripping is a relatively smooth curve, showing a gradual increase in incidence with age after the 40–44 years band.

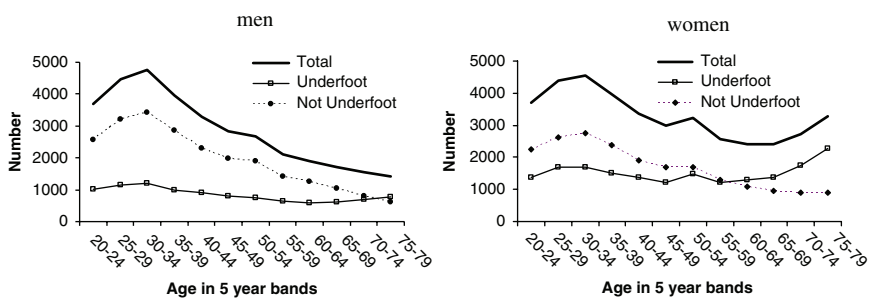


Figure 1. Distribution of accidents by *underfoot* by age (HASS data).

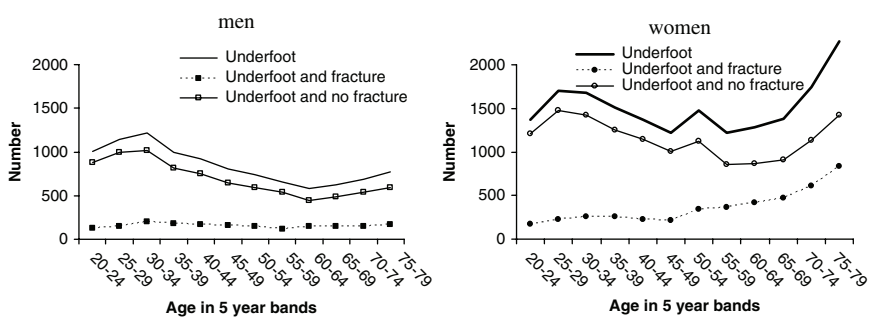


Figure 2. Fractures within underfoot accidents by age (HASS data).

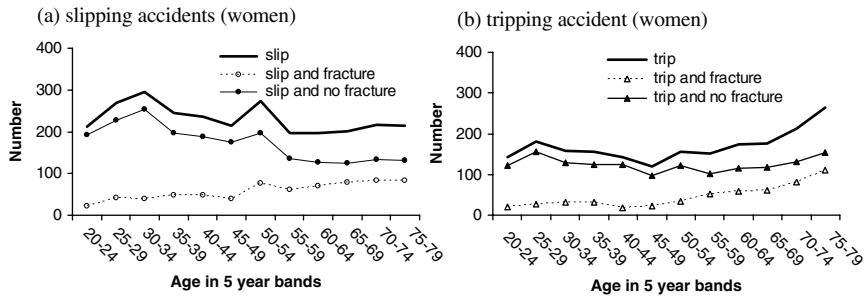


Figure 3. (a) slipping and (b) tripping accidents in women by fracture by age (HASS data).

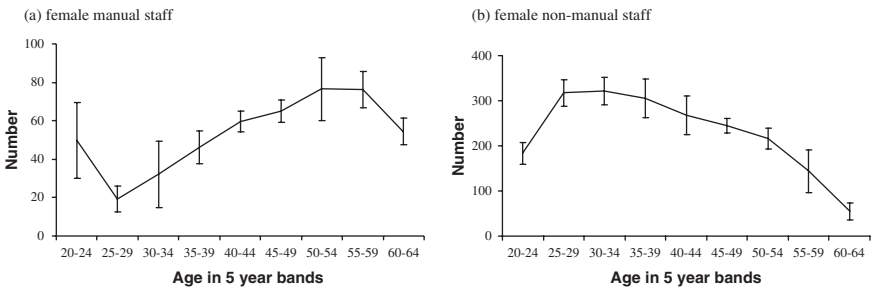


Figure 4. Demographic of female staff, (a) manual and (b) non-manual staff (UoL data).

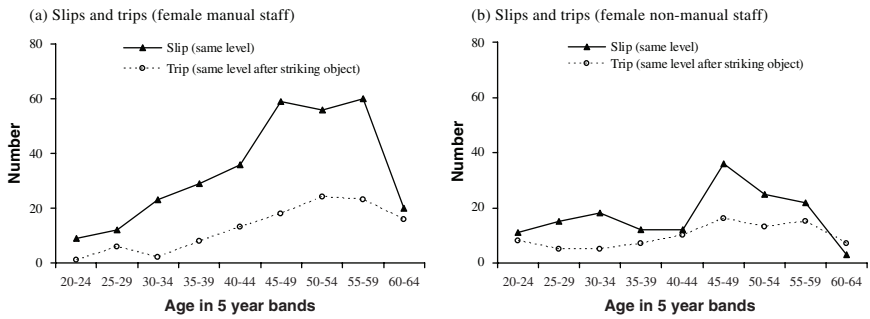


Figure 5. Slips and trips in female staff, (a) manual and (b) non-manual staff (UoL data).

University of liverpool safety office data

Figure 4 shows the demographic data for manual and non-manual female staff at the University of Liverpool in 5 year age bands for the period 1997–2007. The error bars show the standard deviation based on annual counts over the period of the data. (Note the different scales on the y-axes).

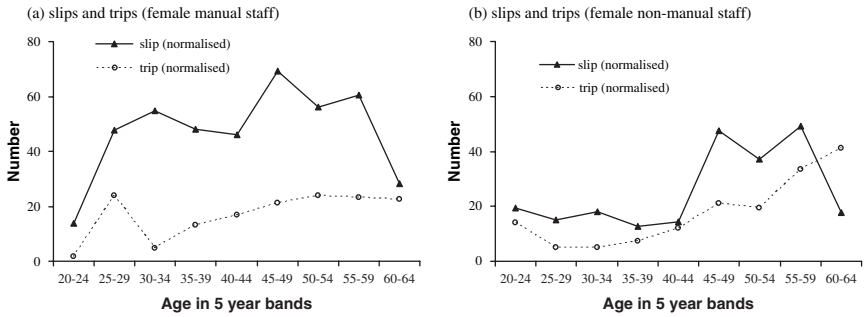


Figure 6. Normalised slips and trips in female staff, (a) manual and (b) non-manual staff (UoL data).

The distributions of slips and trips for female staff are shown in Figure 5. There is a sharp increase in the incidence of slip in both manual and non-manual female staff at the 45–49 years age band. The same data, normalised by the demographic data in figure 4, is shown in Figure 6.

Discussion

Slipping accidents would appear to be especially important near the age associated with the menopause. In the HASS data in Figure 3(a) there is a clear increase in the incidence of slipping accidents at the age band 50–54 years, with an approximate doubling (76 from 40) in the number of fractures compared to the 45–49 years band. In the UoL data in Figure 6 there is a rise in incidence of slipping at the 45–49 years age band. The UoL and HASS datasets both show a rise in the incidence of slipping, although the age bands where this occurs differ. This difference may be accounted for in the severity of injury sustained – the HASS population have injuries of sufficient severity to warrant a hospital visit, and the UoL injuries can be expected to be typically relatively minor.

In both the HASS and UoL datasets the incidence of trips shows a rising trend with increase in age after the 40–44 years age band. Trips would appear to play a more important part in injuries into old age in women.

It is clear that female manual workers at the University of Liverpool have a higher risk of both tripping and slipping when compared with their non-manual colleagues. At the age band 45–49 years, we estimate the increased risk of slipping to be of the order of 6 times. (A count of 59 manual workers with demographic 65 compared with 36 non manual staff with demographic 244).

A recent paper by Liu-Ambrose *et al* (2003) has suggested a link between osteoporosis and postural sway – that the group at risk from fractures is also more likely to fall after an underfoot incident. Cherry *et al* (2005) concluded “The observed excess of fractures in older women falling at work appeared to be explained by greater risk of fracture among those who fell”. We speculate that women in the

HASS data who contributed to the increase in the incidence of slips at the age of 50–54 were more likely to fracture.

Acknowledgements

We are grateful to the Department of Trade and Industry Consumer Safety Unit for providing access to the HASS data. Thanks are also due to Sandra Presland for providing us with the University of Liverpool staff demographic data.

References

Cherry N., Parker G., McNamee R., Wall S., Chen Y., Robinson J. 2005, Falls and fractures in women at work. *Occupational Medicine*, **55**, 292–297

Davies J.C., Manning D.P., Kemp G.J., Frostick S.P. 2001, The rising number of underfoot accidents after the menopause causes both fractures and non-fracture injuries: evidence from a large national database. *Q J Med*, **94**, 699–707

Liu-Ambrose T., Eng J.J., Khan K.M., Carter N.D., McKay H.A. 2003, Older women with osteoporosis have increased postural sway and weaker quadriceps strength than counterparts with normal bone mass: overlooked determinants of fracture risk? *J Gerontol*, **58**, 862–866

McNamee R., Kemmlert K., Lundholm L., Cherry N.M. 1997, Injuries after falls at work in the United Kingdom and Sweden with special reference to fractures in women over 45. *J Occup Envir Med*, **54**, 785–79

THE EPIDEMIOLOGY OF SLIPS, TRIPS, AND FALLS IN THE U.S. ARMY

M. Canham-Chervak, K.G. Hauret & B.H. Jones

U.S. Army Center for Health Promotion and Preventive Medicine, Aberdeen Proving Ground, Maryland, USA

Analysis of medical surveillance and administrative data has demonstrated that slips, trips, and falls (STF) contribute significantly to the burden of injury on the U.S. Army. In 2005, STF were the leading cause of Army injury hospitalizations, accounting for 17.1% (n = 425) of all injury hospitalization visits that received an injury cause code. Rates of STF-related injuries were 9.8 to 11.4 per 10,000 personnel per year from 2000–2005. Among categories of STF (fall/jump from stairs or ladder, fall/jump from different level, fall/jump on same level, twists/turns/slips), hospitalization rates for 'fall/jump from different level' were typically highest, ranging from 3.3 to 4.1 per 10,000 personnel/year, while rates for 'fall/jump from stairs or ladder' were consistently the lowest, ranging from 1.1 to 1.5 per 10,000 personnel/year. Falls/jumps are also the leading cause of non-battle injuries among troops in deployed settings. Among U.S. Army personnel in Operation Iraqi Freedom, one quarter (25.2%, n = 1,476) of non-battle injuries requiring medical air evacuation between March 2003–June 2006 were due to STF. STF from vehicles accounted for 27.9% of non-battle-related falls. Such STF injuries, whether suffered at home or overseas, result in unnecessary costs including medical expenses, lost work time, and manpower reductions. Because of the magnitude and severity of the problem with STF, additional research on risk factors, causes, and interventions to prevent STF among working-age populations is needed.

Introduction

Falls have historically been among the top five causes of injury hospitalizations for U.S. Army personnel, as indicated by military medical surveillance data for 1981–1994 (Gardner, Amoroso et al. 1999). Recent data continue to reflect this trend. In 2005, slips, trips, and falls/jumps (STF) were the leading cause of U.S. Army injury hospitalizations, accounting for 17.1% (n = 425) of all injury hospitalizations that received an injury cause code (Defense Medical Surveillance System 2006).

While the costs of STF-related injuries for the U.S. Army have not been calculated, U.S. civilian workers' compensation data indicate that, among U.S. workers, costs due to falls are second only to motor vehicle crashes and result in an average of $18,838 per claim (National Safety Council 2004). Lost work time due to falls

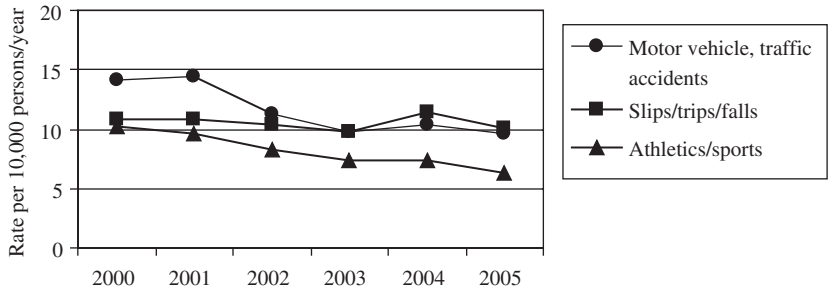


Figure 1. Top 3 causes of injury hospitalizations, active duty U.S. Army personnel, 2000–2005.

among U.S. workers is another important indicator of the injury burden of STF-related injuries; in 2002, the median days away from work for falls to a lower level was 14 days and the median days away from work for falls on the same level was 9 days (U.S. Bureau of Labor Statistics 2002).

This paper summarizes recent epidemiologic data on STF among U.S. Army personnel. Hospitalization data from the Defense Medical Surveillance System (DMSS) are presented to characterize STF among non-deployed personnel for 2000 to 2005. The DMSS, which maintains a longitudinal record of medical encounters for all active duty military personnel (Rubertone and Brundage 2002), records causes of hospitalized injuries using an international military external cause-of-injury coding system (North Atlantic Treaty Organization Military Agency for Standardization 1989).

Data from the U.S. Army Center for Health Promotion and Preventive Medicine's deployment injury surveillance database are presented to describe STF among U.S. Army personnel deployed from March 2003 to June 2006 in support of Operation Iraqi Freedom (OIF). This database contains administrative and medical data obtained from the U.S. Department of Defense Transportation Command (TRANSCOM) Regulating and Command & Control Evacuation System (TRAC^{2}ES). The TRAC^{2}ES data include a free-text patient history field containing a description of circumstances and causes associated with injury incidents. This field is reviewed and coded to enable summaries of causes of deployment-related injuries requiring medical air evacuation from the theater of operations (Hauret, Clemmons et al. 2007).

Falls among non-deployed U.S. Army personnel

From 2000 to 2005, STF were among the top three causes of injury hospitalizations for active duty U.S. Army personnel, with rates ranging from a low of 9.8 per 10,000 personnel/year in 2003 to a high of 11.4 per 10,000 personnel/year in 2004 (Figure 1). In 2003, STF surpassed motor vehicle traffic accidents to become the leading cause of injury hospitalizations for 2003–2005.

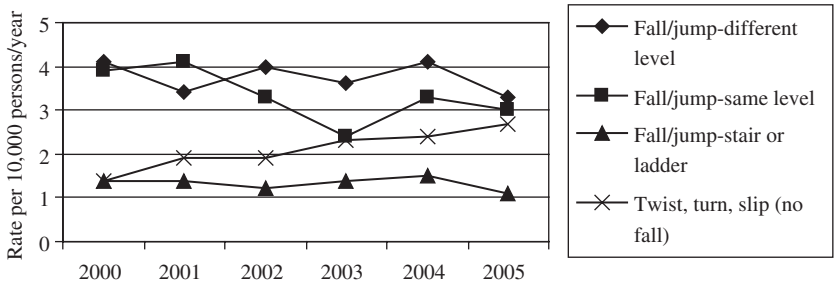


Figure 2. Rates of fall-related hospitalizations by fall type, active duty U.S. Army personnel, 2000–2005.

Table 1. Causes of non-battle injuries in OIF, active duty U.S. Army personnel, March 2003–June 2006.

Cause	Frequency (%)
Falls/jumps	**1,061 (18.1)**
Sports/physical training	1,042 (17.8)
Motor vehicle crashes	875 (15.0)
Crushing or blunt trauma	503 (8.6)
Lifting, pushing, pulling	484 (8.3)
Slips, trips	**415 (7.1)**
Shoes, clothing, body armor	234 (4.0)
Cutting and piercing	183 (3.1)
Instruments of war, own	174 (3.0)
Environmental	167 (2.9)
Twists, turns	53 (0.9)
Other specified	654 (11.2)
Total	5,845 (100.0)

Looking at specific categories of STF, rates for 'fall/jump from different level' were typically highest, ranging from 3.3 to 4.1 per 10,000 personnel/year, while rates for 'fall/jump from stair or ladder' were consistently the lowest, ranging from 1.1 to 1.5 per 10,000 personnel/year (Figure 2).

Falls among deployed U.S. Army personnel

Deployment injury surveillance data show that STF have been a leading cause of non-battle injuries in OIF. Among U.S. Army personnel deployed in support of OIF, falls/jumps were the leading cause of non-battle injuries requiring medical air evacuation between March 2003 and June 2006 (Table 1). STF accounted for 25.2% (n = 1,476) of non-battle injuries requiring medical air evacuation from OIF (Table 1).

Table 2. Type of fall/jump in OIF, active duty U.S. Army personnel, March 2003–June 2006.

Fall/jump type	Frequency (%)
From different level	500 (47.1)
On same level	123 (11.6)
From stairs/ladder	100 (9.4)
Unspecified level	339 (31.9)
Total	1,061 (100.0)

Table 3. Activities associated with non-battle slip, trip, and fall injuries in OIF, active duty U.S. Army personnel, March 2003–June 2006.

Activity	Frequency (%)
On foot (unspecified)	747 (50.6)
From vehicle	412 (27.9)
Walking/march ing/running	186 (12.6)
Climbing (up or down)	86 (5.8)
Other/Unspecified	46 (3.1)
Total	1,476 (100.0)

Of the fall/jump-related injuries medically air evacuated from theater, 47.1% were from a different level, 11.6% were on the same level, and 9.4% occurred on stairs or a ladder (Table 2).

Over half (50.6%) of non-battle STF requiring medical air evacuation from OIF occurred while 'on foot' (not otherwise specified) and another 12.6% occurred while walking, marching, or running (Table 3). Over one-quarter (27.9%) of STF were associated with work on or around vehicles (e.g. while boarding, exiting, or loading equipment).

Conclusions

These data illustrate the continued contribution of STF to the non-deployment U.S. Army injury burden and also suggest that STF are an important contributor to injuries sustained by deployed personnel. More broadly, these data are consistent with U.S. national estimates of non-fatal injuries showing falls to be the leading cause of non-fatal injury among working-age (25–64 years of age) persons (National Center for Injury Prevention and Control 2004).

These data also show that most serious STF-related injuries among deployed and non-deployed U.S. Army personnel are due to falls from one level to another.

Among deployed personnel, falls from vehicles were, unexpectedly, a leading cause of STF-related injuries.

To date, much of the research related to STF has focused on the prevention of STF among the elderly and children. These data indicate a strong need for additional research on risk factors, causes, and interventions to prevent STF among working-age populations. Another area deserving more research is STF from vehicles among persons employed in occupations such as trucking, material handling, and heavy equipment operation.

References

Defense Medical Surveillance System (2006). Causes of U.S. Army active duty injury hospitalizations, 2005 (unpublished raw data). Washington, DC: U.S. Army Medical Surveillance Activity.

Gardner, J. W., P. J. Amoroso, et al. (1999). Chapter 5. Hospitalizations Due to Injury: Inpatient Medical Records Data. *Mil Med* **164**(8 Suppl): 1–143.

Hauret, K. G., N. Clemmons, et al. (2007). *Development of a Deployment Injury Surveillance System using Medical Air Evacuation Data.* USACHPPM report no. 12-HF-056S-07. Aberdeen Proving Ground, MD: U.S. Army Center for Health Promotion and Preventive Medicine.

National Center for Injury Prevention and Control (2004). National estimates of the 10 leading causes of nonfatal injuries treated in hospital emergency departments, United States, 2004 (ftp://ftp.cdc.gov/pub/ncipc/10LC-2004/PDF/10lc-2004-nonfatal.pdf), accessed November 1, 2007. Atlanta, GA: Centers for Disease Control and Prevention.

National Safety Council (2004). Chapter 7: Workers' Compensation. *Injury Facts.* Itasca, IL: National Safety Council.

North Atlantic Treaty Organization Military Agency for Standardization (1989). Standardization Agreement (STANAG): Statistical Classification of Diseases, Injuries, and Causes of Death. STANAG 2050 (Edition 5).

Rubertone, M. V. and J. F. Brundage (2002). The Defense Medical Surveillance System and the Department of Defense Serum Repository: Glimpses of the Future of Public Health Surveillance. *Am J Public Health* **92**(12): 1900–4.

U.S. Bureau of Labor Statistics, Survey of Occupational Injuries and Illnesses (2002). Occupations with the Most Injuries and Illnesses with Days Away from Work, 2002 (http://www.bls.gov/iif/oshwc/osh/case/osch0027.pdf), accessed November 1, 2007. Washington, DC: U.S. Department of Labor.

A PRELIMINARY STUDY ON SLIP POTENTIALS OF STEPPING BAREFOOT ON SLIPPERY FLOORS

Hisao Nagata[1], Masayoshi Kato[2], Hitoshi Watanabe[3], Yukihiko Inoue[4], and In-Ju Kim[5]

[1] *National Institute of Occupational Safety and Health, Kiyose, Tokyo 204-0024, Japan*
[2] *The Center for Better Living, Chiyoda-ku, Tokyo 102-0071, Japan*
[3] *Waseda University, Shinjuku-ku, Tokyo169-8555, Japan*
[4] *Availvs Corporation, Chiyoda-ku, Tokyo 102-0085, Japan*
[5] *School of Health Sciences, University of Canberra, Canberra, Australia*

This study was aimed at measuring and analyzing slip-acceleration, slip-speed, and reaction forces applied to the floor at slip-occurrence. Each subject was asked to step barefoot on the slippery floor from heel to toe. Test subjects were elderly people aged 65 and over. During tests, foot movements, foot velocities, and reaction forces applied to the floor were measured at the beginning of slip occurrence by a motion capture camera, a cord extension meter and a Kisler's piezoelectric type force plate. From the finding of this preliminary study, we expected that the fundamental parameters for identifying slip potentials on surfaces covered with soapsuds could be applied to design a new slip-meter in future.

Introduction

A slip-meter for evaluation of floors in bathing areas must be representative of the conditions during a risky slip. Many parameters such as normal force, velocity, dwell time, etc affect COF measurements. Most slip-testers operate similar way by measuring COF during a phase when normal force, speed and angle are constant. A single value for COF may be insufficient in order to describe the slip resistance nature (Beschorner et al., 2007). The purpose of this study is to evaluate risky slips of flooring materials by slip-experiments with elderly subjects. After obtaining averages and deviations of fundamental parameters for identifying slip potentials on surfaces covered with soapsuds, these data will be applied to design a slip-meter to evaluate safety performance on flooring materials especially in bathing areas. In Japan, in the field of building construction one measuring static friction meter of COF (the O-Y Pull Slip Meter) is popularly used for evaluation of flooring materials including sensory degree of slipperiness and risk of slips of flooring materials (Ono, 1983). For the actual measuring where the O-Y pull slip meter was used, a slider (measuring 7 cm $\times$ 8 cm) onto which a weight of 80 kgf was placed was pulled upwards diagonally at an angle of 18 degrees through the use of a spring. The slip

meter in that study was developed mainly by evaluating slipperiness in sensory tests, in which most test-subjects were students. These slip-experiments were not performed for risky slips, where a fall-arresting harness should be used to prevent falls. A testing methodology for bare feet was applied not only to the evaluation of floor materials used in bathing areas, but also to the evaluation of the sensory slipperiness of floor materials used for sports, such as Judo and Karate which are played in bare feet (Ono et al., 2000). For the slip evaluation in bare feet, the sum of the first peak and second peak of tensile forces obtained using the slip-meter was divided by the weight −80kgf (Ono et al., 1988). The study in this paper discusses risky slips which cause falls through the loss of upright balance due to the sudden movement of the feet at certain accelerations. The slip used in this paper is called "accelerative slip" to distinguish it from slip dealt with in sensory tests. These accelerative slips occur just after foot-landing on slippery floors (Nagata, 2007). Slips with constant velocity were deemed to be of a lesser risky and were not taken up in this study.

Methodology

Instrumentation

Sliding displacements and velocities of heel movements were measured by a cord-extension meter with a pulse encoder and having a measurement precision of 0.02 mm. Output from this meter was processed by a pulse integration counter for displacement-measurements (Cocoresearch CNT-3921) and by a pulse counter with high-speed analogue output for velocity-measurements (Cocoresearch CNT-723). A high speed camera was used in this study. The cord of the meter was attached to the heel of each test subject.

The vertical and horizontal forces at the beginning of slip were measured by Kisler's piezoelectric type force plate (Type 9281B). Measurements of motion pictures of foot movements and forces applied to the floor were triggered by a pair of photo-cells as shown in Figure 1. During the tests, foot motions were measured

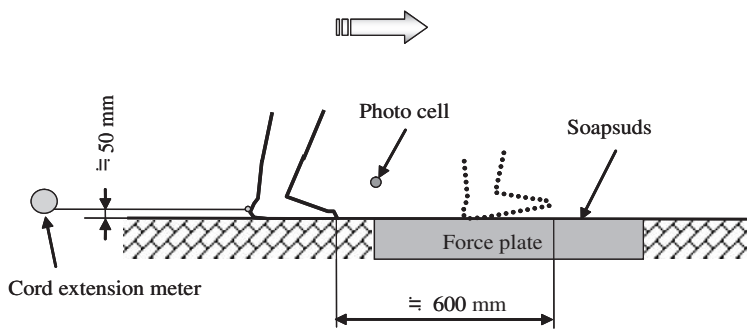


Figure 1. Schematic diagram of experimental setup.

at 1/100 second intervals and vertical forces, horizontal forces and velocities were measured at 1/1000 seconds intervals.

Subjects

Slip tests were conducted on fifteen elderly subjects aged 65 and over. While stepping on slippery floors, each subject was protected from abrupt falls by a fall-arresting harness system. Before slip-tests, every subject was instructed how to step forward. Table 1 shows gender, number of subjects, ages, body weight and stature of elderly subjects. Average contact area of heels of all subjects was 21.8 cm^2 (SD 5.1 cm^2).

Procedure

For the investigation of accelerative slips, all subjects participated in stepping barefoot on a level floor. The surface of force plate was shielded with a smooth plastic sheet (thickness125 μm and Polyethylene terephthalate: PET). For preliminary examinations, surfaces of the force plate were covered with water and soapy water. Floor surfaces were dried quickly during slip-experiments. Film thickness of lubricants on floors also affects friction between bare feet and floors. Dense soapsuds (Alkyl ether sulphate 27% by weight) which do not dry quickly were taken so as to cause actual slips. Sole-marks were described on the force plate to adjust the step length – approximately 600 mm. Each subject was asked to put their right foot on the floor from heel to toe. Subjects were asked to step three times. Before starting experiments, subjects were instructed how to step barefoot and to experience accelerative slips on slippery floors under protection of the fall-arresting harness. Some subjects seemed to be accustomed to risky slips through preliminary experiments. In some cases, accelerative slips and lost-balance could not be observed. Because they held themselves ready to slip and to lose balance while stepping barefoot. Hence, the measurements were only proceeded as clear accelerative slips and lost-balance were observed. One of authors of this paper, suggested how to calculate accelerative slips while subjects walked on slippery floors, when analyzing slip-fall incidents (Nagata, 2007). The velocity change (Δv) and the slip-duration (Δt) were also measured as the difference in the velocity and time from the moment that a slip initiated to the moment when the velocity increased linearly during the slip.

Table 1. Subjects.

		Age (year old)		Body weight (kg)		Stature (cm)	
Gender	Number	Average	SD	Average	SD	Average	SD
Male	8	68.0	2.5	59.9	7.3	168.6	4.1
Female	7	66.7	1.9	58.4	11.4	153.0	3.4

Results

Initial velocity and duration at the occurrence of a slip

Figure 2 shows a typical pattern of the velocity of the foot movements as a function of time when the subjects stepped barefoot. Soon after a foot touches down the floor, velocities of foot movements rapidly and linearly increase. Initial velocities v_0 at the start point of accelerative slips are around 0.43 m/s as shown in Table 1. Averages of slip-duration are around 30 ms for elderly subjects. These values of parameters are very dispersive.

Acceleration and duration at the occurrence of a slip

The slip-acceleration (α) was obtained by dividing the velocity change by the slip-duration ($\Delta v/\Delta t$) as shown in Figure 2. The relationship between slip-duration and slip-acceleration is shown in Figures 4 and 5. According to these results, a slip could result with not only a higher acceleration and shorter duration, but also a lower acceleration and longer duration.

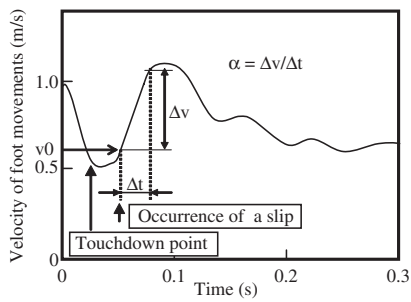


Figure 2. Velocity of foot movements.

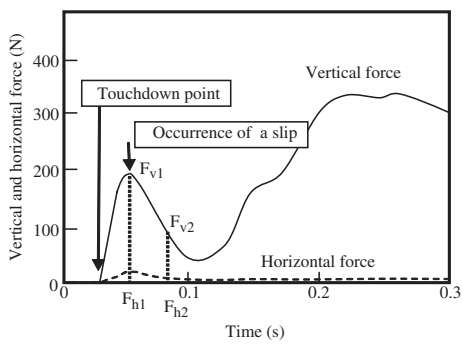


Figure 3. Vertical and horizontal forces.

Reaction forces applied to floor during accelerative slips

The example of floor reaction force after the foot comes into contact with the floor is shown in Figure 3. The average time from the time the foot touches down until the vertical force peaks was about 22 ms (SD 10 ms). The relationship of the vertical force (F_{v1}) and the horizontal force (F_{h1}) at the initial accelerative slip point and the vertical force (F_{v2}) and horizontal force (F_{h2}) at the terminal slip point is shown in Figure 6. From this Figure, it can be seen that the vertical force (F_{v1}, F_{v2}) at the time of the occurrence of the slip are 47N at minimum and 489N at maximum. In calculations to obtain the average value of F_{h1} /F_{v1} and F_{h2} /F_{v2}, the result was 0.06 (SD 0.03). Average vertical forces are 247 N for F_{v1} and 232 N for F_{v2} as shown in Table 3. Average of all vertical forces is 240N (SD 101N).

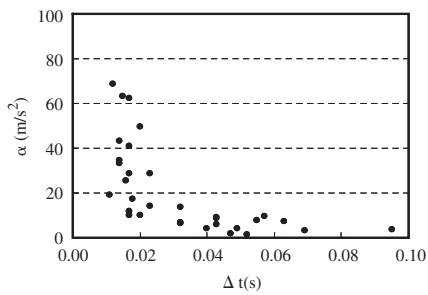


Figure 4. Duration and slip-accelerations.

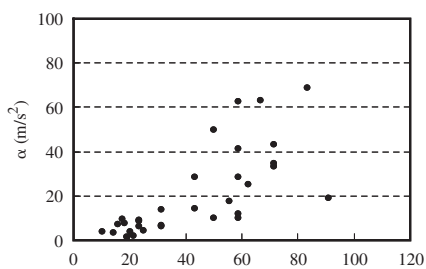


Figure 5. Reciprocal values of Δt and a.

Table 2. Major parameters (Number of data = 32).

Parameter	Mean	SD	CoV
Initial velocity v_0(m/s)	0.43	0.27	63
Slip-duration $\Delta\ t$(ms)	32	20	63
Slip-acceleration α(m/s^2)	20	19	94

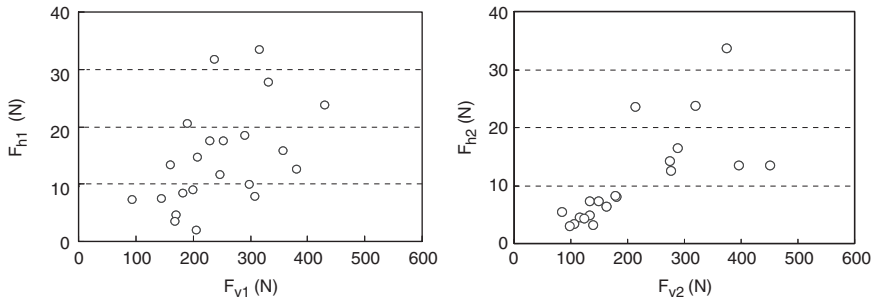


Figure 6. Relation of vertical forces and horizontal forces.

Table 3. Reaction forces (Number of data = 23).

	Reaction forces (N)			
	F_{v1}	F_{h1}	F_{v2}	F_{h2}
Mean	247	14	232	15
SD	81	9	117	14

Conclusion

Our final goal is to learn slip-performances on various flooring materials while stepping barefoot, and to evaluate potential fall-risk of flooring materials in every day life. In this study, flooring materials must be evaluated by examining risky conditions, which produce actual slips. It means that least values of COF at various measuring conditions must be taken into consideration to evaluate risky performance of flooring materials.

Acknowledgment

The collaborative study described in this paper was supported by grants from the Housing Bureau of the Ministry of Land, Infrastructure and Transport of Japan.

References

Beschorner, K.E., Redfern, M.S.,Porter,W.L. and Debski, R.E., 2007, Effects of Slip Testing Parameters on Measured Coefficient of Friction, *Applied Ergonomics*, **38**, 773–780.

Ono, H., 1983, A Study on the Slipperiness of Building Floors and its Methods of Evaluation – Part 3 Design and development of a new slipperiness tester, *Nihon Kenchiku Gakkai Kouzou-kei Ronbun Hokoku-shu (Journal of Structural and Construction Engineering, Transaction of Arichitectual Institute of Japan)*, **333**, 1–26.

Ono, H. and Ochiai, N., 2000, Study on the Evaluating Method of Slipperiness of Floors with Bare Foot, *Nihon Kenchiku Gakkai Kouzou-kei Ronbun Hokoku-shu (Journal of Structural and Construction Engineering, Transaction of Arichitectual Institute of Japan)*, **537**, 21–26.

Ono, H.,Ueno, S.,Yokoyama,Y.,Ohno, R.,Mikami, T., 1988, Study on the Evaluating Method of Slipperiness of Bathroom Floors and Bathtub bottoms from a view point of Safety- Part 1 Study on the measuring method of slip resistance and the relation between the easiness to cross the bathtub's edge and its sectional propotion, *Nihon Kenchiku Gakkai Kouzou-kei Ronbun Hokoku-shu (Journal of Structural and Construction Engineering, Transaction of Arichitectual Institute of Japan)*, **384**, 26–33.

Nagata, H., 2007, Heel Movements during Slip Incidents, Program and Abstracts of XXI International Society of Biomechanics Taipei, 1–5 July 2007, *Journal of Biomechanics*, **40**, Supplement 2, 113.

MODELING SHOE-FLOOR-CONTAMINANT FRICTION APPLIED TO A PIN-ON-DISK APPARATUS

Kurt E. Beschorner[1], Michael Lovell[2], C. Fred Higgs III[3] & Mark Redfern[1]

[1] *Department of Bioengineering, University of Pittsburgh*
[2] *Department of Industrial Engineering, University of Pittsburgh*
[3] *Department of Mechanical Engineering, Carnegie Mellon University*

Slip and fall accidents are a serious health problem causing a substantial economic burden throughout the developed world. Fundamental understanding of friction between shoe-floor-contaminant interfaces could be improved with computational models. We present a shoe-floor-contaminant friction model based on mixed-lubrication that utilizes a number of measurable inputs such as sliding speed, vertical force, shoe material information (geometry, roughness and elastic modulus) and fluid viscosity. A multi-scale modelling approach includes asperity contact and hydrodynamic lift. The asperity contact model determines force on the asperities based on film thickness, shoe roughness and shoe elastic modulus. The hydrodynamic lift model determines load carried by the fluid based on the Reynolds equation, geometry of the shoe material, fluid viscosity and sliding speed. A linear rule of mixtures is used to determine coefficient of friction as a function of the load carried by the fluid and by the asperities. Model data show good agreement with experimental data. As sliding speed increases, load is transferred from the asperities to the fluid causing a decrease in friction coefficient. This computational friction model provides valuable information about the interaction of shoe-floor-contaminant surfaces and shows promise for future development into a useful design tool.

Introduction

Slip and fall accidents represent a substantial economic burden in the developed world and reduce the quality of life for those personally afflicted. Slip and fall accidents were associated with 16% of non-fatal injuries requiring time away from work and 14% of fatal injuries requiring time away from work for the year 2006 in the United States (U.S. D.O.L.S, 2007a, U.S. D.O.L.S, 2007b). Previous research has indicated that slipperiness contributes to approximately 40–50% of falls (Courtney, Sorock et al., 2001).

Coefficient of friction (COF) is well established as contributing factor to slips and falls (Burnfield and Powers, 2005). COF has become an important measure

to evaluate the slipperiness of shoe and floor surfaces. Shoe-floor-contaminant friction is not a very well understood phenomenon and is therefore empirically measured. COF measurements, however, are often dependent on the testing device and the testing parameters (Chang, Gronqvist et al., 2001a; Chang and Matz, 2001; Beschorner, Redfern et al., 2007); a problem made more complicated by the large number of devices (Chang, Gronqvist et al., 2001a). A number of variables are known to affect shoe-floor-contaminant friction and need to be considered in a friction model. Biomechanical variables contributing to shoe-floor-contaminant friction are: sliding speed (Redfern and Bidanda, 1994; Gronqvist, Matz et al., 2003; Beschorner, Redfern et al., 2007), normal force (Gronqvist, Matz et al., 2003; Beschorner, Redfern et al., 2007), shoe-floor angle (Beschorner, Redfern et al., 2007). In addition, friction is known to vary rapidly after heel contact as a result of the squeeze film layer (Chang, Gronqvist et al., 2001b). Environmental factors also contribute to shoe-floor-contaminant friction, such as shoe material (Redfern and Bidanda, 1994; Chang and Matz, 2001), floor material (Chang and Matz, 2001), contaminant (Redfern and Bidanda, 1994; Chang and Matz, 2001; Beschorner, Redfern et al., 2007), shoe roughness (Manning, Jones et al., 1998) and floor roughness (Chang, 2001; Kim, 2004).

Friction within a shoe-floor-contaminant interface is a result of the interaction between the asperities of the shoe and floor material and the lubrication effect provided by the fluid. Within biomechanically relevant condition, coefficient of friction may be sensitive to sliding speed indicating that the lubrication effect of the fluid may be within the mixed lubrication regime of the Stribeck effect (Beschorner, Redfern et al., 2007). Within this regime, the normal force is shared by the asperities and the fluid. Increasing speed within this region causes a transfer of load from the asperities to the fluid, which is the primary reason for a decrease in coefficient of friction within this region.

This paper introduces a computational friction model that includes sliding speed, normal force, geometry of shoe material, material properties of shoe material, roughness of shoe material, and the contaminant. As a starting point, a simple pin-on-disk setup was modelled using a shoe material for the pin and a floor material for the disk. The model is compared against experimental data to evaluate its effectiveness.

Methods

The mixed lubrication model consists of an asperity contact model and a fluid flow model. A multi-scale modelling approach that includes hydrodynamic pressure and contacting asperities is used. The inputs to the model are the shape of the pin (shoe material), roughness of the pin, normal load, viscosity, shoe material properties, sliding speed, asperity COF and fluid COF. The model outputs the percent of the total normal load that is supported by the contacting asperities and the percent of the normal load that is supported by the fluid. The floor was assumed to be rigid and flat. Squeeze film effects were not considered due to the difficulties of capturing this effect with experimental data.

Iterative methods were implemented to solve for film thickness. The condition used to solve the system is that the sum of the forces from both the asperity and fluid model had to equal the total force, applied during experiments.

Experimental setup

All experiments were performed using a pin-on-disk apparatus where floor material representing the disk rotated relative to shoe material representing the pin. The pin was a cylindrical piece (6.4 mm diameter and 3.2 mm thickness) of polyurethane material cut directly from the sole of a shoe. The floor material was a square 51×51 mm piece of commercially available vinyl tile. A diluted glycerol contaminant (50% glycerol/50% water) was generously applied to the floor surface during all trials. Applied normal force was 2.4 N resulting in an average pressure of approximately 73 kPa. Six sliding speeds were tested at levels of: 10, 20, 50, 98, 160 and 200 mm/sec. The average and standard deviation of 5 trials were determined as reference data for the model.

Model inputs

Wherever possible, inputs to the model were measured parameters based on the experimental setup. Physically measured inputs were the normal force, sliding speed, shape of the pin, roughness of the pin, viscosity of the fluid and shoe material elastic modulus. The shoe material was modeled as a 2nd order polynomial fit to profilometry scans:

$$h(x,y) = a * (x^2 + y^2) \tag{1}$$

where, h is the height as a function of the x and y locations on the surface of the pins, and a regression parameter (a).

Roughness of the pin was determined by calculating RMS roughness (R_q) from profilometry scans that were high-pass filtered. Viscosity was measured with a viscometer (Brookfield DV-E) and elastic modulus of the shoe material was determined from compression tests of the shoe material (Instron Material Testing).

Contacting asperity modeling

Asperities were modeled using a Winkler volume pixel (voxel) approach (Terrell, Kuo et al., 2005). The surface of the pin was represented as a pattern of square cuboids, which acted as independent springs. Asperity height was determined by superimposing random asperity heights based on the roughness of the pin onto the pin curvature as determined by Equation 1. The whole pin was offset using the pin clearance/interference distance (d_o):

$$h_{\text{asp}}(x,y) = a(x^2 + y^2) + N(0, R_q^2) + d_0 \tag{2}$$

A voxel size of $64 \times 64 \mu$m was used for a total of 7860 voxels.

Asperities are considered in contact if h is below 0 and the strain on each asperity is calculated from the interference of the pad into the floor and from the thickness

of the pin. Stress on each of the asperities is determined with the stress-strain relationship determined during the compression experiments. Force bourn by the asperities is determined from contact stress integrated over the area.

Hydrodynamic modeling

Hydrodynamic pressure distribution across the surface was determined by implementing numerical methods to solve Reynolds Equation.

$$\frac{\partial}{\partial x}\left[\phi_x \frac{h^3}{\eta}\frac{\partial p}{\partial x}\right] + \frac{\partial}{\partial y}\left[\phi_s \frac{h^3}{\eta}\frac{\partial p}{\partial y}\right] = 12\bar{U}\frac{\partial h}{\partial x} + 12\frac{\partial h}{\partial t} \tag{3}$$

Couette flow factors developed by Patir and Cheng are introduced, ϕx and ϕs, to adjust for flow about asperities due to surface roughness (Patir and Cheng, 1977). Flow factors are a function of surface roughness and surface separation.

Film thickness across the surface of the pin was determined by the shape of the pin (Equation 1) and the interference distance (h_o), a function of d_0, (Equation 2). An asymmetric film thickness was defined to reduce sub-ambient pressures at the trailing edge of the pin. The film thickness for the leading edge of the pin was:

$$h(x,y) = a * \left(x^2 + y^2\right) + h_o \tag{4}$$

The film thickness on the opposite side of fluid entrainment (x > 0) was defined as:

$$h(x,y) = a * \left(y^2\right) + h_o \tag{5}$$

For the numerical solution of the Reynolds equation, the film thickness was meshed with square elements of size 0.13×0.13 mm (1976 total elements).

Friction model

The total friction for the surface was determined as a function of the amount of load bourn by the asperities and the fluid. A linear rule of mixtures was defined such that the friction coefficient for the entire sample was between the limits set by the coefficient of friction for interacting asperities (COF_{asp}) and the coefficient of friction of the fluid (COF_{fl}). The total COF (COF_t) was determined as a function of the proportion of the force bourn by the asperities (F_{asp}/F_t), the proportion of the force bourn by the fluid (F_{fl}/F_t), COF_{asp} and COF_{fl}

$$\mathrm{COF_t} = \frac{F_{\mathrm{asp}}}{F_{\mathrm{t}}}\mathrm{COF_{asp}} + \frac{F_{\mathrm{fl}}}{F_{\mathrm{t}}}\mathrm{COF_{fl}} \tag{6}$$

The values of COF_{asp} and COF_{fl} were chosen to approximately match the model endpoints (i.e. highest and lowest speed) to the experimental data.

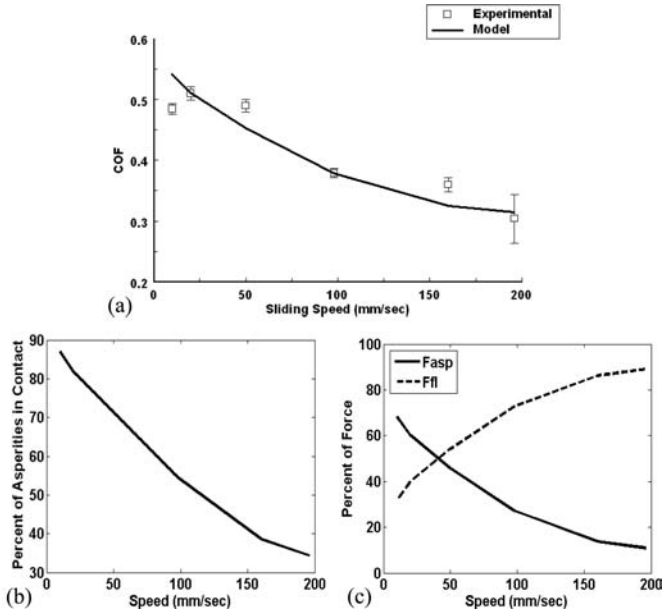


Figure 1. As sliding speed increases, COF decreases (a), fewer asperities are in contact (b) and force is transferred from the asperities (Fasp) to the fluid (Ffl) (c).

Results, Discussion and Conclusions

The model shows good agreement with experimental data (Figure 1a). Experimental data show a non-linear decrease in COF, similar to the mixed-lubrication regime of the Stribeck curve, which is replicated in the model. In addition, asperity and fluid COF values needed to fix the endpoints of the curve are a very reasonable 0.65 and 0.25, respectively. Model and experimental data showed the best agreement at higher sliding speeds. While more complex interactions may be occurring at low speeds, a simple linear rule of mixtures was shown to be an effective friction model.

The computational model provides other outputs, which may be important to understanding the physical interaction within the shoe-floor-contaminant interface. For example, the percentage of the load s by the asperities is shown to decrease as sliding speed increases as expected (Figure 1b). The percentage of asperities in contact is also shown to decrease with increasing speed (Figure 1c).

The developed shoe-floor-contaminant computational friction model provides the opportunity to analyze shoe-floor-contaminant friction at a level not possible with slip-testing devices (at the asperity and fluid flow level). Certain challenges still exist before the developed model can be used as an evaluative or design tool for shoe and floor surfaces. The primary improvement of the model to achieve this goal will be including features of the heel (i.e. the shape of the contact area and tread). Another important improvement would be to add time-dependency to the Reynolds equation by including the squeeze-film term.

References

Beschorner, K.E., M.S. Redfern, et al. 2007, Effects of slip testing parameters or measured coefficient of friction, *Applied Ergonomics* **38**: 773–780.

Burnfield, J. and C. Powers. 2005, Prediction of slip events during walking: An analysis of utilized coefficient of friction and available slip resistance, *Ergonomics* **22** (1): 82–88.

Chang, W.R. 2001, The effect of surface roughness and contaminant on the dynamic friction of porcelain tile, *Appl Ergon* **32** (2): 173–84.

Chang, W.R., R. Gronqvist, et al. 2001a, The role of friction in the measurement of slipperiness, Part 2: survey of friction measurement devices, *Ergonomics* **44** (13): 1233–61.

Chang, W.R., R. Gronqvist, et al. 2001b, The role of friction in the measurement of slipperiness, Part 1: friction mechanisms and definition of test conditions, *Ergonomics* **44** (13): 1217–32.

Chang, W.R. and S. Matz. 2001, The slip resistance of common footwear materials measured with two slipmeters, *Appl Ergon* **32** (6): 549–58.

Courtney, T.K., G. S. Sorock, et al. 2001, Occupational slip, trip, and fall-related injuries–can the contribution of slipperiness be isolated?, *Ergonomics* **44** (13): 1118–37.

Gronqvist, R., S. Matz, et al. 2003, Assessment of shoe-floor slipperiness with respect to contact-time-related variation in friction during heel strike, *Occupational Ergonomics* **3**: 197–208.

Kim, I.-J. 2004, Development of a new analyzing model for quantifying pedestrian slip resistance characteristics: Part I-Basic concepts and theories, *International Journal of Industrial Ergonomics* **33** (5): 395–401.

Manning, D.P., C. Jones, et al. 1998, The Surface Roughness of a Rubber Soling Material Determines the Coefficient of Friction on Water-Lubricated Surfaces, *Journal of Safety Research* **29** (4): 275–283.

Patir, N. and H.S. Cheng. 1977, Average flow model for determining effects of three-dimensional roughness on partial hydrodynamic lubrication, *American Society of Mechanical Engineers (Paper)* (77-Lub-17): 6.

Redfern, M.S. and B. Bidanda. 1994, Slip resistance of the shoe-floor interface under biomechanically-relevant conditions, *Ergonomics* **37**: 511–524.

Terrell, E.J., M. Kuo, et al. 2005, An Approach to Modeling Particle-Based and Contact-Based Wear in CMP, *MRS Bulletin* **5** (V30).

COMPARISON OF SITE-BASED FRICTION MEASUREMENT AND SLIP POTENTIAL

Kevin Hallas, Rick Houlihan, Rob Shaw & Steve Thorpe

Health and Safety Laboratory, Harpur Hill, Buxton, SK17 9JN

This paper is a review of slip resistance data generated over the last 10 years by the Health and Safety Laboratory in situations where concerns regarding the slipperiness of flooring have been raised. This allows the measured pendulum Coefficient of Friction (CoF) to be related to the required CoF predicted by Building Research Establishment. The data presented here are based on over 4300 tests, covering a range of floor surfaces on site. The HSL data suggest a higher level of concern about floors in the moderate Pendulum Test Value (PTV) range than BRE predictions would suggest. Some possible explanations are offered.

Introduction

The Health and Safety Laboratory (HSL) has been involved in the investigation of pedestrian slip accidents for 10 years. This paper is a review of all the slip resistance data that has been generated by HSL and collates the data into a single summary report. This allows the measured pendulum Coefficient of Friction (CoF) to be related to the required CoF predicted by Building Research Establishment (BRE)(Pye & Harrison 2003).

Active and archived reports for all aspects of slip assessment carried out by HSL were systematically reviewed. For each investigation, the reported pendulum test values were input into a spreadsheet to allow collation and analysis.

Method

HSL routinely uses the Pendulum Test to assess the slipperiness of floor surfaces where an accident has occurred or where a Health & Safety Executive (HSE) or Local Authority (LA) inspector has subjective concerns about the slipperiness of a floor surface. Correct characterisation of the floor surface allows meaningful consideration of controls that may be implemented to minimise the likelihood of a slip accident occurring.

For this study, only the data generated using Pendulum Test Slider 96, also known as Four-S rubber (Standard Simulated Shoe Sole) has been considered. The material was developed by the United Kingdom Slip Resistance Group (UKSRG) in collaboration with RAPRA Technology Ltd. as a standard material to assess floor surface

Table 1. Friction requirements for pedestrians (BRE).

Risk. 1 in:	Minimum CoF
1,000,000	0.36
100,000	0.34
10,000	0.29
200	0.27
20	0.24

slip resistance. This slider is widely used to simulate a typical pedestrian shoe heel. Slider preparation was carried out in accordance with the UKSRG Guidelines. The Pendulum Test was calibrated by the British Standards Institution, in accordance with BS7976-3:2002.

Pendulum tests are normally carried out in 3 directions to account for any directionality in the floor surface materials, in both dry and water-wet conditions. Each test consisted of 8 swings, 3 for conditioning and then the 5 test values, the mean of which is the reported value. Data reported from tests undertaken prior to October 2005 are the median rather than mean of the last 5 of 8 readings. Some floor surfaces show a significant difference in PTV between the different directions. For this study, each test result was used rather than taking an average of the 3 values for each surface. This generated a larger data set that should contain more accurate data about the surfaces tested.

Where floors are found to have lower than expected friction under dry conditions (usually less than 55 PTV), cleaning was undertaken by the site investigator using detergent solution. Further pendulum testing was then undertaken to measure the friction under clean conditions.

Guidelines

The BRE data established the required CoF for normal walking. These are used to quantify the risk of a pedestrian slip accident in a normal walking situation, i.e. at a moderate pace, on a level surface, not rushing, pushing, pulling, or turning. From the sample used in the BRE study, a distribution of friction requirements was established, as shown in Table 1.

The UKSRG has been responsible for furthering the understanding of pedestrian slip matters. The group has published a set of guidelines as a methodology for conducting slip resistance testing in relation to pedestrian slipping. The guidelines describe the use of the Pendulum Test, the results from which are reported as a Pendulum Test Value (PTV), which is approximately 100 times the CoF.

Friction requirements vary from one pedestrian to another. A pedestrian slip incident is more likely as the available CoF drops below the required CoF of that individual. In order to apply this to a population, the UKSRG have simplified the above into bands of low, moderate and high slip potential, as shown in Table 2.

Table 2. Extract from UKSRG Guidelines, Issue 3, 2005.

Pendulum Test Value	Slip Potential
≤24	High
25–35	Moderate
≥36	Low

All data was generated in accordance with the UKSRG guidelines applicable at the time, which gives rise to slight differences in the slider conditioning. For this study, the latest version of the guidelines (UKSRG, 2005) was used to interpret the data.

Results

General trends

Most of the data points presented here relate to floors in workplaces where the health and safety inspector is concerned about the slipperiness of the floor. This conclusion is often reached during a walk through the premises, which may be in wet or dry conditions depending on the individual situation. This subjective assessment may be informed by near-miss and accident data held locally.

The relatively high number of floors examined in the region 30–40 PTV suggests that some people find these floors slippery in normal use, which may be due to other factors that increase the friction demand compared to the UKSRG baseline. Where accidents occur at higher PTV, the slipperiness of the floor is probably not the only factor contributing to the accident. Factors such as rushing, turning and manual handling increase the friction demanded of the floor. In these situations, the risk assessment needs to take account of these activities, as PTV should not be considered in isolation. It should also be noted that the Slider 96 rubber used in the test represents a moderate shoe sole, and therefore some shoes will be more slippery than this.

Some of the data represents accident sites, where a pedestrian has slipped in an uncontrolled manner and suffered an injury (either major or over 3 day according to RIDDOR classifications). Since not all slips result in injuries, the actual reported accident rate can be relatively low. This may be further compounded by the embarrassing nature of these accidents, resulting in low near-miss reporting. The floors involved in these incidents make up the bulk of those in the high slip potential category; PTV ≤ 24.

The range of test data presented shows that flooring is available with a very wide range of slip resistance, allowing a suitable floor to be selected for most situations. Where this is not possible, specialist anti-slip footwear can be used to increase the level of friction available to a pedestrian, compared with that measured using Slider 96 rubber.

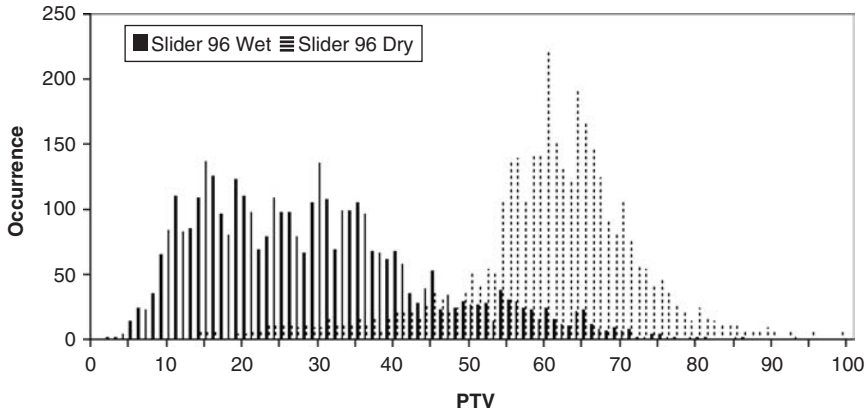


Figure 1. All pendulum test data.

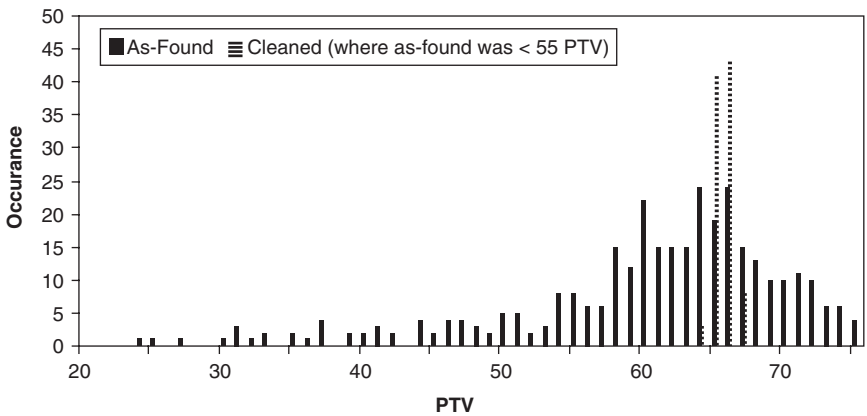


Figure 2. As-found dry & cleaned pendulum test data.

Cleaning

Figure 2 shows a wide range of pendulum test results, which is surprising for a range of apparently dry floors. Usually when slips occur, some form of contamination is present between the shoe and the floor surface. This is most commonly water, though grease or dry contaminants such as flour, talc, fibres or dust can significantly reduce available friction.

Where the as-found dry PTV was less than 55, the floor was cleaned by the investigator using a detergent solution. The effect on the PTV in dry conditions is shown in Figure 2. Where surface contamination has been removed by cleaning, the dry slip resistance is in the region 64–67, as would be expected.

The effectiveness of floor cleaning in increasing the available friction of a dry floor illustrates the importance of appropriate floor cleaning in controlling slip risk. The improvement in slip resistance in wet conditions is often less significant.

However, some examples were noted where a build up of contamination had altered the surface characteristics of the floor, thereby reducing the wet slip resistance. When cleaned effectively, the slip resistance was restored.

Comparison to BRE predictions

The data largely supports the BRE work, the main bulk of the floors which gave cause for concern fall into the CoF range identified by BRE. The HSL data is not pure accident data, as this could not be extracted from the available information. If the data were solely from accidents, direct comparison with the BRE data would be undermined by other factors that increase or decrease the likelihood of a slip in a given situation.

The BRE study was not intended to be used to predict the likelihood of accidents on a floor with a given slip resistance. Consideration would need to be given to the user group, floor condition and exposure to any contamination. For example, where the flooring is slippery when wet, but is rarely wet in use, the overall slip potential may be lower than when the PTV is considered in isolation.

References

P.W. Pye, H.W. Harrison. 2003, *BRE Building Elements: Floors and flooring – performance, diagnosis, maintenance, repair and the avoidance of defects*, BRE Report 460, 2003 (Buildings Research Establishment).

British Standards Institution, BS 7976-3:2002 *Pendulum testers. Method of calibration*, 2002, British Standards Institution, ISBN 0 580 40146 4.

United Kingdom Slip Resistance Group. 2005, *The Measurement of Floor Slip Resistance – Guidelines Recommended by the United Kingdom Slip Resistance Group*, Issue 3 (UKSRG)

NATURE OF THE SHOE WEAR: ITS UNIQUENESS, COMPLEXITY AND EFFECTS ON SLIP RESISTANCE PROPERTIES

In-Ju Kim[1] & Hisao Nagata[2]

[1]*School of Health Sciences, University of Canberra, Canberra, Australia*
[2]*National Institute of Occupational Safety and Health, Tokyo, Japan*

One of the main tribological characteristics on slip resistance properties seems to be a change of shoe sole/heel surfaces, where the surface topography is rapidly modified and continuously produces wear particles and films during every single sliding. Despite the seriousness of this aspect, its features have not been widely explored in this research area. This study was, therefore, focused on identifying wear mechanics and mechanisms of the shoes and their effects on slip resistance properties. In order to investigate the main aim, four shoes were rubbed against two floor surfaces. Initial and worn heel surfaces were quantitatively and qualitatively examined by surface texture measurements and microscopic works, respectively. Measurement results from both the methods clearly showed that adhesive, abrasive. ploughing and fatigue were identified as main wear mechanisms and resulting effects on the slip resistance properties. This study also found that wear developments of each heel surface showed unique patterns, complex interactions and considerably affected slip resistance performances.

Introduction

Slip resistance properties between the footwear and underfoot surface are a great importance for evaluating slip and fall hazards and have been measured as a form of coefficient of friction (COF). Hence, knowledge about friction demands and friction available have been recognized as a prime factor for the fall safety measures (Rowland *et al*., 1996; Manning *et al*., 1998). Since the COF index was adopted to determine whether a slip is to occur, however, there have been great controversies in the accuracy, effectiveness and interpretation of the results. One of the most arguable characteristics in the analysis of slip resistance properties would be a material property of the shoes, where the surface topography rapidly changes and continuously produces wear particles during every single sliding. This aspect indicates that wear development of the heel surface seems to be a major concern and considerably affect slip resistance properties. Surprisingly, however, there has been little analysis on how friction induced wear developments of the shoe heels would affect the slip resistance properties. This study was, therefore, aimed to improve our understanding on friction induced wear mechanics and mechanisms of the shoe surfaces during dynamic friction measurements and its effects on slip resistance

performances. In order to investigate the wear behaviors, four shoes were rubbed against two floor surfaces. The worn surfaces of each shoe heel were quantitatively and qualitatively examined using surface roughness parameters and a scanning electron microscope, respectively. Based on the overall results, an inclusive wear model was suggested for the shoe heels.

Experiments

Dynamic friction tests

Slip resistance measurements were conducted by a pendulum type friction tester that simulates the movements of human foot and the forces applied to the underfoot surface at the moment of heel strike and initial sliding. Normal load of 400 N was applied to the shoe heel and its sliding speed was controlled at a relative speed of 25 cm/sec during the tests. Two polyurethane soles (S1 & S2) and two polyurethane/nitrile compounding soles (S3 & S4) were rubbed against ceramic (F1) and terracotta (F2) plates under clean and dry conditions. Tests were designed by five sub-groups (5, 10, 20, 30 and 50 times of consecutive rubbings) for a systematic analysis.

Surface texture measurements

To obtain full morphological information and detect wear developments, initial and rubbed surfaces of each shoe heel were thoroughly examined by quantitative and qualitative approaches. For the quantitative measures, a number of surface texture parameters were measured by a laser scanning confocal microscope (LSCM, Bio-Rad Microscience). The selected parameters were based on previous studies (Kim and Smith, 2000 & 2003; Kim *et al.*, 2001). Initial and rubbed surfaces were broadly examined by a stereo scanning electron microscope (SEM, XL 30, Philips) in order to facilitate overall wear behaviors such as wear formations, flows and patterns occurring at the floor surfaces during dynamic friction measures.

Results

Slip resistance performances

Figure 1 plots overall slip resistance results amongst the four shoes and two floor surfaces as a function of break-in testing times. As shown, the slip resistance results could be divided into two main categories. That is, the slip resistance performances were significantly dependent upon the shoe materials chosen. Dynamic friction coefficient (DFC) values between the S1 and S2 against the ceramic plate (F1) were largely decreased during the whole tests. Although both the shoes showed high slip resistance capabilities (average DFC > 0.8) until the stage 3, their overall slip resistance performances were drastically dropped (about 30 to 40%) after the whole tests. On the other hand, the other two shoes (S3 and S4) against the terracotta

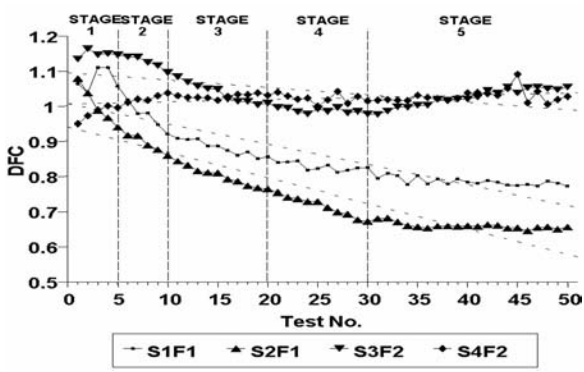


Figure 1. DFC results between the four shoes and two floors under clean and dry conditions.

plate (F2) showed higher slip resistance performances than the S1 and S2 during the entire tests. At the stage 1, the S3 showed the highest slip resistance performance whilst the S4 showed the lowest slip one amongst the shoes tested. The DFC results of the S3 continuously, but slowly decreased until the stage 3. After the stage 3, the DFC reduction of the S3 was stabilized in the stage 4 and started to increase slightly from the beginning of the stage 5. The DFC values of the S4 were also gradually increased from the beginning of the rubbing and kept high levels of DFCs throughout the whole test periods. It was also found that the DFC values of the S1F1 and S2F1 showed negative correlations during the entire tests. This result indicated a fact that single polyurethane-based shoes were less superior than the polyurethane/nitrile compounding ones. However, one conclusive finding was that changes of the initial surface conditions of the shoes would be evident and their variations seemed to significantly affect the extent of slip resistance properties.

Surface roughness parameters

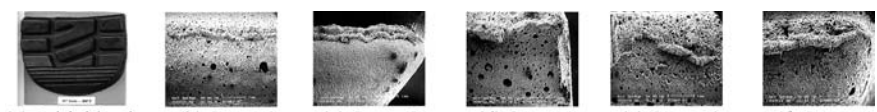
Table 1 summarizes the measured results of surface roughness parameters of each shoe before and after the tests. All the roughness parameters were largely changed during the entire tests. The roughness parameters were largely increased after the initial five rubbings. This increment seemed to be caused by abrasive wear so that the initial layers of the heel surfaces were roughened. With further rubbings, however, the roughness parameters were gradually decreased. This result was due to the removal of the majority of asperity crests in the heel surfaces.

Wear observation

Figs. 2 to 5 show micrographs of the initial and worn surfaces of the four shoes. Initial heel surfaces of each shoe illustrate clean tread patterns without any protruding shape whereas the S2 shows shiny and a number of small porous. After the entire tests, however, the heel surfaces were severely damaged so that their distinctive

Table 1. A summary of the measurement results of roughness parameters for the four shoes.

Shoe type	Test stage	Surface roughness parameters					
		Ra	*Rq*	*Rtm*	*Rpm*	*Rvm*	Δa
Shoe No. 1	Initial	3.048	3.877	17.038	8.198	−8.840	0.082
	Stage 1	6.671	7.997	30.398	17.704	−12.694	0.092
	Stage 2	5.023	6.297	26.995	14.716	−12.279	0.069
	Stage 3	4.529	5.781	25.303	13.430	−11.873	0.064
	Stage 4	4.094	5.051	20.499	11.799	−8.700	0.059
	Stage 5	3.722	4.858	17.860	10.919	−6.941	0.056
Shoe No. 2	Initial	2.860	3.708	15.214	8.228	−6.896	0.069
	Stage 1	4.687	6.197	21.885	13.199	−8.686	0.089
	Stage 2	4.208	5.276	19.388	10.768	−8.620	0.065
	Stage 3	3.521	4.365	17.348	9.323	−8.025	0.056
	Stage 4	3.201	4.045	15.102	7.894	−7.208	0.052
	Stage 5	2.774	3.495	13.665	7.770	−5.895	0.045
Shoe No. 3	Initial	2.897	3.694	14.400	8.671	−5.729	0.124
	Stage 1	4.824	6.165	13.860	5.649	−8.211	0.044
	Stage 2	5.250	6.849	16.650	6.143	−10.507	0.050
	Stage 3	5.459	7.267	17.820	6.813	−11.007	0.051
	Stage 4	5.567	7.375	19.620	8.656	−10.964	0.063
	Stage 5	5.669	7.565	21.960	8.849	−13.111	0.065
Shoe No. 4	Initial	4.990	6.437	25.247	16.698	−8.549	0.038
	Stage 1	7.162	8.079	18.818	9.448	−9.370	0.106
	Stage 2	6.494	7.508	16.677	8.321	−8.256	0.085
	Stage 3	5.883	7.189	14.781	7.654	−7.127	0.081
	Stage 4	5.095	6.197	14.400	6.986	−7.414	0.067
	Stage 5	5.604	6.907	14.577	7.279	−7.298	0.065



(a) Initial heel (b) Stage 1 (c) Stage 2 (d) Stage 3 (e) Stage 4 (f) Stage 5

Figure 2. Micrographs of the wear tracks for the shoe No. 1.

macro- and micro-tread patterns suffered massive changes. The worn surface of the S1show a number of paralleled tearing traces and micro-layered surface textures. Width and depth of the tearing traces were gradually increased with the number of rubbings. In the case of the S2, its original micro-porosities were broken up and formed lots of cavities on its surface layer. This event continuously created new surfaces so that the width and depth of tearing traces were gradually increased with the number of rubbings. Like the case of the S1, the S3 showed a number of paralleled tearing traces and micro-layered surface textures. However, as found in the S2, the worn surface of the S3 was heavily characterized by scratching traces and micro-layered surface textures. In the case of the S4, the heel surface was heavily deformed in the direction of sliding, with evidences of severe abrasion, ploughing,

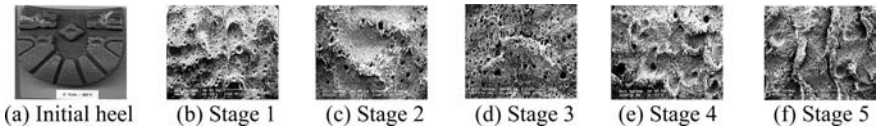


Figure 3. Micrographs of the wear tracks for the shoe No. 2.

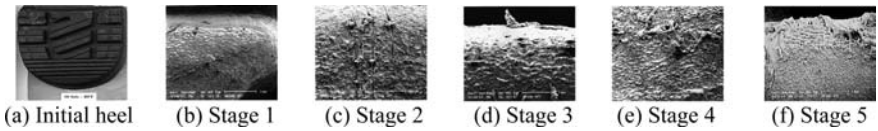


Figure 4. Micrographs of the wear tracks for the shoe No. 3.

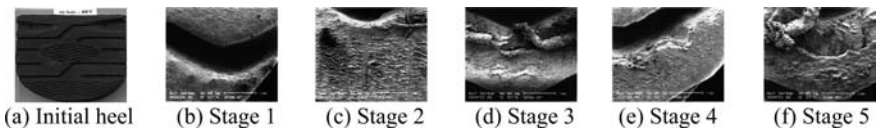


Figure 5. Micrographs of the wear tracks for the shoe No. 4.

and fatigue. Although its basic morphological character was quite similar with the case of the S3, the overall wear mode was severe than the S3.

Discussion

Wear developments of the heel surfaces were thoroughly examined by the micrographs taken with a SEM. As found in the micrographs, wear products from each heel surface showed totally different shapes, sizes and patterns. This result would be a reflection of the fact that the heel surfaces were heavily involved by severe plastic deformations, ploughing and micro-cuttings from abrasion and fatigue of sharp asperities of the flooring surfaces. The wear patterns were widely different and unique according to the shoe types. As described in the experiments, the S1 and S2 were single polyurethane based shoes so that they seemed to be rapidly worn and consisted than the S3 and S4 ones which were polyurethane/nitrile compounding soles. As a result, their wear particles shaped short, round and lumpy style because they should produce lots of particles rapidly and continuously. Whilst in the case of the S3 and S4, they were more resilient and not easily torn so that their wear behaviors were less active and slowly proceeded than the S1 and S2. As a result, the wear patterns showed thin and long striate shapes. Based on the above findings, a comprehensive wear model for the shoes is suggested in Table 2.

Table 2. A summary of an inclusive wear model for the shoes.

Wear mechanisms		Brief descriptions	Surface effects
Adhesive	Adhesive wear	Adhesive wear takes place when relative movements induce breakages of the junction inside the heel rather than at the interface.	Top areas of asperities break off the heel surface
	Material pick-up	Deposits of the flooring material are added to the heel surface peaks.	
Abrasion	Abrasive wear	Abrasive wear is an event by displacement of polymeric materials from the heel in relative sliding motion caused by ploughing effects of hard protuberances on the floor surface.	Polymeric materials are removed, leaving scratches
	Ploughing	Ploughing takes place when abrasion does not include any material removals from the heel surface but only relocation of material.	A furrow and a bulge are formed on the heel surfaces
Fatigue	Fatigue wear	Fatigue wear occurs when the heel surface has been exposed to a large number of alternating tensile and compressive stresses.	Cracks are initiated from the valley areas of the heel surface

Conclusion

In order to investigate wear formations of the shoes and their effects on slip resistance properties, four shoes were rubbed against two commercially available floor surfaces. The initial and worn surfaces of each shoe heel were comprehensively measured and examined by quantitative and qualitative methods. The whole results clearly found that plastic deformations and cuttings by relatively sharp asperities and fatigue by more rounded asperities were identified as major wear mechanisms. Although a fatigue wear processes was not clear at this point, direct microscopic confirmations of crack formations and propagations in the heel surfaces provided strong evidences to ascertain this aspect. In most rubbings, however, this study found that wear processes were a combination of the above two mechanisms, but if any plastic deformations occurred at all, their effects were likely to dominate the magnitude of the total wear.

References

Kim, I.J. & Smith, R. 2000, Observation of the floor surface topography changes in pedestrian slip resistance measurements. *International Journal of Industrial Ergonomics*, **6**(6), 581–601

Kim, I.J., Smith, R. & Nagata, H. 2001, Microscopic observations of the progressive wear on the shoe surfaces that affect the slip resistance characteristics. *International Journal of Industrial Ergonomics*, **28**(1), 17–29

Kim, I.J. & Smith, R. 2003, A critical analysis of the relationship between shoe-floor wear and pedestrian/walkway slip resistance. In M. Marpet & M. Sapienza

(eds.), *Metrology of Pedestrian Locomotion and Slip Resistance*, (ASTM STP 1424, ASTM International), 33–48

Manning, D.P., Jones, C., Rowland, F.J. & Roff, M. 1998, The surface roughness of a rubber soling material determines the coefficient of friction on water-lubricated surfaces. *Journal of Safety Research*, **29**(4), 275–83

Rowland, F.J., Jones, C. & Manning, D.P. 1996, Surface roughness of footwear soling materials: relevance to slip-resistance. *Journal of Testing and Evaluation*, **24**(6), 368–76

LINEAR REGRESSION MODELS OF FLOOR SURFACE PARAMETERS ON FRICTION AT SHOE-FLOOR INTERFACE

Wen-Ruey Chang[1], Simon Matz[1], Raoul Grönqvist[2] & Mikko Hirvonen[2]

[1]*Liberty Mutual Research Institute for Safety, Hopkinton, MA 01748, USA*
[2]*Finnish Institute of Occupational Health, FIN-00250 Helsinki, Finland*

In preventing incidents of slips and falls, friction is widely used as an indicator of surface slipperiness. Surface feature parameters, including surface roughness and waviness, were shown to influence friction by correlating individual surface parameters with the measured friction. However, investigations of the collective input from multiple surface feature parameters as a predictor of friction were not available in the literature. The goal of this study was to develop regression models between the floor surface feature parameters and the measured friction. The dynamic friction between quarry tiles and Neolite using three different mixtures of glycerol and water as contaminants at the interface was measured. The surface texture was quantified with various surface roughness and surface waviness parameters using three different cut-off lengths to filter the measured profiles for obtaining the surface feature parameters. The regression models indicate that the surface parameters are dominant factors in predicting the measured friction coefficient in most of the glycerol concentrations and cut-off lengths in this experiment. The surface roughness parameters contributed significantly to the prediction of the measured friction coefficient when the glycerol concentration was 50% regardless the cut-off length being used. As the glycerol concentration increased to 70%, surface roughness parameters remained dominant at 2.5 and 8 mm cut-off lengths, but surface waviness dominated at the 0.8 mm cut-off length. As the glycerol concentration increased further to 85%, surface roughness parameters remained dominant at the 8 mm cut-off length, but surface waviness dominated at 0.8 and 2.5 mm cut-off lengths.

Introduction

Slips and falls are a serious problem. It was estimated that the annual direct cost of occupational injuries due to slips, trips and falls in the USA exceeded 6 billion US dollars (Courtney et al. 2001). Leamon and Murphy (1995) reported that falls on the same level accounted for 65% of claim cases and 53% of claim costs in total direct workers' compensation for occupational injuries due to slips and falls.

In preventing incidents of slips and falls, friction is widely used as an indicator of floor slipperiness. Surface texture plays a crucial role in determining friction although the contact interface between shoe and floor is complex (Chang et al., 2001a). While the friction measurement related to the problems in slips and falls remains controversial, measurements of surface texture could be a reliable and objective complement to the friction measurement (Chang et al., 2001a).

The majority of the published studies on the effects of surface textures on friction related to slip and fall incidents focused on surface roughness. As reported in the literature, surface roughness parameters R_a, R_{tm}, R_{pm}, Δ_a and R_k, as defined in Table 1, had strong correlations with the friction measured at the shoe and floor interface as summarized by Chang et al. (2004b).

In addition to surface roughness, surface waviness might play a role in surface friction. Chang et al. (2004a) investigated the correlation between surface waviness parameters and the measured friction. Their results indicated that some of the surface waviness parameters on liquid contaminated surfaces could have a stronger correlation with friction when a short cut-off length was used for the surface profile measurements or when the viscosity of the liquid contaminant was high. The cut-off length selected for the surface profile measurements with a profilometer is the filtering length that is used as a reference to properly capture important surface features as indicated by Whitehouse (1994). Among the surface waviness parameters evaluated as listed in Table 1, $W\Delta_a$, W_a and W_{tm} had a strong correlation with the measured friction.

In the previous studies on the effect of surface feature parameters on friction, the main focus was to investigate the relationship between individual surface feature parameters with the measured friction. Based on the results of linear correlation coefficients, surface parameters with higher correlation coefficients with friction were identified. However, the combined effects of these surface feature parameters on the measured friction were not investigated. The objective of this study was to investigate linear regression models of floor surface feature parameters, including surface roughness and waviness, on the friction between quarry tiles and Neolite. Through regression models, the surface feature parameters that contribute the most to the friction measured could be identified. The data used in the previous publications (Chang et al., 2004a and 2004b) were used as the basis in this investigation. The data for the surface feature measurements were collected with three cut-off lengths, while the friction was measured with three different glycerol concentrations.

Methods

The footwear material used in this experiment was Neolite, a common test material in the USA. The dynamic friction between Neolite and quarry tile was measured with an apparatus known as the slip simulator, developed by Grönqvist et al. (1989). A commercially available profilometer (Rank Taylor Hobson, Leicester, UK, model S3F) was used to measure the surface profiles of quarry tiles. Details

of the apparatus were previously described (Chang et al., 2004b). Glycerol mixed with water at 85%, 70% and 50% ratios by weight was used as contaminants in this experiment.

The movement of the slip simulator foot attempts to resemble a slip during a short time duration of contact between the shoe and floor surfaces right after a heel strike (Grönqvist et al., 1989). The traditional dynamic friction measures resistance to a steady state sliding, but the dynamic friction measured with the slip simulator, also called transition friction (Chang et al., 2001b), represents the available friction in a slip initiation. This type of dynamic friction is critical in determining whether a slip incident will occur, especially when contaminants are present (Chang et al., 2001b).

The heel of the shoe used in the friction measurements was replaced with a piece of Neolite liner which was lightly sanded near the heel tip to resemble a slightly worn shoe. Whenever there was a change in tile or glycerol concentration during friction measurements, the Neolite surface was sanded with 180 grit abrasive paper using an orbital sander while maintaining the worn shape in order to maintain a consistent surface condition.

The surface texture on the unglazed quarry tiles was systematically altered by sand blasting as described in the previous paper (Chang et al., 2004b). Eight tiles were prepared under each of five different sand blasting processes, resulting in a total of 40 tiles.

At the center of each tile, six parallel surface profile measurements, 10 mm apart, were made in the direction of friction measurements. The assessed length for the surface parameter calculations was 4 times the cut-off length. Three cut-off lengths of 0.8, 2.5 and 8 mm, a band width of 100, and the 2CR PC filter, a recursive filter with a phase correction, were used. Background information about these terminologies and surface profile measurements is available in the literature (Whitehouse, 1994).

The test shoe was mounted to the slip simulator with an artificial foot. For normal and shear force measurements at the shoe and tile interface, the tile was attached to a force plate. The horizontal velocity of the shoe was maintained at a constant speed of 0.40 m/s. The shoe, initially positioned slightly above the floor surface, was lowered onto the surface at a speed of 0.1 m/s. The data collection used in the final analyses began 100 ms after the heel touch-down, which was the instant when the normal force reached 100 N, for a time interval of 50 ms to average the normal and friction forces, friction coefficient, and sliding velocity. The applied normal force and a shoe-tile contact angle were 700 N and 5°, respectively. The validities of this test method and test conditions were summarized by Grönqvist et al. (1989).

The measurement sequence with glycerol concentrations and sand blasting processes was randomized. All of the 40 tiles were used repeatedly with different glycerol concentrations. The tiles were sequentially cleaned in an ultrasonic bath with water, and then wiped with a 50% ethanol solution right before the friction measurements. The amount of contaminant applied to the tile surfaces was the maximum allowed by the surface tension and the covered area was much larger than the area in contact with the shoe during friction measurements. Six repeated friction measurements were taken for each tile and glycerol concentration, but the first friction measurement was discarded in the subsequent analyses since the shoe

did not have contaminant on its heel area and might differ from the other five. The tiles were cleaned in the ultrasonic bath immediately following the friction measurements.

The average of five repeated friction measurements from each tile under each glycerol concentration was used to represent the measured friction from the tile and glycerol concentration. Similarly, the average of six surface profile measurements from each tile and cut-off length was calculated for each surface parameter.

The definitions of all surface parameters used in this study, including 20 surface roughness parameters and 14 surface waviness parameters, can be found in the previous publications (Chang et al., 2004a, 2004b). Graphic illustrations of most of the surface roughness parameters are available in the literature (Chang et al., 2001a). Illustrations of the equivalent surface waviness parameters are similar to those of surface roughness since similar definitions are applied to the surface waviness profiles. All these surface parameters were first divided into the two groups: surface roughness and waviness. Correlation between surface parameters was calculated separately in each group. Then, they were each subdivided into different categories of parameters calculated by surface vertical heights based on moments from the mean such as root mean square, skewness and kurtosis, peak to valley, peak height from the mean and valley height from the mean, and by horizontal dimensions such as average spatial wave lengths as well as hybrid parameters involving amplitude and spacing such as surface slopes. One or two parameters from each category in each group were put into the initial calculations of the linear regression models based on their individual correlation with the measured friction and also their correlation with other parameters in the same category. Additional surface parameters were eliminated from the models based on their levels of statistical significance and impact on the adjusted R^2 value of the models. This process was repeated several times in order to reduce the number of parameters in the final analyses due to a small data sample size available ($n = 40$). The final parameters used were R_a, R_{pm} and Δ_a in roughness and W_{tm} and $W\Delta_q$ in waviness, defined in Table 1.

Results and discussion

The results of the linear regression model in the last iteration for each glycerol concentration and cut-off length are shown in Table 2. Five surface parameters, three in surface roughness and two in surface waviness, remain in the final models. The p value of each surface parameter indicates that the surface roughness parameters such as Δ_a and R_{pm} contributed significantly to the prediction of the measured friction coefficient when the glycerol concentration was 50% regardless the cut-off length being used. As the glycerol concentration increased to 70%, surface roughness parameters remained dominant at 2.5 and 8 mm cut-off lengths, but surface waviness dominated at the 0.8 mm cut-off length. As the glycerol concentration increased further to 85%, surface roughness parameters remained dominant at the 8 mm cut-off length, but surface waviness dominated at 0.8 and 2.5 mm cut-off lengths. The transition from the surface roughness to surface waviness

Table 1. The definitions of surface roughness and waviness parameters.

Surface Roughness Parameters:

R_a – arithmetical average of surface heights or the center line average of surface heights
R_k – kernel roughness depth
R_{tm} – average of peak to valley height in each cut-off length
R_{pm} – average of the maximum height above the mean line in each cut-off length
Δ_a – arithmetical mean of surface slope

Surface Waviness Parameters:

W_a – arithmetical average of surface heights or the center line average of surface heights
W_{tm} – average of peak to valley height in each cut-off length
$W\Delta_a$ – arithmetical mean of surface slope
$W\Delta_q$ – root mean square of surface slope

Table 2. The *p* values for the surface parameters and R^2 in the regression models.

Cutoff mm	Glycerol concentration	Adjusted R^2	R_a	R_{pm}	Δ_a	W_{tm}	$W\Delta_q$
0.8	50	0.8424	0.9058	**0.0011**	**<0.0001**	0.4836	0.6246
	70	0.8101	0.1666	0.0993	0.1939	**0.0005**	**0.0163**
	85	0.4563	0.6247	0.9862	0.6888	**0.0105**	0.2908
2.5	50	0.8784	**0.0007**	**0.0005**	**<0.0001**	0.589	0.0969
	70	0.8822	0.7292	**0.0036**	0.164	0.5907	0.6063
	85	0.7243	0.0979	0.2157	0.1218	**0.0176**	**0.0293**
8	50	0.8232	**0.0101**	**0.0113**	**<0.0001**	**0.0029**	**0.0074**
	70	0.9516	0.7555	**0.0304**	0.0696	0.6376	0.4829
	85	0.7274	0.3918	0.3557	**0.0409**	0.6681	0.9857

parameters depends on the cut-off length. For the cut-off length of 0.8 mm, the surface roughness parameters dominated the model at 50% glycerol concentration, but the surface waviness dominated at 70% and 85% concentrations. For the cut-off length of 2.5 mm, the surface roughness parameters dominated at 50% and 70% glycerol concentrations, but the surface waviness dominated at 85% glycerol concentration. For the cut-off length of 8 mm, the surface roughness parameters dominated all three glycerol concentrations used in this experiment. As the cut-off length was increased, some of the components in the waviness were moved into the roughness which confirmed the results obtained with the regression models. At the 50% glycerol concentration, the surface parameters with a statistical significance remained significant as the surface parameters in the models were reduced in the final iterations. However, the surface parameters with a statistical significance changed as the surface parameters in the models were reduced in the final iterations at 70% and 85% glycerol concentrations.

The adjusted R^2 values in the regression models shown in Table 2 represent collective contributions from these surface parameters. Based on the R^2 values, the cut-off length of 2.5 mm, on average, could generate surface parameters that

could better predict the measured friction. The results were consistent with the results based on the correlation obtained from single surface parameters reported in the previous publications (Chang et al., 2004a, 2004b). Beside the condition of 85% glycerol concentration and 0.8 mm cut-off length, the R^2 values appear to remain quite high, greater than 0.71, for all the conditions. Also, as the number of parameters was reduced in the final analyses as more surface parameters were eliminated from the models, the R^2 values remained quite steady.

The results obtained in this study indicate that one cannot find a single surface parameter as an indicator of friction for all the glycerol concentrations and cut-off lengths in this experiment. In order to improve the output of the regression models, one will need to customize the surface parameters in the model for each glycerol concentration and cut-off length.

References

Chang, W. R., Grönqvist, R., Hirvonen, M. and Matz, S. 2004a, The effect of surface waviness on friction between Neolite and quarry tiles, *Ergonomics*, 47 (8), 890–906.

Chang, W. R., Hirvonen, M. and Grönqvist, R. 2004b, The effects of cut-off length on surface roughness parameters and their correlation with transition friction, *Safety Science*, 42 (8), 755–769.

Chang, W. R., Kim, I. J., Manning D. and Bunterngchit, Y. 2001a, The Role of Surface Roughness in the Measurement of Slipperiness, *Ergonomics*, 44 (13), 1200–1216.

Chang, W. R., Grönqvist, R., Leclercq, S., Myung, R., Makkonen, L., Strandberg, L., Brungraber, R., Mattke, U. and Thorpe, S. 2001b, The role of friction in the measurement of slipperiness, Part 1: friction mechanisms and definition of test conditions. *Ergonomics*, 44 (13), 1217–1232.

Courtney, T. K., Sorock, G. S., Manning, D. P., Collins, J. W. and Holbein-Jenny, M. A. 2001, Occupational slip, trip, and fall-related injuries – can the contribution of slipperiness be isolated? *Ergonomics*, 44 (13), 1118–1137.

Grönqvist, R., Roine, J., Järvinen, E. and Korhonen, E. 1989, An apparatus and a method for determining the slip resistance of shoes and floors by simulation of human foot motions, *Ergonomics*, 32 (8), 979–995.

Leamon, T. B. and Murphy, P. L. 1995, Occupational slips and falls: more than a trivial problem, *Ergonomics*, 38, 487–498.

Whitehouse, D. J. 1994, *Handbook of Surface Metrology* (Institute of Physics Publishing, Bristol, UK).

TRANSPORT

COMBINED POSTURE AND VIBRATION INVESTIGATIONS AMONG OCCUPATIONAL DRIVERS

Nastaran Raffler, Ingo Hermanns, Rolf Ellegast & Siegfried Fischer

Institute for Occupational Health and Safety of the German Social Accident Insurance, BGIA, Sankt Augustin, Germany

Musculoskeletal diseases or permanent damages can be caused by simultaneous workloads as high Whole Body Vibration (WBV) exposure and awkward body postures. CUELA (Computer-assisted registration and long-term analysis of musculoskeletal load) measuring technique for seated occupations has been used to investigate the combined exposures among occupational forklift and automobile drivers in a pilot study. Although the WBV exposures were similar, the combined exposure analysis resulted in higher risk for the forklift driver than the automobile driver. Hereby the influence of the body posture on the WBV is clearly demonstrated.

Introduction

The additive workload of the body posture to the WBV exposure has recently become more important in the occupational research. Several observational methods as RULA, OWAS and video taping have been applied to investigate the influence of the body posture on WBV exposure (Hoy *et al*, 2005). However, these methods provide rarely objective assessment about continuous movements of the subject by means of body angles. The CUELA is used for motion and posture capture in field analysis since several years (Ellegast and Kupfer, 2000). A new version of the system is specially designed for analysis of seated postures in combination with WBV measurements. In this pilot study the combined posture and vibration exposures among occupational forklift and automobile drivers have been exemplary measured and analysed.

Methods

Combined WBV and postural work load of a forklift driver (male, 1.87 m, 100 kg) has been investigated during his routine occupational operations. As reference, an automobile driver (male, 1.81 m, 92 kg) has been examined by the combined exposures during his drive. Both subjects were in good health. The observed tasks are summarised in Table 1.

WBV measurements were conducted in three orthogonal axes, according to ISO 2631-1 (1997), DIN EN ISO 8041 (2005). Vibration acceleration was measured

Table 1. Occupational tasks of the investigated vehicles.

Automobile (Ford Mondeo)		Forklift (Kalmar)	
Task	Measurement duration	Task	Measurement duration
Stopping/idling	00:02:53 (2%)	Transporting and driving, backward	00:23:40 (28%)
Driving forward (city traffic)	00:24:05 (20%)	Positioning the fork	00:23:35 (27%)
Driving forward (highway)	01:34:35 (78%)	Transporting and driving, forward	00:38:59 (45%)
Total tasks	**02:01:33 (100%)**	**Total tasks**	**01:26:14 (100%)**

Figure 1. CUELA measuring system for seated subjects.

by two sensors on the seat surface and at the seat mounting point (sampling rate 48 KHz). In order to combine vibration and postural workloads instantaneously, the vibration exposure is considered as the vector sum a_v, calculated by Eq. 1 of the frequency weighted accelerations defined by health criterion.

$$a_v = \sqrt{(1.4a_{wx})^2 + (1.4a_{wy})^2 + a_{wz}^2} \tag{1}$$

Due to the lack of standards concerning the evaluation of vector sum, three categories are suggested to assess the instantaneous vibration workload. The low and middle vibration categories are assigned for the vibration vector sum smaller than $0.5\,m/s^2$ and between $0.5\,m/s^2$ and $1.5\,m/s^2$, while the high category defines vector sum greater than $1.5\,m/s^2$.

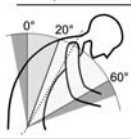
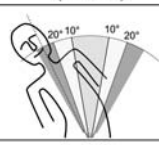
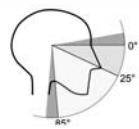
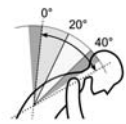
Body posture as joint angles were recorded by extended CUELA system for seated subjects (Hermanns *et al*, in press). The measuring system consists of accelerometers and gyroscopes, as well as a miniature data storage unit with a flash memory card (sampling rate 50 Hz). The battery driven system is attached to the subject's clothing (Figure 1). Therefore, subject's movements can be recorded during occupational operations without any hindrance.

In Table 2 the recorded body angles and their degrees of freedom are specified. Following ISO 11226 and EN 1005-4 body angles were classified in three categories: neutral, moderate and awkward (Table 3). The categories for vertical and

Table 2. Body angles and degrees of freedom measured with the CUELA system.

Joint or body region	Degree of freedom
Head	Sagittal inclination
Cervical spine	Flexion/extension
Thoracic spine	Sagittal and lateral inclination at Th10
Lumbar spine	Sagittal and lateral inclination at L5
Hip joint	Flexion/extension
Knee joint	Flexion/extension

Table 3. Body posture categories.

	Body angle				
	Trunk inclination (forward)	Trunk inclination (lateral)	Vertical head inclination	Back flexion (forward)	Back flexion (lateral)
					
neutral	0° to 20°	0° to 10°	0° to 25°	0° to 20°	0° to 10°
moderate	20° to 60°	10° to 20°	25° to 85°	20° to 40°	10° to 20°
awkward	<0° or >60°	>20°	<0° or >85°	<0° or >40°	>20°

lateral flexion of the trunk have been suggested. Using back support during the measurements any extension of the trunk was attributed to neutral posture.

An additional synchronised video recording provides a control feature associating the tasks and specific work situations. CUELA software application shown in Figure 2 allows for simultaneous access to the synchronised vibration, posture and video data.

In addition to video and 3D computer figure the software provides graphic representation of joint angles and vibration data. Also the categories for each exposure can be represented by the panels shown in Figure 2.

A schematic graph combines the statistical results of each body angle and vibration exposures as percentage of the measuring time. Figure 3 shows an exemplary task concerning back flexion versus vibration.

Three risk groups are defined to assess the combined workloads: Low Risk (dotted sections), Risk possible (dashed sections) and High Risk (crossed sections).

Results and discussion

The two investigated machines have been analysed by the tasks observed during their operations. The daily WBV exposure (8h) evaluation for the automobile resulted

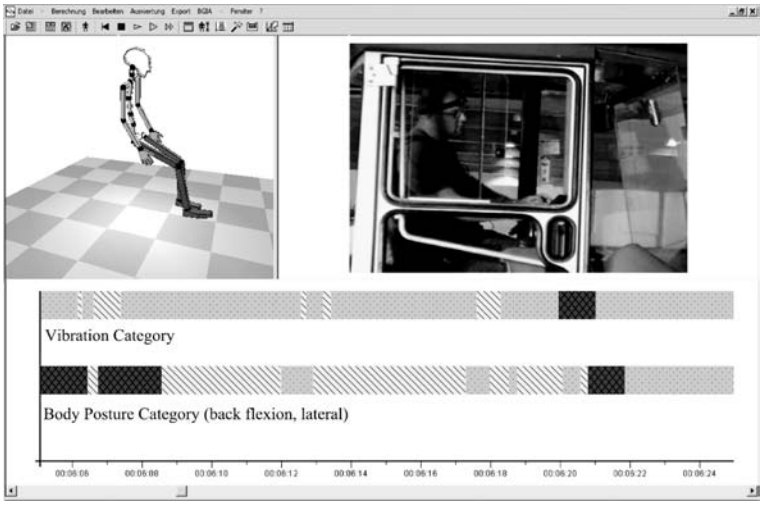


Figure 2. The CUELA evaluation software.

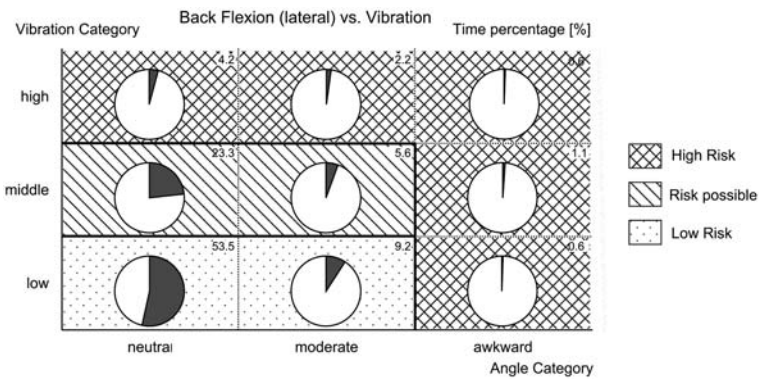


Figure 3. The combination of body angle and vibration categories.

in 0.16, 0.10 and 0.36 m/s^2 (x, y and z direction respectively), while the values for forklift resulted in 0.25, 0.17 and 0.27 m/s^2.

Figure 4 shows the vibration distribution among each task considering vector sum 1.4 by means of box plots (5th, 25th, 50th, 75th and 95th percentile). Besides stopping/idling task, the diagram in Figure 4 demonstrates that the vibration exposure of the two machines is very similar.

On the contrary, the posture analysis (lateral back flexion in this case) indicates the difference of the two machines. Since both drivers were leaning back while operating, the back extension has been excluded from awkward posture analysis. Figure 5 shows the percentage of each risk class concerning the combined posture and WBV assessment.

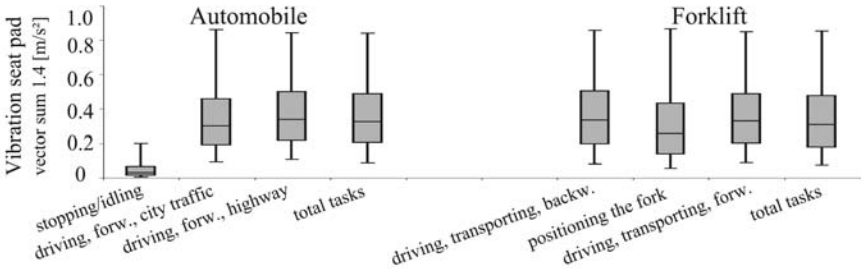


Figure 4. WBV distribution considering vector sum 1.4.

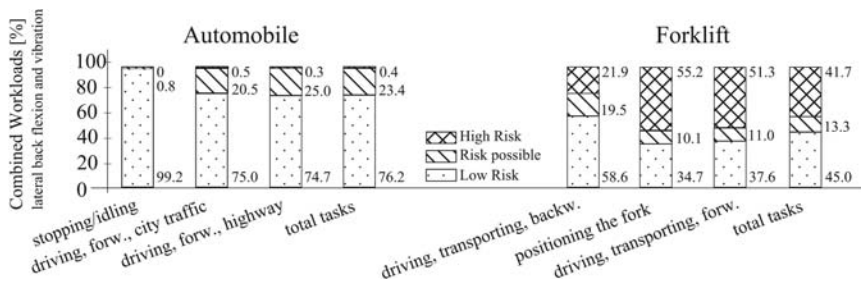


Figure 5. Combined lateral back flexion and vibration workloads.

Statistical analysis of the automobile operations indicates that total tasks attributed 0.4% of the measuring time to the high risk group and 23.4% to the risk possible group. The automobile driver spent 72.6% of the measuring time during low vibration and almost neutral posture.

The total tasks of the forklift driver have been analysed to be 45% of the measuring time in the low risk, 13.3% in the risk possible and 41.7% in the high risk group.

A comparison between the two observed analyses clarifies that the forklift driver has been more exposed to awkward posture than the automobile driver. Considering the tasks in details, the forklift driver adopts awkward lateral back flexion while looking at the fork located in the front of the machine. Also the set up of the machine influences the posture of the subject. The automobile driver mostly uses the side mirrors during the operations, while the forklift driver mostly uses the front and side windows.

The main goal of this pilot study is to highlight the important influence of the posture on the WBV. Furthermore, epidemiological studies are needed correlating the combination of the workloads to medical investigations. Therefore, more combined measurements among many different machine drivers are needed to clarify the influence of the posture.

References

Hermanns I., Raffler N., Ellegast R., Fischer S. and Goeres B. 2007 (in press), Simultaneous field measuring method of vibration and body posture for assessment of seated occupational driving tasks, *Industrial Ergonomics*, **37**

Hoy J., Mubarak N., Nelson S. and Sweert M. 2005, WBV and posture as risk factors for low back pain among forklift truck drivers. *J Sound Vibration*, **284**, 933–946

Ellegast R. and Kupfer J. 2000, Portable posture and motion measuring system for use in ergonomic field analysis. In K. Landau (ed.) *Ergonomic Software Tools in Product and Workplace Design* (Institut für Arbeitsorganisation, Stuttgart), 47–54

SAFE AND FUEL EFFICIENT DRIVING: DEFINING THE BENCHMARKS

Mark S. Young[1], Stewart A. Birrell[1], Neville A. Stanton[1], Neil Thorpe[2] & Mark Fowkes[3]

[1]*Human-Centred Design Institute, School of Engineering and Design, Brunel University, Uxbridge, Middlesex UB8 3PH, UK*
[2]*Transport Operations Research Group, Newcastle University, Newcastle-upon-Tyne, NE1 7RU*
[3]*MIRA Ltd. Nuneaton, Warwickshire, CV10 0TU*

In many complex systems, there is often a tension between the pressures of production and protection, but modern society increasingly expects an additional level of performance – protecting the environment. One of the key political and societal targets for the ecologically-friendly movement is the transport industry. The Foot-LITE project is an effort to encourage drivers to adopt safer and more fuel-efficient behaviours by providing feedback and advice on driving style. The work at Brunel University is focused on the ergonomics of the system – determining the driver behaviour parameters and designing an interface to optimise performance. In this paper, we describe a review of the literature on safe and fuel efficient driving. Ergonomics has proven itself in contributing to safety and performance, now it has an opportunity to contribute to the environment.

Introduction

The aim of the 'Foot-LITE' project is to develop a system which will encourage drivers to drive in a s*afer* and *greener* (i.e., more environmentally friendly) manner through on-board advice and off-line feedback. In order to achieve this it is necessary to ascertain the driver behaviours pertaining to both safe and green driving, to determine whether there are any conflicts between these goals, and to establish the task structure of driving and information requirements for the driver. The following paper is therefore structured to first discuss the effect of driving style on fuel consumption and emissions and then safety, before examining where the two goals overlap and where they compete. The paper then closes with a discussion of the driver's information requirements, including a summary of the positive and negative aspects of Advanced Driver Assistance Systems (ADAS) relevant to Foot-LITE.

Eco-driving

In 2006, within the European Union, road transport was responsible for nearly 20% of total greenhouse gas emissions. In September 2007 the Commission for

Integrated Transport (CfIT) suggested that fuel savings from 'smart' or 'eco-driving' can lead to average fuel savings of 5 to 10%. This equates to a saving of approximately £100 a year to the average motorist.

Although changing the way vehicles are driven can have an impact on fuel consumption and emissions, the effects are relatively small when compared to longer-term strategies such as infrastructure or technological change. Nevertheless, studies suggest that simply asking drivers to drive economically can have a beneficial effect on fuel consumption. A simulator study by van der Voort *et al* (2001) found a 10% decrease in fuel consumption when participants drove economically compared to their normal driving style. Similarly, a study by Waters and Laker (1980) predicted that appropriate changes in driving style could result in fuel savings of up to 15%. These reductions may be further enhanced by offering advice or feedback to support a fuel efficient driving style. Such advice could influence both current behaviour as well as potentially inducing longer term behavioural adaptations.

When drivers are asked to drive more efficiently, they generally interpret this as to drive slower. It is true that fuel efficiency is at its maximum between 60 and 80 km/h (Andre and Hammarstrom, 2000; Haworth and Symmons, 2001). El-Shawarby *et al* (2005) reported that cruising at 104 km/h used 54% more fuel compared to driving at 72 km/h, while reducing speed to 56 km/h increased fuel consumption by 15%. At speeds lower than 60 km/h, fuel consumption increases dramatically, with less marked increases for speeds higher than 80 km/h.

However, reducing speed is not the only – nor is it the optimal – strategy for optimising eco-driving, especially considering the implications for journey time. Studies indicate that speed itself does not cause large environmental problems in urban traffic. Instead, the focus should be on changing individual driver behaviour, environments and vehicles so as not to promote heavy acceleration, power demand or high engine speeds. These interventions would have the most impact on reducing fuel consumption and vehicle emissions. Numerous driving factors have been linked to eco-driving, the common consensus includes:

- Plan ahead to avoid stopping
- Use moderate engine speeds and a uniform throttle for steady speeds
- Change gear up as soon as possible using positive (but not heavy) acceleration
- Avoid sharp braking
- Use engine braking for smooth deceleration

Whilst significant improvements can be made by changing policy or infrastructure, efforts to modify driver behaviour can also lead to significant reductions in fuel consumption and emissions. Adopting an eco-driving style can reduce fuel consumption and emissions by around 10–15% – with significantly higher savings when compared to aggressive driving styles. These effects are more evident with urban driving, with petrol engines, and – if modern driving techniques are adopted (that is, not just reducing speed) – come at little cost to journey time.

Safe driving

Despite reductions in accident rates European roads in recent years, the statistics for killed and seriously injured (KSI) remains unacceptably high. In 2001 the European Commission White Paper on transport policy set a target for a 50% reduction in road fatalities by 2010. It is suggested that this cannot be achieved by the introduction of new technologies alone, but combined with improvements to infrastructure and enforcement of current safety measures. Of the three causative factors of road accidents (user, environment, vehicle), road users contribute 95% of the total (Sabey and Taylor, 1980). Thus encouraging drivers to drive safer, adhere to speed limits, avoid excessive accelerations and maintain appropriate headway could potentially help the trend to cut accidents and fatalities on our roads.

Throughout the literature, driving speed consistently emerges as the main correlate of crash risk and crash severity (Haworth and Symmons, 2001; Taylor *et al*, 2002). Taylor *et al* (2002) suggested that accident frequency increases with driving speed to the power of approximately 2.5. In other words, a 10% increase in mean speed results in a 26% increase in the frequency of all injury accidents. This increases to 30% when considering just KSI accidents (Taylor *et al*, 2002). Similar research cited in Haworth and Symmons (2001) suggests that for every 1 km/h reduction in speed, a 3% drop in accidents is observed.

Robinson and Campbell (2006) reported that 'exceeding the speed limit' or 'going too fast for the conditions' were contributory factors in 15% of all accidents on UK roads in 2005. When considering the severity of an accident, speed factors contributed to 26% of fatal accidents – which in turn accounted for 28% of all road fatalities (793 deaths). The number one cause of all accidents in this study was 'failure to look properly'.

Eco-driving vs. safe driving

Whilst much has been discovered about the effects of driving style on fuel economy and emissions, there is a relative dearth of scientific research on how driving style affects safety outcomes. One exception is a recent project evaluating the effectiveness of the Institute of Advanced Motorists (IAM) driver coaching programme (Stanton *et al*, 2007). A rigorously controlled study found that IAM coaching improved driver skills, knowledge and attitudes. Other research cited by Stanton *et al* (2007) suggests that IAM coaching also reduces frequency of traffic offences, as well as frequency and severity of accidents.

Further evidence on safe driving styles emerges from the research into eco-driving programmes. Haworth and Symmons (2001) cited evidence for a decrease of around 35% in accidents rates with two different driving fleets following EcoDriving training. This was in addition to the observed benefits for fuel consumption (11%) and emissions (23–50%). In another case study of the effectiveness of economical driver training for van fleets, Hedges and Moss (1996) showed that accident rates dropped by 40% after training.

Safe and efficient driving styles are not always compatible, however. Some factors that improve efficiency or safety may have confounding effects on the other.

For instance, maintaining a constant speed through avoiding braking may compromise headway; travelling in the highest possible gear may adversely affect vehicle control. Moreover, whilst large heavy vehicles may perform better in crashworthiness tests, these same characteristics make them less environmentally friendly than smaller vehicles.

Advanced Driver Assistance Systems

Advanced driver assistance systems (ADAS) are primarily aimed at improving safety, comfort and convenience (e.g., Richardson *et al*, 1997) by either supporting the driver or taking over certain driving tasks. Such systems include lane tracking, adaptive cruise control, collision warnings, GPS, intelligent speed adaptation, or enhanced vision. ADAS has become more popular over the past few years as consumer awareness for safer cars has driven market forces. Several existing and emerging ADAS devices purport to affect safety and/or fuel economy.

Ericsson *et al* (2006) investigated the possibility of satellite navigation systems optimising a route for lowest fuel consumption. The analysis suggested that using the fuel optimised navigation system could have reduced fuel consumption by 8.6%, corresponding to a 4% reduction across all journeys. The authors suggested existing navigation systems can offer a fuel saving potential, since 82% of the most fuel efficient routes were the same as the standard shortest route option. Ericsson *et al* (2006) also explored the potential for further fuel savings if the navigation systems included real-time traffic congestion monitoring. It was found that for 76% of 'disrupted' journeys, a more fuel efficient route existed, which on average would lead to a further 7.6% reduction in fuel consumption (a 5.8% reduction overall).

Given the consequences of traffic jams for fuel consumption, van Driel *et al* (2007) investigated the impact of a congestion assistant (CA) on driving behaviour and acceptance in a driving simulator. The CA consisted of three aspects: congestion warning and information (warning of traffic jams ahead), an active gas pedal (counterforce is applied to the pedal when approaching the jam if the system judges speed to be too high), and 'stop and go' (the longitudinal driving task is taken over by the system when in the jam). Overall, it was concluded that the active gas pedal and stop and go functions improved safety (by reducing speed on approach to a jam), enhanced traffic efficiency (due to reduced headway times), and reduced emissions (through lower accelerations; van Driel *et al*, 2007). The congestion warning, on the other hand, had no impact on driving behaviour since it was activated when the traffic jam was still 5 to 1.5 km away. Nonetheless, driver acceptance of the system was good, with drivers appreciating how the stop and go function relieved them of the task. The active gas pedal was the least accepted function by drivers in this study, a finding in common with research suggesting that systems which restrict driver's control are less likely to be accepted (van der Laan *et al*, 1997, cited in van Driel *et al*, 2007).

More specific eco-driving tools have also been investigated. A prototype fuel efficiency advice tool was evaluated in a driving simulator by van der Voort *et al* (2001). The tool gave feedback in the form of advice (e.g., shift earlier) or extended

advice (e.g., shift earlier from 2nd to 3rd gear). In addition, a series of LEDs reflected the extent to which actual driving deviated from optimal driving in terms of fuel economy. Results showed that fuel efficiency extended advice reduced fuel consumption by 16% over normal driving, and by 7% against unsupported economy driving (i.e., no advice tools). The benefits of an advice tool were particularly pronounced in an urban environment, with reductions of 14% compared to normal driving, and 23% compared to aggressive driving. As with other eco-driving techniques, there was no significant difference in trip time between any of the trials. Crucially, though, drivers without the advice tool later reverted to normal patterns of driving, whereas those with extended advice showed further reductions in extreme negative accelerations.

Conclusions and future research

There is evidence that eco-driving techniques can also lead to reductions in road accidents. Whilst the link between eco-driving and safe driving is not well established, and there are distinct scenarios in which safety and economy are in conflict, generally the driving-related factors are compatible. Various types of existing ADAS device can have ancillary benefits for both safety and eco-driving, while specific products for eco-driving assistance are beginning to emerge. The evidence suggests that these tools are most effective in urban environments, and that – unlike training programmes – the benefits are maintained over time. A system such as Foot-LITE could, therefore, provide real-time driver feedback based on advanced or eco-driving techniques to improve both safety and efficiency. However, the human factors associated with the driver's task and information needs must be taken into account when designing such a system in order to ensure that the positive benefits are gained while avoiding any negative impacts on workload or distraction. The next steps in the Foot-LITE project are to develop a set of candidate interface designs using established human factors methods, and to test these in the Brunel University Driving Simulator.

Acknowledgements

Foot-LITE is sponsored by the EPSRC, the DfT, and the Technology Strategy Board under the Future Intelligent Transport Systems initiative. The Foot-LITE consortium is comprised of: MIRA, TRW, RDM, Hampshire County Council, the Institute of Advanced Motorists, Nissan, Ricardo, Sussex Police, TfL, Zettlex, the University of Southampton, the University of Newcastle, and Brunel University.

References

Andre, M. and Hammarstrom, U. 2000, Driving speeds in Europe for pollutant emissions estimation, Transportation Research Part D, 5(5), 321–335

El-Shawarby, I., Ahn, K. et al. 2005, Comparative field evaluation of vehicle cruise speed and acceleration level impacts on hot stabilized emissions, Transportation Research Part D, 10, 13–30

Ericcson, E., Larsson, H. et al. 2006, Optimizing route choice for lowest fuel consumption – Potential effects of a new driver support tool, Transportation Research Part C, 14, 369–383

Haworth, N. and Symmons, M. 2001, The relationship between fuel economy and safety outcomes, (Monash University, Victoria, Australia) 1–57

Richardson, M., Barber, P., King, P., Hoare, E. and Cooper, D. 1997, Longitudinal driver support systems. Proceedings of Autotech '97 (IMechE, London) 87–97

Robinson, D. and Campbell, R. 2006, Contributory factors to road accidents, (Department for Transport, London) 1–16

Sabey, B. E. and Taylor, H. 1980, The known risks we run, (Transport and Road Research Laboratory) report no. SR 567

Stanton, N. A., Walker, G. H., Young, M. S., Kazi, T. A. and Salmon, P. 2007, Changing drivers minds: The evaluation of an advanced driver coaching system, Ergonomics, 50(8), 1209–1234

Taylor, M., Baruya, A. et al. 2002, The relationship between speed and accidents on rural single-carriageway roads, (TRL, Wokingham) 1–32

van der Voort, M., Dougherty, M. et al. 2001, A prototype fuel efficiency support tool, Transportation Research Part C, 9, 279–296

van Driel, C., Hoedemaeker, M. et al. 2007, Impacts of a Congestion Assistant on driving behaviour and acceptance using a driving simulator, Transportation Research Part F, 10, 139–152

Waters, M. and Laker, I. 1980, Research on fuel conservation for cars, (Transport and Road Research Laboratory, Crowthorne) 1–38

DETECTION OF MEMORY LAPSE DURING FLIGHT: A STUDY OF BACK-UP SYSTEMS AND FEEDBACK

Kristina Enflo

The Royal Institute of Technology, KTH, Stockholm, Sweden

Memory lapses during a normal flight are investigated to determine from which support systems the pilots receive feedback on performance. The role of feedback is analysed by decision making theory and illustrated with a new model. Results from field studies performed at a Swedish airline include descriptions of memory strategies, checklists and a recently introduced Electronic Flight Bag (class 1, type B). By these methods it is possible to receive feedback early in a decision sequence, and prevent human error from becoming dangerous and causing confusing information.

Introduction

In about 70 per cent of all aircraft accidents the original cause can be traced to human error (Shappell and Wiegmann, 1996). A human error can be described as all those events in which a planned action fails to achieve its intended outcome. Slips and lapses are errors which can be described as failures in the execution of an action. Such errors are however not necessarily unsafe acts, if an unsafe act is defined as an error made in the presence of a potential hazard (Reason, 1990). According to Reason, very few unsafe acts result in actual damage or injury, even in relatively unprotected systems. Also, in highly protected systems, the various types of defence can only fail through the combination of several different causal events. The most effective method to prevent accidents is thus to establish a feedback loop that influences the state of the system early in the accident sequence.

Purpose and methods

The purpose of this study is to investigate how the pilots receive feedback from different support systems during a normal flight. A simple form of an EFB, referred to as “the handheld device”, is investigated to examine its potential in providing feedback on performance.

The study started with 24 observations in cockpit which were made to explore the pilot work process (see Enflo and Barchéus, 2006). In addition, 8 semi-structured pilot interviews were conducted in order to get the pilots’ descriptions of decisions and support systems. Feedback on results was received from the director of flight operations at the airline.

Theoretical framework

To understand the effect of actions and errors, as well as characteristics in feedback, dynamic decision theory has been studied. Dynamic decision making can be defined as several decisions made in a sequence to achieve control over a dynamic real-time situation (Brehmer and Allard, 1991; Edwards, 1962). Earlier decisions produce information which is relevant to later decisions and there is, thus, a dependency between the decisions. During a sequence of decisions, the surrounding world can change either as a consequence of the decisions or independently of the decisions, or both. Furthermore, in a dynamic situation the time when a decision is made is of significance (Brehmer and Allard, 1991). Although a memory lapse may not be defined as inadequate decision making, the consequences of forgetting something may become the same as the consequences of a decision not to act.

Dynamic decision making can be clearer described by introducing vectors. Vector a_d refers to a decided action, and i_d, is the decision maker's information about the state of the world. The vectors are finite and variables of the number of the decision n, and the time. Hence, the information that the decision maker has, $i_d(n+1)$, depends on earlier information $i_d(n)$ and the actions $a_d(n)$ (Lind *et al.*, 1984). To control a dynamic decision situation, the decision maker has to learn the relation between the decided action and the information, F_1 (Brehmer & Allard, 1991), where $i_d = F_1(a_d)$, see Figure 1. Moreover, to describe the dependency between the decisions a function G_n is defined. Thus, $a_d(n+1) = G_n(a_d(n), i_d(n))$.

Feedback delays in a model of dynamic decision making

Experts need to make sure that their actions correspond to their intentions and this can be achieved with feedback (Reason, 1990). Brehmer and Allard (1991) introduced the terms of feedback quality and feedback delays since they had detected the decision makers' difficulties in understanding delayed feedback. As a result, they stressed the importance of finding solutions that can help people cope with such delays. Feedback quality and delays are here interpreted as terms that describe the information the decision maker has about the state of the world, i_d. By developing the vector i_d, quality and delay are described by three vectors, thus $i_d = i_r + i_u - i_m$, where i_r is the required information to maintain the perfect system, i_u is unnecessary information and i_m is missing information, see Figure 1.

The airplane accident in Teterboro, New Jersey, February 2nd 2005 can be seen as an example of the effect of missing feedback or feedback delay. Before starting the take-off procedures, the flight crew failed to ensure that the airplane was loaded within the weight and balance limits. The captain did not perform the balance calculation as he was expected to do. As the centre of gravity was situated too far to the front of the plane, the airplane failed to rotate (lift) at the intended rotation speed during take-off. Consequently, it ran off the end of the runway where it hit a building. Both pilots were injured as well as one of the passengers. The accident could have been prevented if the co-pilot had cross-checked the captain's operations and thus detected the missing balance calculations. The normal appearance of the

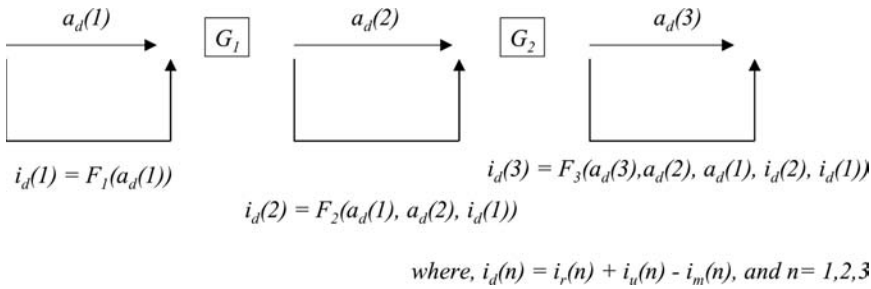


Figure 1. Changes in feedback along the decision sequence. G_n is the dependency function, a_d is the decided action, i_d is the information and n is the decision number.

airplane's acceleration during take-off also contributed to the delayed and missing feedback at the time. Even after the accident, the captain did not understand that the centre of gravity had been inappropriate located, since the acceleration seemed normal (NTSB, 2006).

In Figure 1 the author has illustrated a dynamic decision sequence which contains feedback changes (based on Lind et al., 1984). The model shows how complex the feedback information becomes in a dynamic system. The airplane accident described above can be illustrated with Figure 1: The first decided action, $a_d(1)$, is to calculate needed weight and balance at take-off. However, the actual action, $a(1)$, is that the balance calculations are not made. In addition, because of problems in the system, the pilot does not get the correct information about the state of the world, there is missing information. The pilot accelerates the aircraft which can be considered as the second decision, $a_d(2)$, and the information shows him that the state is normal, thus there is still missing information. As the pilot gets to the third decision, which is to rotate at the rotation speed, he receives the feedback that the airplane cannot lift. The feedback, $i_d(3)$, is in this case affected by the lacking balance calculations, the first action. However, since the pilot already is on the third decision, he has difficulties understanding this feedback. The information says only that the airplane cannot lift, $i_d(3)$; it does not say that the balance is not calculated, $i_d(1)$. Missing feedback information can be either completely lost or only delayed. In the accident described above the information is missing until after the accident, when it is too late. In Figure 1, it is shown how the feedback information changes along the decision sequence from $i_d(1)$ to $i_d(3)$. The complex information $i_d(3)$ that the pilot finally receives may thus be difficult to understand.

By concluding the results from theories and developed theories, it is understood that to support operators in a dynamic environment there should be dynamic systems that provide feedback on performance early in the decision sequence and display what information is missing. Thus, the question is: does such a support system exist in reality?

Results from the field studies

Many of the decisions during flight may be described as discrete decisions with a limited number of alternatives. The requirements, rules and procedures, as well as the critical environment and the time pressure, may result in such decisions. Some of the decisions made at every flight could even be described as binary decisions with two choices: "on or off" or, "go or no go". Moreover, the studies indicated that the pilots deal with dynamic decisions: many of the decisions are dependent on earlier ones, i.e. the decision whether or not to land depends on earlier decisions regarding alternative destinations and amount of fuel.

The most common explanation when the pilots were asked how they remember procedures was that they remember because of their experience. Other explanations were education, simulator training, routines, rules, checklists, flow patterns, mnemonics, triggers and memory items. The memory items were described as certain tasks performed at trigger points, which are events during flight. Later, the most critical tasks of the memory items are checked off on a checklist. Flow patterns and mnemonics are used to support the remembering of a certain order of the items, and this order standardises the work procedures to facilitate cooperation between pilots. According to the pilots, they were very dependent on the checklists before they gained experience.

Detection during flight

According to the interviewed pilots, they all very occasionally make memory lapses. The non-critical tasks are the ones that could be forgotten and never the fundamental items such as "flaps". A contributing factor may be that most often the non-critical items are memory items not checked by a checklist. Other reasons for memory lapses are stress, lack of sleep, and hunger. The most common explanation, however, was interruptions while performing the memory items. A reason for a lapse could also be a recent change in the procedures. There are however "systems" that could

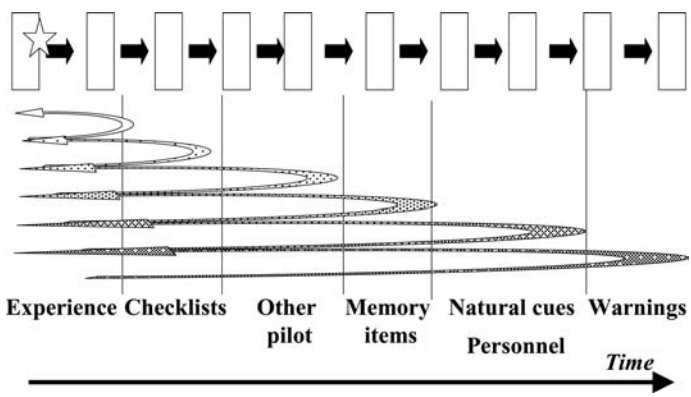


Figure 2. Feedback loops from the backup systems. The order of the back-up systems may be different than the illustrated one.

detect when something is forgotten. Only with experience, do pilots develop a sense for what should be done. They described this as automatically sensing that something is forgotten. If they do not sense it, the pilots usually detect the memory lapse by themselves when performing the checklists. Alternatively, the other pilot may detect that something is forgotten. Another way to detect a memory lapse is when the next flow of memory items comes; it is then realised that the last flow was not executed. Moreover, natural cues in the environment or other personnel may indicate the lapse. If the forgotten item is not yet detected, aircraft technology intervenes with warnings. However, the lapse is most often detected earlier and there are no warnings. In Figure 2 the flight tasks are shown where the star illustrates an error committed in the beginning of the flight. The arrows illustrate feedback loops received from different backup systems, based on a timeline.

The potential of the handheld device

Recently a simple form of an EFB[1] (class 1, type B) has been introduced at the airline to support aircraft performance as well as aircraft weight and balance calculations performed by the pilots. The device has the format of a mobile telephone equipped with computing functions and can be connected to a database. In the tasks where the device is used it may be seen as a support to remember the needed information since it shows the required information on the display. In the performance calculations the device may ask for missing information if the pilots forget to register information. The device also shows the accuracy of the balance calculations since presents if the results are within certain limits. However, there is no function that shows if the input data is registered correctly (Enflo, 2007).

Discussion

Although the methods can be seen as triangulation, the results could be further strengthened and evaluated by for example experimental studies where the pilots are asked to think aloud.

One way to help operators cope with feedback delays is to present missing information, as the handheld device does in some of the functions. The results indicated however that the most important sources of detecting a memory lapse were by experience, the checklists and by the other pilot. The new technology may help in some situations but since there are situations where it cannot provide feedback to the operator, the role of the other pilot is essential.

In normal flights with decisions made with a lot of routine, it seems that the aim with a support system should be to keep the pilots in a correct loop by validating actions and provide feedback in case of an error. The handheld device appears to have the potential of being such a support system; however, it should be used within co-operation with the other pilot. Furthermore, such a device appears more dynamic

[1] A Nokia Communicator

than the memory strategies and the checklists presented since it is adjustable for new situations.

References

Brehmer, B. and Allard, R. 1991, Dynamic Decision Making: The Effects of Task Complexity and Feedback delay. In J. Rasmussen, B. Brehmer and J. Leplat (ed.), *Distributed Decision Making: Cognitive models for cooperative work*, (Johan Wiley & Sons Ltd, Chichester), 319–334.

Edwards, W. 1962, Dynamic Decision Theory and Probabilistic Information Processing. *Human Factors*, **4**, 59–74.

Enflo, K. and Barchéus, F. 2006, *Interruptions in preflight – jump seat observations of communication in the cockpit.* HFN conference on Human Factors and Economic Aspects of Safety, Linköping, Sweden, Swedish Human Factors Network.

Enflo, K. 2007, *Support of decisions in the preflight phase.* In Proceedings: the 8th International Conference on Naturalistic Decision Making, Pacific Grove, CA.

Lind, M., Allard, R. and Brehmer, B. 1984, *Dynamiskt Beslutsfattande*: *Dynamic decision making*, FOA Rapport No. C 53015-H2, (Försvarets forskningsanstalt, Stockholm).

Loukopoulos, L.D., Dismukes, R.K. and Barshi, I. 2001, Cockpit interruptions and distractions: A line observation study. In *Proc. 11th International Symposium on Aviation Psychology*.

National Transportation Safety Board, 2006, *Runway Overrun and Collision*, Aircraft Accident Report No. NTSB/AAR-06/04, (NTSB, Washington, D. C.).

Reason, J. 1990, *Human Error*, (Cambridge University Press, Cambridge).

Shappell, S. and Wiegmann, D. 1996, U. S. naval aviation mishaps 1977–92: Differences between single- and dual-piloted aircraft. *Aviation, Space, and Environmental Medicine, 61*, **1**, 65–90.

ERGONOMICS ISSUES IN A JET CAR CRASH

Mike Gray

Health and Safety Executive, Edgar Allen House,
Sheffield, S10 2GW, UK

A crash occurred in September 2006 during filming for a television programme of a jet car being driven at speed. As part of the ensuing investigation by the Health and Safety Executive, the role of some of the human factors issues was assessed.

Introduction

During filming of the driving of a jet powered car in September 2006 for the BBC TV programme Top Gear the car left the track while travelling at high speed. The driver, presenter Richard Hammond (RH), was fortunate to survive the crash. The Health and Safety Executive (HSE) carried out an investigation into the circumstances surrounding the incident and as part of this the human factors issues were studied.

The incident occurred at Elvington airfield where the jet car "Vampire" was driven down the runway at speeds up to 300 MPH. The car was a purpose built dragster powered by a Rolls Royce Orpheus jet engine; it held the current Outright British Land Speed record of 300.3 mph.

In preparation for the filming RH had received a written briefing describing the arrangement of the car's controls. On the morning of the day of the incident he received training in the form of a series of briefings and in-car instruction. In the afternoon a series of runs with increasing jet power were then undertaken. It was on the seventh run at about 5:25 pm that the crash occurred, initiated by a front tyre failure.

Ergonomics issues

The investigation was lead by two HSE inspectors with assistance from North Yorkshire Police. They collected information from the various parties involved in planning and carrying out the filming of the jet car runs. They produced a report (HSE 2007) analysing the events that lead up to the incident and recommending actions to help minimise risks in the future. Human factors input to the investigation was provided by the author. This concentrated on the physical arrangements of the driver's cab, the operation of the controls, the time available for action and the training requirements.

The arrangements for driving this jet-powered car are somewhat different from those found in conventionally powered vehicles. During any run the left foot has to be kept pressed on the hold to run pedal or the engine will cut out. The right

foot operates a brake pedal but it is primarily used for holding the car on the start line, while the engine is started by the crew using a separate starter. A hand throttle is set by the driver during starting and adjusted to a predetermined engine speed setting once the jet engine is running. Higher engine speeds were used as the day progressed. While the engine is coming up to speed the driver has to keep pressing on the foot brake to ensure the car stays at the start. For afterburner use it is necessary to depress a switch attached to the steering wheel for a second or so. The run then commences when the right foot is released from the brake. The driver then has to look forward out of the cockpit and use the steering wheel to maintain a straight course down the runway. This in practice means keeping an angle of about 30° on the steering wheel with the right hand high, apparently this counters the built in tendency for the car to veer to the right.

Once the car reached the area of the runway designated with a green cone on either side the driver should pull the parachute release lever back, this deploys the parachute while simultaneously cutting the engine. The parachute provides a significant deceleration and the driver keeps steering while speed is lost.

The overall operation of the car is thus to hold the brake on the line, keep the hold to run pedal depressed, adjust the throttle for start and then increase to the required power. Hold the steering wheel, activate the after burner switch. The car will then rapidly accelerate during which time the driver has to steer in a straight line. The driver needs to view the runway to maintain a straight line and activate the parachute when the green cones are reached. The time from the start to the release of the parachute is around 20–30 seconds.

Cockpit layout

The layout of the area where "Vampire's" driver sat was assessed. This was based on a visual examination and measurement of the car in the garage to which it was moved after the crash had occurred. The roll cage had been cut from the car at the scene of the accident and the steering column disconnected but the positioning of pedals, levers, seat etc. was reasonably clear. The compact layout of the cockpit of Vampire means that the various hand-operated controls should be within reach of most people. Adequate reach is important because the driver is restrained by a full 5-point harness; this removes the possibility of leaning forward to gain extended reach. There was no provision for adjustment of either the seat or the controls. The width of the cockpit is around 485 mm that will allow most people to fit in but it will be a tight fit for larger individuals.

The parachute deployment lever is positioned about 780 mm from the back of the seat, at the side of the cockpit just inboard of the chassis tubing. To operate the lever the driver has to remove his hand from the steering wheel and move it sideways and slightly forward. The shaped end of this control makes it readily distinguishable by touch. The hand should readily be able to grip the end of the lever, a 38 mm diameter ring, the driver needs to pull the lever back at least 200 mm to move the rods which release the sprung cap over the parachute and cut off the fuel to the

engine. About 70% of men would be able to fully grip the lever when restrained in the car, but the majority of women would find this reach difficult.

The positioning of the foot controls are such that virtually all people should be able to fully reach the pedals. Taller people will find that there is insufficient space for their legs.

Generally, the lack of adjustment for the position of the seat, pedals and steering wheel means that tall or heavy people would have difficulty fitting in the car. Fortunately RH at about 1.7 m tall is of similar stature to the regular driver of the car and consequently should be able to reach and operate the important controls while strapped into the seat, though the parachute deployment lever could be toward the end of his reach.

Driver expectations

When driving a car people have expectations as to how the controls used will operate. Because of this, standardized arrangements of pedals and hand controls have been developed which facilitate someone moving from one vehicle to another. Thus, the accelerator is provided to the right of the brake pedal. People can get used to other arrangements but will be more likely to select the wrong action if they have to deal with a novel arrangement, especially in an emergency. Systems, which take account of people's expectations and are tolerant of errors are more likely to be used successfully. (see (HSG48) – HSE (1999)).

There have been examples of accidents involving fork lift trucks where the layout of the pedals was different to the normal car arrangement. Some designs of fork lift truck use separate forward and reverse accelerator pedals which require a driver to use a pedal on the right to go backwards while the pedal to its left accelerates forward. A brake pedal is available for use by the left foot. If the driver is reversing and wishes to stop, then moving the foot to the left and pushing that pedal would cause the truck to slow but then accelerate forward. The failure of the truck design to conform to common expectations has resulted in some fatal accidents. (Gray 2000).

The jet car has many similarities with normal car operation, especially single seater cars. However from a control point of view they are quite different. There is no foot control for increasing speed, the left foot is continually required to push down, the steering requires quite a lot of lock to keep in a straight line and the stopping mechanism is hand operated as a single event rather than the continuing action required by a foot brake. In addition, there is no real control of speed; the driver will not feel as if they are in the control loop in the same way as with most other cars. The noise, vibration, acceleration and the knowledge that record breaking speeds were being reached provide additional factors which would be expected to influence human performance of a task.

A set sequence of actions have to be performed during each run. The expert who normally drove the car talked RH through each of the pre-run checks and ensured they were done in order. On early runs RH's visor steams up and he complains that he cannot see. Subsequently he keeps the visor open until the last few seconds to

try and keep it clear. A hand signal is given to remind him to put the visor into place once the engine is running at full throttle. Ensuring that actions are completed and in the right order requires concentration without distraction. Typically people use checklists to ensure they do not lose their place in a sequence. It appears from the video footage that verbal direction as well as the briefing sessions are used to make sure RH does things correctly. Once the foot brake is released however he has to deal with things by himself.

The main action he takes once the run has commenced is to try and steer in a straight line while maintaining pressure on the left pedal. It is evident from the audio recording that he finds it difficult to see well out of the cockpit, mainly because of the misting of the visor, though this is said to be better on the later runs. The seating arrangements appear to give RH a clear view forwards though the g force when being accelerated forwards will push the head backwards which may make it more difficult to gain clear sight of the runway and cones. There also appears to be some uncertainty about where he is meant to align the car on the runway. On an early run he says that he is trying to get midway between the cones, this however is not the centre line of the runway, which is cambered and so not the best line – one cone is moved after the first two runs to help with alignment. The telemetry for the penultimate run indicates a steering correction at high speed; this apparently is because the car was deviating from the centre line. A driver's ability to maintain a straight line at high speed will increase with experience of the steering system, briefing sessions and training while stationary cannot provide this type of experience.

The action of pulling the parachute lever to stop the car is one that is peculiar to this type of car. RH is given a series of sequential instructions during the briefing session which requires him to pull the lever back when the green cones are reached. Thus for normal driving he has to keep left foot depressed, keep the car on a straight line using the steering and watch for the green cones. He has to contend with noise, vibration (recorded comment that it is very bumpy), and the g force. The instruction to pull the lever at the end of the run is clear and the sequence of events leading up to it is straightforward.

It is however when the driver has to get involved with another task that the progression of actions can become disturbed. Once the car starts to move away from the straight line and not respond to the steering as it has previously then the driver will focus on dealing with that situation. An experienced driver will have amassed a number of strategies for dealing with this type of situation, and in what is perceived as an emergency a competent driver will bring these to bear. Less experienced people might well freeze and not be able to take action. RH, as someone who has driven many fast cars on tracks, is likely to have developed a feel for how a car that is going out of control can be brought back under control. This will involve the use of the steering and foot brake.

Bailey (1996) in discussing human error suggests that "People are frequently the victims of unwanted releases of automatic performance", i.e. a well practiced sequences of actions can be undertaken unintentionally if an event which seems like the start of a sequence occurs. Where a habit is strong enough then cues that even only partially match the situation can trigger its performance. Thus while RH is driving the jet car and it is behaving as expected he is dealing with a novel situation

but one for which he has been given a set of rules. As soon as the situation changes to dealing with a loss of control his habits of trying to feel the situation with the steering and bring the speed under control by use of a foot brake can be expected to take over. If the car does not seem to be responding to these controls then a good driver will re-evaluate and try to find other means for dealing with the situation.

Time aspects

The maximum speed on the final run is 288 MPH, a 100 metres would be covered in only 0.7 seconds at this speed.

The time interval between becoming aware of a dangerous situation and taking defensive action against it is the reaction time. Many studies (e.g. Green 2000, Triggs and Harris 1982) on reaction times have been carried out which have explored the factors which affect reaction time, and the components which make it up. Reaction time can be broken down into a number of elements:

- sensation time (detecting something),
- perception time (recognising the meaning of the sensation),
- response selection time (deciding what to do),
- movement time (time for muscles to move the body) e.g. moving foot onto the pedal and pushing,
- device response time (e.g. for the brakes to start to slow the vehicle).

The action of removing the right hand from the steering wheel and feeling for a lever to pull back is quite different from the response for keeping control of a skidding car. Overcoming a practiced response will take further time. Deciding to take such action will probably require the driver to process information from the sound and feel of the car, as well as any visual cues. In any case deploying the parachute is not a fast action in the same way as using a foot brake.

According to telemetry the car was going well until 14.25 s from the start when a change in vertical velocity is detected. 0.35 seconds later the vehicle starts to turn right at 2.1g, this continues to build for a further 0.66 seconds to 5.1g. The elapsed time from the indication of a problem until a large impact is 1.25 s, a further 0.67 s elapses before telemetry is lost.

A skilled driver might be able to apply the foot brake in about half a second, even faster if anticipating the need to brake. The time taken to operate the parachute would be appreciably longer, partly because of the additional mental processing and partly because of the additional physical movement required to relocate the hand and pull back the lever in order to activate the parachute and fuel shutdown controls. An experienced jet car driver who is practised at using the parachute to stop the car may be able to operate the control in about 1 second after an alerting event. A less experienced driver, working in attentive processing mode, could take an additional 0.5 to 1 seconds to carry out this action.

Video of earlier runs suggests that the parachute takes about 1 second to inflate from the time the cover over the drogue chute is released. However even if the

driver is able to activate the parachute control within 1 second of an alerting event substantial braking will not occur until 2 seconds after the event. The fact that the parachute was deployed during Richard Hammond's crash but did not have time to fully inflate suggests that he must have moved to activate the parachute control very soon into the sequence of events.

Conclusions

Although the car setup was not adjustable it was probably suitable for someone of Richard Hammond's size.

The means for controlling the jet car are different to those of a conventional track car. The differences are such that the skills built up by an experienced driver might result in inappropriate actions in the event of an emergency.

The time available for a driver to take appropriate action is short and an inexperienced jet car driver might travel around 100 metres further than an experienced driver while deploying the parachute.

References

Bailey R.W., 1996, Human Performance Engineering, 3rd Edition, (Prentice Hall, New Jersey).

Gray M.I., 2000, Controls, Displays and Ergonomics Standards – Ensuring Health and Safety in Machinery Design, Ergonomics for the New Millennium. Proceedings of the XIVth Triennial Congress of the IEA, San Diego, California, USA, July 29–August 4, 2000.

Green M., 2000, "How long does it take to stop?" Methodological analysis of driver perception-brake times, Transportation Human Factors, 2(3), 195–216.

HSE, 2007, Investigation into the accident of Richard Hammond, http://www.hse.gov. uk/foi/releases/richardhammond.pdf

HS(G)48, 1999, Reducing error and influencing behaviour, Health and Safety Executive.

Triggs T.J. & Harris W.G., 1982, Reaction time of drivers to road stimuli, Human Factors report No. HFR-12, Monash University Accident Research Centre, Victoria.

DRIVERS AND DRIVING: A SEGMENTATION BASED ON BEHAVIOURS AND ATTITUDES

Patrick W. Jordan, Fang Chen & Anders Lindgren

Chalmers University of Technology, Department of Computer Science and Engineering, SE-412 96 Göteborg, Sweden

A segmentation of drivers is presented, based on driving style, behaviours and attitudes. Four main segments are identified and described. The segmentation was complied using the Delphi method and is based on the synthesis of the views of a range of leading experts in the field.

The segmentation is an essential first step in understanding drivers and creating solutions which will enhance driving behaviour.

Introduction and approach

Statistics show that there are a wide variety of different types of accidents and a variety of behavioural causes. These reflect drivers' attitudes towards driving and the approaches and behaviours that they bring to the road (Lancaster and Ward 2002).

This paper reports a study in which drivers were segmented according to attitudes and behaviours. The approach was based on a technique known as the Delphi method (Linstone and Turoff 1975) This involves interviewing experts in a particular field – in this case driving and driver behaviour – and coming to a conclusion based on the common ground between them.

In this case representatives of the following institutions were interviewed: Institute of Advanced Motorists, Royal Society for the Prevention of Accidents, Transport Research Laboratory, Brunel University, Chalmers University, Driving Standards Authority, Driving Instructors' Association.

The outcome was the segmentation described below. This doesn't necessarily represent the views of any of the organisations listed above, but rather a composite of their views based on the professional judgement of the author.

The segmentation

Drivers were largely defined by their attitudes towards speed and their attitudes towards safety. These two dimensions formed the primary variables for the segmentation and drivers positioning on these dimensions was broadly predictive of not only their attitudes, but also their driving behaviour.

A range of secondary variables associated with each of the segments was also identified. These included:

Demographics: age, gender, socio-economic status etc.
Accidents: the frequency with which drivers were likely to have accidents as well as the nature and causes of the accident.
Examples of Behaviour: the way that drivers behave when they are on the road. This includes the risks that they may take and their attitudes and behaviour towards other road users.
Enjoyment of Driving: is driving something they enjoy, is it a chore or is it something they actively fear?
Distances: the average mileage covered per year.
Distances: the average mileage covered per year.
Type and Purpose of Journey: the reason that people make journeys and the characteristics of their journeys in terms of road type etc.
Car Type: the types of car that people in each of the segments are most likely to drive.

Four segments were identified as follows:

Reckless drivers

This segment loves speed and gives little thought to safety. Their characteristics are as follows.

Love Speed. Enjoy speed and sometimes show off to others by driving fast.
Unaware of Risks. Many in this group are inexperienced drivers who are unaware of the risks which driving holds. They may have the feeling that they are invincible.
Aggressive. They can be aggressive towards more cautious drivers who get in their way.
Poor Hazard Perception. Again this is due to inexperience. Evidence still suggests that hazard awareness is comparatively poor in new drivers.
Speed Violations. May in this segment consistently drive above the speed limit.
Follow too Closely. Following too closely is a major safety hazard with this segment.
High Accident Rate. This segment is by far the most prone to accidents, with an accident rate out of all proportion to the size of the segment.
Overestimate Ability. Drivers in this segment are likely to overestimate their ability. They think that because they are young and have good reaction times that they are safe, but this is more than offset by the lack of perception and judgement.
Younger, (<25). Most of this segment are aged under 25 years old and have, at the most, only a few years of driving experience.
Male $++$. Although some women do fall into this segment it is very predominantly male.
Lower Mileage. This group has a comparatively low annual mileage.
Urban, A-Roads. A high proportion of their driving is to and from leisure activities and much of it is also just for the fun of driving.

Enjoy Driving. As well as finding driving exciting it is, for many of them, a rite of passage. Driving is also a social activity for some and they may enjoy driving their friends around or talking about cars and driving with their friends.
Super-mini, Small Family, Pre-owned, Modifications. This segment most commonly drives small cars – in the majority of cases these will be pre-owned. Some have also modified their vehicles, adding power and fitting loud exhaust pipes.
Best: 'Youthful Exuberance'. The majority of people in this category can best be described as displaying 'youthful exuberance' which they will grow out of as they get older.
Worst: 'Dangerous Sociopath'. Some drivers in this group can be an extreme danger to themselves and others displaying a recklessness that endangers themselves and other road users. Some may be killed or kill others before they grow out of being in this segment.

Progressive drivers

The people in this segment also like to drive fast but this is tempered with a concern for safety. Most of them have also got considerable driving experience and many are highly skilled drivers.

Enjoy Speed. This segment enjoys speed, but is actually more focussed on making progress than on the speed itself. They want to reach their destination quickly.
Drive Purposefully. They are decisive and progressive drivers. They have confidence in their decision making and driving ability and do not dither over driving decisions.
May be Aggressive. Some in this segment may be intolerant of other drivers and be inclined to remonstrate with or intimidate drivers who they feel have got in their way or are driving poorly. In particular they can get irritated with those who drive slowly.
Underestimate Danger. They may underestimate the dangers associated with fast or aggressive driving.
Possible Speed Violations. Their desire to make rapid progress can sometimes lead to them driving in excess of the speed limits.
May Follow too Closely. A common characteristic of drivers in this segment is a tendency to follow too closely behind the vehicle in front.
Higher than Average Accident Rate. This segment has a higher than average accident rate. Speed violations and aggressive driving can be contributory factors in some cases.
Overestimate Ability. Most in this segment regard themselves as a better driver than average. However, accident statistics would suggest that the segment as a whole is not.
Wide Age Range (25–65). This segment covers a wide age range and is fairly evenly spread over this range.
Male +. There are more men than women in this segment, although the gender skew is far less than in the Reckless Drivers segment.

Higher Mileage. This is the segment that does the highest annual mileage. A significant proportion of the segment have to drive as part of their jobs, for example to see clients.
Motorways, A-Roads, B-Roads, Urban. They drive on all types of road and are the segment who do the most motorway miles.
Generally Enjoy Driving – Pride in Driving, Pride in Car, Talking Point. Although they can get frustrated with the hold-ups and congestion on the roads many in this segment get enjoyment from driving and take a pride in their car and their driving.
Family, Compact Executive, Executive. Many will buy their cars new or at least nearly new. For many, their car is somewhat of a status symbol.
Best: 'Knight of the Road'. At their best drivers in this segment can be 'knights of the road – courteous, progressive, efficient and considerate.
Worst: 'Road Hog'. At worst they can be 'road hogs' displaying discourteous, impatient and even bullying behaviour.

Diligent drivers

This group puts safety first and also do their best to comply with traffic laws. This doesn't necessarily mean they drive slowly – many in the segment will drive up to the speed limit. However it is safety rather than speed that is the first priority.

Safety Conscious. This segment has safety as their top priority, although that doesn't necessarily mean that they drive slowly.
Aware of Risks. They have a realistic understanding of the risks associated with driving.
Not Aggressive. They rarely show aggression and are generally courteous.
Good Hazard Awareness. They are a group who learn as they gain experience and have a good understanding and awareness of hazards.
Obey Traffic Laws. They take care to drive within the law and try to avoid violations.
Follow at Sensible Distance. This segment appreciates the dangers involved in tailgating and will generally follow at a safe distance from the car in front.
Low Accident Rate. Diligent Drivers have the lowest accident rate of all the segments.
Occasional Misjudgements. When an accident does occur it is most often the result of a misjudgement rather than a violation or aggressive behaviour.
Wide Age Range (25–65). Like the Progressive Drivers this group span a wide age range and are fairly evenly distributed across it.
Female +. There are more women than men in this group. The skew is approximately the inverse of the Progressive Drivers segment.
Average Mileage. They are about average when it comes to mileage – more than Reckless Drivers and Cautious Drivers, but less than Progressive Drivers.
A-Roads, B-Roads, Urban. Drive on a variety of roads but do less motorway driving than the Progressive Drivers. Many of their miles are done with their children in the car.

Utilitarian View of Driving – Facilitating Life Activities, Safety, Efficiency. Driving is a means to an end for this group. It is not something which they necessarily like or dislike.
Supermini, Family, MPV, SUV. Most are not particularly interested in cars and will pick a vehicle for its practical qualities rather than its performance, image or status.
Best: 'Efficient and Courteous'. The best drivers in this segment are efficient, courteous and look out for their own safety and that of other road-users.
Worst: 'Selfish and Inconsiderate'. At their worst they look out for their own safety, but can be selfish and have little regard for the safety and convenience of others.

Cautious drivers

Many in this segment dislike driving and will only do it if they have to. Safety is the top priority and some may drive rather slowly. Sometimes their over-caution can lead to danger.

Cautious and Nervous. The drivers in this segment have a cautious attitude towards driving, in some cases to the point of being nervous.
Perceive Driving as Risky. They are aware of the dangers associated with driving. In some cases they may have an exaggerated perception of these.
Drive Cautiously. They tend to be very cautious drivers. Sometimes this is safer but at other times over-cautiousness can lead to dangerous indecisiveness.
Variable Hazard Awareness. They are conscious of the need to be aware of hazards, however they are not always adept at spotting them.
Obey Traffic Laws. Drivers in this segment generally have respect for the law. They regard the traffic laws as being for their own safety and the safety of others.
Follow at Sensible Distance. They do not tailgate or show aggression towards others.
Moderate Accident Rate. Drivers in this group may have accidents from time to time.
Concentration and Judgement Lapses. Lapses in judgement and concentration tend to be the main source of accidents caused by drivers in this segment.
Older, (<65). The average age of this segment is higher than that of any of the others, although by no means all of the drivers in this segment are over 65.
Female +. There is also a slight skew towards female drivers in this segment.
Lower Mileage. This segment does a lower average mileage than average.
Urban, A-Roads, B-Roads. Most of their journeys are relatively short local trips.
Dislike Driving – Stressful, No Interest. Because driving can be stressful for them they take little pleasure in it and find it to be of little or no interest.
Super-mini, Small Family. They tend to see cars simply as a means of getting from A to B. Good value for money and reliability will be their priorities when selecting a car.
Best: 'Slow but Methodical'. At their best drivers in this segment drive in a thorough and methodical manner, even if they may be somewhat slower than average.
Worst: 'Irritating Ditherer'. At worst they can be indecisive causing inconvenience and possibly even danger to other road users.

Discussion/Conclusions

An important point to note is that at this stage the segmentation has yet to be validated statistically. Work on this is currently in progress. However, for now, they offer a good overview of driving attitudes and behaviours and help to make sense of the issues associated with different drivers.

The segments are not equal in terms of the levels of road safety associated with each. Diligent Drivers appear to be the safest segment, suggesting that there may be significant road safety benefits in persuading other drivers to mimic or adapt their driving behaviours.

Public information programs, legislation, law enforcement and training are potential routes to achieving this, as are the opportunities afforded by new in-car safety technologies. These include systems such as Adaptive Cruise Control that keep the car a safe distance from the vehicle ahead; Electronic Brake Assist which takes over control of the brakes from the driver in the event of an emergency; and various other systems which can make driving safer and less stressful.

However, the effectiveness of any attempts to improve driving safety are likely to be highly dependent on gaining the acceptance of the drivers at which they are aimed. Evidence suggests that many drivers do not learn from their mistakes and that those with a high accident rate are the least likely to see the value of driving more safely (Home Office 1997).

Tailoring messages and technologies to gain the acceptance and change the behaviour of high-risk groups is a major challenge for those working in this field. Understanding the characteristics of the segments is an essential first step in meeting that challenge.

References

Home Office, 1997, *Roadcraft: The Police Drivers Handbook* (Stationary Office Books)

Linstone, H.A. and Turoff, M., 1975, *The Delphi Method* (Addison-Wesley)

Lancaster, R. and Ward, R., 2002, *The Contribution of Individual Factors to Driving Behaviour* (Health and Safety Executive, Research Report 020)

DIFFERENT RIDERS ON DIFFERENT MOTORCYCLES: A QUASI-EXPERIMENTAL APPROACH TO RIDER COMFORT

Alex Stedmon[1], Simon Hargreaves[2], Martin Fitz-Gibbons[2] & Chippy Wood[2]

[1]*Centre for Motorcycle Ergonomics, University of Nottingham, Nottingham UK*
[2]*'Bike' Magazine, Emap Automotive, Peterborough UK*

A quasi-experiment was designed in which five participants with different anthropometric characteristics rode eight different motorcycles. Results indicated that riders perceived the comfort of motorcycles even before they sat on them. Other results illustrated that a number of measures can be taken that provide a consistent insight into motorcycle comfort.

A quasi-experiment to investigate rider comfort

What makes a motorcycle comfortable to ride? Part of the answer lies in the physical design hopefully matching the physical characteristics of the rider. Features such as controls, displays, seating and wind protection all depend on the relative size and location of the rider to be effective (Roberston & Minter, 1996). Generally, the physical size of a rider dictates the type of motorcycle they might find most comfortable. It would be simple to match riders to specific motorcycles however riders differ in their perceptions of what makes a motorcycle comfortable and in their expectations of how comfortable different motorcycles might be.

Whilst motorcycles are sold in one size, it is clear that 'one size does not fit all' and different sized riders will have very different experiences of comfort on the same motorcycle. In order to explore the subtle relationship of rider comfort and motorcycle design a quasi-experimental trial was designed in which a number of riders with different anthropometric characteristics rode different styles of motorbikes and provided comfort ratings across a number of dimensions. Quasi-experiments present the opportunity of conducting research in the 'real world' rather than a laboratory setting, although there is a degree of compromise in the control over variables (Cook & Campbell, 1979).

Participants

Five participants (4 male and 1 female) volunteered for the comfort trials. The participants were aged between 30–46 years old (mean age = 39.2 years, SD = 6.3 years). The participants were chosen to represent different anthropometric characteristics (tall, short, average, heavy and female). The riders all held full motorcycle licences with an average riding of 8 hours per week, 9,900 miles per year and 9.6 years' riding experience.

Apparatus

The comfort trials were held at Bruntingthorpe test track on 30 August 2007. Eight types of motorcycle were used: **supersports** (Honda CBR600RR), **naked road bike** (Kawasaki Z750), **large enduro** (BMW R1200GS), **small enduro** (BMW G650X Country), **faired road bike** (Suzuki Bandit 1200), **tourer** (Yamaha FJR1300), **sports-tourer** (Triumph Sprint) and **cruiser** (Yamaha Midnight Star). Questionnaires were used to collect demographic data and to assess rider perceptions and experiences of comfort.

Design & procedure

The combination of five riders and eight motorcycles produced a within-subjects matrix of 40 randomised comfort trials. The following data were collected: rider perceptions and experience of comfort, anthropometric data, body part discomfort rated after each ride, motorcycle comfort ratings and motorcycle design ratings. Participants were invited to volunteer through 'Bike' magazine's database of riders. Riders rode each motorcycle around the test track and were instructed to ride as they would on a real road. After 30 minutes, riders returned to be interviewed by magazine staff before completing ergonomics questionnaires. Each rider completed eight riding trials (one on each motorcycle).

Results

Rider perceptions and experience of comfort

Pre-ride perceptions and post-ride experiences of comfort were analysed. Prior to the trials the large enduro and sports tourer were perceived as most comfortable whilst it was almost unanimous that the supersports bike appeared to be the most uncomfortable. After completing the comfort trials the sports tourer was regarded as most comfortable whilst the cruiser and small enduro were regarded as least comfortable.

Anthropometric data

As only 1 female rider took part in the trial descriptive statistics are presented for male data in Table 1. 'Bike' data from the comfort trials was compared to UK 50th percentile population data for 19–65 years (from: Pheasant & Haselgrave, 2006, based on a 1981 OPCS survey).

Body part discomfort

Primary and secondary regions of discomfort were collated according to criteria from a previous motorcycle comfort study (Stedmon, 2007). Primary areas of discomfort were apparent in the wrists, neck, shoulders, hands, mid-to-lower back,

Table 1. Descriptive statistics for male motorcycle rider anthropometry (n = 4).

	Weight (kg)	Stature (mm)	Acromion grip (mm)	Inner leg (mm)	Hip width (mm)	Knee height (mm)	Buttock to knee (mm)	Trunk height (mm)	Thigh circumference (mm)
'Bike'									
mean	109.4	1800	627.5	763.8	381.3	575.0	580.0	642.5	588.8
(SD)	(50.4)	(11.6)	(87.6)	(38.2)	(53.6)	(58.0)	(29.4)	(99.5)	(131.6)
UK									
mean	55	1740	n/a*	n/a*	360	545	595	n/a*	n/a*
(SD)	(12)	(70)			(29)	(32)	(31)		

* non-standard measures where data are not available in Pheasant & Haslegrave (2006)

Table 2. Overall comfort ratings (0 = very uncomfortable, 10 = very comfortable).

	Riders					Comfort	
Motorcycle type	Tall	Short	Average	Heavy	Female	Total	Rank
Sports tourer	9	10	9	8	10	46	1
Tourer	8	10	8.5	9	5	40.5	2
Large enduro	7	8	7	8.5	9	39.5	3
Naked road bike	8.75	4	9	7	9	37.75	4
Faired road bike	8	7.5	9	8	5	37.5	5
Supersports	8	4	7	6	10	35	6
Cruiser	4	8	3	3	7	25	7
Small enduro	4	3	0	2	7.5	16.5	8

bum, thighs, and feet. Secondary areas of discomfort were found in elbows, forearms, upper back, and knees. When body part discomfort data were analysed for each type of motorcycle the sports tourer was the most comfortable and small enduro was the least comfortable.

Motorcycle comfort ratings

The overall comfort ratings for each motorcycle are presented in Table 2.

The results illustrate that the sports tourer was rated the most comfortable motorcycle whilst the small enduro was rated the least comfortable motorcycle.

Motorcycle design ratings

The ratings for the most and least comfortable motorcycles are detailed in Figure 1.

The ratings illustrated the difference between the most comfortable (sports tourer) and the least comfortable (small enduro) motorcycles.

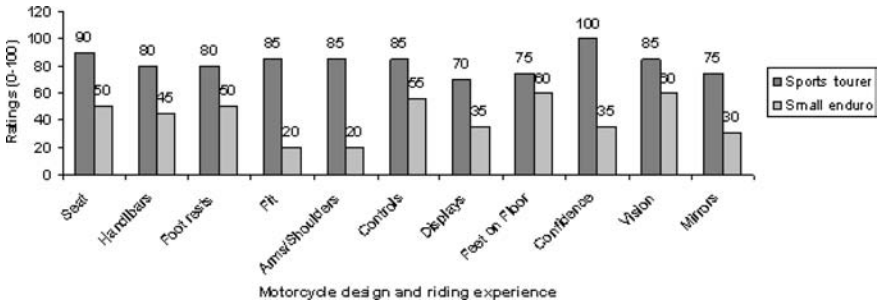


Figure 1. Motorcycle design ratings.

Discussion

Different riders assessed different types of motorcycles:

- **Naked road bike** – offers little weather protection and riders tend to sit upright
- **Faired road bike** – similar to the naked road bike but with added weather protection
- **Supersports** – with the greatest aerodynamic styling riders have to adopt to adopt a low profile posture
- **Tourer** – tends to offer the greatest levels of comfort and weather protection
- **Sports tourer** – this motorcycle offers a compromise between a supersports and tourer motorcycle
- **Enduro (Large and small)** – similar to touring motorcycles but usually sit higher on the road and can be more affected by the weather
- **Cruiser** – are the traditional 'custom-styled' motorcycles offering a low, relaxed riding position characterised with the rider sitting back on a low, small, seat

Riders form impressions about motorcycle comfort even before they sit on them. Some riders might expect a certain degree of discomfort with certain types of motorcycles (the supersports motorcycle was rated by most riders as being potentially uncomfortable). In the trial, after riding the motorcycles, riders altered their perceptions and even the largest rider was surprised by the comfort of a supersports motorcycle. Based on riding experience in the trials the sports tourer was rated the most comfortable whilst the small enduro was the least comfortable.

The anthropometric data supported the findings of Robertson & Minter (1996) illustrating that the motorcyclists were generally larger than the general UK population for stature. It must be noted however, that with such a small sample and each of the riders chosen to represent specific anthropometric characteristics (e.g. tall, short, average, heavy and female) no statistical analysis was conducted on the data (for a recent investigation into motorcyclist anthropometrics please refer to Stedmon et al, 2008).

The body part discomfort ratings supported a general trend as reported by Stedmon (2007) where the wrists, lower arms, neck and bum were affected by riding.

In the current study however, there were more body areas identified than the previous study illustrating a greater degree of variance in the ratings. This may have been due to the relatively short time riders had on each motorcycle (approximately 30 minutes). This may have caused some contamination of data with symptoms from previous rides carrying over to subsequent rides if there was not enough time for them to fully develop/fade. When body part discomfort was analysed for each type of motorcycle the sports tourer was generally the most comfortable and the small enduro was the least comfortable. This finding supported the post-trial rider perceptions and was reinforced with the motorcycle comfort ratings (illustrated in Table 2). From the cumulative ratings that each rider gave the motorcycles the sports tourer style of motorcycle gained an almost maximum score. The small enduro styled motorcycle was rated the lowest. This may have been an artefact of the quasi-experiment in that the test track had a 2 mile straight on it which made sustained riding difficult on a physically small motorcycle and may have affected rider perceptions.

From the data a general trend emerged with riders finding the sports tourer most comfortable and the small enduro least comfortable. This was further supported with the results of the motorcycle design ratings. Figure 1 illustrates that the sports tourer did not score less that 70 on any of the dimensions (it scored 80+ on 7 of the 11 dimensions and a full 100 for rider confidence) whereas the small enduro did not score higher than 60 on any dimensions. The two motorcycle types illustrate the value of developing a more sensitive measure of comfort than the single score for motorcycle comfort rating.

This experiment incorporated a number of measures for investigating different aspects of motorcycle comfort. These were presented to the participants on simple questionnaires and provided a method of communicating and exploring the issues in a number of ways. It has also provided a means of developing an understanding of the relationships between the different measures. There appears to be some consistency in the same types of motorcycles generally being rated as the most or least comfortable, which provides a degree of validity to the method.

Conclusion

Comfort is a very subjective concept as some riders will be affected by it to greater or lesser degrees. What this quasi-experiment has shown, however, is that a number of measures can be taken that provide a consistent insight into the many dimensions that affect the perception and experience of motorcycle comfort.

Acknowledgement

This work was conducted alongside work developed for a ‘Bike’ magazine feature (‘Are You Sitting Comfortably’ Nov 2007).

References

Cook, T.D. & Campbell, D.T. (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Rand McNally: Chicago

Pheasant, S.T. & Haslegrave, C.M. (2006). *Bodyspace: Anthropometry, Ergonomics and the Design of Work*. Taylor & Francis Ltd. London.

Robertson, S. & Minter, A. (1996). A study of some anthropometric characteristics of motorcycle riders. *Applied Ergonomics* 27(4) 223–229.

Stedmon, A.W. (2007). Motorcycling is a pain in the neck, numb bum, aching wrists – or all these things! A rider comfort survey. In, P. Bust (ed). *Contemporary Ergonomics 2007*. Taylor & Francis Ltd. London.

Stedmon, A.W., Robertson, S., Yates, T. & Fitzpatrick, B. (2008). Profiling a new user group: anthropometric data for motorcyclists. In, P. Bust (ed). *Contemporary Ergonomics 2008*. Taylor & Francis Ltd. London.

PROFILING A NEW USER GROUP: ANTHROPOMETRIC DATA FOR MOTORCYCLISTS

Alex Stedmon[1], Sandy Robertson[2], Tom Yates[1] & Ben Fitzpatrick[1]

[1]*Centre for Motorcycle Ergonomics, University of Nottingham, Nottingham UK*
[2]*Centre for Transport Studies, University College London, London UK*

This study builds on previous research examining motorcycle rider anthropometry (Robertson, 1987; Robertson & Minter, 1996) which found evidence that UK motorcycle riders were taller than the general population. A sample of 57 motorcycle riders was obtained and a range of dimensions were measured. Male motorcycle rider stature was significantly greater than the general population which supported the findings of previous research and provide a strong indication that motorcyclists may be a unique user group that requires further investigation to build a larger data set in the future.

Fitting the rider the to the motorcycle

Motorcycles, by their design, tend not to 'fit the user' and instead riders are expected to 'fit the motorcycle' (Robertson 1987; Stedmon, 2007). This goes very much against user-centred design principles and often dictates the market for different types of motorcycle based on the physical size of the rider. Standard motorcycles are often very limited in their adjustability and design features such as controls, displays, seating and wind protection all depend on the relative size and location of the rider to be effective (Roberston & Minter, 1996). In a recent comfort survey it was clear that 'one size does not fit all' and different sized riders will have very different experiences of the same motorcycle (Stedmon, et al, 2008). Given the limited adjustment of motorcycles, it is important to understand the physical characteristics of the user population if design solutions are to fully support them.

Rider anthropometrics

Previous measurements of samples from the population of UK motorcycle riders have shown that they are consistently taller than the normal UK population (Robertson 1987; Robertson & Porter, 1987; Roberston & Minter, 1996). In the study by Robertson (1987) self-reported stature data proved a useful data collection exercise however there were some concerns about the reliability of the data. These were addressed in a follow up study by Roberston & Minter (1996) where measurements were taken by ergonomists using an anthropometer (accurate to 1 mm) under

controlled conditions. The distributions of the stature data from both these studies were remarkably similar and illustrated that the motorcyclists were proportionally taller by approximately 35 mm than the general UK population. Part of the reason for Robertson & Minter's study was the debate resulting from a UK Government initiative to develop 'leg protectors' as a safety device for motorcyclists. A critical feature of these protectors was the small tolerance required for the location of the rider's lower leg/knee behind the protector. The results of the study illustrated a high degree of variability in knee position against the side of the tank (as a function of upper and lower leg size relative to sitting posture). Prior to this little was known about the relative positions of various parts of the body against the motorcycle when riding.

Since 1996 no follow up work on rider anthropometrics has been reported. There have been major changes in motorcycle styling, particularly with sports bikes, where seat height has risen by as much as 60 mm in the last 25 years with the result that shorter riders find it more difficult to place their feet on the ground (Hargreaves, 2005). From the design of many motorcycles it is not clear if any consideration has been given to the different sizes of people who may wish to ride them and it has been suggested that motorcyclists are taller than the general population due to a form of 'natural selection' occurring as shorter riders are excluded because they are not big enough to fit many motorcycles (Robertson, 1987). With no recent anthropometric data, this provided the basis for the current investigation.

Method

The methodology used in the current study followed that of Robertson & Minter (1996) including the use of the same annual British Motorcycle Federation (BMF) event to obtain a sample of riders which facilitated comparisons between the data sets.

Participants

A self-selected sample of 59 motorcyclists (48 male and 11 female) volunteered to be measured. The participants were aged between 16–67 years old (mean age = 44.69 years, SD = 9.33 years). Male motorcyclists ranged from 20–67 years (mean age = 45.19 years, SD = 8.89 years) and female motorcyclists ranged from 16–60 years (mean age = 42.55 years, SD = 11.28 years).

Apparatus

An anthropometer (accurate to 1 mm) was used to measure the participants and digital scales were used to weigh them.

Procedure

Participants were recruited at the BMF show during 19–20 May 2007. They were asked if they had taken part in the previous anthropometric study in 1993 and invited

to take off jackets and footwear. They were weighed and then standing on a flat surface the following measurements were taken: stature, acromion grip and inner leg. With the participant seated (back straight, thighs horizontal and knees at 90 degrees) further measures were taken for: hip width, knee height, buttock to knee, trunk height, and thigh circumference. Trunk height and thigh circumference were included as extra measures in this study as they are important in understanding the 'fit' of the rider to the motorcycle. After the measurements were taken participants were entered into a prize draw.

Results

Anthropometric data

Tables of descriptive statistics for each gender are presented below illustrating the data collected at the in the current study (BMF data) compared to UK 50th percentile population data for 19–65 years (from: Pheasant & Haselgrave, 2006, based on a 1981 OPCS survey).

Table 1. Descriptive statistics for male motorcycle rider anthropometry (n = 48).

	Weight (kg)	Stature (mm)	Acromion grip (mm)	Inner leg (mm)	Hip width (mm)	Knee height (mm)	Buttock to knee (mm)	Trunk height (mm)	Thigh circumference (mm)
BMF									
mean	91.55	1760.6	634.4	796.0	352.9	566.5	614.2	612.9	550.8
(SD)	(22.44)	(8.25)	(42.6)	(53.0)	(36.3)	(35.2)	(40.9)	(45.1)	(66.2)
UK									
mean	55	1740	n/a*	n/a*	360	545	595	n/a*	n/a*
(SD)	(12)	(70)			(29)	(32)	(31)		

Table 2. Descriptive statistics for female motorcycle rider anthropometry (n = 11).

	Weight (kg)	Stature (mm)	Acromion grip (mm)	Inner leg (mm)	Hip width (mm)	Knee height (mm)	Buttock to knee (mm)	Trunk height (mm)	Thigh circumference (mm)
BMF									
mean	70.64	1650.0	598.2	769.1	343.6	529.1	579.1	584.5	50.73
(SD)	(10.92)	(88.1)	(31.6)	(50.3)	(28.7)	(32.7)	(40.9)	(40.9)	(35.2)
UK									
mean	63	1610	n/a*	n/a*	370	500	570	n/a*	n/a*
(SD)	(11)	(62)			(38)	(27)	(30)		

* non-standard measures where data are not available in Pheasant & Haslegrave (2006)

Statistical analyses

Independent sample T-test analyses were conducted on the BMF data compared with data from Pheasant & Haslegrave (2006). Significant effects were observed for weight: male riders ($p < 0.001$) were heavier (91.55 kg) than the general population (55.0 kg) and female riders ($p < 0.05$) were heavier (70.63 kg) than the general population (63.0 kg). A significant effect was observed for stature: male riders ($p < 0.05$) were taller (1760.6 mm) than the general population (1740.0 mm). Significant effects were observed for knee height: male riders ($p < 0.001$) were larger (566.5 mm) than the general population (545.0 mm) and female riders ($p < 0.01$) were larger (529.1 mm) than the general population (500.0 mm). A significant effect was observed for buttock to knee: male riders ($p < 0.001$) were larger (614.2 mm) than the general population (595 mm). As with Robertson & Minter (1996), the BMF stature data were also compared to car driver data (Haslegrave, 1980). A significant effect was observed for male riders ($p < 0.05$) illustrating that male riders are taller (1760.6 mm) than car drivers (1738.1 mm). No significant effect was observed for female riders ($p > 0.05$).

Discussion

The significant effects that were observed illustrated that both male and female riders are significantly heavier and lower leg measurements are significantly larger than the general UK population. Given the styling of modern motorcycles and the emphasis on higher seats, it could be that riders with proportionally longer legs 'fit' motorcycles better than those with shorter legs. The current stature data is consistent with the earlier research (Robertson, 1987; Robertson & Minter, 1996), illustrating that male riders are taller than the general population. Figure 1 below illustrates the

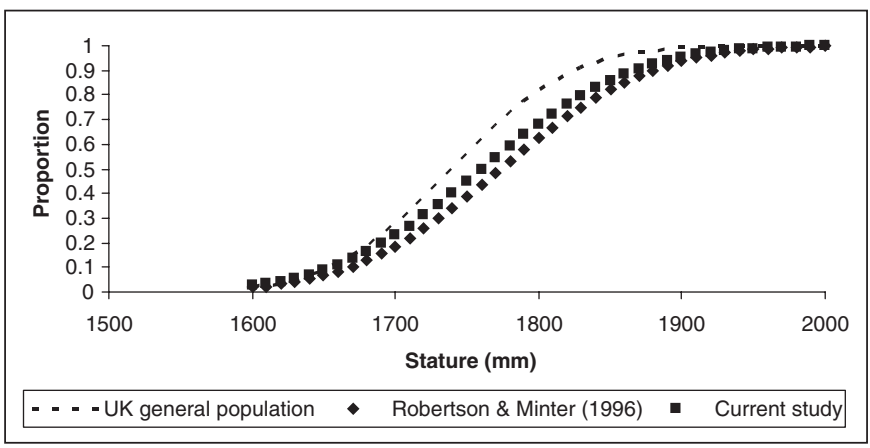


Figure 1. Distributions of stature based on reported mean and standard deviation.

distribution of stature data of male riders for the UK general population, BMF data reported in the current study and that from Robertson & Minter (1996).

The graph illustrates that whilst the current stature data was lower than Robertson & Minter's data, both data sets are significantly higher than the general population. The combined data of Robertson & Minter (1996) and the current study only accounts for 157 participants from two self-selected samples and some degree of caution should be considered in generalising the significant effects to the total population of motorcyclists. However, the two data sets collected over 14 years apart both illustrate a significant difference to the UK general population. They therefore provide a strong indication that motorcyclists may be a unique user group that requires further investigation to build a larger data set in the future.

If the argument that modern motorcycle styling has led to increased seat height is correct (Hargreaves, 2005) and that some form of 'natural selection' excludes shorter people from riding motorbikes (Robertson, 1987) then the current study might have been expected to show an increase in the stature of riders over the earlier data. However, this was not the case as the current stature data is slightly, but not significantly, shorter (1760.6 mm) than the previous study (1774.0 mm). This is still consistent with the hypothesis that there is a form of self-selection taking place by the user population due to the physical size of the motorcycle. However, the samples included different age distributions and possibly different motorcycle types. This, coupled with the lack of information about the dimensions (seat height) of the motorcycles that riders rode in both samples means that a direct relationship between changes in motorcycle design and stature of the population cannot be effectively investigated. At this stage, the issue remains open and further investigation is required.

Conclusion

This study was a direct follow up to Roberston & Minter's previous research and illustrated similar findings 14 years after the first study was conducted. There is evidence that the male motorcycle rider population is still taller than the general population and car driver population, which supports the notion that motorcyclists are a unique user group that requires further investigation and systematic measurement to build a larger data set in the future.

Acknowledgement

The authors would like to thank Triumph Motorcycles and the British Motorcyclists Federation for their support.

References

Hargeaves, S. (2005) Wanted: A sportsbike for big blokes. *Bike Magazine* (May) 83–88.

Pheasant, S.T. & Haslegrave, C.M. (2006) *Bodyspace: Anthropometry, Ergonomics and the Design of Wor*k. Taylor & Francis Ltd. London.

Robertson, S.A., (1987) *Exploratory Motorcycle Ergonomics*. Universities Transport Studies Group Conference, London, 1987, Paper TSU ref. 406.

Robertson S.A. & Porter J.M. (1987) Motorcycle ergonomics: an exploratory study. In, E.D. Megaw (ed.) *Contemporary Ergonomics 1987*. Taylor & Francis: London.

Roberston, S.A. & Minter, A. (1996) A study of some anthropometric characteristics of motorcycle riders. Applied Ergonomics **27**, 223–229.

Stedmon, A.W. (2007) Motorcycling is a pain in the neck, numb bum, aching wrists – or all these things! A rider comfort survey. In, P. Bust (ed.). *Contemporary Ergonomics 2007*. Taylor & Francis: London.

Stedmon, A.W., Hargreaves, S., Fitz-Gibbons, M. & Wood, C. (2008) Different riders on different motorbikes: a quasi-experimental approach to rider comfort. In, P. Bust (ed) *Contemporary Ergonomics 2008*. Taylor & Francis: London.

SPATIAL DIFFERENCES IN SENSIBLE AND LATENT HEAT LOSSES UNDER A BICYCLE HELMET

Guido De Bruyne[1], Jean-Marie Aerts[1], Georges Vander Perre[2], Jan Goffin[3], Ignace Verpoest[4] & Daniel Berckmans[1]

[1]*M3-BIORES, KULeuven, Leuven, Belgium*
[2]*Biomechanics and Engineering Design, KULeuven, Leuven, Belgium*
[3]*Experimental Neurosurgery and Neuroanatomy Section, KULeuven, Leuven, Belgium*
[4]*Metallurgy and Materials Engineering, KULeuven, Leuven, Belgium*

This research aims at quantifying spatial gradients in skin temperature and sweat production under a bicycle helmet. Distribution of sweat production, skin temperature and air temperature was measured at different positions under a bicycle helmet on five male and four female test persons. Male test persons performed a physical exercise of 100 Watt during the first 20 minutes (low effort level) and 150 Watt during the next 20 minutes (high effort level). Female test persons performed a physical exercise of 80 Watt during the first 20 minutes (low effort level) and 120 Watt during the next 20 minutes (high effort level). Two zones with significantly different skin temperatures were found during low effort level and three zones where found during high effort level. The average temperature difference between frontal and temporal area was 4.6°C at low effort level and 5.5°C at high effort level. Sweat production was found to be not significantly different ($P > 0.05$) for the four different positions on the head during low and high effort level due to the large differences in sweat production between the different test subjects. This was shown by the six out of nine test persons that had a significant difference in sweat production between the frontal and the parietal region. Finally, two zones of significantly different air temperature were found between head and helmet for high and low effort level. Average temperature was 2.1°C warmer at the occipital area compared with the frontal area at high effort level. The results of this research can be used in future to help designing helmets with better thermal comfort.

Introduction

Every year, many cyclists die as a consequence of an unfortunate tumble or a traffic accident in which they suffer from a severe head injury. Most of these head injuries could be avoided or strongly reduced by wearing a cyclist crash helmet (Burke),

Figure 1. Side view of windtunnel.

but many cyclists still do not like to wear a crash helmet because of the thermal discomfort that comes along with wearing it (Wardle and Iqbal). The human head plays an important role in the thermoregulation of the body. Up to 50% of the produced latent and sensible heat can be transferred via the head, causing high temperatures and sweat production on the head (Rasch et al.). Research has shown that, on whole body scale, the amount of sweat production differs from body part to body part (Johnson et al.) and that the distribution of the total amount of latent heat losses changes with changing effort level and environmental conditions (Desruelle and Candas).

Liu and Holmer as well as Bruhwiler et al. investigated the ventilation efficiency of bicycle helmet with sweating manikin heads. Their methods are useful for controlled experiments, but spatial gradients in local and sensible heat loss are neglected.

Many researchers (for example Coleman and Mortagy, Burns, Gisolfi et al., John and Dawson, and Davis et al.) have investigated the thermal response that comes along with wearing a bicycle helmet or sports helmet, but none of them have investigated gradients in sensible and latent heat loss under a helmet. It was expected that latent and sensible heat losses differ significantly on various positions between head and helmet.

Materials and methods

Test installation

The test setup consisted of an ergometer (Tunturi, type T300), placed in a small wind tunnel that was placed in a climatic chamber.

The test setup consisted of two basic units: the climatic chamber, the ergometer and the sensors. The climatic chamber had two units: climate control, and wind tunnel.

Three types of variables were measured: sweat production, skin temperature and air temperature. Sweat production was acquainted by a SKINOS SKD-4000. The device offered four probes for sweat measurements (mg/min, cm^2) (Fig. 2). Eight thermocouples (type T) were used to measure air temperature distribution between

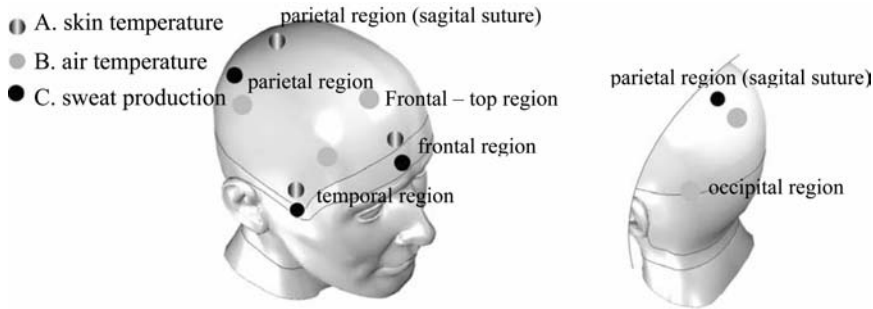


Figure 2. Positions of the measurement positions.

head and helmet (Fig. 2) and skin temperature (Fig. 2) (°C). A standard commercial helmet was used for the test.

Experiments

Five male and four female test persons were used in the test. Average age was 26,7 years (SD 3,5). Ambient air temperature was set at 20°C and air velocity was set at 2.5 m/s at 50 mm in front of the frontal area of the test subject. Air velocity was low (2.5 m/s) due to our interest in thermal comfort for recreational cyclists and children. Two effort levels were used during each test. A first period (low effort level) of twenty minutes was set at 100 Watt for men and 80 Watt for women. The second period (high effort level) of twenty minutes was set at 150 Watt for men and 120 Watt for women.

Statistical analysis

Statistical analysis was performed using ANOVA1 to test if the differences between the sensor positions for each variable were significant. A multiple comparison test was used to visualize sensor positions with significantly different results, enabling a classification between sensor positions. All variables were analyzed from minute 15 until 20 and from minute 35 until 40. The null hypothesis stated that there was no significant ($\alpha = 0.95$) difference between the variables. The alternative hypothesis stated that there was ($\alpha = 0.05$) a significant difference between the variables.

Results

Skin temperatures

Skin temperature was measured at three locations as can be seen in Figure 2. P-values far below 0.05 were calculated for both low effort level and high effort (5.91E-06 for low effort level and 1.15E-06 for high effort level). The results clearly indicate that significant gradients in skin temperature at three positions on the head

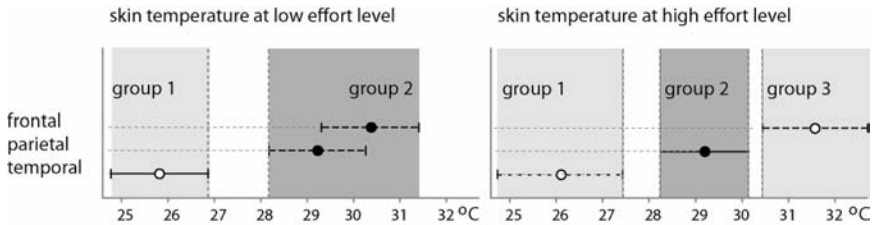


Figure 3. The results of the multiple comparison test. Average skin temperature with a 95% confidence interval is plotted per position.

were observed when a high effort level (150 Watt for men and 120 Watt for women) was applied when subjects wore a bicycle helmet.

Figure 3 shows the results of the three positions where skin temperature was measured. Two groups could be observed when effort level was low (100 Watt for men and 80 Watt for women). Frontal and parietal region, forming group 2, had results that are not significantly different from each other. Both positions had results that were significantly different compared to the temporal region, forming group 1. Three significantly different groups could be observed when effort level was high. This shows that all positions, where skin temperature was measured, had temperatures that were significantly different from each other when effort level was high.

When effort level was low, average temperature (during the last five minutes of that effort level) was 1.2°C warmer at the frontal area compared to the parietal area. Additionally, average temperature was 4.6°C warmer at the frontal region compared to the temporal region and average temperature of the parietal region was 3.4°C warmer compared to the temporal region. When effort level was high, average temperature was 2.4°C warmer at the frontal compared to the parietal region and 5.5°C compared to the temporal. Average skin temperature varied 0.7°C between high and low effort level. The results show that the temporal region was the coldest region that was measured for both low and high effort level. The frontal and parietal regions were the warmest regions at low effort level. The frontal region was the warmest area at high effort level.

Air temperature

P-values far below 0.05 were calculated for variations in air temperature at both low and high effort level (0.001 at low effort level and 0.004 at high effort level). The alternative hypothesis, stating that the five positions were air temperature was measured between head and helmet, are part of different groups, was accepted for both high and low effort level. The results clearly indicate that significant gradients in air temperature at five positions between head and helmet were observed when a high effort level was applied when subjects wore a bicycle helmet. Figure 4 shows the results of the five positions where air temperature was measured between head and helmet. Two groups can be observed when effort level is low and when effort level is high. However, only the frontal top region is part of group 1: both when effort level is low and high. Parietal, parietal (sagital suture) and occipital region

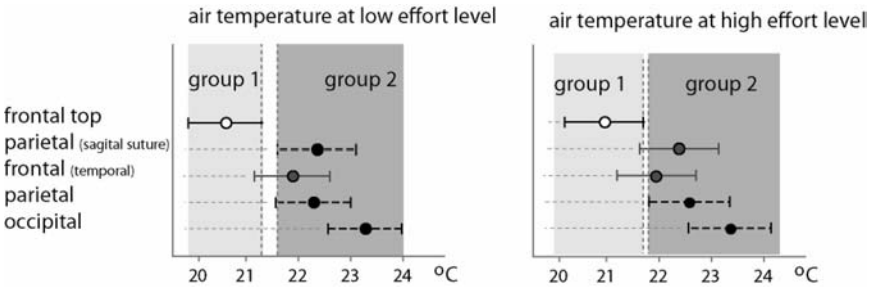


Figure 4. The results of the multiple comparison test. Average air temperature with a 95% confidence interval is plotted per position.

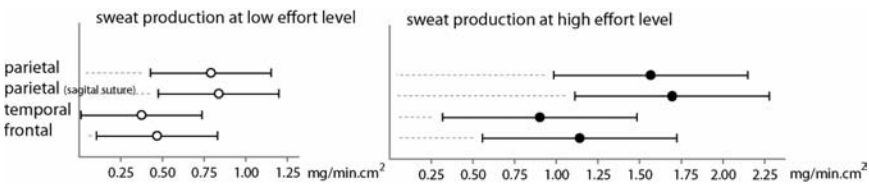


Figure 5. The results of the multiple comparison test. Average sweat production with a 95% confidence interval is plotted per position.

are part of group 2 when effort level is low. When effort level is high the parietal and occipital region are part of group 2. The results show that air temperatures at the front of the helmet are cooler as elsewhere under the helmet. Mean air temperature was up to 2.40°C warmer at the occipital region compared to the frontal top region when effort level was high and 2.66°C when effort level was low.

Sweat production

The results of the ANOVA1 test show that the 4 positions where sweat production was measured were not part of different groups ($P > 0.05$, $P = 0.24$ at both low and high effort level). This shows that gradients observed during the test were not significant. However, a significant difference ($P < 0.05$) in sweat production between frontal and parietal (sagital suture) region was observed for both high and low effort level for six individual test subjects when an ANOVA1 test was performed per test subject. This shows that sweat production can be spatially different for an individual person, but that no different zones of sweat production on the head could be identified for the nine test subjects that were investigated.

Results

This research quantified local thermal aspects under bicycle helmets. Distribution of air temperature, skin temperature and sweat production was measured under a bicycle helmet on five male and four female test persons during an ergometer test

with a step in effort level. Skin temperatures under the helmet were found to be not spatially homogeneous ($P < 0.05$). For low effort level (100 Watt for men and 80 Watt for women), two zones of skin temperatures were identified and average temperature was up to 4.6°C warmer at the frontal region compared to the temporal. Three zones of skin temperature were identified for high effort level and average temperature within the zone was up to 5.5°C warmer between the frontal and the temporal region. Also spatial gradients in air temperatures under the helmet were observed ($P < 0.05$). Two zones could be identified at low and high effort level. At high effort level, average temperature was 2.1°C warmer at the occipital area compared with the frontal top area that is near to the inlet. Finally, also sweat production was found to be not homogeneous over the head ($P < 0.95$), but these results were not significantly different ($P > 0.05$) due to the large variations in sweat production between the different test subjects. However, six test persons out of nine did have a significant difference in sweat production between the frontal and the parietal region.

This research can help (virtual) thermal manikins to become more accurate compared to real life experiments.

References

Bruhwiler PA, Buyan M, Huber R, Bogerd P, Szitman J, Graf SF, and Rosgen T. 2006. *Heat transfer variations of bicycle helmets*. Journal of Sports Sciences 24(9), 999–1011.

Burke ER. 1988. *Safety standards for bicycle helmets*. The Physician and Sportsmedicine 16(1), 148–153.

Burns R. 1975. How hot the hat? Bike World Magazine, 30–31.

Coleman AE and Mortagy AK. 1973. Ambient head temperature and football helmet design. Medicine and Science in Sports 5(3), 204–208.

Davis GA, Edmisten ED, Thomas RE, Rummer RB, and Pascoe DD. 2001. *Effects of ventilated safety helmets in a hot environment*. International Journal of Industrial Ergonomics 27, p. 231–329.

Desruelle AV and Candas V. 2000. *Thermoregulatory effects of three different types of head cooling in humans during a mild hyperthermia*. European Journal of Applied Physiology and Occupational Physiology 81(1–2), p. 33–39.

Gisolfi CV, Rohlf DP, Navarude SN, Hayes CL, and Sayeed SA. 1988. *Effects of wearing a helmet on thermal balance while cycling in the heat.* The Physician and Sportsmedicine 16(1), p. 139–146.

John D and Dawson B. 1989. *The effects of wearing two different cycling helmets on thermoregulatory responses to prolonged submaximal exercise in hot, dry conditions.* Journal of Human Movement Studies 16, 203–214.

Liu XX. and Holmer I. 1997. *Evaluation of evaporative heat-transfer characteristics of helmets.* Applied Human Science 16(3), 107–113.

Rasch W, Samson P, Cote J, and Cabanac M. 1991. *Heat loss from the human head during exercise.* Journal of Applied Physiology 71(2), p. 590–595.

Wardle S and Iqbal Z. 1998. *Cycle helmet ownership and wearing: results of a survey in South Staffordshire.* Journal of Public Health Medicine 20(1), 70–77.

Author index

Abbas, M.Y. 303, 309, 315, 321
Abbott, R. 496
Aerts, J-M. 785
Ahmad, R.K. 215
Alcántara, E. 687
Algora, E. 687
Allen, P.H. 597, 607
Anderson, J. 8
Aujla, M. 327

Bainbridge, L.C. 425
Baines, R. 333, 338, 343
Baird, A. 34, 441
Balfe, N. 552
Barr, D. 485
Barrett, L. 75
Beards, S. 23
Bell, J.L. 693
Benedyk, R. 389
Beschorner, K.E. 717
Bird, F. 589
Birrell, L.N. 425
Birrell, S.A. 502, 749
Birtles, M. 453
Blechko, A. 145
Bowes, C. 699
Brace, C. 197
Buckle, P. 3
Burton, A.K. 425
Bust, P. 197
Bust, P.D. 221

Canham-Chervak, M. 705
Cardoso, C. 291
Carey, M. 570
Carhart, M. 669
Carse, A. 139
Castelló, P. 179
Catchpole, K. 357
Chan, J.W. 127
Chandler, E. 29
Chang, C-C. 675
Chang, W-R. 693, 735
Chen, F. 767
Chen, Y. 375
Chesters, R. 333, 338, 343
Chiou, S. 693
Ciccarelli, L. 669
Clark, R. 626
Clarke, S. 273
Clarke, T. 546, 552, 564, 577
Clarkson, J. 267, 279, 297
Clarkson, P.J. 273, 291
Clayton, R.J. 49
Cobb, S. 92
Cohen, H.H. 621
Coleman, J. 496
Collins, J.W. 693
Cooper, L.H.K. 381
Courtney, T.K. 693
Cummings, R. 61

Dadashi, Y. 546
Darker, I. 145
David, G. 23
David, H. 239
Davies, J.C. 699
De Bruyne, G. 785
Devereux, J. 406
Dianat, I. 431
Dickinson, C. 511
DiDomenico, A. 675
Di Jorio, S. 663, 669
Dingwall, P. 564
Downes, D. 127
Dunham, P. 106
D'Cruz, M. 92

Edmunds, J. 164
Edwards, D.J. 191
Ellegast, R. 743
Ellegast, R.P. 400
Enflo, K. 755
Ergs S, F. 101
Evanoff, B. 693
Evans, J. 471
Evans, P. 127
Fain, N. 69
Farrington-Darby, T. 577
Faruk, M. 303, 309, 315, 321
Ferreira, J. 453, 459
Ferreras, A. 179
Fischer, S. 743
Fitz-Gibbons, M. 773
Fitzpatrick, B. 779
Fowkes, M. 749

Gale, A. 145, 285, 375, 381
Gale, A.G. 363
García, C. 179
Geddes, N. 261
Gibb, A.G.F. 197, 221
Gibson, H. 158
Gibson, W.H. 164
Goel, V.K. 40
Goffin, J. 785
Golightly, D. 92
Goodman-Deane, J. 273
Gray, M. 459, 761
Gray, R. 663
Gregson, W. 485
Grönqvist, R. 693, 735
Gyi, D.E. 411
Gálvez, J. 179

Hadjev, N. 657
Hadjisimou, M. 389
Haines, V. 261
Halford, V. 621
Hallas, K. 723
Han, D.R.C. 395
Hancock, P.A. 417
Hancock, R. 40
Hansen, J.P. 285
Hansson, B. 215
Hargreaves, S. 773
Harris, S. 133, 139
Haslam, R.A. 502
Haslegrave, C.M. 431
Hassana, R.A. 112

Hatanaka, S. 644
Hauret, K.G. 705
Heaton, N. 471
Heller, M. 663, 669
Hermanns, I. 400, 743
Higgs III, C.F. 717
Hignett, S. 351
Hirvonen, M. 735
Hokkirigawa, K. 644
Hollowood, J. 92
Horton, K. 8
Houlihan, R. 723
Hsiao, H. 185
Hull, S. 29
Hutchings, J. 523

Inoue, Y. 710

Jackson, A. 153
James, L.T.C. 417
Johnson, C. 613
Johnson, G. 121
Jones, B.H. 705
Jordan, P.W. 81, 767

Kato, M. 710
Kaufman, L. 127
Kemp, G.J. 699
Kendall, N.A.S. 425
Kennedy, G.A.L. 49
Kennedy, R. 164
Kenvyn, F. 540
Kim, I-J. 710, 728
King, G. 577
Kirwan, B. 158, 164
Kleinert, J. 400
Koh, I. 395
Kooijman, A. 69
Kreckler, S. 357
Kuzel, M. 663, 669

Langdon, P. 267, 273, 279, 291, 297
Leaviss, J. 221
Lehane, P. 626
Lewis, T. 297
Lim, K.Y. 395
Lindgren, A. 767
Lombardi, D.A. 693
Lovell, M. 717
Lowe, E. 523, 558
Lowten, M. 517
Lyons, M. 369

Mackie, E. 613
MacLennan, H. 303, 309, 315, 321
Mansfield, N.J. 40, 191, 203
Manson, R. 406
Massicotte, A. 657
Matz, M. 693
Matz, S. 735
McCulloch, P. 357
McGorry, R.W. 675
McGovern Narkevicius, J. 529
McKenzie, J. 471
McNeill, M. 75
Merlo, L. 417
Mishra, A. 357
Mitchell, J. 333, 338, 343
Moellenbach, E. 285
Moes, N.C.C.M. 69
Molloy, E-M. 49, 55
Moncrieff, K. 92
Morris, N. 479
Murphy, P. 534

Nagata, H. 710, 728
Narayanamoorthy, R. 40
Nelson, A. 693
Nery, D. 267
Newell, G. 203
Newman, M. 106
Newton, R. 303, 309, 315, 321
Nolan-McSweeney, M. 523
Notini, L. 191, 203

Oltra, A. 179
Ormerod, M. 303, 309, 315, 321
Ow, S.Z.Y. 395
Oxley, L. 465
Oztug, O. 3
O'Regan, S. 209

Page, Á. 179
Parsons, C. 436, 447
Patel, H. 92
Payá, F. 687
Pearce, B.G. 425
Perre, G.V. 785
Persad, U. 279
Pettitt, M. 92
Pickup, L. 558
Piedrabuena, A. 179
Piek, J. 496
Poirier, P. 651, 657
Pollock, C. 496
Portolés, J. 687
Punchihewa, H.K.G. 411

Quirion, F. 651, 657

Raffler, N. 743
Redfern, M. 717
Reilly, T. 485, 491
Richardson, M. 34
Richter, M. 170
Riley, C. 121
Riley, D. 459, 465
Roberts, G. 517
Robertson, S. 779
Rowe, I. 638
Rowlinson, S. 227
Roys, M. 632, 681
Ryan, B. 523, 564

Schock, A. 564
Scott, A. 583
Scott, H. 375
Sharples, S. 92, 133, 139, 546, 552
Shaw, R. 723
Siemieniuch, C. 55
Siemieniuch, C.E. 49
Sinclair, M. 55
Sinclair, M.A. 49
Smith, A. 602
Smith, A.P. 597, 607
Solisch, T-O. 170
Sorock, G.S. 693
Stammers, R.B. 252
Stanley, L. 453, 459
Stanton, N.A. 749
Stedmon, A. 92, 133, 139, 327, 773, 779
Straker, L. 496
Straughn, S. 663
Stubbs, D. 3, 23
Sutton, L. 491

Tainsh, M. 101
Tanuj, D. 395
Tapley, S. 465
Taylor-Phillips, S. 363
Thorpe, N. 749
Thorpe, S. 723
Tortosa, L. 179
Trigg, M. 471
Tyers, C. 209

Umbers, I. 164

van Doorn, E. 69
Vassiliades, V. 127
Verpoest, I. 785

Wadsworth, E.J.K. 597, 607
Waller, S. 291
Wallis, M. 363
Wang, H-F. 14, 86, 246
Wang, K-L. 86
Ward, J. 121
Watanabe, H. 710
Waterson, P. 351
Weatherill, W. 523
Weber, B. 400
Wellman, H.M. 693
Wilkinson, J. 153
Wilson, J. 92, 133, 139
Wilson, J.R. 546, 552, 564
Wolf, L. 693
Wood, C. 773
Woodcock, A. 106, 112, 447, 613
Woods, V. 3, 23
Wright, M. 632, 681

Yamaguchi, T. 644
Yates, T. 779
Yip, B. 227
Young, M.S. 749
Yu, S-Y. 14
Yung, R. 395

Zamora, T. 687

Contemporary Ergonomics 2008

Editor: Philip D. Bust

The broad and developing scope of ergonomics — the application of scientific knowledge to improve peoples' interaction with products, systems and environments — has been illustrated over two decades by the books which make up the Contemporary Ergonomics series. Presenting the proceedings of the Ergonomics Society's annual conference, the series embraces the wide range of topics covered by ergonomics. Individual papers provide insight into current practice, present new research findings and form an invaluable reference source. The volumes provide a fast track for the publication of suitable papers from international contributors. These are chosen on the basis of abstracts submitted to a selection panel in the autumn prior to the Ergonomics Society's annual conference held in the spring.

A wide range of topics are covered in these proceedings, including: occupational health and safety; slips trips and falls; health and well being of construction workers; inclusive design; hazardous industries; HCI; rail ergonomics; complex systems; ergonomics 4 schools; and applications of ergonomics.

As well as being of interest to mainstream ergonomists and human factors specialists, Contemporary Ergonomics will appeal to all those who are concerned with people's interactions with their working and leisure environment — including designers, manufacturing and production engineers, health and safety specialists, occupational, applied and industrial psychologists, and applied physiologists.

ISBN 0415465753

9 780415 465755

THE Ergonomics society