Volume 16

AHFE International

Neville Stanton
Steve Landry
Giuseppe Di Bucchianico
Andrea Vallicelli. Editors

Advances in Human Aspects of Transportation

Part II

Open Access Science in Human Factors Engineering and Human Centered Computing

Issue 16 2021

Advances in Human Aspects of Transportation

Part II

Advances in Human Aspects of Transportation

Part II

Edited By

Neville Stanton
Steven Landry
Giuseppe Di Bucchianico
and
Andrea Vallicelli

Published by AHFE International ©

Published by AHFE International ©

No claim to original U.S. Government works

Published in the United States of America

Version Date: 20210710

ISSN 2771-0718 (Online)

Applied Human Factors and Ergonomics International

International Standard Book Number: 978-1-4951-2098-5 (eBook)

Proceedings of AHFE International Conference (AHFE 2021), 25-29 July, 2021, NY, USA.

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the AHFE Web site at
http://www.ahfe.org

Table of Contents

Section 11: Road and Rail - Driver State Detection and Simulated Driving: Drawbacks and Opportunities

Evolution of a new adjustable motorcycle test rig design for measuring motorcyclist fatigue during prolonged riding 3
H. Rashid, M. Ma'arof, A. Omar, S. Abdullah and Roseleena Jaafar, Malaysia

Motorcycling: Awkward posture is the best posture! 12
M. Ma'arof, H. Rashid, A. Omar, S. Abdullah, I. Ahmad, R. Jaafar and S. Karim, Malaysia

Application of galvanic vestibular stimulation for the evaluation of vehicle settings in a fixed-base driving simulator 21
M. Wiedemann, W. Remlinger and K. Bengler, Germany

Comparison of real and simulated driving for a static driving simulator 29
M. Krause, L. Yilmaz and K. Bengler, Germany

How do immersive driving environments influence user performances and experiences? 41
D. Reich, J. Wittkugel and R. Stark, Germany

Section 12: Road and Rail - Highly Automated Driving -Aspects of Driver Vehicle Interaction II

Hands and feet free driving: Ready or Not? 55
V. Banks and N. Stanton, UK

Taking over control from highly automated vehicles 64
C. Gold and K. Bengler, Germany

The "Methods for Designing Future Autonomous Systems" (MODAS) project: Developing the cab for a highly autonomous truck 70
S. Krupenia, A. Selmarker, J. Fagerlonn, K. Delsing, A. Jansson, B. Sandblad and C. Grane, Sweden

End-users' acceptance and use of adaptive cruise control systems 82
N. Strand, M. Karlsson and L. Nilsson, Sweden

Section 13: Road and Rail - Human Mobility Increase as a Key Element in Designing Innovative Transportation Systems

Ergonomic issues in design of innovative means of transportation and transportation systems 95
I. Grabarek and W. Choromanski, Poland

Lever wheelchair for disabled people: An answer to the increasing demand for wheelchair mobility 107
K. Fiok and W. Choromanski, Poland

Analysis of ride comfort on the stairs climbing wheelchair 114
G. Dobrzynski, W. Choromański and I. Grabarek, Poland

Traffic fatalities and serious injuries in Europe – A study of association with cultural, demographic and income variables 126
I. Barata and D. Coelho, Portugal

Simulation study of wheelchair's maneuverability
G. Dobrzynski, W. Choromański and K. Fiok, Poland 137

Section 14: Road and Rail - Perception and Control Issues in the Design of Advanced Driving Assistance Systems

Augmented reality head-up-display for advanced driver assistance systems: A driving simulation study 149
L. Halit, A. Kemeny, H. Mohellebi, S. Garbaya, F. Merienne, S. Michelin and V. Albaut, France

Eco-driving with Electric cars: Effect of continuous and on-demand feedback on driving behavior and safety 158
S. Beloufa, B. Vailleau, J.-M. Boucheix, A. Kemeny and F. Merienne, France

Effects of interface complexity of head-up display on driver's driving performance and self-perception 165
S. Li, D. Wang and W. Zhang, China

Influence of information flows on the results of a driver's activity in the system "driver – car – road – environment" 176
V. Dolya, S. Kobiakov and I. Afanasieva, Ukraine

Section 15: Road and Rail - Transport Data and Analysis

Deconstructing the whole journey traveller experience 186
A. Woodcock and J. Osmond, UK

Comparative analysis of driving maneuvers using 3D body posture data 194
A. Kondyli, V. Sisiopiku, L. Zhao and A. Barmpoutis, USA

Driving behavior at intersections equipped with red light cameras – A Monte Carlo simulation based on driving simulator data 204
J. Cornu, K. Brijs, S. Daniels, T. Brijs, E. Hermans, and G. Wets, Belgium

Electric vehicle drivers overcome range anxiety prior to their first trip 216
S. Mansbridge, M. Burgess and M. Harris, UK

Parking behaviours for inductive charging 224
S. Birrell, D. Wilson, G. Dhadyalla, C. P. Yang, A. McGordon and P. Jennings, UK

Section 16: Road and Rail - Systems Thinking in Road and Rail Transport

Combining human factors methods with transport data recordings 236
G. Walker and A. Strathie, UK

Key players and their interactions within the young driver road safety system: Perspectives and Partners 246
B. Scott-Parker, P. Salmon, N. Goode and T. Senserrick, Australia

Pathways to failure? Using work domain analysis to predict accidents in complex systems 258
P. Salmon, M. Lenne, G. Read, G. Walker and N. Stanton, Australia/UK

How much route guidance is enough? 267
G. Walker and A. Manson, UK

Life in the wild? Using ergonomics systems methods and field of safe travel theory to evaluate intersection designs 277
M. Cornelissen, P. Salmon, N. Stanton and R. McClure, Australia/UK

Extracting design information from the outputs of systems analysis: A case study in public transport ticketing 286
G. Read, P. Salmon, M. Lenne and D. Jenkins, Australia/UK

Section 17: Road and Rail - Electric Vehicles: New Challenges for Human Machine Interaction

Social media monitoring as a tool for the development of disruptive innovations 300
A. Keinath and R. Vilimek, Germany

Does range anxiety exists? Experiences from the ELVIRE project 307
M. Nilsson, Sweden

Attitude towards electric vehicles in the Czech Republic 314
P. Zamečnik, M. Šucha, P. Řezač and K. Bohmova, Czech Republic

New features on noise, vibration and harshness posed by the next generation of vehicles 319
N. Campillo Davo, M. Sanchez-Lozano, R. Peral-Orts, B. Pluymers and W. Desmet, Spain/Belgium

Pushing the limits? Regenerative braking in takeover situations 326
C. Witzlack, P. Cocron and J. Krems, Germany

Examining user-range interaction in battery electric vehicles: A field study approach 334
T. Franke, M. Gunther, M. Trantow, J. Krems, R. Vilimek and A. Keinath, Germany

Section 18: Road and Rail - System Design and Evaluation II

Advanced in-vehicle applications to mitigate driver aggression 348
A. Soro, S. Wollstadter and A. Rakotonirainy, Australia/Germany

Experimental studies on driver's expectations regarding the positions of display and control elements in the car cockpit 359
M. Schwalk, S. Hetzinger and T. Maier, Germany

Evaluation of telematic applications for information provision in public transport 371
A. Szarata, O. Kokkinou, S. Basbas and I. Politis, Poland/Greece

A driving simulator study on the effect of transversal rumble strips located nearby dangerous curves under repeated exposure 381
C. Arien, K. Brijs, W. Ceulemans, G. Vanroelen, E. Jongen, S. Daniels, T. Brijs and G. Wets, Belgium

Development of a concept for ergonomic design and work organization of remote lock operation centers 394
J. Butzler, F. Motz, A. Mertens, H. Horoufchin, C. Broehl, N. Jochems and C. Schlick, Germany

Section 19: Road and Rail - Driver Distraction and Inattention

Perceived and real impacts of digital advertising billboards on driving performance 408
V. Sisiopiku, M. Islam, S. Wittig, S. Welburn and D. Stavrinos, USA

Recording and evaluation procedure of drivers' distraction in accordance with driver's characteristics in high speed arterials 420
E. Misokefalou and N. Eliou, Greece

Driving maneuver prediction based on driver behavior observation 427
F. Diederichs and G. Pöhler, Germany

Impact of music tempo on simulated driving performance 439
J. Navarro and E. Reynaud, France

Section 20: Aviation - Human Factors Issues in Air Traffic Management II

Development and application of a hybrid control theory model to quantify human-machine interaction problems on the flight deck 450
C. Surakitbanharn and S. Landry, Japan/USA

How can procedures and checklists help pilots in abnormal flight situations? 456
A. Haslbeck, P. Gontar and E. Schubert, Germany

Using system-wide trust theory to analyze passenger loss of trust in aircraft automation 462
S. Winter, S. Rice and K. Reid, USA

Fuzzy model of air traffic controller attitude to the risk during tactical and immediate decision making 468
A. Reva and S. Borsuk, Ukraine

Section 21: Aviation - Human Factors Issues in Air Traffic Management III

Air traffic controller workload impact by integration of Datalink system in an advanced airport controller working position 480
M. Ellejmi, S. Dubuisson, M. Bonnier and R. Lane, France

Air traffic controller support by speech recognition 492
O. Ohneiser, H. Helmke, H. Ehr, H. Gurluk, M. Hossl, T. Muhlhausen, Y. Oualil, M. Schulder, A. Schmidt, A. Khan and D. Klakow, Germany

What about the next generation? Assessing experts' judgments of human abilities required for working in a future ATC environment 504
D. Schulze-Kissing and H. Eisfeldt, Germany

Evaluation of an arrival coordinator position in a terminal metering environment 517
M. Kupfer, J. Mercer, T. Callantine, V. Gujral and A. Gomez, USA

Identifying markers of performance decline in air traffic controllers 529
T. Edwards, S. Sharples, B. Kirwan, J. Wilson and N. Balfe, UK/France/Ireland

Preface

Human Factors and Ergonomics have made a considerable contribution to the research, design, development, operation and analysis of transportation systems which includes road and rail vehicles and their complementary infrastructure, aviation and maritime transportation. This book presents recent advances in the Human Factors aspects of Transportation. These advances include accident analysis, automation of vehicles, comfort, distraction of drivers (understanding of distraction and how to avoid it), environmental concerns, in-vehicle systems design, intelligent transport systems, methodological developments, new systems and technology, observational and case studies, safety, situation awareness, skill development and training, warnings and workload.

This book brings together the most recent human factors work in the transportation domain, including empirical research, human performance and other types of modeling, analysis, and development. The issues facing engineers, scientists, and other practitioners of human factors in transportation research are becoming more challenging and more critical.

The common theme across these sections is that they deal with the intersection of the human and the system. Moreover, many of the chapter topics cross section boundaries, for instance by focusing on function allocation in NextGen or on the safety benefits of a tower controller tool. This is in keeping with the systemic nature of the problems facing human factors experts in rail and road, aviation and maritime research– it is becoming increasingly important to view problems not as isolated issues that can be extracted from the system environment, but as embedded issues that can only be understood as a part of an overall system.

In keeping with a system that is vast in its scope and reach, the chapters in this book cover a wide range of topics. The chapters are organized into 30 sections over three volumes.

Part I:

Section 1: Aviation - Human Factors Issues in Air Transportation, Aviation Safety and Risk Analysis
Section 2: Aviation - Human Factors Issues in Air Traffic Management I
Section 3: Maritime - Design Tools and Methods
Section 4: Maritime - Communication and Cognitive Performances
Section 5: Road and Rail - Road Infrastructure, Design and Safety
Section 6: Road and Rail - System Design and Evaluation I
Section 7: Aviation - Recent Methodological Developments and Results from Psychophysiology in Ergonomics (PIE)
Section 8: Road and Rail - Highly Automated Driving - Aspects of Driver Vehicle Interaction I
Section 9: Road and Rail - Human Factors at Level Crossings
Section 10: Road and Rail - Risk and Driving

Part II:

Section 11: Road and Rail - Driver State Detection and Simulated Driving: Drawbacks and Opportunities
Section 12: Road and Rail - Highly Automated Driving -Aspects of Driver Vehicle Interaction II
Section 13: Road and Rail - Human Mobility Increase as a Key Element in Designing Innovative Transportation Systems
Section 14: Road and Rail - Perception and Control Issues in the Design of Advanced Driving Assistance Systems
Section 15: Road and Rail - Transport Data and Analysis
Section 16: Road and Rail - Systems Thinking in Road and Rail Transport
Section 17: Road and Rail - Electric Vehicles: New Challenges for Human Machine Interaction
Section 18: Road and Rail - System Design and Evaluation II
Section 19: Road and Rail - Driver Distraction and Inattention
Section 20: Aviation - Human Factors Issues in Air Traffic Management II
Section 21: Aviation - Human Factors Issues in Air Traffic Management III

Part III:

Section 22: Maritime - Human Diversity
Section 23: Maritime - Training and Work Simulations
Section 24: Road and Rail - Eco-Driving, Public Transport and Pedestrians
Section 25: Road and Rail - Human Factors in Rail Systems
Section 26: Road and Rail - Accident Analysis and Prevention
Section 27: Road and Rail - Individual Differences in Driving
Section 28: Road and Rail - Human Factors in Rail and RLX
Section 29: Road and Rail - System Design and Evaluation III
Section 30: Road and Rail - Naturalistic and Simulator HF Driving Communities

This book will be of interest and use to transportation professionals who work in the road and rail, aviation and maritime domains as it reflects some of the latest Human Factors and Ergonomics thinking and practice. It should also be of interest to students and researchers in these fields, to help stimulate research questions and ideas. It is my hope that the ideas and studies reported within this book will help to produce safer, more efficient and effective transportation systems in the future.

We are grateful to the Scientific Advisory Board which has helped elicit the contributions and develop the themes in the book. These people are academic leaders in their respective fields, and their help is very much appreciated, especially as they gave their time freely to the project.

Road and Rail

K. Bengler, Germany
G.Burnett, UK
P. Chapman, UK
F. Chen, Sweden
D. Coelho, Portugal
L. Dickson-Bull, USA
L. Dorn, UK
I. Glendon, Australia
I. Grabarek, Poland
J. Groeger, Ireland
R. Happee, Netherlands
S. Jamson, UK
D.Kaber, USA
J. Krems, Germany
M. Lenne, Australia
F. Mars, France
D.McAvoy USA
A. Mills, UK
R. Risser, Austria
P. Salmon, Australia
S. Sharples, UK
G. Walker, Scotland
K. Young, Australia

Aviation

A. Alexander, USA
H. Davison Reynolds, USA
M. Draper, USA
M. Feary, USA
B. Gore, USA
B. Hooey, USA
D. Kaber, USA
M.Kupfer, USA
K. Latorella, USA
A. Majumdar, UK
N. McDonald, Ireland
J. Mercer, USA
E. Rantanen, USA
C. Samms, USA
A. Sebok, USA
S. Verma, USA
T. von Thaden, USA
K. Vu, USA
B. Willems, USA

Maritime

A. D. Alkan, Turkey
D.Andrews, UK
M. Barnett, UK
M. Grootjen, Netherlands
T. Koester, Denmark
S. N. MacKinnon, Canada
M. Musio Sale, Italy

Neville A. Stanton

University of Southampton

United Kingdom

Steven J. Landry

Purdue University

USA

Giuseppe Di Bucchianico and Andrea Vallicelli

University of Chieti-Pescara

Italy

Editors

Section 11

Road and Rail:

Driver State Detection and Simulated Driving - Drawbacks and Opportunities

Evolution of a New Adjustable Motorcycle Test Rig for Measuring Motorcyclist Fatigue during Prolonged Riding

Helmi Rashid, Muhammad Izzat Nor Ma'arof, Abdul Rahman Omar, Sukarnur Che Abdullah and Roseleena Jaafar

Motorcycle Engineering Test Lab (METAL)
Faculty of Mechanical Engineering, Universiti Teknologi MARA
40450 Shah Alam, Selangor Darul Ehsan, Malaysia

ABSTRACT

Studies concerning motorcycle have been an overwhelming area of research interest since the past few years. This trend of studies focuses on many issues relating to motorcycle including design, safety, accident preventions, fatalities statistics, ergonomics and many more. In performing these studies, researchers have used many methodologies, tools and also equipment in acquiring their data. In this study, the evolution of a new adjustable motorcycle test rig is introduced. Combination of computer aided design (CAD) and finite element analysis (FEA) software made it possible to design and simulates the test rig's conceptual design before being fabricated. The test rig was designed to provide full adjustability for researchers to replicate established motorcyclist riding postures based on the Riding Posture Classification (RIPOC) system. The test rig setup also provides elements of environmental effects to give a more real riding experience and better fidelity to motorcyclist respondents during experiments. By having such test rig that is coupled with electronic data logger and telemetry devices, studies concerning riding postures and fatigue can be done in a much proper method and controlled experimental setup with better safety for both the respondents and researchers. Adjustable attributes provided on the test rig makes it a unique design of its own and is being patented to protect the author's Intellectual Property Rights (IPR) on the test rig design.

Keywords: Motorcycle accidents, simulators, motorcycle test rig, motorcycle ergonomics, motorcyclist fatigue, prolonged riding

INTRODUCTION

The increase of motorcycle accidents every year (Organization, 2013) had made this a very alarming issue amongst motorcycle researchers to study more on how to prevent and reduce those accidents. Many factors has been previously identified to cause such motorcycle accidents including age and gender, the use of safety devices, weather and environment, road and highway designs, motorcycle and rider characteristics, influence of alcohol and drug abuse and also other road users fault (Keall & Newstead, 2012; Pinto, Cavallo, & Saint-Pierre, 2014; Schneider, Savolainen, Van Boxel, & Beverley, 2012). Apart from these factors, aspect of the motorcyclist or human factors contributing to such motorcycle accidents were also been studied previously (Ampofo-Boateng, Muzlia, Parnabas, Abdullah, & Sueb, 1999; Hurt & Dupont, 1977; Macdonald, 1985) and now becoming more significant to motorcycle researchers.

Not like other vehicle simulators that have been long established by many institutions and agencies, whether for academic, industrial or commercial purpose, the bibliography of motorcycle simulators is quite lacking (Nehaoua, Arioui, & Mammar, 2011). However, current trend has changed where a number of the studies have started to use motorcycle simulators in collecting their data as a mean of replicating real motorcycle riding experience for their respondents (Cossalter, Doria, & Lot, 2004; Crundall, Crundall, & Stedmon, 2012; Crundall, Stedmon, Saikayasit, & Crundall, 2013; Huth, Biral, Martin, & Lot, 2012; A. W. Stedmon, Brickell, Hancox, Noble, & Rice, 2012). Besides that, most of the current motorcycle simulators mainly offer one type of riding posture based on their design or type of motorcycle being used on the simulator (see Figure 1, Figure 2 and Figure 3).

Figure 1. DIMEG riding simulator (Huth et al., 2012)

Figure 2. MotorcycleSim simulator (A. W. Stedmon et al., 2012)

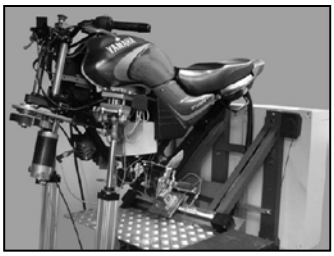


Figure 3. IFSTTAR motorcycle riding simulator (Benedetto et al., 2014)

These types of motorcycle simulators were mainly focused on studying the interactions between the motorcyclist and the displayed surrounding environment according to different setups. Moreover, the motorcyclist were also tested for their behaviors and response towards road hazards besides testing for new accident precaution measures in preventing motorcycle accidents.

Besides the aforementioned motorcycle related studies, aspect of ergonomics especially involving motorcyclist fatigue can also best suit the use of motorcycle simulators. In a previous study reported by the Department for Transport, London, (Horberry, Hutchins, & Tong, 2008), they concluded that there is a significant paucity of motorcyclist fatigue study especially studies that employed appropriate experimental control. The few studies found mainly relied on self-report data and more research to investigate correlation of motorcyclist fatigue and motorcycle accidents are essential (Perez-Fuster, Rodrigo, Ballestar, & Sanmartin, 2013). This was mainly caused by the inconsistent terminology and understanding of the motorcyclist fatigue itself. Despite this gap, efforts in investigating this ergonomics aspect of motorcyclist were pioneered by several researchers but with the use of a real motorcycle with different methodologies (Balasubramanian & Jagannath, 2014; Velagapudi, Balasubramanian, Adalarasu, Babu, & Mangaraju, 2010). The aims of the studies were identifying motorcyclist fatigue through several on road testing using a surface electromyography (sEMG) device and procedures. However, the motorcycle used in the studies limit the researchers to perform various riding postures to be tested on their respondents and exposes their respondents to real road hazards.

Hence, to investigate motorcyclist fatigue by using a motorcycle test rig that can be adjusted to suit various riding postures is essential in narrowing this research gap. In contrast to a motorcycle simulator cost of development, a motorcycle test rig provides lower cost of investment to develop. However, both means of replicating a real motorcycle riding experience faces the same fidelity and validity issues when using either a simulator or a test rig (Benedetto et al., 2014; Huang & Gau, 2003; Jamson & Jamson, 2010; A.W. Stedmon, M.S.Young, & Hasseldine, 2009). In compromising this issue, a new adjustable motorcycle test rig is introduced by researchers from the Motorcycle Ergonomics Test Lab, Faculty of Mechanical Engineering, Universiti Teknologi MARA Malaysia to provide better fidelity of motorcycle riding experience with a controlled environment effects for motorcycle ergonomics research purpose especially for prolonged motorcycle riding.

CONCEPTUAL DESIGN AND STRUCTURAL ANALYSIS

Motorcycle Test Rig Chassis Design

The primary design stage of this adjustable motorcycle test rig started with outlining several design concepts from current motorcycle models. With the basis of different motorcycle provide different riding postures, the main reference for designing this test rig chassis was based on the established Riding Posture Classification (RIPOC) system (Ma'arof & Ahmad, 2012). Four types of riding postures namely as RIPOC Type 1 (forward leaning), Type 2 (upright), Type 3 (seatback-leg-forward) and type 4 (double forward) were respectively considered for the test rig design concepts to be adjusted according to those riding postures.

Initial sketches of the chassis concept were brainstormed to generate ideas and finally being finalized using Pugh decision matrix selection method with specific concepts criteria (Kuppuraju, Ittimakin, & Mistree, 1985; Thakker, Jarvis, Buggy, & Sahed, 2009). Scores and ratings were determined and the highest total average score gained by one of the design concepts was selected as the preliminary design. The preliminary design was modeled using CATIA V5R20 computer aided design (CAD) software for better visualization (see Figure 4).

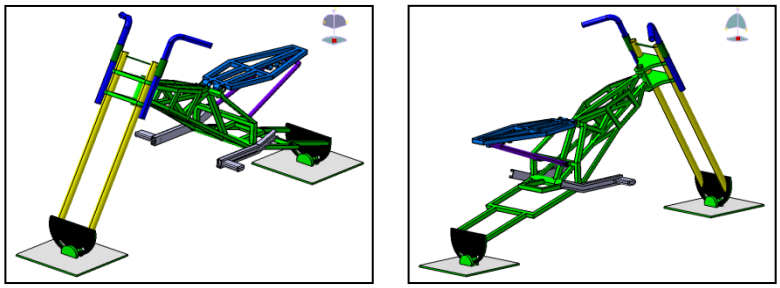


Figure 4. Initial CAD design concept of adjustable motorcycle test rig

Adjustability features were considered for the handle bar, seat and foot peg designs whereby the positioning of these main parts on a motorcycle generally influences different riding postures on different models of motorcycles (Ma'arof & Ahmad, 2012). Besides that, the chassis also advanced from the concept of human's rib cage and spine having it to be named as *RISC*^{TM} (Rib and Spine Chassis) that gives room for an innovative collapsible motorcycle tank to be included on it. The aforementioned characteristics of the test rig narrows down the fidelity issues of replicating various motorcycle riding postures onto only one test rig.

Structural Finite Element Analysis (FEA)

This adjustable motorcycle test rig chassis went through a series of revisions to best suit the functionality and fabrication constraints of the test rig. A detail design of the test rig chassis was modeled using the CATIA V5R20 CAD software to generate a finite element model for further structural FEA simulations. Concurrently, a free body diagram (FBD) of the test rig chassis was produced to calculate all reacting forces acting on it and to identify boundary conditions for limiting the degree of freedom (DOE) during the FEA simulations. The finite element model was assigned with mild steel for the material property and meshed using a linear tetrahedron mesh before being loaded with a 90kg weight for the FEA simulations using one of the Computer Aided Engineering (CAE) packages provided in the CATIA V5R20 CAD software (see Figure 5).

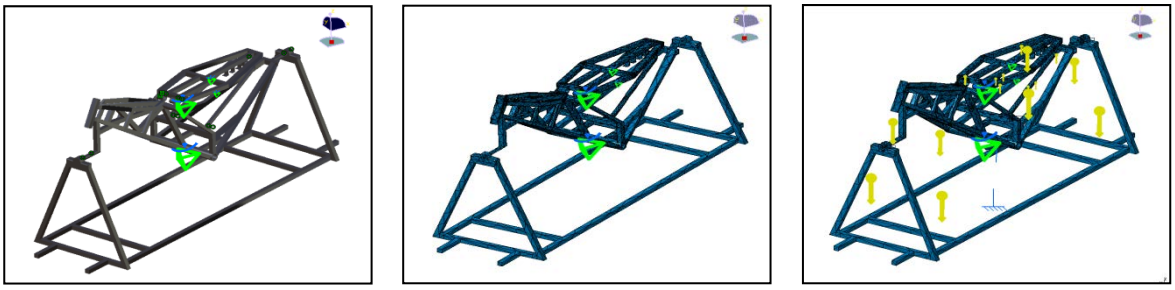


Figure 5. Finite element model for the FEA simulations

The structural FEA simulations were done to predict the behavior of the test rig chassis under certain loading of the motorcyclist. This enabled the research team to identify possible maximum stresses and deflections that may lead to structure failure (Mingzhou, Qiang, & Bing, 2002) during operation. By this, an optimum design of the test rig chassis can be developed with better factor of safety. Results from the FEA simulations shown a slight Von Mises stress on both end supporters of the test rig with minimal deflection and a 1.07 acceptable factor of safety (see Figure 6).

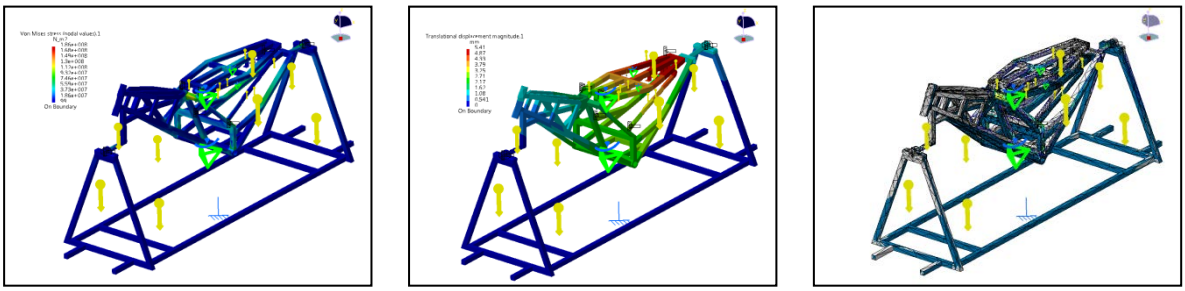


Figure 6. Maximum Von Mises stress, maximum deflection and deformation results from FEA simulations

FABRICATING THE PROTOTYPE

Acquiring of Material

The material used in the fabrication works to build the adjustable motorcycle test rig prototype was mild steel. It was the same properties of mild steel that was used to simulate the FEA in CATIA V5R20 CAD software (see Table 1). Other materials used range from different types of bolts and nuts, ready stock motorcycle seat, handle bar clip on and foot peg, and other few motorcycle accessories as an effort to provide better replication and fidelity.

Table 1: Mild steel material properties (Riley, Sturges, & Morris, 2006)

Structural Mechanical Property	Value
Young Modulus	2.1×10^{11} N/m^2
Poisson Ratio	0.28
Density	7800 kg/m^3
Thermal Expansion	1.4×10^{-5} Kdeg
Yield Strength	200×10^{6} N/m^2

Fabrication Works

The adjustable motorcycle test rig was fabricated in-house at the Machine Shop, Faculty of Mechanical Engineering, Universiti Teknologi MARA Malaysia to protect its IPR. Prior to the fabrication, series of FEA simulations and design revisions were done to confirm the strength and sturdiness of the test rig chassis. Mild steel was extensively used to build the first prototype of this test rig due to its strength and low cost consideration. Weight of the test rig was not a substantial factor as for the design was not meant to be mobile. Metal inert gas (MIG) welding was used to bond the structures together giving the prototype a rigid and robust chassis quality (see Figure 7).

Figure 7. Fabrication works of the test rig prototype

During the fabrication works, the test rig design also went through some design revisions due to fabrication constraints. The revised designs were analyzed again using FEA to observe any obvious changes to its maximum Von Mises stress and deflection on the structure. It was found that the revisions on the design did not affect much on the strength of the structure but only a minimal increase in total weight that was not substantial to the design criteria.

HUMAN-MACHINE-ENVIRONMENT-INTERFACE (HMEI) SETUP

Environmental Elements

Amongst the risk factors that contributes to motorcycle accidents is the surrounding weather and environment (Jones, Gurupackiam, & Walsh, 2013; Keall & Newstead, 2012; Schneider et al., 2012). Environmental effects including wind gust, heat, and noise are elements that motorcyclist directly experience when riding motorcycles. Hence, it is crucial to include these effects in the adjustable motorcycle test rig setup to provide a near-to-real riding experience for the respondents during the experiments. This effort will also narrow the fidelity issue by providing a better replication of on-road motorcycle riding experience into an indoor simulation. Apart from that, including the elements of environment into the test rig setup will complete the interfaces between human-machine interfaces (HMI) into another extent of human-machine-environment interfaces (HMEI) (Ma'arof, Ahmad, Abdullah, & Karim, 2012).

Environmental elements that were considered to be included into the test rig setup are elements of visual, wind gust, heat and noise. A projector and a screen are used as the visual aid to project real video of motorcycle riding activity that was recorded in prior to the setup. To acquire this, a JVC Adixxion action camera was used to capture real footage and sound on a real motorcycle. The projector is placed on an overhead projector stand and the screen is located in front of the test rig providing a clear visualization of what does a motorcyclist sees in front of him/her when riding a real motorcycle.

To replicate the wind gust and heat to the motorcyclist on the test rig, industrial fan blower and heater are installed into the test rig setup. Wind speed and heat intensity are based on motorcycle speed and also the typical surrounding climate and weather of Malaysia. The overall layout concept was drafted using Google SkecthUp 2013 to have an overview of how should the placements of the test rig setup looks like (see Figure 8).

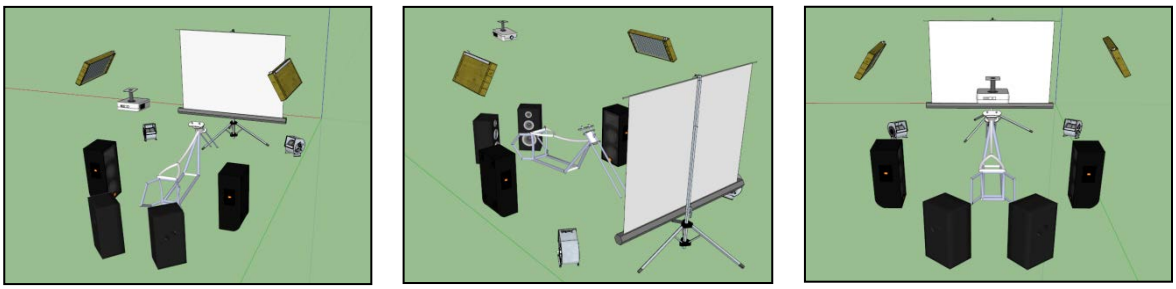


Figure 8. Layout concept of the test rig setup

Integration of Adjustable Motorcycle Test Rig and Surrounding Elements

To create another near-to-real riding experience on the test rig and to give the respondents the feel of controlling the test rig like handling a real motorcycle, the throttle on the test rig acts like a real motorcycle throttle. Simplified electronic controller made it possible to use the throttle as the fan blower and video speed controller. By having this, when the respondent's starts to rotate the throttle (accelerate), the speed of the fan blower will start to build up supplying sufficient wind gust to the respondents and the video will also start to speed up. The throttle also controls the sound level of the surrounding speakers that includes the sounds of other vehicles, motorcycle engine, exhaust and other sounds captured in the footage. These conditions replicate how wind gust starts to blow back the respondent as if when accelerating on a real motorcycle and the video shows that the motorcycle is in a moving condition, while the surrounding sounds add up the near-to-real riding experience on the test rig setup. Heat from the heater is controlled and monitored during experiments to give the right heat intensity as if the respondents are really riding under the sun.

MEASURING MOTORCYCLIST FATIGUE

It is known that body posture, muscle fatigue and musculoskeletal disorder (MSD) are three main elements that are related to one another especially in a sitting condition (Abdul Hadi Abdol Rahim & Ibrahim Othman, 2010; Balasubramanian & Jagannath, 2014; Balasubramanian, Jagannath, & Adalarasu, 2014; Karmegam et al., 2009; Mörl & Bradl, 2013; O'Sullivan et al., 2010; O'Sullivan, McCarthy, White, O'Sullivan, & Dankaerts, 2012; Srinivasan & Balasubramanian, 2007; A. W. Stedmon, 2007; Velagapudi et al., 2010). However, studies related to motorcyclist fatigue that employs proper method and experimental setup are lacking (Horberry et al., 2008). Hence, introducing this adjustable motorcycle test rig is significant to help researchers to understand more on this motorcyclist fatigue matter.

As for muscle fatigue is best being measured using a surface electromyography (sEMG) device, motorcycle researchers has started to use this device in acquiring muscle activity patterns on motorcyclist (Balasubramanian & Jagannath, 2014; Velagapudi et al., 2010). In those studies, different procedures were used to measure their respondent's muscle activity using the sEMG device, but the motorcycle that they used apparently limited to one type of motorcycle model. This limit the exploration of muscle activity patterns using various riding postures that constitutes from different types of motorcycle models. However, it is still possible if they have the access to various types of motorcycle models to be used and tested. This will require a big group of motorcyclist to lend their motorcycle and will surely be a big constraint to the researchers to gather such a big group together for this purpose (A.W. Stedmon et al., 2009). Henceforth, using this test rig makes it possible to replicate various riding postures based on various motorcycle models. This gives motorcycle researchers the opportunity to study muscle activity patterns using various riding postures that relates to muscle fatigue.

Pilot tests have been done using this test rig setup coupled with a sEMG device where results of the muscle activity show different patterns when assuming different riding postures (see Figure 9). Further tests were done to validate the muscle activity patterns consistency when sitting on the test rig with certain riding posture compared to sitting on a real motorcycle with the same riding posture. Prior to these sEMG data collection and experiments, issues concerning ethics in the study have been approved by the UiTM Research Ethics Committee.

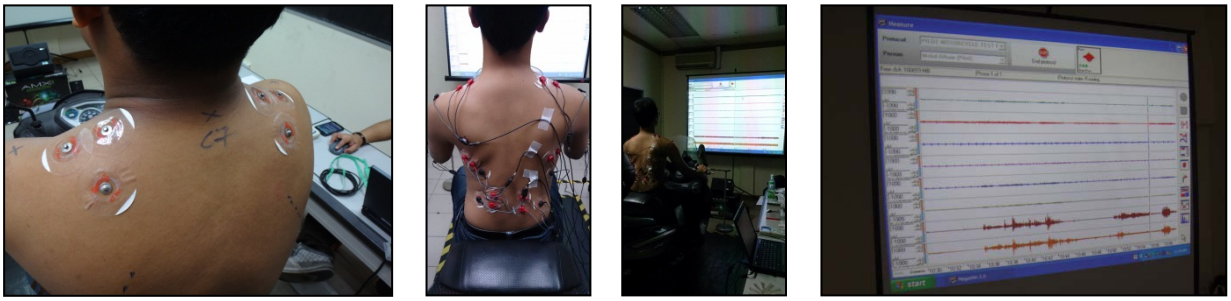


Figure 9. Pilot tests in measuring muscle activity patterns

CONCLUSIONS

Risk factors contributing to such motorcycle accidents have been identified previously listing from many factors including the motorcyclist themselves (Keall & Newstead, 2012; Pinto et al., 2014; Schneider et al., 2012). In studying the aspect of motorcyclist that includes their behavior, response, physiology and interactions during riding a motorcycle, human factors or ergonomics studies relating to these aspects are essential. It is now becoming an interesting research area to explore in an effort to help reducing motorcycle accidents that are keeping on increasing every year.

Henceforth, researchers at the Motorcycle Ergonomics Test Lab, Faculty of Mechanical Engineering, Universiti Teknologi MARA Malaysia had made another breakthrough in motorcycle ergonomics studies by introducing a new adjustable motorcycle test rig for studying motorcyclist posture and fatigue matters. As part of the experiment equipment to be used in another study framework that focuses on measuring motorcyclist muscle fatigue in prolonged riding, this test rig provides another alternative of an experimental setup that is much proper within a

controlled environment. The unique attributes that the test rig comprises enabled it to be patented to protect the authors Intellectual Property Rights (IPR) upon the test rig's design.

The idea of developing a means of platform to measure motorcyclist fatigue that causes motorcycle accidents started to evolve from this study. Knowing that cost of development to build a motorcycle simulator is highly expensive for such studies, the authors opt to develop only a motorcycle test rig that is much lower in cost. Benchmarking and referring to current motorcycle simulators that are available, apparently most of them only provide one possible riding posture that is based on the type of motorcycle model that they acquire in their simulators. This opens another opportunity for the authors to design and develop a motorcycle test rig that can be adjusted to suit various riding postures, mainly the riding postures established by the RIPOC system (Ma'arof & Ahmad, 2012). Although the test rig can be adjusted, maximum and minimum adjustments to the handle bar, seat and foot pegs are still based on real motorcycle dimensions to avoid any odd riding postures that are impossible to be performed on existing motorcycle type of models.

Fidelity of the test rig in replicating a near-to-real motorcycle riding experience is sustained through the environmental element effects included in the test rig setup. Effects of wind gust, heat and noise during real motorcycle riding are created with the aid of various equipment and control system. Besides that, the test rig chassis itself is being specifically designed to look like a motorcycle that uses several motorcycle accessories on it to give a better feel of motorcycle riding experience to the respondents. By having all these attributes, a better fidelity of a motorcycle test rig was successfully developed for the use of various potential motorcycle ergonomics studies.

However, ongoing studies are still being done by the researchers at the Motorcycle Ergonomics Test Lab to enhance the capability of this test rig to another extent. Elements of vibrations will be included in the test rig setup as another package of creating another near-to-real riding experience on the test rig. This decision will need a big investment of cost to implement. Sophisticated control systems coupled with high tech hardware involving hydraulics and pneumatics systems will contribute most to the high investment cost of implementation. However, it will be worth it to invest in such a big cost as for more exploration on motorcyclist interactions towards their surrounding and body response can be performed. Besides that, the test rig interface is also being considered to be upgraded to use a programmable and fully interactive system for instance the STISIMS Drive system and interface. By using such system, the respondents riding the test rig have the opportunity to have better control and interaction with the displayed interface. This system has been proven to be very supportive and good in simulating road conditions with the capability of setting up own hazards preferences for any condition (Crundall et al., 2013; A. W. Stedmon et al., 2012). The throttle control, brakes and clutch can also be synchronized with the interface making it another big advantage to have such system integrated into the test rig. Thus, it does not only limit the use of the test rig for studies concerning motorcyclist postures and fatigue, but also opens other opportunities for further motorcycle ergonomics studies in the mere future.

However, to replicate a rainy day riding environment using the test rig setup is impossible and very risky as for the electronics devices used on the test rig setup is not meant to be used outside during raining condition. Hence, the environment being replicated is only limited to non-raining condition and an assumption is made that riding motorcycle in a rainy day increase the risk of the motorcyclist to get involved in a road accident. Therefore, it is clear that this risk is high enough to forbid motorcyclist to ride during raining condition and allows researchers not to consider raining condition in their studies. Nevertheless, how high does raining condition risk the lives of motorcyclist on the road that contributes to motorcycle accidents can be further studied and explored.

ACKNOWLEDGMENT

The authors would like to acknowledge the Ministry of Education (MOE) Malaysia for providing the research fund for this research study through the Fundamental Research Grant Scheme (FRGS/1/2014/TK01/UITM/02/3). Special thanks to the research team at the Motorcycle Ergonomics Test Lab, Faculty of Mechanical Engineering, Universiti Teknologi MARA Malaysia especially to Muhammad Hafizudin Kamal, Muhammad Firdaus Azhar, Wan Muhammad Syahmi Wan Fauzi, Mohd Afham Azmi and also Muhammad Adzri Adzman. Special thanks also to Rohani Haron, expert physiotherapist from Faculty of Health Sciences, Universiti Teknologi MARA Malaysia, the technicians at the Machine Shop, Faculty of Mechanical Engineering, Universiti Teknologi MARA Malaysia and also to the Human Factors and Ergonomics Society of Malaysia (HFEM) for their continuous support upon this study.

REFERENCES

Abdul Hadi Abdol Rahim, Abdul Rahman Omar , Isa Halim, Alias Mohd Saman,, & Ibrahim Othman, Mas Alina, Suhadiyana Hanapi. (2010, 5-7 December 2010). *Analysis of Muscle Fatigue associated with Prolonged Standing Tasks in Manufacturing Industry.* Paper presented at the 2010 International Conference on Science and Social Research (CSSR 2010).

Ampofo-Boateng, Kwame, Muzlia, Syarifah, Parnabas, Vincent, Abdullah, Nabilah, & Sueb, Rosilawati. (1999). Human Factors and Motorcycle Accidents in Malaysia: Universiti Teknologi MARA.

Balasubramanian, V., & Jagannath, M. (2014). Detecting motorcycle rider local physical fatigue and discomfort using surface electromyography and seat interface pressure. *Transportation Research Part F: Traffic Psychology and Behaviour, 22*, 150-158.

Balasubramanian, V., Jagannath, M., & Adalarasu, K. (2014). Muscle fatigue based evaluation of bicycle design. *Applied Ergonomics, 45*(2, Part B), 339-345.

Benedetto, Simone, Lobjois, Régis, Faure, Vérane, Dang, Nguyen-Thong, Pedrotti, Marco, & Caro, Stéphane. (2014). A comparison of immersive and interactive motorcycle simulator configurations. *Transportation Research Part F: Traffic Psychology and Behaviour, 23*, 88-100.

Cossalter, Vittore, Doria, Alberto, & Lot, Roberto. (2004). *Development and validation of a motorcycle riding Simulator.* Paper presented at the World Automotive Congress, Barcelona, Spain.

Crundall, E., Crundall, D., & Stedmon, A. W. (2012). Negotiating left-hand and right-hand bends: a motorcycle simulator study to investigate experiential and behaviour differences across rider groups. *PLoS One, 7*(1), e29978.

Crundall, E., Stedmon, A. W., Saikayasit, R., & Crundall, D. (2013). A simulator study investigating how motorcyclists approach side-road hazards. *Accident Analysis and Prevention, 51*, 42-50.

Horberry, T., Hutchins, R., & Tong, R. (2008). Motorcycle Rider Fatigue: A Review. London: Department for Transport.

Huang, Jiung-Yao, & Gau, Chung-Yun. (2003). Modelling and designing a low-cost high-fidelity mobile crane simulator. *International Journal of Human-Computer Studies, 58*(2), 151-176.

Hurt, H.H., & Dupont, Cj. (1977). Human Factors in Motorcycle Accidents. *SAE Technical Papers*.

Huth, V., Biral, F., Martin, O., & Lot, R. (2012). Comparison of two warning concepts of an intelligent Curve Warning system for motorcyclists in a simulator study. *Accident Analysis and Prevention, 44*(1), 118-125.

Jamson, Samantha L., & Jamson, A. Hamish. (2010). The validity of a low-cost simulator for the assessment of the effects of in-vehicle information systems. *Safety Science, 48*(10), 1477-1483.

Jones, S., Gurupackiam, S., & Walsh, J. (2013). Factors influencing the severity of crashes caused by motorcyclists: Analysis of data from Alabama. *Journal of Transportation Engineering, 139*(9), 949-956.

Karmegam, K., Ismail, M.Y., Sapuan, S.M., Ismail, N., Bahri, M.T. Shamsul, Shuib, S., & Seetha, P. (2009). A Study on Motorcyclist's Riding Discomfort In Malaysia. *Engineering e-Transaction, 4*(1), 39-46.

Keall, M. D., & Newstead, S. (2012). Analysis of factors that increase motorcycle rider risk compared to car driver risk. *Accident Analysis and Prevention, 49*, 23-29.

Kuppuraju, Nagesh, Ittimakin, Prawit, & Mistree, Farrokh. (1985). Design through selection: a method that works. *Design Studies, 6*(2), 91-106.

Ma'arof, M Izzat Nor, Ahmad, Ismail N, Abdullah, Nik R, & Karim, Shamsury A. (2012). *Motorcycling Riding Issues: Understanding the Phenomenon and Development of Ergonomics Intervention in Improving Perceived Comfort for Prolonged Riding.* Paper presented at the International Conference on Design and Concurrent Engineering (iDECON 2012), Melaka.

Ma'arof, M Izzat Nor, & Ahmad, Ismail N. (2012). *Proposed Standard Method for Motorcycle Nomenclature System.* Paper presented at the 2012 Southeast Asian Network of Ergonomics Societies Conference (SEANES), Langkawi.

Macdonald, Wendy A. (1985). Human factors and road crashes: A review of their relationship (D. o. M. Engineering, Trans.): University of Melbourne.

Mingzhou, Su, Qiang, Gu, & Bing, Guo. (2002). Finite element analysis of steel members under cyclic loading. *Finite Elements in Analysis and Design, 39*(1), 43-54.

Mörl, Falk, & Bradl, Ingo. (2013). Lumbar posture and muscular activity while sitting during office work. *Journal of Electromyography and Kinesiology, 23*(2), 362-368.

Nehaoua, L., Arioui, H., & Mammar, S. (2011, 20-23 June 2011). *Review on single track vehicle and motorcycle simulators.* Paper presented at the Control & Automation (MED), 2011 19th Mediterranean Conference on.

O'Sullivan, Kieran, O'Dea, Patrick, Dankaerts, Wim, O'Sullivan, Peter, Clifford, Amanda, & O'Sullivan, Leonard. (2010). Neutral lumbar spine sitting posture in pain-free subjects. *Manual Therapy, 15*(6), 557-561.

O'Sullivan, Kieran, McCarthy, Raymond, White, Alison, O'Sullivan, Leonard, & Dankaerts, Wim. (2012). Can we reduce the effort of maintaining a neutral sitting posture? A pilot study. *Manual Therapy, 17*(6), 566-571.

Organization, World Health. (2013). Global Status Report on Road Safety 2013 - Supporting A Decade of Action: World Health Organization.

Perez-Fuster, P., Rodrigo, M. F., Ballestar, M. L., & Sanmartin, J. (2013). Modeling offenses among motorcyclists involved in crashes in Spain. *Accident Analysis and Prevention, 56*, 95-102.

Pinto, Maria, Cavallo, Viola, & Saint-Pierre, Guillaume. (2014). Influence of front light configuration on the visual conspicuity of motorcycles. *Accident Analysis & Prevention, 62*, 230-237.

Riley, William F., Sturges, Leroy D., & Morris, Don H. (2006). *Mechanics of Materials* (6th ed.): Wiley.

Schneider, W. H. th, Savolainen, P. T., Van Boxel, D., & Beverley, R. (2012). Examination of factors determining fault in two-vehicle motorcycle crashes. *Accident Analysis and Prevention, 45*, 669-676.

Srinivasan, J., & Balasubramanian, V. (2007). Low back pain and muscle fatigue due to road cycling-An sEMG study. *Journal of Bodywork and Movement Therapies, 11*(3), 260-266.

Stedmon, A. W. (2007). *Motorcycling is a pain in the neck, numb bum, aching wrists - Or all these things! A rider comfort survey*.

Stedmon, A. W., Brickell, E., Hancox, M., Noble, J., & Rice, D. (2012). MotorcycleSim: A user-centred approach in developing a simulator for motorcycle ergonomics and rider human factors research. *Advances in Transportation Studies*(27), 31-48.

Stedmon, A.W., M.S.Young, & Hasseldine, B. (2009). Keeping It Real or Faking It: The Trials and Tribulations of Real Road Studies and Simulators in Transport Research *Contemporary Ergonomics 2009*: Taylor & Francis.

Thakker, A., Jarvis, J., Buggy, M., & Sahed, A. (2009). 3DCAD conceptual design of the next-generation impulse turbine using the Pugh decision-matrix. *Materials & Design, 30*(7), 2676-2684.

Velagapudi, S. P., Balasubramanian, V., Adalarasu, K., Babu, R., & Mangaraju, V. (2010). Muscle fatigue due to motorcycle riding. *SAE Technical Papers*.

Motorcycling: Awkward Posture Is the Best Posture!

Muhammad Izzat Nor Ma'arof[1], Helmi Rashid[1], Abdul Rahman Omar[1], Sukarnur Che Abdullah[1], Ismail Nasiruddin Ahmad[1], Roseleena Jaafar[1] and Shamsury Abd Karim[2]

[1]Motorcycle Engineering Test Lab (METAL),
Faculty of Mechanical Engineering, Universiti Teknologi MARA,
40450 Shah Alam, Selangor Darul Ehsan, Malaysia
[2]MASMED, Universiti Teknologi MARA,
40450 Shah Alam, Selangor Darul Ehsan, Malaysia

ABSTRACT

A study has documented that there are only four possible working postures for motorcycling, yet, all of these postures are argued by another study to be generally "unfitting" for human and hence a negative option as a mode of daily transportation. Even so, contradictory to this statement, motorcycle as a mode of transportation is becoming even more popular in Asia. The objective of this study is to show that the aforementioned "unfitting" value of the motorcycling working posture is a must in motorcycling. The significance of this study is to share a new perspective in ergonomics. Three studies were conducted: First, a descriptive within-subject design research of real world motorcycling experience using participant observation method. Second, an ethnographic controlled observation to support the findings. Third, an intrinsic case study on motorcyclists who experienced comfort related issues. Results showed that motorcyclists are relatively satisfied with the riding posture since it makes it possible for them to perform the motorcycling work tasks efficiently, thus, ensuring their safety. A too comfortable riding posture actually results in the deficit of the motorcyclist's riding performance. In short, motorcycling requires a certain level of "discomfort" in order to ensure safely.

Keywords: Motorcycle ergonomics, accident, safety, awkward posture, fatigue

INTRODUCTION

Ergonomics has been defined by numerous authors from various perspectives since its inception. Even so, the fundamental aim of ergonomics has always been idealistic – to ensure the human operator is working in a safe physical and psychological manner by 'fitting' the job to the human (Kromer et al., 2000b). This goal is also achieved by encouraging the human operator to work as close as possible to their neutral joint posture (Ergonomics Guideline for Manual Handling, 2010).

To achieve this fundamental aim, Kromer et al., (2000a) noted that there are three goals of ergonomics which need to be accomplished according to their hierarchy. First is to ensure the human operator is safe – both in term of mortality and health, by providing "tolerable" working conditions. The second goal to be fulfilled is to provide "acceptable" working conditions for the human operator to perform the work. Finally, after the fulfillment of the first two goals, generating "optimal" working conditions for the human operator is warranted.

In short, the outcome of ergonomics knowledge aims for the human operator to work within a safe environment while performing the work at their "natural" state – "natural" from the perspective of human biomechanics, physical and psychological characteristics and traits (Health and Safety Executive, 2013; Bosch Rexroth AG, 2012; Travail Sècuritaire, 2011; Berry et al., 2009; Openshaw and Taylor, 2006; Handbook of Human Factors and Ergonomics

Methods, 2005; Kroemer et al., 2000). The hierarchical goals noted by Kromer et al., (2000) will be the main guideline for this study. This study would discuss the result on an exploratory research which could be contradictory to the fundamentals or laws of ergonomics. This study aims to prove that though the motorcycle riding posture is awkward, it is still the best working posture for motorcycling.

GLOBAL STATISTICS OF MOTORCYCLE ROAD ACCIDENTS

Various literatures such as Ma'arof and Ahmad (2012a), Ministry of Transport Malaysia (2011), Chen and Chen (2011), Velagapudi (2010), Karmegam et al., (2009 and 2012) have all noted that motorcycle is a universally popular human mode of transportation. In various Asian major cities, motorcycles strive beyond other fuel-powered vehicles due to being efficient (example: maneuverability through congested areas, amount of parking space required), yet, economical (example: fuel expenditure, maintenance cost) (Chen and Chen, 2011). The demand of this vehicle has shown escalating growth (Shell.com, 2013) and predicted to annually increase in the near future by this study – although experiencing the ever soaring global fuel price (Young, 2013) and the Honda Annual Report (2012) which reported the increment of 9.7% in global motorcycles sales of the brand from the previous fiscal year. As of the year 2011, 9,985,308 motorcycles were officially registered in Malaysia alone (Ministry of Transport Malaysia, 2011).

Nevertheless, corresponding to the global increasing demand for motorcycles, the number of road accidents involving this vehicle also increased nationwide. Figure 1 summarizes the percentage of road users' deaths involving rider of motorized 2- or 3-wheeled vehicle in the South East Asian region. Global Status Report on Road Safety (2013) reported that Thailand and Laos recorded the highest percentage at 74%, followed by Cambodia (67%), Malaysia (59%), Singapore (46%), Indonesia (36%), China (35%), India (32%) and Myanmar (23%). (Note: The data for Brunei, East Timor, Philippines and Vietnam are not available in the report). Various literatures had noted poor behavior, the non-usage of helmets, alcohol abuse, age factor, lack of motorcycling education, night-to-day riding and riding without license as among the contributing factors towards motorcycle road accidents (Keall and Newstead, 2012; Manan and Várhelyi, 2012; and Jou et al., 2012). Besides that, fatigue experienced by motorcyclist during riding was also found to be a contributing factor towards motorcycle fatalities (Horberry et al., 2008; Velagapudi et al., 2010; Marina et al., 2011 and Williamson et al., 2011).

This study agrees with the findings by these literatures. Nevertheless, this study would like to share the following quote: "…horsepower wasn't the cause before, and was unlikely to be the cause now, of crashing motorcycles. Rather it is rider errors in applying core technical skills that cause them to go down. That was my story, and now I have proof, so I'm sticking to it" (Code, 2010). (Note: Keith Code is a former motorcycle racer, distinguished motoring writer and the founding father of the California Superbike School). This quote clearly highlights the strong connection between road accidents and the motorcyclist – the human factor.

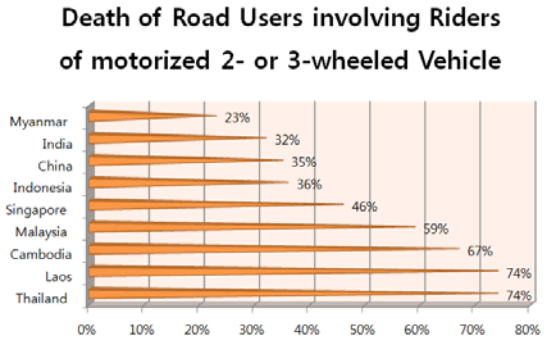


Figure 1. Percentage of Road Users' Deaths involving Riders of Motorized 2- or 3-Wheeled Vehicle in The South East Asian Region (source: Adapted from Global Status Report on Road Safety, 2013)

The "core technical skills" mentioned by Code (2010) is closely related to the "work task of motorcyclist" noted by Ma'arof and Ahmad (2012a). During real world motorcycling, the motorcyclist is required to use the whole body and each body part has a specific and unique work task. For instance, only the right hand would control the throttle while the left hand would manage the clutch lever (for manual transmission motorcycle). Thus, the questions here are: How could a motorcyclist be capable of "applying the core technical skills" without error? How could a motorcyclist maintain from making any error? How could the motorcyclist make any necessary corrective actions if

error does occur? There are several answers for these questions; nevertheless, all of the answers could be traced back to the very root of motorcycling – the riding posture. (Note: the term "error" here is referring to "error based on skills" as noted by Cañas et al., n.d)

The riding posture could also be treated as a sitting posture, only that it is performed on a motorcycle. Ma'arof and Ahmad (2012b) has outlined the "riding posture establishment conditions" in describing how a riding posture is assumed. The following section would give more in depth discussion on motorcycling riding posture.

MOTORCYCLE RIDING POSTURE VS IDEAL SITTING POSTURE

From the "riding posture establishment conditions", Ma'arof and Ahmad (2012b) has classified and simultaneously identified all possible working postures for motorcycling via the "Riding Posture Classification System" or the "RIPOC" System. The riding postures are: the "RIPOC Type 1: Forward lean riding posture", "RIPOC Type 2: Upright riding posture", "RIPOC Type 3: Seatback-leg-forward riding posture" and "RIPOC Type 4: Double forward riding posture" (Ma'arof and Ahmad, 2012b). In this study, these riding postures would be acknowledged as one, that is, from a general perspective of a working posture for motorcycling without specifying on a specific type of riding posture – unless noted. Now, how does the motorcycling riding posture stand before the definition of an "ideal" or "neutral working posture"?

There are various definitions in describing an "ideal" or "neutral working posture" and in this case, the focus is for an optimal "sitting" posture. The debates on determining optimal sitting posture still continue as noted by O'Sullivan et al., (2012 and 2010) and Openshaw and Taylor (2006). Several literature noted that an optimal sitting posture is when it is comfortable and this is achieved when the following conditions are satisfied as shown in Table 1.

Table 1: Comfortable sitting posture establishment conditions (O'Sullivan et al., (2012) and 2010; Berry, 2009; Openshaw and Taylor, 2006; Dul and Weerdmesster, 2003; Cal/OSHA, 1999)

Body Region/ Other Notes	Conditions
1. Spine	i. Maintaining the natural S- or standing shape of the spine similar to the shape while standing ii. Slight lumbar lordosis and thorax is relaxed iii. Sufficient lumbar support is provided iv. No slouching or hunching over
2. Muscles	i. No excessive muscle tone is required
3. Joints	i. Least possible extent from its natural positioning
4. Body regions exposed to surface contact and pressure distribution (e.g. buttocks, arms)	i. No significant stress is made on these body regions
5. Feet	i. Placed on a support (e.g. floor surface)
6. Other notes	i. Sitting posture could be varied and not being assumed for prolonged session

From the outline given for comfortable sitting posture, could the noted "riding postures" (Ma'arof and Ahmad, 2012b) be categorized as ideal working postures? Ma'arof et al., (2012) has concluded "the riding posture (sitting position) demonstrated by a motorcyclist on any given motorcycle is both an awkward and poor posture". The study argued that the riding postures are "constrained", "static", "cramped" and "poor". Henceforth, if all the motorcycling riding postures are deemed to be awkward, uncomfortable and physically and physiologically compromising, is it still relevant for motorcycling? Could it be replaced? Could we mimic the sitting characteristics of driving (four wheeled vehicle) to motorcycling? So, these are the research questions that would be answered by this study.

RESEARCH METHODOLOGY

To answer the research questions, the following qualitative research methodologies were formulated.

Field Research

The field research was divided into the following research method:

i. Non-structured observation
   A descriptive within subject research design (participant as observer type of observation) on motorcycling was performed by the authors. The data was recorded via data collection sheet.
ii. Ethnographic controlled observation
   Minor ethnographic controlled observation was also conducted to support the findings. The authors spent time close to two years with motorcyclists from various background and motorcycling experiences.

Intrinsic Case Study

Case studies were also conducted to further comprehend the effects of comfort during motorcycling. Semi-structured interview was selected as the data collection method. Interview sessions were conducted with three motorcyclists who were experiencing issues regarding comfort.

RESULTS

Field Research

The following results were obtained via the field research:

i. Comfort in motorcycling is very subjective and highly dependent on the motorcyclist. Factors such as riding styles, for instances, leisure riding or speed riding would commonly determine the motorcyclist's personal definition of "comfort" with respect to motorcycling. Both extremes of interests (high regard and disregard) exist in term of describing comfort. Senior motorcyclists usually prefer high emphasis on comfort.
ii. Significant difference in the motorcyclist capability in controlling and maneuvering the motorcycle was apparent upon assuming different riding postures.
iii. Duration of time in assuming, controlling and holding the riding posture were critical in determining the types of physical and physiological discomfort and fatigue experienced by the motorcyclist.
iv. Trends or patterns of physical and physiological discomforts differ with different riding postures being assumed, maintained and controlled.
v. Motorcyclists are relatively satisfied with the riding posture of their respective motorcycles.

Intrinsic Case Study

Case Study 1

A motorcyclist was hospitalized for 3 days following an accident due to falling asleep while on a solo riding journey from Skudai, Johor Darul Takzim to Shah Alam, Selangor Darul Ehsan (total of 327 km). The accident took place in Batu Pahat, Johor Darul Takzim (108 km from Skudai) at about 10 o'clock in the morning after travelling for approximately 1 hour. The motorcyclist clearly noted that he had prepared for the journey. He had sufficient sleep and taken light meal prior to the journey. The motorcyclist was riding a Suzuki X7 250 motorcycle and cruising at velocities within the range of 80 -100 km/h. The subject was in his early twenties and healthy. The following details were given by the subject in describing the accident:

i. He was too relaxed and comfortable. He was under very minimal time constraint since he was scheduled to reach Shah Alam at 2pm in the afternoon.
ii. The comfortable riding condition was further elevated upon a mist of rain that occurred a few minutes prior to the accident. The subject seriously emphasized that the accident occurred was not due to the rain.
iii. The comfortable and soothing motorcycling condition results in him getting drowsy, hence, slowly drifting to the side of the road and crashed. The subject only realized that he was involved in an accident when he woke-up on a hospital bed.
iv. The subject survived the accident with almost no external and internal injuries except experiencing minor amnesia on the event. The motorcycle was also hardly damaged.

At the end of the interview session, the subject further added that he agreed that certain degree of discomfort is a necessity for safe motorcycling. It is interesting to note that an awkward riding posture such as the "RIPOC Type 1:

forward lean riding posture" (Ma'arof and Ahmad, 2012b) was noted by the subject to be better suited for motorcycling, although discomforting in prolonged session.

Case Study 2

A motorcyclist almost involved in an accident upon experiencing micro sleeps while on a solo riding journey from Penang to Shah Alam, Selangor Darul Ehsan (total distance of 372 km). From experiencing the micro sleeps, the motorcyclist stopped and took breaks (twice) with approximately 15 minutes per session. The motorcyclist noted that he had prepared for the journey by having sufficient sleep and meal prior to the journey. Nevertheless, the subject also highlighted that he was feeling rather excited the night before the journey. The subject was in his early twenties, healthy and riding a Honda EX5 and cruising below 100 km/h. The following details were given by the subject in describing the event:

i. The micro sleeps were the result from riding fatigue. The subject emphasized that muscular and mental fatigue were the main types of fatigue experienced.
ii. The motorcyclist varies his riding posture throughout the journey. He chose to assume a pseudo-"RIPOC Type 1: forward lean riding posture" (Ma'arof and Ahmad, 2012b) to improve his riding experience.

At the end of the interview session, the subject strongly emphasized that assuming an awkward riding posture actually leads to comfort in motorcycling and agrees with the notion that certain degree of discomfort is required for overall safety in motorcycling.

Case Study 3

A motorcyclist frequently experienced micro sleeps during his routine daily commute to work that requires him to travel from Kajang to Shah Alam covering a distance of approximately 50 km on a one way route. The micro sleeps mostly occurs during the morning on the way to work and sometimes during the way back. The motorcyclist was in his early twenties, healthy and rode a Yamaha 110ss. At the peak of the micro sleeps the motorcyclist was almost involved in a motorcycle crash. This nearly happened when the motorcyclist fell into the micro sleeps for a few seconds without noticing that the motorcycle had steered to left side of the road. Upon regaining consciousness from the micro sleeps, the motorcyclist found himself on the other side of the road and fortunately missed the middle road barrier. The following details were given by the subject in describing the event:

i. Muscle fatigue was the main cause of the micro sleeps as claimed by the motorcyclist due to maintaining a static riding posture throughout the route to work.
ii. "RIPOC Type 2: Upright riding posture" (Ma'arof and Ahmad, 2012b) was maintained by the motorcyclist in a 50 km non-stop commute to work.
iii. Wind gust generated from the speed of the motorcycle approximately 80-90 km/h pushed the motorcyclist's body back and requires the motorcyclist to grip the throttles much harder and exhausted the body that leads to muscle fatigue.

The motorcyclist stresses that maintaining a certain riding posture on a motorcycle and constraining the body to the motorcycle for better control may lead to fatigue. When the body is in a fatigue and exhausted condition, micro sleeps and lost of control at the bends took over. Subsequently, this resulted in inappropriate decision making.

DISCUSSION

Motorcycling is a very unique vehicle operating experience. It resembles bicycling especially for the small displacement model (less than 250cc). Nevertheless, motorcycling in general requires a higher degree of motor-control than bicycling. However, it is significantly different from operating a four-wheeled vehicle. The most prominent criteria that separates driving a four-wheeled automobile and riding a motorcycle is the "Uno Body Motion" or "UBM" (Ma'arof et al., 2012) – that is the "condition where the whole body of the motorcyclist has to act as one, that is, synchronized" (Ma'arof et al., 2012) to effectively and safely steer the motorcycle during motion. The term "UBM" (Ma'arof et al., 2012) summarizes the "core technical skills" and "work task of motorcyclist" mentioned by Code (2010) and Ma'arof and Ahmad (2012a).

Failure to perform the "UBM" (Ma'arof et al., 2012) at any time during real world motorcycling could results in "errors based on skills" (Cañas et al., n.d). Hence, this could potentially lead to road accident. From the results given by the field research and reviewing the motorcycling bible, Code (1993), the one thing that could assure "UBM" (Ma'arof et al., 2012) could be performed to the highest efficiency is the riding posture. For the hands-on real world riding experience assessment, the author had the opportunity to vary the riding postures while riding on the same motorcycle. The experimental riding was performed for the duration of 10 hours (with several rests taken) through

variable types of roads. The results showed significant differences in term of the effort required to be exerted by the motorcyclist in maneuvering and controlling the motorcycle. This is because of the following reasons:

i. In general, the riding posture provides the motorcyclist with the ease (muscular effort) of covering various range of motions of certain body joints and limbs, for instances, the arms and legs.
ii. Different riding posture provides variable degree of the ease for the motorcyclist in covering the aforementioned range of motions.
iii. The riding posture also enables the motorcyclist to utilize the leverage provided by the design of the motorcycle in performing the "core technical skills" (Code, 2010).

In short, the riding posture in general makes it easier and possible for the motorcyclist to efficiently perform the "UBM" (Ma'arof et al., 2012). Nevertheless, the results here are viewed and discussed only from the perspective of the human operator. The same issue on maneuvering and controlling the motorcycle, viewed from the perspective of the motorcycle would be discussed in another study. The assessment showed that the higher the degree of discomfort in assuming, controlling and holding the riding posture, the better it is for the motorcyclist to perform the motorcycling work task. "Discomfort" in this context is very subjective, yet, closely associated with the study by Ma'arof et al., (2012) in acknowledging the riding posture as "cramped", "constrained" and "poor". Certain range of motions for instances the wrist extension could be varied to neutral once the desired throttle control is achieved. While others such as spinal flexion and shoulder abduction are usually assumed for longer duration – this is what was meant by Ma'arof et al., (2012) in acknowledging the riding posture as "static".

From the ethnographic observation, the author gathered the following results:

i. Each motorcyclist has a unique riding preference. Hence, their riding styles are different too. This result in variable definitions and perceptions made on comfort in motorcycling.
ii. Motorcyclists would usually purchase motorcycles that suit their personal style of riding. Hence, they are relatively satisfied with the riding posture for their respective motorcycles. Nevertheless, they are still interested in increasing the level of comfort during real world motorcycling in order to improve the experience. The authors were notified with various aftermarket products for various types of motorcycle (hence, various types of riding postures) that improve the ergonomics of the motorcycle. Commonly the aftermarket products are associated with the following trinity: handlebar, seat and foot-pegs.
iii. Senior motorcyclists (age) usually prefer a relatively more comfortable riding posture since this type of riding posture places less stress on their physical and physiological system – especially at the joints such as the wrists and ankles.

For experimental studies on investigating the motorcycling riding posture, the best reference frame is the duration of time spent in assuming, controlling and holding the riding posture. Distance traveled by the motorcyclist is not the best reference frame. For instance, a more powerful motorcycle could travel at a further distance in comparison with a less powerful motorcycle within the same time frame. Thus, exposure time to the risk factor (the riding posture) is the best reference time for motorcycling. The case studies give a perspective on the importance of assuming in the overall safety aspect of motorcycling. The summary is as follows in Table 2.

Table 2: Result summary of case studies

Case Study	Result Summary
Case study 1	i. High level of comfort from both physical and psychological aspect while motorcycling could potentially lead to accident ii. Certain degree of discomfort is required for overall safety in motorcycling iii. Based on the motorcyclist's anthropometric characteristics and the motorcycle ridden, the motorcyclist assumed the "RIPOC Type 2: upright riding posture" (Ma'arof and Ahmad, 2012b)
Case study 2	i. Assuming the riding posture for prolonged session results in various forms of fatigue with muscular and mental fatigue being the most significant. ii. Based on the motorcyclist's anthropometric characteristics and the motorcycle ridden, the motorcyclist assumed the "RIPOC Type 2: upright riding posture" (Ma'arof and Ahmad, 2012b). In varying his riding posture, the subject chose a more awkward riding posture - mimicking the "Type 1" Ma'arof and Ahmad, 2012b) – since it is more suited for motorcycling in term of aerodynamics and fuel expenses. iii. The duration of time in assuming the riding posture would result in various outcomes on the motorcyclist such as physical and physiological discomforts and fatigue. Hence, the need for further investigations.

Case study 3	i. Maintaining a certain riding posture in a prolonged ride may lead to fatigue. ii. Fatigue causes motorcyclist to be in an exhaustive state due to constraining the body on the motorcycle for better control. iii. However the prolonged duration builds up a certain level of boredom and may cause the motorcyclist to lose focus and judgment due to micro sleeps and drowsiness.

All the results from the field research and the case studies prove that awkward riding posture is a necessity for overall safety in motorcycling. A too comfortable riding posture would make it difficult to be physically mobile during real world motorcycling. This study strongly emphasizes that comfort is inversely proportional to the overall safety in motorcycling.

Indeed motorcycling experience could be improved with additional comfort, but, the concern here is to control the level of comfort in motorcycling. How much comfort is sufficient enough for motorcycling would be studied in another study. For this study, the authors anonymously agreed that based on the results and literature – comfort could be continuously supplemented in the motorcycling experience (e.g. using thicker seat-pan, additional lumbar support) for as long as the motorcyclist is capable in performing the "core technical skills" at optimal level (Code, 2010). Once the motorcyclist is incapable of performing the necessary task at the optimal safe level the amount of comfort should be decreased.

CONCLUSION

Conclusively, this study is not radically contradictory to the fundamental laws of ergonomics. It only advocates a different opinion on the niche area. This study strongly highlights that the notion of "the ideal human operator working posture" is limited to the human operator's safety performance. This is specifically critical when the work performed is mortally hazardous – such as real world motorcycling where the human operator is constantly at high risk.

The most crucial aspect of real world motorcycling is for the motorcyclist to be constantly capable of performing the fundamental motorcycling work tasks at the highest level of efficiency – and at times even demanding the motorcyclist to contradict with the natural survival instincts (Code, 1993). Any failure to perform the work task efficiently could drastically place the motorcyclist at mortality risk. Hence, this would violate the law of ergonomics: to ensure the human operator's safety while performing their work tasks (Health and Safety Executive, 2013; Bosch Rexroth AG, 2012). Based on the results, it is proven that the motorcyclist requires certain degree of "discomfort" in term of the working posture (the riding postures) to ride safely. Each of the riding posture represents varying degree of discomfort to the motorcyclist, hence, varying degree in the motorcyclist's capability in performing the motorcycling work task. Nevertheless, all the riding postures generally facilitate the motorcyclist in efficiently performing the work tasks. This finding is supported by literature such as Karmegam et al., (2012 and 2009), Velagapudi et al., (2010) and Code (1993).

The findings also prove that the current motorcycling riding postures are still the best working postures for motorcycling. Removing or eliminating the riding postures could seriously jeopardize the motorcyclist's safety. The riding postures could only be improved to promote its ergonomics qualities, yet, under a very close bound of safety. The study by Karmegam et al., (2012) is one of the best examples for such innovative study. The study managed to improve the riding posture – evidently the "RIPOC Type 2 Upright Riding Posture" (Ma'arof and Ahmad, 2012a) without significantly compromising on the motorcyclist's capability in performing the work tasks.

Indeed, assuming awkward riding postures could leads to the development of various physical, physiological and psychological discomforts. Independent or co-dependently, these discomforts would leads to fatigue. Thus, there is a need to address this issue for only then countermeasures and necessary preventive actions could be taken to minimize the unwanted physiological effects. However, how riders fatigue are defined has yet to be outlined because there is a significant paucity of motorcyclist fatigue research and those available are mostly self-reported data without proper design studies and methodologies (Horberry et al., 2008). Finally, the development of the noted physiological disorders is controlled by the duration of time being spent by motorcyclist in assuming, controlling and holding the riding posture. If the duration of time in which the motorcyclist could safely and healthily assume, hold and control the riding posture could be identified and the development of unwanted physical, physiological and psychological effects could also be controlled.

ACKNOWLEDGEMENT

The authors would like to express their utmost gratitude to Rohani Haron, Muhammad Fauzi Mohd Talib, Muhammad Firdaus Azhar, Muhammad Hafizudin Kamal, Wan Muhammad Syahmi Wan Fauzi, Mohd Afham Azmi, Muhammad Adzri Adzman, the staffs of the Faculty of Mechanical Engineering and the Institute of Graduate Studies in Universiti Teknologi MARA and Human Factors and Ergonomics Society Malaysia who have, directly or indirectly, contributed to this research. The authors would also like to acknowledge the Ministry of Education (MOE) Malaysia for providing the research fund for this research study through the Fundamental Research Grant Scheme (FRGS/1/2014/TK01/UITM/02/3).

REFERENCES

Berry, C., McNeely, A., Beauregard, K., Dickens, H., Geddie, J. E. (2009), "*A Guide To Ergonomics*", Occupational Safety and Health Division, NC Department of Labor, pp. 1-26.

Bosch Rexroth AG, (2012), "*Ergonomics Guidebook for Manual Production Systems*", Bosch Rexroth AG 2012, pp. 1-16.

Cal/OSHA, (1999), "*Easy Ergonomics*", Department of Industrial Relation, 1999, pp. 2.

Cañas, J.J, Velichkovsky, B. B. and Velichkovsky, B.M., (n.d.), "*Human Factors and Ergonomics*", pp. 6-7.

Chen, C. F. and Chen, C. W. (2011), "Speeding for fun? Exploring the speeding behavior of riders of heavy motorcycles using the theory of planned behavior and psychological flow theory", Accident Analysis and Prevention 43, pp. 983-990.

Chen, H-C., Chen, W-C., Liu, Y-P., Chen, C-Y., and Pan, Y-T. (2009), "*Whole-body vibration exposure experienced by motorcycle riders – an evaluation according to ISO 2631-1 and ISO 2631-5 standards*", International Journal of Industrial Ergonomics 39, pp. 708-718.

Code, K. "Code Break – Electronic Rider Aids. How A Chip Can Save Your Bacon". Motorcyclist July 2010, Retrieved August 22, 2013, from Sport Rider official website: http://goo.gl/hviXs8

Dul, J. and Weerdmesster B., (2003), "*Ergonomics for Beginners A Quick Reference Guide*", Second Edition, Taylor and Francis e-Library, pp.5-41.

Health and Safety Executive. (2013), "*Ergonomics and Human Factor At Work*", HSE, pp. 1-3.

Horberry, T., Hutchins, R. and Tong, R. (2008), "*Motorcycle Rider Fatigue: A Review*", Department for Transport, London 78, 2008.

Kroemer, K., Kroemer, H., and K. Kroemer-Elbert, K. (2000a), "*Ergonomics how to design for ease and efficiency (second edition)*", Prentice Hall International Series in Industrial & System Engineering, 2000a, pp. 1.

Kroemer, K., Kroemer, H., and K. Kroemer-Elbert, K. (2000b), "*Ergonomics how to design for ease and efficiency (second edition)*", Prentice Hall International Series in Industrial & System Engineering, pp. 398.

Keall, M. D., and Newstead, S. (2012), "*Analysis of factors that increase motorcycle rider risk compared to car driver risk*", Accident Analysis and Prevention, vol. 49, pp. 23-29.

Ma'arof, M. I. N., Ahmad, I. N., Abdullah, N. R., Karim, S. A. (2012), "*Motorcycling Riding Issues: Understanding the Phenomenon and Development of Ergonomics Intervention in Improving Perceived Comfort for Prolonged Riding*", iDECON.

Ma'arof, M. I. N., and Ahmad I. N. (2012), "*Proposed Standardized Method for Motorcycle Nomenclature System*". Southeast Asian Network of Ergonomics Societies Conference (SEANES), Road User Ergonomics, Paper 60.

Manan M. M. A., and Várhelyi, A. (2012), "*Motorcycle fatalities in Malaysia*", IATSS Research 36, pp. 30–39, 2012

Marina, M., Porta, J., Vallejo, L. and Angulo, R., (2011), "*Monitoring hand flexor fatigue in a 24-h motorcycle endurance race*", Journal of Electromyography and Kinesiology, vol. 21, pp. 255-61,

Ministry of Transport Malaysia, Transport Statistic, 2011, pp. 28.

O'Sullivan, K., O'Dea, P., Dankaerts, W., O'Sullivan, P., Clifford A., and O'Sullivan, L. (2010), "*Neutral lumbar spine sitting posture in pain-free subjects*", Manual Therapy 15, pp. 557-561.

O'Sullivan, K., O'Sullivan, P., O'Sullivan, L., and Dankaerts, W., (2012), "*What do physiotherapists consider to be the best sitting spinal posture?*", Manual Therapy, 2012, pp. 1-6.

Openshaw, C., Taylor, E. (2006), "*Ergonomics and Design A Reference Guide*", Allsteel Inc., pp.15.

OSHA, (2000), "*Ergonomics: The Study of Work*", U.S. Department of Labor Occupational Safety and Health Administration, 2000, pp. 1-10.

Jou, R. C., Yeh, T. H., and Chen, R. S. (2012), "*Risk factors in motorcyclist fatalities in Taiwan*", Traffic Inj Prev, vol. 13, pp. 155-62.

Sècuritaire, T. (2011), "*Ergonomics Guidelines for Manual Material Handling 2nd Edition*", WorkSafeNB pp. 1-18.

Shell. (2013), "*Shell Responds To Boom In Asian Motorcycle Market".* Retrieved on 22nd August 22, 2013 from Shell Global official website: http://goo.gl/bXADNe

Velagapudi, S.P., Balasubramanian, V., Adalarasu, K., Babu, R. and Mangaraju, V. (2010), "Muscle fatigue due to motorcycle riding", SAE Technical Papers, 2010.

W. H. t. Schneider, P. T. Savolainen, D. Van Boxel, and R. Beverley, (2012), "*Examination of factors determining fault in two-vehicle motorcycle crashes*", Accident Analysis and Prevention, vol. 45, pp. 669-76.

Williamson, A., Lombardi, D.A., Folkard, S., Stutts, J., Courtney, T.K. and Connor, J.L., (2011), "*The link between fatigue and safety*", Accident Analysis Prevention, vol. 43, pp. 498-515.

Young, E. (2013), "*Petrol Price: Why are they rising?*" Retrieved on 22nd August 22, 2013 from BBC News Official Website: http://www.bbc.co.uk/news/business-21547684.

Application of Galvanic Vestibular Stimulation for the Evaluation of Vehicle Settings in a Fixed-Base Simulator

Matthias Wiedemann [1], Wolfram Remlinger [1] and Klaus Bengler [2]

[1] AUDI AG
Department for Processes and Methods in R&D, I/ET-84
85045 Ingolstadt, Germany

[2] Institute of Ergonomics
Technische Universität München
85747 Garching, Germany

ABSTRACT

Driving simulation is a well probed and accepted technology for virtual development in automotive engineering. In the usual driving task visual perception is the most important factor. However, for the evaluation of vehicle settings in the field of driving dynamics it is essential, not only to generate realistic visual, but also vestibular cues for the driver. Therefore, dynamic driving simulators are usually applied, though they are a costly investment.
Galvanic Vestibular Stimulation (GVS) creates the impression of accelerations by influencing the human equilibrium organ. In flight and driving simulation successful implementations of GVS systems already exist, mainly with the focus on reducing simulator sickness or to improve driver's adaption to the simulator.
Here, a concept for linking a GVS system with a fixed-base driving simulator is introduced. First, the adaption of the GVS system to individual preferences is conducted. In a second step, the evaluation in terms of objective driving behavior and subjective survey data is investigated to describe the influence of the GVS induced acceleration perception. Experiments are run for the driver as well as the passenger role. As a result the experiment should prove, whether the user is enabled to differentiate between various vehicle settings in the field of driving dynamics, as it would be possible in a dynamic driving simulator.

Keywords: Galvanic Vestibular Stimulation, Driving Simulation, Driving Dynamics, Human Perception

INTRODUCTION

Virtual development is assumed to be the key factor for further success in automotive engineering during the next years. Ongoing diversification to meet customers' expectations of an individual, personalized car increases the demand of testing procedures. Physical experiments however require the time-consuming and cost-intensive assembling and testing of prototype vehicles. The application of virtual tools has the potential to meet this increasing challenge in many ways.

First conceptual questions can already be discussed in the very early states of the development process in which prototype vehicles are not available at all. Furthermore positive influences are generated for following steps in

development. In foreseeable future there will still be the need of physical testing procedures. However, by an improved system knowledge gained in the virtual world, the number of tests can significantly be reduced, while quality and informative value rise.

Today's development process in automotive engineering is focused on customer relevant properties of the vehicle. Several hundred features and qualities of a new model have to be appraised by its designing engineers. The appropriate virtual methods last from computational fluid dynamics simulation to finite element methods. As the features of a car often are highly emotional, engineers need the possibility to get a subjective impression as well. Hence, not only objective calculation methods are needed but also tools for subjective experience have to be developed. High resolution 3D-models are state of the art for the visualization of design relevant questions. A high number of other vehicle properties can be illustrated in a driving simulator.

HUMAN PERCEPTION OF DRIVING DYNAMICS

Visual Perception

Certainly, visual perception is the most important information in the driving task. Most references mention numbers around 90%, but it is uncertain, whether the ratio between visual and other cues can be quantified at all (Sivak, 1996).

Vision is especially needed to percept the absolute position of the vehicle (e.g. speed, yaw angle or side slip angle), the relative movement on the road (e.g. relative speed, yaw angle offset or track offset) and to interact with other traffic participants (Wolf, 2009; Schimmel, 2010). According to Schimmel (2010) for evaluating properties in the field of driving dynamics, speed, radius of curvature, side slip angle, yaw angle and the steering wheel angle are the most important information, that have to be presented as an optical cue.

For a simulator this means, that special effort has to be made in the configuration of displays or projection systems for features as brightness, contrast, field of view or resolution. In simulators, that are state of the art, most of these challenges are already solved. For instance in terms of resolution typical simulator values already are close to the limit of human visual acuity (Greenberg and Blommer, 2011).

Auditory Perception

In usual road traffic the aural sense should not be underestimated. The sound of other traffic participants especially in city conditions can help to give orientation in complex traffic situations. Also warning strategies of driving assistance systems often use auditory cues. In the context of vehicle development auditory perception is especially important for questions of sound design or to avoid noise as a consequence of vibrations.

However, in the field of driving dynamics auditory perception seems to have a subordinated role. Certainly, the wheel and driving noise can help engineers, but Schimmel (2010) remarks that track engineers usually do not take them into account.

Haptic Perception

Haptic perception denotes the combination of proprioceptive and tactile perception. The first one is important to control the input in human machine interfaces, primary the steering wheel, throttle, clutch or brake pedal. Latter for instance helps to perceive vibrations in the vehicle with high sensitivity. For the evaluation of driving dynamics haptic cues are an important information for the driver. Differences in steering wheel angle, steering wheel torque as well as braking response cannot be assessed without them.

To generate the best possible feeling of immersion in a driving simulator, today force feedback steering wheels or hardware-in-the-loop simulations of braking pedals are state of the art.

Vestibular Perception

In a vehicle longitudinal, lateral and vertical as well as roll, pitch and yaw accelerations occur. The human equilibrium organ consists of three semicircular canals to detect angular movement and the otolith organ to perceive linear accelerations and gravity. This complex system is very sensitive in detecting variations. For angular accelerations threshold values around 2-4°/s² are mentioned in a wide range of experiment settings (Tomaske, 1983). In the case of lateral accelerations threshold values even as low as 6.5cm/s² can be found (Kingma, 2005).

Accelerations as a response of steering wheel or brake inputs are important cues to evaluate cornering abilities and braking performance. Schimmel (2010) points out, that especially lateral, longitudinal and yaw accelerations have a strong influence on the perception of vehicle properties.

DRIVING SIMULATION IN THE VIRTUAL DEVELOPMENT PROCESS

Demands on Simulators

Driving simulation is an accepted and well probed technology within the virtual development process. In the fields of driver assistance systems or functional interior design mainly fixed-base simulators are used. For analyzing driving dynamics or comfort, such as steering or suspension settings, more sophisticated, dynamic simulators with a motion platform are applied.

Depending on these diverse use cases different approximations to reality are necessary. Negele (2007) uses two generally accepted models to arrange a classification. The first is the three-layer model of the driving task, in which the primary task itself is divided in the navigation, vehicle guidance and stabilization layer (Donges, 1978; Bubb, 2003). The combination with Rasmussen's model of human behavior (Rasmussen, 1983) leads to different demands on driving simulators.

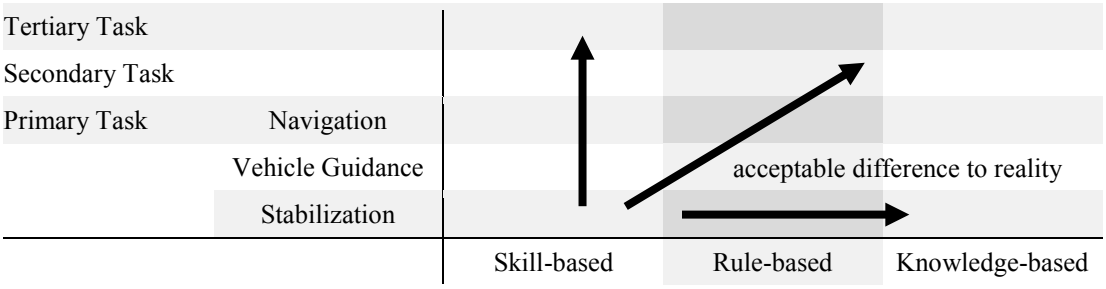


Figure 1: Scheme to classify driving simulators. (Adapted from Negele, 2007)

Whenever there is enough time or the task is not directly linked to driving, some simplifications can be made, as there is the possibility of mental compensation for the driver. For time-critical and skill-based tasks however the whole range of cues for human perception has to be reproduced in high quality. (Negele, 2007; Advani and Hosman, 2001)

In the case of the evaluation of driving dynamic properties this means, that almost no simplifications of reality are possible.

Overview of Technical Solutions

As pointed out before, for the evaluation of driving dynamics properties vestibular, proprioceptive and tactile cues

are important. For each of them a wide range of technical solutions in simulators was created, lasting from large and cost intensive systems to small but still very effective improvements.

From the very beginning the aim of driving and flight simulation was to represent movements as realistically as technically possible. With the invention of the Stewart platform motion in six degrees of freedom (DOF) was made applicable (Stewart, 1966). As humans do not recognize tilt angles up to 25°, as long as the movement is slow enough, it is possible to use a part of gravity for the impression of accelerations, e.g. centrifugal accelerations by tilting the motion platform (Negele, 2007; Hosman and Advani, 2001). Nevertheless this technique is not always able to reach the values of the real acceleration. Therefore scaling factors in the motion cueing algorithms, which control the platform movement, have to be implemented.

Today's large simulators use redundant systems to provide best possible vestibular cues. The probably best known example is the combination of the Stewart platform with lateral and/or longitudinal displacement systems like in the Renault, Toyota or National Advanced Driving Simulator. Those simulators in research and automotive development are able to present lateral and longitudinal accelerations at least during the normal driving task without scaling. These abilities of course come with the disadvantage for the operating company to be a costly invest.

Usually Stewart platforms are only suitable for representing motion at low frequencies. Therefore other systems additionally have to be integrated in a simulator to close the gap to reality. Noise Vibration Harshness (NVH) describes occurring vibrations in the vehicle in the transition range between auditory and haptic perception, around 20 to 100 Hz. In driving simulation this usually undesired effect is necessary for a better immersion. Berberich, Gebhard, Bohlen, Danninger and Lienkamp (2012) introduce a technical solution for high and low frequencies by integrating actuators into the steering column and seat rail. Their results show that in the range from 20 up to 60 Hz realistic excitations can be achieved.

Latency Time

Entire immersion cannot be achieved by presenting perfect cues for each sensory channel itself. Coordination and synchronization of the different pieces of perceptual information is of outstanding significance. A perfect steering wheel feeling with realistic forces, restoring torque and elasticity becomes worthless if the image generation is not able to show the effects on the screen in real time without noticeable latencies.

Stoner, Fisher and Mollenhauer (2011) assume latency times to be the "most important measure of simulator performance." At the same time it seems to be complex to calculate the effect on the driver, as there is a lack of explicit numbers (Kemeny, 2001; Stoner, Fisher and Mollenhauer, 2011).

GALVANIC VESTIBULAR STIMULATION

Functionality

As explained before, the human equilibrium organ is divided into the semicircular canals for detecting circular motion and the otolith organ to percept linear accelerations. In the canals the movement of hair cells is essential for the transduction of the stimulus into a neuronal signal. The hair cells' receptors are situated along the circumferential axis of the particular canal. Depending on the spatial orientation of the semicircular canals the acceleration perception results in the specific direction. (Fitzpatrick and Day, 2004; Goldstein, 2008)

While the horizontal canal responds to yaw, the other two canals (anterior and posterior canal) are tilted at 45° to the sagittal (frontal) axis and therefore detect both, roll and pitch movements (Fitzpatrick and Day, 2004).

Galvanic Vestibular Stimulation (GVS) creates a virtual signal of head movement by applying a small current to the head at the position of the mastoid with the help of two electrodes. The physiological effect is a bypass of the transduction in the hair cells resulting in a direct activation of the nerve. (Fitzpatrick and Day, 2004)

By this method however, all three semicircular canals and the otolith organ are stimulated at the same time, as "unlike a natural stimuli, GVS has no direction" (Fitzpatrick and Day, 2004). As a consequence of the special spatial orientation of the semicircular canals and the mirror inverted structure of the left and right equilibrium organ some of the effects neutralize each other. In result, Fitzpatrick and Day (2004) describe an impression of rotation around an axis of 27.1° from a horizontal line as the calculated effect for the semicircular canals. The effects for the otolith organ in terms of linear accelerations however show no clear results.

Still their conclusion is that test persons during GVS use will percept strong roll and slight yaw accelerations towards the side of the cathode.

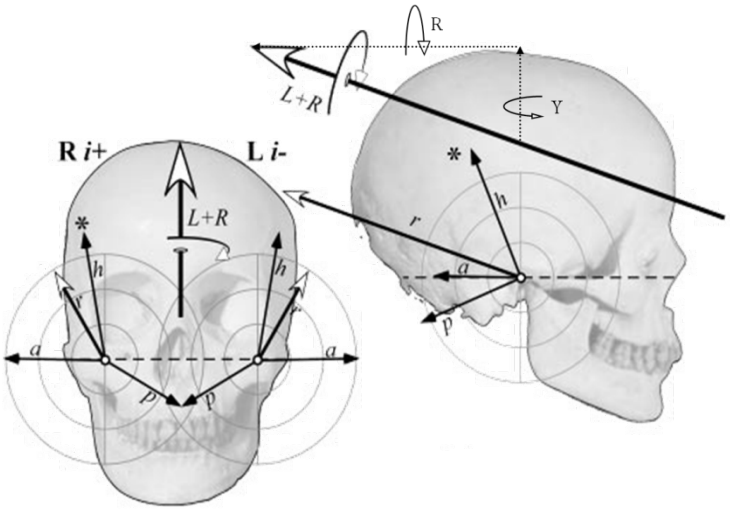


Figure 2: Resulting vectors during Galvanic Vestibular Stimulation with the cathode on the left (L i-) and the anode on the right (R i+) side. Vectors of anterior, posterior and horizontal canal (a, p, h), resulting vectors of left and right canals (r), vector sum of left and right side (L+R) with its roll (R) and yaw (Y) components. (Adapted from Fitzpatrick and Day, 2004)

In the meantime even more complex systems were realized. In 2011 a GVS system with the ability to simulate accelerations in all directions was protected by patent (Cevette, Stepanek and Galea, 2011). This is made possible by varying the number and positions of the electrodes.

Application in Flight and Driving Simulation

The possibility to create impressions of acceleration with comparatively low effort, makes this technology interesting for the application in flight and driving simulation.

Moore, Dilda and MacDougall (2011) describe the use of GVS in a spaceflight simulator to create the feeling of spatial disorientation after a spaceflight. The previous problem in this simulator was the lack of realism during the landing approach of the space shuttle. While in real conditions landing speed usually was too high, simulator landings were conducted perfectly by the pilots. Therefore a GVS system was used to rebuild the impression of reentering the earth's gravitational field. Pseudorandom stimuli influenced the sense of balance and lowered the pilots performance, resulting in significantly more realistic landings.

Another well-known use case is the reduction of the simulator sickness syndrome by GVS. The application of GVS in this context was already patented by Mark in the United States in 1985 (Mark, 1985).

Reed-Jones, Reed-Jones, Trick and Vallis (2007) evaluated a GVS system in different conditions by measuring subjective as well as objective data. Their finding was that the use of GVS significantly reduces motion sickness and

disorientation. In terms of objective driving behavior it could be shown that steering variability is reduced. Furthermore there is an increase in speed during curved driving. According to Reed-Jones et al. (2007) this effect points out that the task of curve driving is easier to handle for the test persons with GVS than in usual simulator conditions and therefore shows that driving behavior becomes more realistic.

CONCEPTUAL DESIGN OF THE APPLICATION OF A GALVANIC VESTIBULAR STIMULATION SYSTEM

Aims of the Project

As several studies show, GVS is able to generate a realistic feeling of accelerations. The aim here is not to investigate details of physiological function of the stimulation as in medical research, but the possibility of using it in the virtual development process of automotive engineering.

Improvements by the help of GVS can only be achieved, when the technology offers the possibility of reliable and valid decisions on vehicle settings. Therefore it is especially important to investigate the driver's feeling in terms of subjective driving impression and the ability to identify and separate different settings of the vehicle. Furthermore the analysis of objective driving behavior and comparison to data from real vehicles is necessary.

Experimental Design

Fig. 3 illustrates the experimental design. After some pretests the evaluation is conducted for the passenger as a well as the driver role. The analysis of threshold values can already give evidence for the reliability of the GVS system. Validity in terms of driving behavior is tested by comparison with real vehicle data.

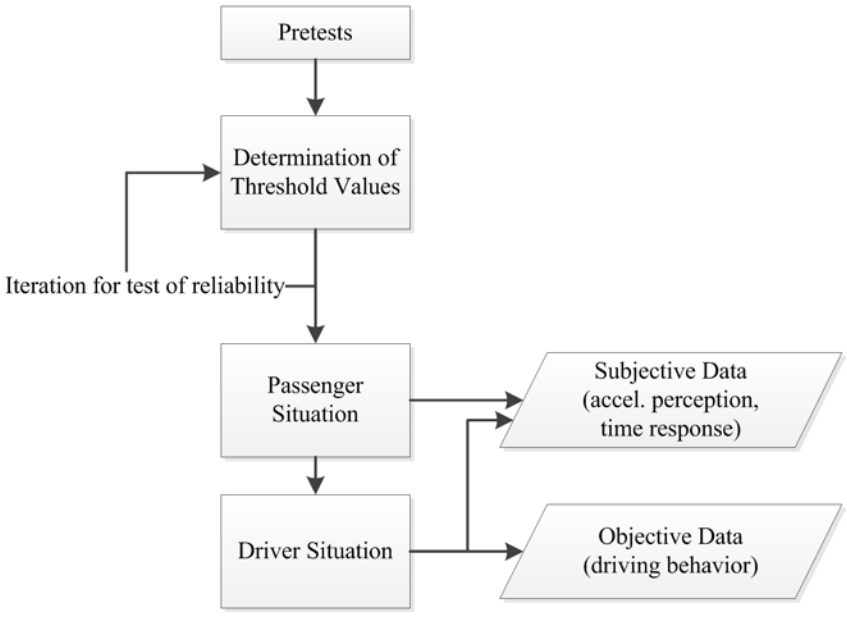


Figure 3: Experimental design and resulting data

Technical Implementation

The hardware of the used GVS system consists of two electrodes, which are connected to a receiver unit. The associated transmitter is connected to a PC-system. The current can be varied by the help of a rotary potentiometer

at the receiver unit (max. 1.6mA, DC). The Audi simulators use the software veDYNA (TESIS DYNAware) for calculating vehicle dynamics. The occurring lateral accelerations are read out by the GVS software tool, low-pass filtered and send to the transmitter unit in real-time.

Conclusions from Pretests

For the realization of the experiment pretests with four participants were conducted. First, the basic question, whether centrifugal forces can be felt in the right way was investigated. The test persons sat on the passenger seat and were asked to name the direction of the curve while having their eyes closed. Previous to the test drive the highest possible lateral acceleration of 10 m/s² was applied to the GVS system. The participants then had the possibility to adapt the intensity of current to personal preferences.

Correct answers were given in 64.3% of the trials. However there were no mismatches, all other curves just stayed unperceived. The directions of curves with maximal lateral accelerations of 5m/s² or higher were entirely detected. This leads to the conclusion that the chosen intensity of current and thereby stimulation was possibly too weak. None of the participants used a higher current than 0.5 mA. In comparison Reed-Jones et al. (2007) report to have used a current of 0.6-1.25 mA. The main reason for the test persons not to decide for a higher intensity was a prickle and an irritation of the skin at the position of the electrode. Other phenomena were a metallic taste or the feeling of disorientation.

As a consequence the fixation of the electrodes has to be optimized to improve the current flow. For further testing, different types of electrodes, varying in size and form, are applied.

Subjective Evaluation

As the pretests already show, the use of GVS has a strong personal component. The same current intensity does not correspond to the same intensity of acceleration perception. This implicates the question how experiment results especially in the case of subjective answers can be made comparable.

First, it is important to adapt the system to the individual threshold values. For safety reasons the current intensity has to stay beyond the pain threshold. As the highest occurring acceleration in the driving dynamics model is limited to 10 m/s², this value is used when test persons are asked to set the highest tolerable current intensity. The perception threshold is determined by a staircase procedure for the variation of the current intensity at an acceleration of 2 m/s². As discussed before, the perception threshold for lateral accelerations in clinical testing conditions is much lower. Otherwise, the threshold in a vehicle probably lies higher as there are several other interference factors. On the simulated test track in each curve lateral accelerations of at least 2 m/s² occur at normal driving speed.

Beside the threshold values other individual influences have to be investigated. It is uncertain whether for each test person the intensity of left or right site stimulation is the same. Consequently, all measurements must take this into account. Furthermore, the manner how the acceleration feeling occurs has to be tested. As discussed above, GVS mainly creates a feeling of roll and slight yaw movement. The conformity with real centrifugal forces must be analyzed in a survey. In addition, the time response between the acceleration perception and the physically correct point in time can influence the level of realism. As well this effect is subject to research.

Objective Evaluation

Objective data for the evaluation of driving behavior is recorded. As the aim of the project is to improve the drivers' ability to differentiate vehicle settings in a simulator, realistic steering wheel or braking inputs are a basic requirement.

The used test track is a precise, laser scanned simulation of a real track, hence several data sets of real driving exist. These can be used to compare for example steering behavior like steering variability or frequencies and specify the

level of realism. With the help of the findings of Reed-Jones et al. (2007) it is as well possible to cross check both results.

CONCLUSIONS

Galvanic Vestibular Stimulation is a well investigated technology in medical research. First successful applications in the field of flight and driving simulation support the thesis of a higher level of realism in driving behavior. Pretest results show individual differences in the perceived strength of the stimulation. Therefore, as a first step, individual threshold values have to be determined. In a second step the subjective and objective data is analyzed in both, passenger and driver situation. If an improvement in comparison to a usual fixed-base simulator is observed, GVS systems can be a helpful technology in the virtual development process in automotive engineering.

REFERENCES

Advani, S. and Hosman, R. (2001), “Integrated motion cueing algorithm and motion-base design for effective road vehicle simulation”, Proceedings of the driving simulation conference 2001, pp. 263-271.

Berberich, T., Gebhard, P., Bohlen, S., Danninger, O., Lienkamp, M. (2012), “A new approach to the presentation of vibration phenomena in vehicles“, Proceedings of the FISITA 2012 World Automotive Congress, F2012-J08-002.

Bubb, H. (2003), “Fahrerassistenz primaer ein Beitrag zum Komfort oder für die Sicherheit?“, in: Der Fahrer im 21.Jahrhundert, VDI-Berichte Nr. 1768, Duesseldorf.

Cevette, M., Stepanek, J., Galea, A. (2011), “Galvanic vestibular stimulation system and method of use for simulation, directional cueing, and alleviating motion-related sickness”, United States Patent Application Publication, US 2011/0029045 A1.

Donges, E. (1978), “Ein regelungstechnisches Zwei-Ebenen-Modell des menschlichen Lenkverhaltens im Kraftfahrzeug”, in: Zeitschrift für Verkehrstechnik 24, pp. 98-112.

Fitzpatrick, R., Day, B. (2004), “Probing the human vestibular system with galvanic stimulation”, in: Journal of Applied Physiology 96: 2301-2316.

Goldstein, E.B. (2008), “Wahrnehmungspsychologie: Der Grundkurs”, Spektrum Akademischer Verlag, Berlin, ISBN 978-3827417664.

Greenberg, J., Blommer, M. (2011), “Physical Fidelity of Driving Simulators”, in: Handbook of Driving Simulation for Engineering , Medicine and Psychology: 7-1 – 7-24.

Kemeny, A. (2001), “Recent developments in visuo-vestibular restitution of self-motion in driving simulation”, Proceedings of the driving simulation conference 2001, pp 15-18.

Kingma, H. (2005), “Thresholds for perception of direction of linear acceleration as a possible evaluation of the otolith function”, BioMedCentral Ear, Nose and Throat Disorders 5:5.

Mark, H. (1985), “Vestibular Stimulation Method”, United States Patent, 4558703.

Moore, S.T., Dilda, V., MacDougall, H.G. (2011), “Galvanic vestibular stimulation as an analogue of spatial disorientation after spaceflight“, Aviation, Space, and Environmental Medicine, Vol.82, Num.5, pp. 535-542.

Negele, H.J. (2007), “Anwendungsgerechte Konzipierung von Fahrsimulatoren für die Fahrzeugentwicklung”, doctoral thesis, Technische Universität München.

Rasmussen, J. (1983), “Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models”, IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-13, No. 3, pp. 257-266.

Reed-Jones, R.J., Reed-Jones, J.G., Trick, L.M., Vallis, L.A. (2007), “Can Galvanic Vestibular Stimulation Reduce Simulator Adaptation Syndrome?”, in: Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design: Driving assessment 2007, Iowa City, University of Iowa, Public Policy Center, pp. 534-540.

Schimmel, C. (2010), “Entwicklung eines fahrerbasierten Werkzeugs zur Objektivierung subjektiver Fahreindrücke“, doctoral thesis, Technische Universität München.

Sivak, M. (1996), "The information that drivers use: is it indeed 90% visual?", Perception 25(9), pp. 1081-1089.

Stewart, D. (1966), “A platform with six degrees of freedom”, The Institution of Mechanical Engineers Proceedings 1965-66, 180 Part 1 (15), pp. 371-386.

Stoner, H., Fisher, D., Mollenhauer, M. (2011), “Simulator and Scenario Factors Influencing Simulator Sickness”, in: Handbook of Driving Simulation for Engineering, Medicine and Psychology: 14-1 – 14-24.

Tomaske, W. (1983), “Einfluss der Bewegungsinformation auf das Lenkregelverhalten des Fahrers sowie Folgerungen für die Auslegung von Fahrsimulatoren”, doctoral thesis, Hochschule der Bundeswehr Hamburg.

Wolf, H. (2009), “Betrachtung des Lenkgefühls beim Führen eines Kraftfahrzeugs unter besonderer Berücksichtigung ergonomischer Erkenntnisse und Methoden“, doctoral thesis, Technische Universität München.

Comparison of Real and Simulated Driving for a Static Driving Simulator

Michael Krause, Levent Yilmaz and Klaus Bengler

Institute of Ergonomics
Technische Universität München
Garching, D-85747, Germany

ABSTRACT

A traffic light assistant on a smartphone is used in real and simulated driving in an experiment. The speeding behavior of the test subjects is assessed as metric of travelled distance with more than 5 km/h above the speed limit, subjective scores by NASA-RTLX, general fuel reduction and comparison of occurring accelerations. The speeding reduction showed numerically not the same values for real and simulated driving, but the reduction effect was present in both; in real with the smaller effect. According to subjective ratings the simulator driving is judged in tendency more demanding. A fuel reduction effect for the green wave coordination found in the simulator could also be found in real traffic, but also with lower effect. The experiments revealed no absolute validity for any metric, but relative validity for some metrics.

Keywords: In-vehicle Information, Traffic Light Assistant Driving Simulation, Simulated Driving, Real Driving, Simulator Study, Driving Behavior

INTRODUCTION

The project KOLIBRI was started in April 2011 and lasted until June 2013. The aim of the project was to reduce stops at traffic lights and thus cut down on pollutants such as noise and noxious emissions resulting from avoidable decelerations and accelerations. The project focused on traffic lights outside urban areas (arterial roads). In the project, two test sites were modernized, so the traffic lights could transmit their process data to a central server. Based on this historical data, traffic engineers from TRANSVER GmbH (Munich) were able to estimate the behavior of traffic lights at special times in the future. These estimated traffic light switching times were transmitted to a demonstration car of BMW and smartphones from the Technische Universität München to inform the driver in advance about the state of the next traffic light.

The experiment published in this paper was conducted during an intermediated step in the project on one of the test sites in the north of Munich. The traffic lights were programmed with a fixed coordinated switching scheme (green wave). With the fixed switching scheme, it was possible to program the behavior of the traffic lights into smartphones without the need of the communication infrastructure between the traffic lights and central servers, which were just being implemented at that time. Thus, the smartphones in the test vehicles had enough information about the (exact) time behavior of the traffic lights, the GPS position of stop lines and the velocity and position of the moving car via GPS. From former experiments, we already had a model of the road section for our driving simulator to test early prototypes of the Human Machine Interface. It was also possible to program the green wave into the driving simulator. Thus, we had the opportunity to test the interface on the real road and in the simulator. Therefore, we had the idea to let the KOLIBRI smartphone system be evaluated by the same test subjects (within design) in real and simulated driving.

This idea was originally not part of the project, but due to the possibilities offered by modern smartphones it could be easily implemented even in a short two-year project such as KOLIBRI. This is an example of how fast evolving

consumer electronics stand in contrast to long-term developments and the product lifecycles of automobiles or even infrastructure components.

Different institutes have a long tradition in driving simulator (validation) studies, a recommendable review can be found by Blana (1996). A recent overview of this field can be obtained by Mullen et al. (2011) and Ahlström et al. (2012).

The understanding of validities, especially relative and absolute validity, in this paper is that of Mullen et al. (2011). So to cite Mullen et al. (2011): "While absolute validity requires that the two driving environments produce the same numerical values, relative validity is established when the differences between the two environments are in the same direction, and of the same or similar magnitude."

While former studies mainly found relative validity: "Absolute validity is rarely established in driving simulator studies, […]" Mullen et al. (2011). Ahlström et al. (2012) are one of the rare cases; they achieved absolute validity in real and simulated driving for speed.

Traffic light assistance is also no new topic. There are several projects and papers around dealing with this issue in theory or computer simulations. We were more interested in projects that have a strong link to real experiments. A short overview of different approaches for real traffic light assistance was given in Krause and Bengler (2012). Here we will especially refer to Pauwelussen et al. (2008). They implemented a green wave support system in their driving simulator. As did Wooldridge, Trayford and Doughty, (1984), who tested a traffic light assistant in real traffic. Both projects made thoughts about behavior and human factors engineering.

METHODS

Experimental Design

The experiment had a within-design, so every subject drove each condition and every subcondition. Two main blocks, afterwards referred to as **conditions**, were:

- Driving on a real road (**REAL**)
- Driving in the simulator (**SIM**)

Each of these conditions was further subdivided into four **subconditions:**

- Rural road driving, uncoordinated traffic light switching scheme without the KOLIBRI traffic light assistant (**NCO**)
- Rural road driving, coordinated traffic light switching scheme (green wave) without the KOLIBRI traffic light assistant (**GW**)
- Rural road driving, coordinated traffic light switching scheme (green wave) with the KOLIBRI traffic light assistant on a smartphone (**KO**)
- Suburban driving (**URB**)

The subconditions are further described in the following sections.

Procedure

The experiment lasted about two hours per test subject. Due to time restrictions, only two persons drove first in the simulator and then on the road. The rest followed the sequence: first real driving, than the simulator. The order of the subconditions was randomized. The direction of driving (north-south or south-north) was also randomized, but controlled between simulator and real driving. The real tracks are about ten minutes from the institute, the test subject drove this distance to accommodate to the real car. For accommodation to the driving simulator, the subjects drove 5 minutes. The experiment was carried out in June/July 2012. The KOLIBRI Human Machine Interface was verbally described to the subjects, while standing still.

Tracks

In the conditions NCO, GW and KO, the same road section was used. The road section is a 6.5 km part of the federal road B13 in the north of Munich. On this section, seven traffic-light-controlled intersections were passed. The subject drove straight forward from 48°17'56 N 11°34'36 E to 48°14'48 N 11°36'7E. The road has two lanes per direction separated by guide rails (see Figure 1). The subjects randomly drove the conditions in direction north-south or south-north. The volume of traffic under real circumstances is below 500 vehicles/hour/direction during the day at the time of the experiments. In the simulator, some cars in front and behind were added, that never interacted with the driver (locked in place at a certain distance). The speed limit on these road sections is most of the time 100 km/h and reduced to 70km/h in traffic light areas. The street was modeled for the simulator (see Figure 1). The focus was not on every single visual detail; instead, that stop lines and speed limits were at the right place, in line with GPS positions. The traffic light scheme (green wave) especially designed for the KOLIBRI project was only active at noontime. During the time before and after this period, the traffic lights were uncoordinated and operating on a traffic response scheme by car detectors. For the simulator, the green wave scheme was programmed into the simulated traffic lights. For the uncoordinated scenario in the simulator, an uncoordinated backup scheme of the real traffic lights was set up. On the actual road, every single traffic light is equipped with this backup scheme, which is activated when car detectors are not functioning. As can be seen in Figure 1, the road section has a highway-like appearance.

Figure 1 B13 track
Top: Real road (slightly different view angle, camera at radiator grill)
Bottom: Simulator track

For suburban driving on a real road, a suburban district near the B13 was used (see Figure 2 and 3). The route through this area was approx. 2 km long, mainly with speed limit of 30 km/h. The examiner specified the route and announced turns.

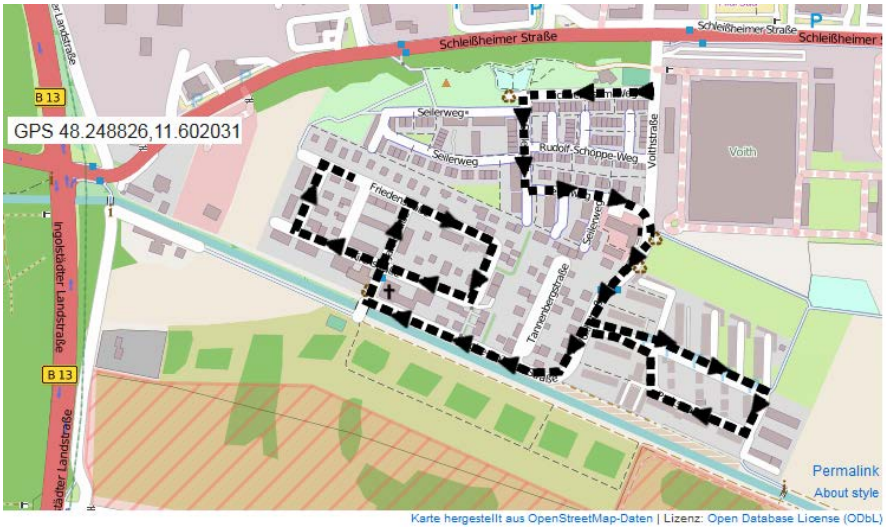


Figure 2 Real track through a suburban area

Figure 3 Real track suburban driving environment

For the simulator, an innercity scenario was used (see Figure 4) that was originally implemented for another project (Helmbrecht, 2012). The scenario was not a model of the real suburban district. The distance to travel was 2.3 km, restricted to 50km/h or 30km/h. The track through this city (turns) was communicated to the driver via an automated navigation voice.

Figure 4 Simulator track suburban driving (Helmbrecht, 2012)

Road Vehicle

For the experiment on the road, a BMW 520d (year of manufacture 2011, type F10) diesel-engined and with automatic gear was used. The velocity was tracked with a GPS Trip Recorder 860E at 1 Hz. Fuel consumption was manually read from the board computer and reset following each condition. According to personal communication with BMW technicians, the measurement of fuel consumption in the board computer had not been manipulated. The test subject drove the car without the Head Up Display (HUD).

Driving Simulator

The static driving simulator of the Institute of Ergonomics consists of three projectors (1400x1050 resolution), which offers an almost 180-degree front view on three 3.4 m x 2.6 m screens. Three other projectors displayed images to screens behind the BMW mockup, visible to the driver via the car mirrors. For driving simulation, the software SILAB V3.0 from WIVW GmbH, Würzburg was used. CarSim V7.11 from Mechanical Simulation, Ann Arbor, was used for the car model to calculate fuel consumption and dynamics. The software was loaded with a 150kW sedan model. An active steering wheel with software from Simotion, Munich provided the driver an adequate feeling for steering. The simulator logged all available data with 120 Hz.

KOLIBRI Traffic Light Assistant – Human Machine Interface (HMI)

The traffic lights were operating in the subcondition GW and KO on a coordinated green wave, so the switching times were predefined and could be programmed into a time-synced smartphone. The KOLIBRI traffic light assistant system on a smartphone gave the driver speed recommendations (within the allowed speed limit) to make it to the next traffic light while still green (Figure 5, left) or it would advise that it would not be possible to make the next traffic light while still green (Figure 5, middle). If the driver drove more than 10 km/h too fast, a speed alert was shown (Figure 5, right).

Figure 5 KOLIBRI MMI
Left: Speed recommendation. Green area indicates speeds to reach the next traffic light at green.
Mid: It is not possible to make the next traffic light at green (Prepare to stop).
Right: The vehicle is more than 10 km/h faster than allowed (Speed Alert).

The smartphone was mounted right beside the steering wheel next to the air vents (see Figure 6).

Figure 6 Mounting of the smart phone in the simulator (left) and the vehicle for the field test (right)

NASA-RTLX

In the experiment, the well-known task load questionnaire NASA-TLX (Hart and Staveland, 1988) was used in an unweighted respectively equally-weighted version, sometimes called NASA-RTLX (R = raw). The German translation stems from Seifert (2002). The questionnaire was filled in after each experimental subcondition.

Participants

Twenty subjects participated in the experiment (19 male, 1 female). No participant quit as a result of simulator sickness. The subjects were between 25 and 35 years of age, the average age was 28 years of age. All subjects had a valid driver's license and drove on average 12,000 kilometers per year (SD: 7,300 km). 90% had driven a simulator before and 80% had specific previous knowledge of the static driving simulator at the Institute of Ergonomics. All participants had experience with an automatic gear. Eight persons were familiar with the KOLIBRI traffic light assistant. Four of these persons had used the assistance system before in other simulator experiments. The other four participated in an experiment conducted in parallel, in which the assistant was also used on the real road.

Statistics

Unless otherwise reported, the alpha level is set 0.05.

The critical value for an r-coefficient to become significant on level of p=0.05; df=18 is **0.444**. In tables, this value is indicated in light gray.

The level of p=0.01;df=18 results in a critical value for the correlation of **0.561,** often connected to the term “highly significant” is shown in tables as dark gray.

The default reported error bar is the 95% Confidence Interval (CI).

RESULTS

Driving Behavior – Speeding

Figure 7 and Table 1 contain information about the metric, i.e. percentage of travelled distance with more than 5 km/h above speed limit. In former studies (Krause and Bengler, 2012; Krause et al. 2014a), we used the percentage of time that the subjects were traveling above the speed limit, in relation to the time when the car is moving (>5 km/h). We assumed that expressing speeding in percentage of travelled distance and with a little tolerance of 5 km/h above speed limit would be slightly more accurate and easier to imagine,.

As can be seen in Figure 1, the subjects drive in all scenarios faster in the simulator than on a real track. This difference is the lowest for the suburban condition (URB). While for the highway conditions (NCO, GW), the quickly travelled, distance is doubled in the simulator compared to real driving.

The comparison of GW with the KOLIBRI driver information system (KO) reveals that in simulated driving, KOLIBRI can reduce the fast travelled percentage of distance from 40% to 22.7%. On the real track, this effect is still present, but due to the lower speeding behavior, KOLIBRI reduced speeding from 21.4% to 14.9%. A one sided t-test for repeated measurement between REAL_GW and REAL_KO would be significant p=0.02 (df=19; t =2,19).

These findings are in line with former KOLIBRI simulator studies (Krause and Bengler, 2012; Krause et al. 2014). These former papers included time-based speeding percentages without a 5 km/h limit tolerance. We reassessed the data from these experiments with the metric used in this paper and found for the data of Krause and Bengler (2012) a reduction from 48.9% speeding distance to 20.8% with the KOLIBRI assistance. And from 42.5% to 22.3% for Krause et al. (2014a). Thus, the effect is stable and reproducible. It is also worth noting that the first KOLIBRI implementations for the simulator experiment (Krause and Bengler, 2012) did not included a speed alert. So this indicates that the main speed reduction of the KOLIBRI system is not an enforcement by negative feedback (warning) of the speed alert, else a positive compliance of the test subjects to the speed recommendations of the KOLIBRI system. The speed recommendations of the KOLIBRI system are always within the allowed speed limit. From the similarity with former experiments with simulator driving only, we also conclude that the sequence in this experiment (simulator driving mainly after real driving) seemed not to change the speeding behavior.

Table 1 contains the correlation coefficients for the experimental conditions. The highest correlation of all conditions is between the use of the KOLIBRI system in real and simulated driving (0.805). Therefore, the use of the KOLIBRI system in the simulator is a good predictor for the individual speeding behavior with the KOLIBRI system on the road (and vice versa). On the other hand, simulated driving with the driver information system KOLIBRI seems to change the behavior of some subjects, which is a good indication. Thus, this condition has a weak correlation to other simulated scenarios. The simulated scenarios have high inter-correlation, except for the before-mentioned KOLIBRI system.

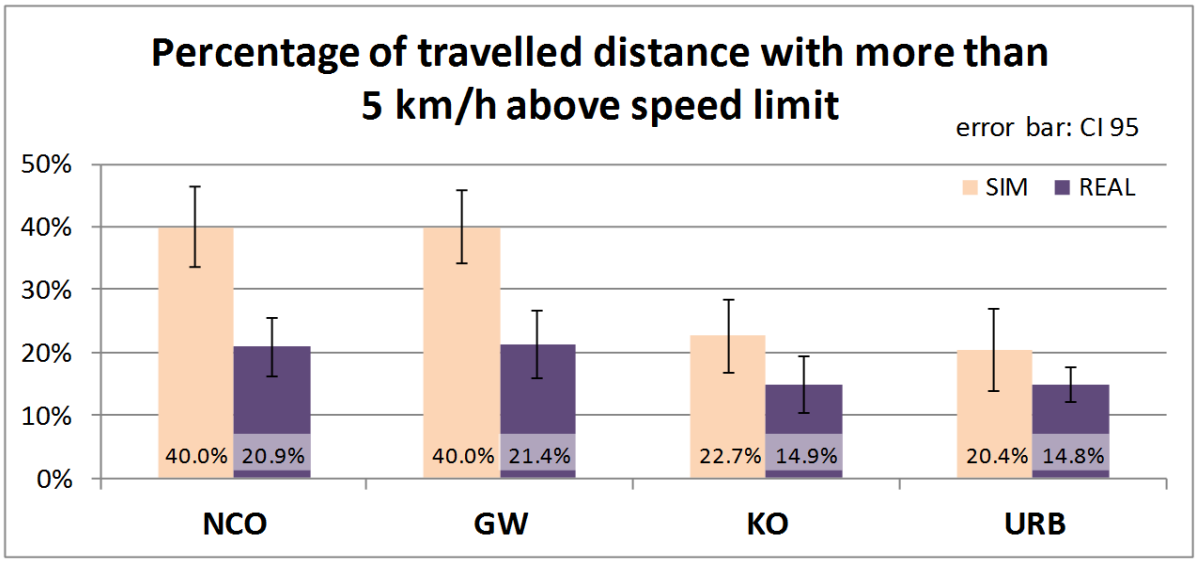


Figure 7 Average percentage (N=20) of travelled distance with more than 5 km/h above speed limit

Another interesting finding is that driving in a real suburban area seems a good predictor of speeding for all assessed conditions in the simulator (and vice versa). Real driving with the KOLIBRI system shows similar values. For the correlations between the real driving conditions, there is only one significant finding: Speeding in suburban areas seems correlated with speeding while using the KOLIBRI system.

The coefficients for SIM_NCO x REAL_NCO, SIM_GW x REAL_GW, SIM_KO x REAL_KO and SIM_URB x REAL_URB have a special meaning. These directly connect real and simulated driving. The values all show a positive correlation, but only two are highly significant. Therefore, it is interpreted, that the simulator setup is well suited to test speeding behavior in slow suburban driving and the influence of driver assistance systems like KOLIBRI on a highway. On the other hand, the simulator setup seems to have reduced capability to get information

for uninfluenced speeding behavior on a highway.

Table 1 Correlation of percentual distance travelled with more than 5 km/h above speed limit

	REAL_NCO	REAL_GW	REAL_KO	REAL_URB	SIM_NCO	SIM_GW	SIM_KO	SIM_URB
REAL_NCO	1.000							
REAL_GW	0.379	1.000						
REAL_KO	0.132	0.210	1.000					
REAL_URB	0.013	-0.061	0.520	1.000				
SIM_NCO	0.314	0.532	0.431	0.489	1.000			
SIM_GW	0.174	0.357	0.469	0.580	0.653	1.000		
SIM_KO	0.186	0.158	0.805	0.516	0.432	0.378	1.000	
SIM_URB	-0.178	0.272	0.490	0.598	0.473	0.700	0.252	1.000

The used speeding metric revealed values between min. 0.04% and max. 45.03% for driving on the real road and min. 2.07% and max. 68.17% in the simulator. Therefore, we obtained values above 0% even for the most defensive driver under each condition.

Driving Behavior – Accelerations

The accelerations from the GPS tracks and driving simulator data are calculated in Figure 8. The calculation used the differential quotient of the GPS velocity (1Hz). The driving simulator data was downsampled to 1Hz and handled the same way. In the figure, all subconditions (regardless of highway or sub-urban) are combined into one histogram. For positive accelerations, the curves show good congruity. For decelerations, the real driving shows a small side peak at around -0.3m/s². This is a typical deceleration for coasting. The side peak around -0.6m/s² for the simulated driving has the same reason (coasting) and the simulated car model would also have a typical coasting value around -0.3m/s², but due to a mechanical or electrical problem with the brake, the simulator applied an additional deceleration of around -0.3m/s². We assessed the data of the former experiments; these did not include the brake problem.

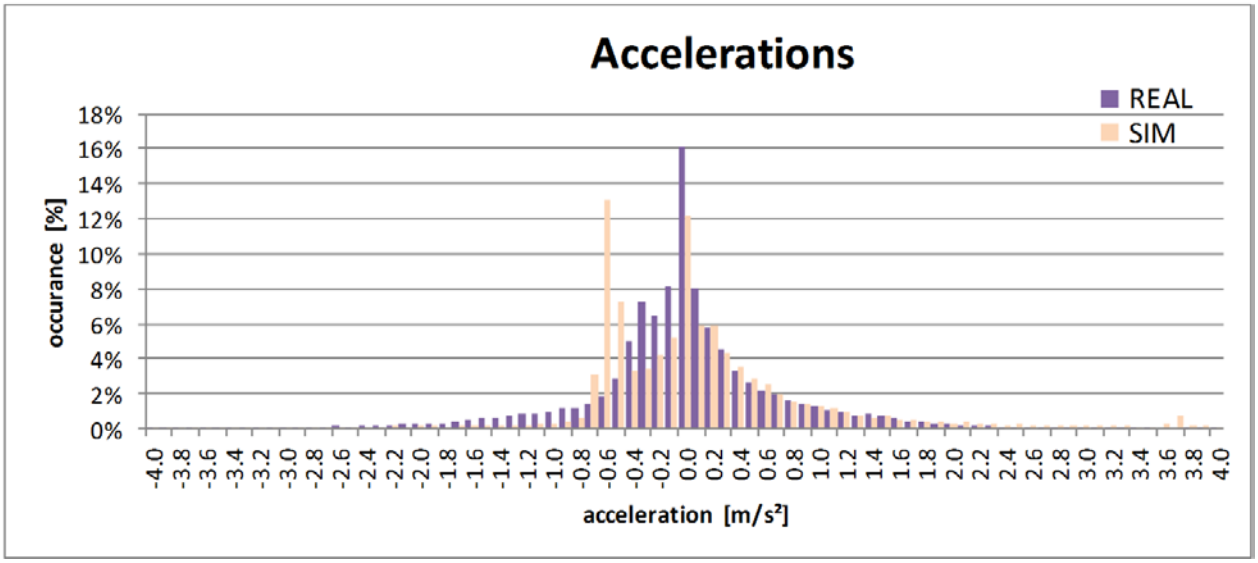


Figure 8 Accelerations from real and simulated driving

Driving Behavior – Fuel Consumption

The individual fuel consumption for the subcondition NCO was set to 100%. Thus, a driver with a fuel consumption of 7 l/100 km in NCO and 5.6 l/100 km in GW would have a consumption of 80%. These individual values are averaged for Figure 9. In the simulator, the coordination of the traffic lights could lower fuel consumption by 12% and the use of the KOLIBRI system on the coordinated route by another six percentage points. This reduction of about 18% is in line with findings in Krause and Bengler (2012), where the use of the KOLIBRI system on an uncoordinated route in the simulator reduced fuel consumption by 20%. The fuel reductions for the coordinated scheme in real traffic is visible in both conditions (GW and KO) with around 7%. A repeated ANOVA for the absolute fuel values from SIM with Bonferroni-corrected comparison revealed significance between all subconditions. The same analysis for REAL showed no significance between GW x KO and even NCO x KO ($p = 0.055$).

The findings from the simulator, i.e. that the KOLIBRI system can lower fuel consumption even on a coordinated green wave road could not be found in real traffic. Table 2 contains the averages and deviations of the absolute fuel values for all subconditions.

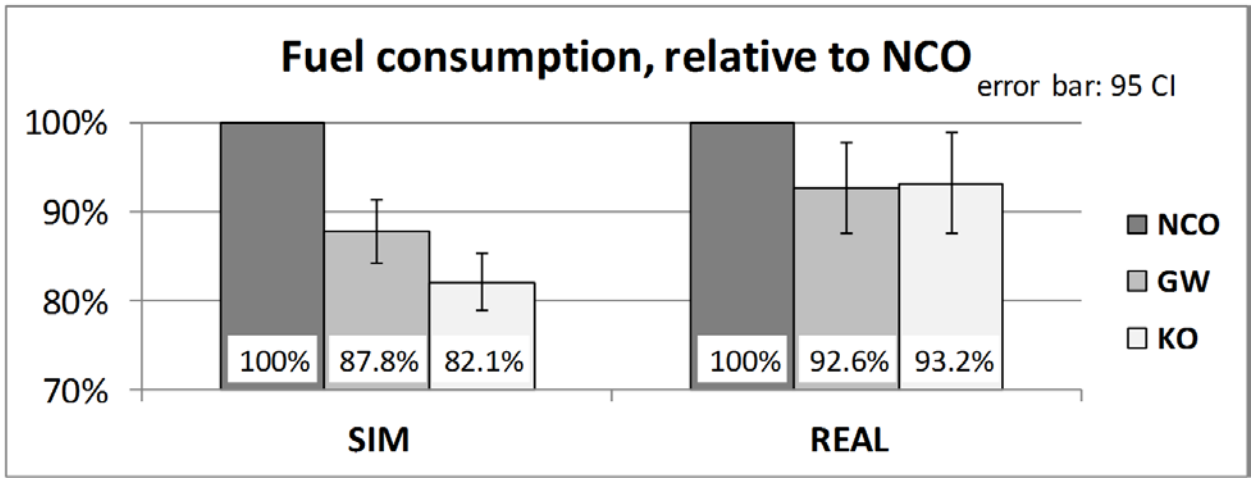


Figure 9 Fuel consumption in relation to the uncoordinated subcondition (NCO)

Table 2 Mean and standard deviation of absolute fuel values (liter/100 km) for the subconditions

	SIM_URB	SIM_NCO	SIM_GW	SIM_KO	REAL_URB	REAL_NCO	REAL_GW	REAL_KO
Mean	9.4	8.6	7.5	7.0	8.4	6.5	5.9	6.0
SD	0.7	0.9	0.8	0.5	0.8	0.9	0.5	0.6

Subjective Ratings

The NASA-RTLX values (score between 0 to 100) are reported in Figure 10. The values in simulated driving are always higher than the scores from real driving. This is in line with most of the findings reported in Blana (1996, p.50). Figure 11 shows at the example of subcondition GW, the NASA-RTLX sub-dimensions. These reveal that most of the difference stems from the subscales “Performance” and “Frustration”. Thus, it seems not very satisfying for subjects to drive in a simulator, compared to real road driving. At first glance, the scores for real road driving KOLIBRI nearly double from around 14 (GW) to 26 (KO). In another experiment, we gathered objective data with a Detection Response Task (Knott, 2013), which showed that KOLIBRI did not induce measurable additional cognitive load. Thus, our interpretation of the NASA-RTLX workload scores is slightly different, as might be expected: Driving on a straight two-lane highway in off-peak hours seems to be encountered as boring, with NASA-RTLX values around 14. These under load values are shifted with KOLIBRI to a value of 26 that is similar to driving in a quiet suburban area. Rather than making the boring situation more interesting by e.g., driving faster, the driver probably directs his available attention at the longitudinal control of the vehicle, motivated by KOLIBRI. This simple “gamification” presumably results in a reduced speeding behavior

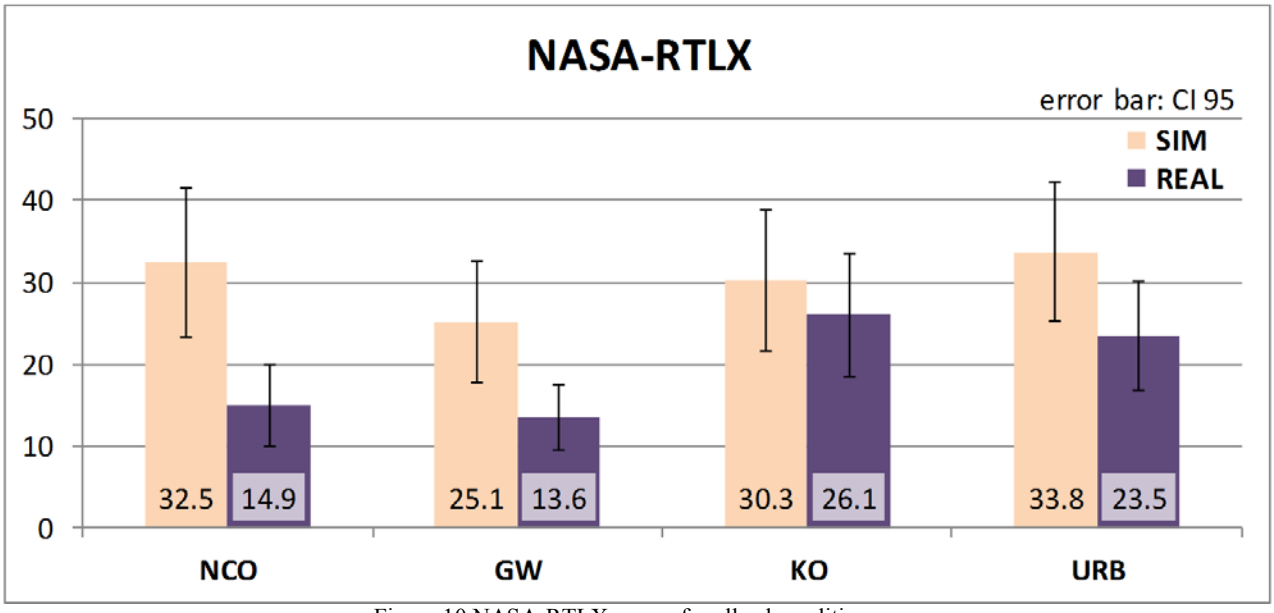


Figure 10 NASA-RTLX scores for all subconditions

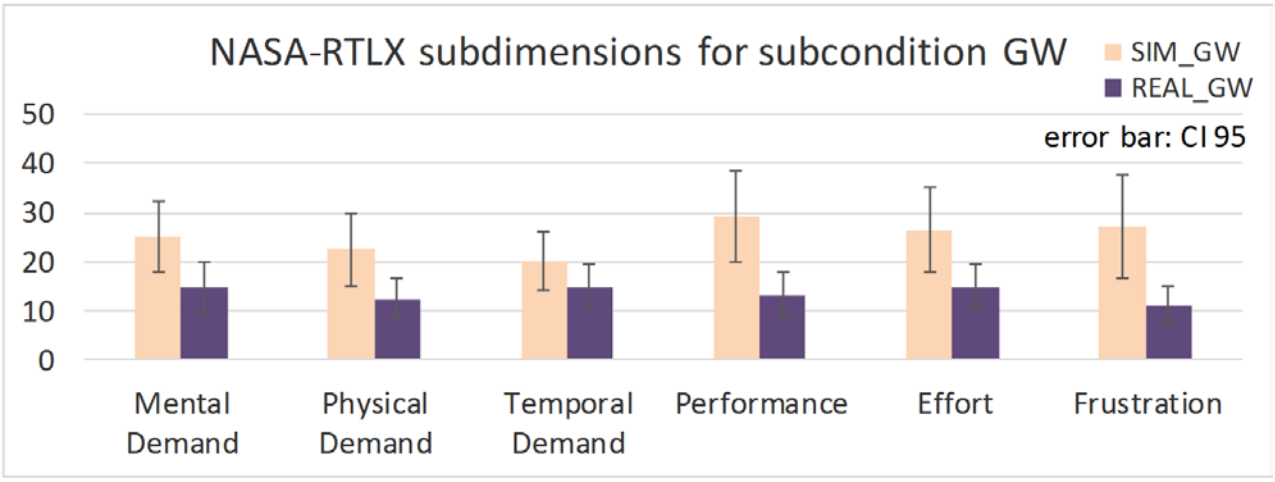


Figure 11 NASA-RTLX subdimension at the example of subcondition GW

DISCUSSION & CONCLUSIONS

It is known that a possible countermeasure of test subjects to compensate for an increased workload by an information system, is to reduce the driving speed. So one might argue that the test subjects with KOLIBRI might tend to drive more within the speed limit (see Figure 7), due to an increased workload by the KOLIBRI system. The subjective ratings (see Figure 10) of an increased workload would also support this hypothesis. We conducted an experiment in real traffic to obtain objective data for the workload with the KOLIBRI assistant with a variant of a Detection Response Task: The Tactile Detection Task (TDT). The objective values from this TDT found no increased mental demand for the test subjects (Knott et al., 2013). Thus, it is likely that the reduced speeding can be attributed to the intended purpose of the KOLIBRI system: Guide the driver to the next traffic light and indicate at an early stage whether the traffic light will be red or green.

It is worth noting that traffic engineers designed the green wave coordination by analyzing the driving behavior of uninformed drivers of light vehicles on the road section. Thus, a kind of human factors engineering. The results show that an informed driver with the KOLIBRI system in real traffic does not seem to have an advantage in terms of fuel consumption (see Figure 9) for a modern fuel-efficient car. However, Figure 7 shows that the informed driver can achieve this result with a more relaxed manner of driving, with less speeding behavior. Nevertheless, we assume that diver information on upcoming traffic light conditions are especially useful for uncoordinated traffic lights.

The authors are surprised by the outcome of the acceleration histogram (Figure 8). The simulator changed its characteristics slightly; maybe due to repair activities, aging of components, or even other reasons. A short checklist provided to the examiners by the simulator specialists would be helpful as a quality-management benchmark. This would enable them to quickly perform a rough check of key characteristics of a simulator. Examples: If car model X is active, the car accelerates from a standstill to 100 km/h in typically X seconds. When the car drives100 km/h and the accelerator pedal is released, the car decelerates to 90 km/h in about X seconds. If the steering wheel is suddenly turned by 90° the car needs X seconds to drive a 360° turn.

Maybe one can even think about a more sophisticated technical procedure. So the simulator is checked periodically for altered characteristics in visual, auditory or anthropometric domain.

Driving in real traffic revealed a lower speeding behavior. We consider regulatory issues to be partially responsible for this: In Europe (75/443/EEC) car speedometers are not allowed to show speeds lower than the real vehicle speed. To account for this, car manufacturers have a tolerance level to show a lower speed than the actual one. Therefore, if the speedometer indicates a speed of 100 km/h it may only be 95 km/h, or even 88 km/h. This “detuning” is not implemented in any of our three static driving simulators. It can be seen as a lack of physical or face validity. Ahlström et al. (2012) accounted for the altered speedometer behavior and used an advanced vibration system for the road surface in the simulator and thus achieved even absolute validity for speed in real and simulated driving.

OUTLOOK

During the KOLIBRI project, we conducted three additional experiments on the real road for human factors engineering, beside the one reported in this paper. One assessed the mental demand by objective measures (Knott et al., 2013), another the gaze behavior in real traffic (Krause et al., 2014b) and the last, how subjects react to wrong information by the traffic lights assistance system. The data of the last experiment are currently under examination and are yet to be published.

One drawback of the experiment reported here was that the display content (speed recommendation) of the smartphone in real driving situations was not logged. So we could not assign behavior (e.g. decelerating, accelerating, speed level) to the display content. Logging was performed in the last real road experiment, which we are currently analyzing. With the data of the last experiment a between subject comparison of driving behavior in real and simulated driving of former experiments regarding the specific display content of the traffic light assistant will be possible.

Most partners of the KOLIBRI project are also trying to organize a follow-up project, to lead the traffic lights assistance over into regular operations, with an operator model.

In the meantime, the institute has decided to use an electric vehicle as staff car. Therefore, it would be tempting, to validate the simulator in the special field of electric driving.

ACKNOWLEDGMENTS

The project was funded by Bayerische Forschungsstiftung (Bavarian Research Foundation). The authors would like to thank the project partners TRANSVER, BMW Group, Board of Building and Public Works in the Bavarian Ministry of the Interior, especially the Bavarian Road Administration, Center for Traffic Management.

REFERENCES

Ahlström, C., Bolling, A., Sörensen, G., Eriksson, O., & Andersson, A. (2012). Validating speed and road surface realism in VTI driving simulator III (VTI rapport No. 745A).

Blana, E. (1996). Driving simulator validation studies: A literature review. ITS Working Paper 480

Council Directive of 26 June 1975 on the approximation of the laws of the Member States relating to the reverse and speedometer equipment of motor vehicles, European Commission.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), Human mental workload. Amsterdam: North Holland Press; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co.

Helmbrecht, M. (2012) Ein vorausschauendes Fahrerassistenzsystem in komplexen Verkehrssituationen. Diploma Thesis at the Institute of Ergonomics, Technische Universität München

Knott, V., Krause, M., & Bengler, K. Einsatz des KOLIBRI-Ampelassistenten im realen Straßenverkehr – Eine kognitive Belastung für den Fahrer? pp.437-440. GfA Press, Dortmund.

Krause, M., & Bengler, K. (2012). Traffic Light Assistant - Driven in a Simulator. In R. Toledo-Moreo, L. M. Bergasa, & M. Á. Sotelo (Eds.), Proceedings of the 2012 International IEEE Intelligent Vehicles Symposium Workshops. IEEE Intelligent Vehicle Symposioum (IV-2012), Madrid, June 3-7, 2012.

Krause, M., Rissel, A., & Bengler, K. (2014a). Traffic Light Assistant - What the User Want: AHCI Conference March 2014, Barcelona. IARIA XPS Press. (in press).

Krause, M., Knott, V., & Bengler, K (2014b). Traffic Light Assistant – Can Take My Eyes Off of You. In HFES Europe Chapter Annual Meeting 2014, Turin . (in review).

Mullen, N., Charlton, J., Devlin, A., & Bedard, M. (2011). Simulator Validity: Behaviors Observed on the Simulator and on the Road, pp 13-1 – 13-18.

Pauwelussen, J., Hoedemaeker, M., & Duivenvoorden, K. (2008). CVIS Cooperative Vehicle-Infrastructure Systems: Deliverable D.DEPN.4.1b Assess user acceptance by small scale driving simulator research.

Seifert, K. (2002). Evaluation multimodaler Computersysteme in frühen Entwicklungsphasen. Ein empirischer Ansatz zur Ableitung von Gestaltungshinweisen für multimodale Computer-Systeme (Dissertation). Technische Universität Berlin, Berlin. Retrieved from http://opus.kobv.de/tuberlin/volltexte/2002/432/pdf/seifert_katharina.pdf

Wooldridge, M. J., Trayford, R. S., & Doughty, B. J. (1984). Urban energy saving through dynamic advisory speeds: The progress of a research project / M.J. Wooldridge, R.S. Trayford, B.J. Doughty: [S.l.] : CSIRO. Division of Energy Technology.

How Do Immersive Driving Environments Influence User Performances and Experiences?

Diana Reich[a], Hanna Jorinde Wittkugel[b] and Rainer Stark[c]

[a]Research training group prometei
Technische Universität Berlin
Berlin, 10587, Germany

[b]Department of Human Factors
Technische Universität Berlin
Berlin, 10587, Germany

[c]Department of Industrial Information Technology
Technische Universität Berlin
Berlin, 10587, Germany

ABSTRACT

Prospective evaluations of human-vehicle-interactions during early prototyping stages are important (Mayhew, 1999) to ensure safety and usability for innovative solutions. To do so, highly realistic appearing test environments will help to provide reliable and valid findings. The high-end version of a realistic test environment is a real car driving study, of course. Nevertheless, they are difficult to control, manipulate and replicate and thus to standardize. They are also more time consuming and expensive. Therefore, one economizing suggestion is the implementation of immersive (driving) environments within simulator studies to provide users a more realistic feeling. This paper discusses research investigating the influence of different levels of immersivity within driving environments. Two important influencing factors were used to examine different levels of immersivity: visual parameters and auditory parameters. Objective data and subjective user impressions were measured and analyzed. Twenty participants took part in the driving simulator study and performed the Lane Change Task within different immersivity conditions. Objective and subjective data have shown advantages for the most immersive driving environment and provide evidence to suggest a more aware and realistic perception of the driving situation. Therefore, higher immersive driving environments are suggested regarding evaluations of prospective human-vehicle-interactions.

Keywords: Immersivity, Driving Task, Presence Experience

INTRODUCTION

Prospective evaluations of human-vehicle-interactions during early prototyping stages are beneficial, important and useful, to ensure safety, usability and customer satisfaction for innovative solutions, e.g. in-car devices (Mayhew, 1999). To do so reliable and valid evaluation methods and verification tools need to be developed to support car

(component) industry. Innovative in-car devices should increase driving comfort, and ensure safety of oneself and other road users. Thus, an ongoing evaluation during the development process regarding safety, usability and driving comfort is necessary. For such dual-task (driving and interacting simultaneously) evaluations a huge quantity of methods are available for developers (see Breuer, Bengler, Heinrich & Reichelt, 2003 for an overview). These methods differ significantly in the amount of time and money. Up to now, dual-task scenarios are evaluated mostly with driving simulator experiments. How big the amount of expenses in terms of time and costs while evaluating in-car devices are, depend on the type of applied driving simulator. Driving simulators differ regarding the grade of reality between each other and compared to real driving situations (Schindler, Kelsch, Heesen, Dziennus, Temme & Baumann, 2013). Driving simulator requirements may differ substantially depending on the research question and experimental set-up. Knappe, Keinath und Meinecke (2006) summarized advantages and disadvantages at the use of driving simulators, which should be taken into account regarding the choice of dual-task evaluation method. First, potentially dangerous driving scenarios can be realized risk-free for the driver. Infrequent and uncommon happenings can easily be replicated by simple configurations of the concerning traffic situations. Additionally, driving simulator experiments enable precise replication of concerning traffic situations as often as one likes. The implementation of more frequent unidirectional and bi-directional traffic volume is easier compared to real car driving experiments. Moreover, limiting factors of influence during the test drive, like weather changes or differences regarding lighting conditions because of different daytimes, don`t need to be taken into consideration while designing the experiment. All experimental simulator test drives are characterized by high efficiency and comparability. Additionally on top of that, an often used basis of assessment for in-car devices are objective measured variables, like number of lane deviations. Data collection doesn`t require such complicated sensors compared to real car driving experiments and for that reason data recording is more robust with simulator experiments.

Unfortunately, driving simulator lack regarding user's perceived reality and presence experience. Thus, reliability and validity of the findings are reduced. Various factors are known to affect drivers' workload (Mehler, Reimer & Zec, 2012). One of these factors is the context in which the driver is operating the vehicle, e.g. drivers might feel more stressed during heavy rain or snow on a jammed highway than on a sunny day driving along an empty road. But it is difficult to represent these risky situations to the driver and make them feel present and aware with the potential hazard. For example, Schneegass, Pfleging, Broy, Heinrich and Schmidt (2013) "(…) believe that workload cannot easily be assessed in a simulator study". Indeed, with immersive driving environments the presentation and experience of risky scenarios and bad conditions of visibility when foggy, raining or snowing could be improved (Shahrokhi & Bernard 2004). Immersion is a metaphoric use of the experience of submersion applied to representation, or simulation. The success of immersing the user by presenting an (driving) environment can actually immerse the user is dependent on many factors, e.g. visibility, surround sound and interactive user-input (Nechvatal, 2009).

Furthermore, although many psychological studies have been focused on dual-task scenarios (e.g. Levy & Pashler, 2001; Meyer & Kieras, 1997a, 1997b), there are only few systematical investigations of realistic dual-task scenarios (Salvucci & Taatgen, 2008; Taatgen, 2005). Salvucci (2005) refers to the disregard of the relation to reality and emphasizes particularly ecological validity. That implies more realistic appearing driving environments will help to provide more reliable and valid findings.

Of course, the high-end version of a realistic test environment is a real car driving experiment. Nevertheless, if you are able to conduct one, the early prototyping stage would have been completed already. Real car driving experiments are also difficult to control, manipulate and replicate and thus to standardize. Moreover, by conducting real car driving experiments developers and researchers are strictly limited in (1) the way of manipulating physical characteristics of the vehicle (e.g. car cockpit, position of the monitor) and (2) the number of achievable stages of expressions (e.g. different vehicle types). Finally, if the vehicle or vehicle parts need to be rebuilt or new-built, real car driving experiments end up a lot more time consuming and expensive.

Therefore, one economizing suggestion is the implementation of immersive (driving) environments for early prototyping evaluations, to provide users a more realistic experience. The aim of this research is to combine advantages of driving simulator experiments (e.g. repeatability) with positive effects of immersivity (e.g. presence experience). To create a sense of full immersion, the five senses (sight, sound, touch, smell, taste) must perceive the digital environment to be physically real (Nechvatal, 2009). Immersive technology can perceptually fool the senses through: Panoramic 3D displays (visual), surround sound acoustics (auditory), haptic and force feedback (tactile), smell replication (olfactory), and taste replication (gustation). Due to highly realistic appearing test environments

more reliable and valid findings are expected. Moreno and Mayer (2002) postulated, that the more immersive an (driving) environment is, the more realistic perceived by users, because of higher presence experience. Through increasing immersivity more realistic driving simulator experiments could be implemented and due to that fact presence experience can be improved. The level of immersivity is modifiable by manipulating the vehicle or mock-up themselves (e.g. geometry, surfaces, brightness), the driving environment (e.g. representation of real-existing streets) and/or the interaction space between human and vehicle (e.g. sensomotoric in- and output). Any desired type of vehicle and driving scenario is implementable through 3D-visualization software very fast by simple drag`n drop features. Additionally, this immersive evaluation approach enables more complex experimental designs and is characterized by high controllability and changeability. According to that, there is no difficulty to implement multi-factorial experimental designs with several repeated measurement factors (Weber & Wetzel, 2013).

This paper discusses research investigating the influence of different levels of immersivity within a standardized driving environment. Two important influencing factors were used to examine different levels of immersivity: Visual parameters (2D vs. 3D) and auditory parameters (no sound vs. car sound). It is assumed, that these types of immersivity cues can make a significant difference in perception of the environment and presence experience (Nechvatal, 2009). Objective data and subjective user impressions were measured and analyzed.

METHOD

Design

The experiment was realized as completely crossed 2x2 within-subjects design to investigate the influence of two immersivity parameters (visual and auditory). As independent variables the visual representation of the driving environment (2D vs. 3D) was adapted on the one hand, on the other hand the auditory representation of driving sound (no sound vs. car sound) was varied. The sample was divided into four groups in order to process the driving task under randomized conditions. Each group started with different conditions: (A) 2D; no sound, (B) 2D; car sound, (C) 3D; no sound and (D) 3D; car sound. Therefore, condition (A) represents a very *low immersive* condition without any immersivity parameter and condition (D) represents a very *high immersive* condition with two immersivity parameters. Each group had to pass all four conditions and answer all questionnaires. Objective and subjective data were captured as dependent variables. The following objective data were collected: Driving parameters of the LCT for determining driving performance for each condition, eye-tracking data and results of the SAGAT questions for measuring situational awareness. Subjective dependent variables were assessed via different questionnaires. These constitute subjective workload which was raised on the NASA-TLX, perceived driving performance (PDP) measured by the questionnaire for self-assessment of driving performance, and the subjects' perceived quality of experience of virtual reality, which was assessed by the questionnaire for Presence and Immersive tendency (PIT) in virtual realities.

Participants

The sample was composed of 20 participants (female = 8; male = 12) with an average age of 28.65 years (SD = 5.95). Approximately three-quarters of the subjects were students (N = 16), the remaining four were employed. Three participants were left-handed. One person wore glasses, another wore contact lenses. More than half of the participants (N = 12) were already familiar with the use of driving simulators ("little" = 3, "medium" = 3, "a lot" = 2, "very much" = 4) and eight participants had no experiences with driving simulators. Seven participants had experiences with eye-tracking systems („medium" = 2, "a lot" = 3, "very much" = 2), the remaining 13 participants had never used an eye-tracker before. The experiment took about 75 minutes. The subjects participated voluntarily and received an expense allowance of ten euros. A valid driver's license was required.

User Tasks

The primary user task was driving and performing the Daimler Chrysler Lane Change Task (LCT; Mattes, 2003). It is a low-cost and highly validated measure for driver distraction caused by secondary tasks. It was developed by vehicle manufacturers and represents a standardized and ISO-normed tool for detecting driver distraction (ISO, 2008). The LCT comprises driving simulation software and analysis software. The task consists in a sequence of

lane change maneuvers while driving with a fixed speed of 60 km/h on a three-lane road. The lane change is displayed on signs on both sides of the road. Once the subject recognizes the instruction on the respective sign in front the appropriate lane change must be executed fast and efficiently to keep deviations from the optimal driving lane at a minimum. One task takes about three minutes. The parameter *mean deviation* was extracted from the LCT to calculate the deviation of the subjects' driven lane from the lane given as a reference within the analysis software (the normative model) in meters over the whole track. To perform the LCT simple hardware (PC and game steering wheel) is sufficient. In this study it was controlled via Logitech Driving Force GT force feedback wheel system which included a steering wheel and gas and brake pedal. This was connected to the driving simulator FESTO Airmotion ride. The simulator was merely used as an interface for steering as well as accelerating and braking. (The latter were not part of the driving task but served to interrupt the ride while answering the SAGAT questions). Speed regulation or changing gears was not necessary due to the predetermined constant speed of 60 km/h. The LCT was displayed on a 1,80 m x 1,10 m holobench in front of the driving simulator (see Figure 1).

Figure. 1. Driving simulator and holobench at the VRSC

The situation awareness task was measured by SAGAT (Situation Awareness Global Assessment Technique, Endsley, 1995a) during each LCT condition. The participant was asked to stop once per run at varying times. The screen (driving environment) was turned off and the participant was asked to state the actual lane (1) as well as the previously driven lane (2). Additionally, the participant was asked for a number that was displayed in the right top of the driving environment before (3). After answering the questions the screen was turned back on and the subject completed the LCT. The answers could only be right or wrong, which determined the corresponding values of 1 or 0 as results. To indicate the level of situation awareness the values were added. A higher number represents a higher number of wrong answers and thus less situational awareness. For analysis the number of wrong answers was counted (0-2).

Requirements

View and Sound

The driving environment was presented on the Barco TAN Holobench (see Figure 1), a projection table consisting of two orthogonal projection surfaces (each 1,80 m x 1,10 m). For this study, only the vertical screen was used to display the driving environment (three-lane road, signs, background). An active shutter 3D technology allows a stereoscopic (3D) representation of the scene. A projector (controlled by the rendering client) alternately displays the image for the left and the right eye using a video refresh rate of 100 Hz. Shutter glasses worn by the subject synchronously block the respective eye by polarization of the Liquid Christal Display (LCD) glasses. Consequently the image displayed is synchronized with left and right eye image, similarly to the screen with a frequency of 100 Hz. Thus, each eye perceives 50 frames per second, which creates a 3D view for the subject by merging the two images. The synchronization is generated by the computer rendering the LCT. The signal is transmitted to the shutter glasses via infrared transmitter. The change between monoscopic (2D) and stereoscopic (3D) view can be evoked by switching the holobench to active stereo mode. Therefore, shutter glasses need to be turned on/turned off simultaneously.

Car sound was the second parameter, besides 3D view, to increase immersivity of the driving environment. The

driving sound was generated automatically by starting the LCT and was reproduced by speakers connected to the LCT computer and holobench. Switching between car sound/ no car sound conditions was realized by turning on and off the speakers.

Eye-tracking/ Shutter glasses

Visual fixations and eye movements were recorded by Ergoneers' Dikablis eye-tracking system. Within this study participants had to wear LCD shutter glasses to enable stereoscopic (3D) view. In order to realize the gaze data recording and stereoscopic view at the same time, the participants wore a self-built combination of shutter glasses and eye-tracking system (see Figure 2). Therefore, the camera of the eye-tracking system was attached to the shutter glasses to record eye and pupil movements. For recording the field of vision the video signal from the holobench was used. Therefore, graphic card output of the holobench was connected via an Ergoneers adapter to the Ergoneers laptop. The adapter was specifically designed for this experiment. The video signal provided an image of the depicted scene. The software mapped the eye movement directly to the scene.

A calibration of the eye-tracking system was necessary before each run. Therefore, camera position was adapted by the experimenter and pupil detection was configured using the Dikablis software. Subsequently, within the calibration process the mapping of eye camera images and the image of the scene was executed. The participants were asked to look at each of the four corners on the screen (holobench). The experimenter confirmed every point by clicking on the referenced point within the image of the scene on the PC. The mapping between the image of the eye camera and the image of the scene was adopted by the system. For each run recording of eye-tracking had to be started and stopped manually. From recorded visual fixations and eye movements the parameters *duration* and *number of glances* regarding three areas of interest (AOI). By means of the trigger function three AOIs were defined: (1) road, (2) sign on the left and (3) sign on the right. The glances on each AOI were analyzed by the software Dikablis Analysis. Each analyzed sequence started after reaching the start sign of the LCT and ended after passing the last sign. The break to answer the SAGAT questions was not included. Glance durations and frequencies for the three AOIs were recorded and evaluated.

Figure 2. Self-built combination of eye-tracking system and shutter glasses

Questionnaires

The German questionnaire for presence and immersive tendency in virtual realities (PIT; by Scheuchenpflug, 2001) was used to analyze differences regarding user experiences within different immersive environments. It is based on the English Presence Questionnaire (PQ) which measures the presence of subjective experience, as well as the Immersion Tendency Questionnaire (ITQ) for the evaluation of general willingness to empathize with synthetic and virtual environments (Witmer and Singer, 1998). Overall, the immersion experience is determined by means of five dimensions. Three dimensions of subjective presence experience are captured: (1) Spatial presence (21 items), (2) quality of the interface (11 items) and (3) involvement (6 items). Two further dimensions measure the general immersive tendency, i.e. how strongly the subject generally feels involved in the situation: (4) emotional involvement (6 items) and (5) degree of involvement (6 items). Each of the 50 items is ordinally scaled as a Likert scale from 1 ("very much / very good / very real / ... ") to 7 ("very little / very poor / very unreal / ..."). For analysis

means of the respective items were summarized for each dimension. The average was calculated. Some items had to be inverted.

For the assessment of subjective workload the NASA Task Load Index (NASA-TLX) was used (Hart & Staveland, 1988). The standardized questionnaire represents a subjective measure for mental workload regarding the evaluation of human-machine systems. The multi-dimensional rating procedure is based on six subscales. The following dimensions are assessed: (1) Mental demand, (2) physical demand, (3) temporal demand (perceived time pressure), (4) task performance (satisfaction with level of task performance, i.e. tracking, speed and spacing), (5) effort (overall mental and physical strain while driving) and (6) stress (perceived stress while driving). The items were ratio scaled on a continuous scale from "low" to "high" (0 to 100).

The self-assessment questionnaire for perceived driving performance (PDP) investigates how the participants estimate their own performance in the driving task within the respective level of immersivity. It comprises eight items: (1) Tracking, (2) maintaining speed, (3) spacing, (4) concentration, (5) overall judgment of maintaining speed, distance and tracking, (6) subjective sense of security, (7) relaxation and (8) general impression. The items are interval scaled ranging from "very good" (5) to "very poor" (0). For the evaluation the average was calculated using all scores. The higher the value, the better the self-estimated driving performance. Values can range between 0 (worst self-estimated driving performance) and 40 (best self-estimated driving performance).

The Simulator Sickness Questionnaire (SSQ; Kennedy, Lane, Berbaum & Lilienthal, 1993) captures the subjects' current condition as a checklist before and after the experiment. Therefore 16 symptoms of nausea and dizziness were queried: General unease/ discomfort, fatigue, headache, eyestrain, difficulties concerning visual acuity, increased salivation, transpiration, sickness/ emesis, difficulties in concentrating, feeling of pressure in the head, blurred vision, dizziness (with eyes open) dizziness (with eyes closed), balance disturbances, upset stomach, belching. The subjective condition was entered on an ordinal scale as a four-point Likert scale from (1) "not at all" to (4) "severe". For the evaluation the scores of all items were added. The higher the value, the worse the condition of the subject. The lowest value of 16 (each item scores one point) represents the best condition and the value of 46 represents the worst condition (4 points for each item).

Procedure

The study was performed in the VRSC (Virtual Reality Solution Center) at the Fraunhofer IPK Berlin. After the welcome by the experimenter and the registering of demographic data by means of a demographic questionnaire, the current condition of the subject was determined via SSQ. It was followed by a test drive in a 2D environment without shutter glasses or eye-tracking to make the subject feel comfortable with the LCT, the test environment and driving simulator. After the test drive the calibration of the eye-tracking system was executed. Thereafter, the participants had to pass the LCT under randomized driving conditions (levels of immersivity). The eye-tracking system was recalibrated before each run. In order to assess the situation awareness each run was interrupted once for SAGAT. After completion of each run (condition) the questionnaires NASA-TLX, questionnaire for self-assessment of driving performance and the questionnaire for presence and immersive tendency in virtual realities were completed by the participants. After the last run the subjects' condition was re-interrogated by SSQ. After receiving the expense allowance the participants were released.

RESULTS

In the following the results of the study are presented. Data from the four experimental conditions were analyzed regarding dependent variables. To investigate statistically significant differences weighted scores were compared across conditions using a within-subjects repeated measures analysis of variances (ANOVA) with an alpha level of .05. For the pairwise comparison a Bonferroni correction with SPSS-adjusted significance level of .05 and corresponding paired t-tests were used. Afterwards, relationships between dependent variables were examined using Pearson`s correlation coefficient. IBM SPSS v21.0 was used in the statistical analysis. Every assumption for an ANOVA was met in the reported data.

Driving Performance

The scores of different driving parameters of the LCT were analyzed and compared across conditions using a within-subjects repeated measure ANOVA with a Greenhouse-Geisser correction.

No significant differences for driving performance were found. The four conditions did not differ significantly from each other. Indeed, lowest *lane deviation* was found for the most immersive condition (M=1.21; SD=.39) and highest *lane deviation* for the least immersive condition, 2D without sound (M=1.29; SD=.35).

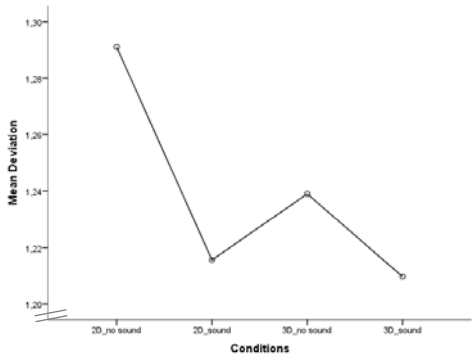


Figure 3. Driving performance LCT by immersive conditions

Eye-tracking Data

Unfortunately, due to recording failures and implementation issues eye-tracking data from 10 participants had to be removed later. For the remaining 10 participants the overall mean times in seconds and the number of glances on the road and signs (right, left) for each of the conditions were compared across conditions using within-subjects repeated measures ANOVA with Greenhouse-Geisser corrections. No significant differences for eye-tracking data were found, though.

SAGAT

The scores of the SAGAT data were compared across conditions using a within-subjects repeated measures ANOVA with a Greenhouse-Geisser correction. No significant differences were found among all conditions.

Presence and Immersive Tendency

The scores of the PIT questionnaire within virtual environments were compared across conditions using within-subjects repeated measures ANOVA with Greenhouse-Geisser corrections. The questionnaire consists of three dimensions *(a) spatial presence, (b) quality of the interface and (c) involvement.* Figure 4 shows differences between the immersive conditions regarding the three dimensions of the presence and immersivity questionnaire.

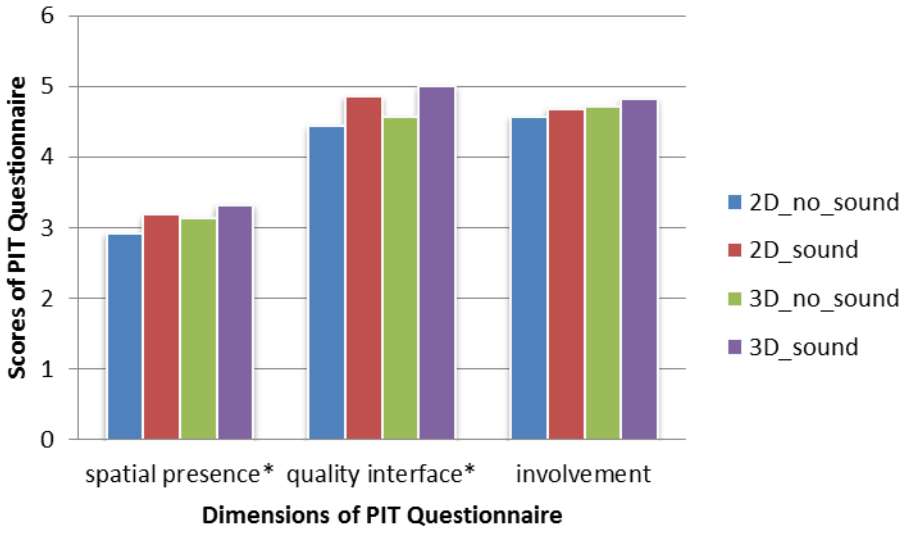


Figure 4. PIT scores by immersive conditions

The ANOVA revealed a statistically significant difference in ratings of *spatial presence* among the four immersive conditions, $F(3, 57) = 3.13$, $p = < .05$, $n^2part = .141$. Significant pairwise comparisons regarding *spatial presence* were found between the two extreme immersive conditions (2D; no sound vs. 3D; car sound), $F(1, 19) = 5.71$, $p = < .05$, $n^2part = .231$.

Corresponding planned t-tests showed that spatial presence was rated significantly minor within low immersive condition, 2D no sound (M=2.92; SD=.84) compared to the high immersive condition, 3D car sound (M=3.32; SD=.83), $t(19) = -2.39$, $p = .027$.

The ANOVA revealed also a statistically significant difference in ratings of *quality of interface* among the four immersive conditions, $F(3, 57) = 4.17$, $p = < .05$, $n^2part = .180$. Significant pairwise comparisons regarding *quality of interface* were found between the two extreme immersive conditions (2D; no sound vs. 3D; car sound), $F(1, 19) = 7.63$, $p = < .05$, $n^2part = .287$.

Corresponding planned t-tests showed that *quality of interface* was rated significantly minor within low immersive condition, 2D no sound (M=4.44; SD=.71) compared to the high immersive condition, 3D car sound (M=5.0; SD=.66), $t(19) = -2.90$, $p = .009$.

No significant differences were present in the *involvement* data. Indeed, Figure 4 shows that the overall involvement rating for each immersive condition showed the highest scores for the most immersive condition (M=4.82; SD=.93) and the lowest scores for the least immersive one (M=4.57; SD=.86).

Cognitive Workload

Weighted NASA-TLX scores were compared across conditions using a within-subjects repeated measures ANOVA with a with Greenhouse-Geisser correction. Figure 5 shows differences between the immersive conditions regarding the six dimensions of the NASA-TLX questionnaire.

The ANOVA revealed a statistically significant difference in ratings of *temporal demands* among the four immersive conditions, $F(3, 57) = 4.07$, $p = < .05$, $n^2part = .176$. Figure 5 shows highest scores for the least immersive condition, 2D no sound (M=36.75; SD=27.35), compared to 2D with car sound (M=26.5; SD=21.34), 3D no sound (M=20.00; SD=21.70) and 3D with car sound (M=26.5; SD=26.71). Significant pairwise comparisons regarding *temporal demands* were found for 2D; no sound and 3D; car sound, $F(1, 19) = 5.41$, $p = < .05$, $n^2part = .222$.

No other significant differences were present in the workload data. Indeed, the overall rating showed highest total workload scores for the low immersive condition (M=33.17; SD=18.62) and lowest scores for the high immersive

condition (M=32.08; SD=18.50).

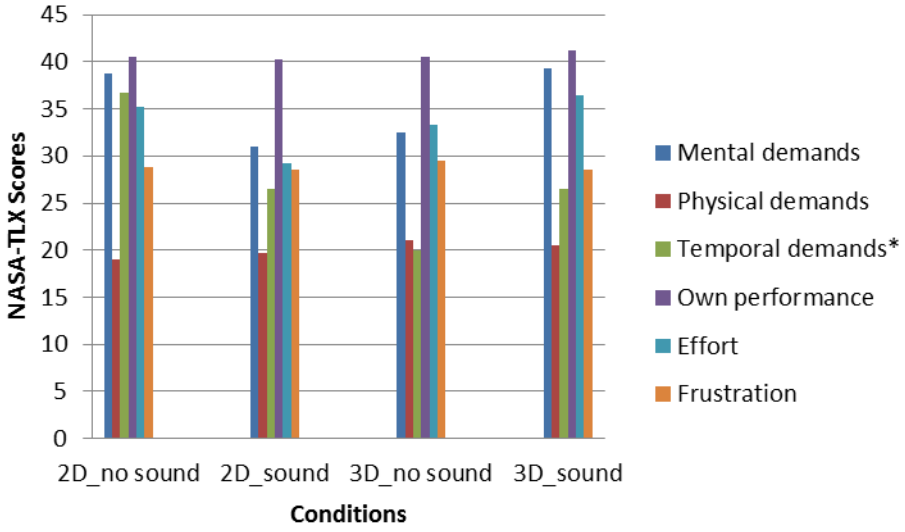


Figure 5. NASA-TLX scores by immersive conditions

Perceived Driving Performance

The scores of the self-assessment questionnaire PDP were compared across conditions using a within-subjects repeated measures ANOVA with a Greenhouse-Geisser correction. No significant differences were found among all conditions. Indeed, the overall rating of perceived driving performance showed highest scores for the most immersive condition (M=5.03; SD=.67) and lowest scores for the least immersive condition (M=4.82; SD=.86).

Simulator Sickness Questionnaire

The scores of the SSQ were compared using a within-subjects repeated measures ANOVA with Greenhouse-Geisser corrections. The ANOVA revealed a statistically significant difference in the pre-post comparison, $F(1, 19) = 5.71$, $p = < .05$, $n^2part = .231$. Participants` ratings of physical condition on the four-point Likert-like scale were higher after the driving experiment (M=21.0; SD=5.59) compared to their arrival (M=19.0; SD=3.73)

Relationship of Dependent Variables

In an attempt to investigate underlying relationships in the study, a few post hoc correlations test were performed. When investigating the relationship between driving performance and subjective impressions, a significant negative correlation was found for NASA-TLX and lane deviation within immersive driving environments (2D; car sound), $r = -.54$, $p < .05$ and condition 3D; no sound, $r = -.52$, $p < .001$. It can be assumed that there is a relationship between cognitive workload and driving performance. Although, it doesn`t mean that the higher spatial presence the more demanding, because the most and least immersive condition did not reveal a positive correlation.

Another negative correlation was found between NASA-TLX and PDP for each of the four immersive driving conditions. See Table 1 for exact results. These correlations suggest that as the better the own driving performance perceived the less demanding the driving task.

A similar positive correlation was found between NASA-TLX and spatial presence (PIT) for each immersive condition, except the most immersive one. The results are listed in Table 1. It can be assumed that there is a relationship between cognitive workload and increasing spatial presence. Although, it doesn`t mean that the higher spatial presence the more demanding, because the most immersive condition didn`t reveal a positive correlation. No other significant correlations were found.

Table 1: Correlations between subjective data

	NASA-TLX_2D_no_sound	NASA-TLX_2D_sound	NASA-TLX_3D_no_sound	NASA-TLX_3D_sound
PDP_2D_no_sound	-.64**			
PDP_2D_sound		-.67**		
PDP_3D_no_sound			-.58**	
PDP_3D_sound				-.80**
PIT_2D_no_sound	.49*			
PIT_2D_sound		.51*		
PIT_3D_no_sound			.52*	
PIT_3D_sound				n.s.

Notes. *p < .05; **p < .001; PDP=perceived driving performance; PIT=presence and immersive tendency

DISCUSSION

In this paper, we investigated the influence of different levels of immersivity within a standardized driving task environment. Therefore, visual parameters (2D vs. 3D) and auditory parameters (no sound vs. car sound) were used to examine different levels of immersivity. These types of parameters can make a significant difference in perception of the (driving) environment and presence experience. Objective data and subjective user impressions were measured and analyzed.

The results showed no significant differences in driving performance (LCT) between the four immersive conditions. But it`s a matter of common knowledge, that it is sometimes difficult to find (significant) differences with the LCT, particularly without a secondary task, even if the LCT has a high internal validity compared to more realistic driving simulations. Although, no significant differences for driving performance were found, overall lowest lane deviation was measured for the most immersive condition (3D; car sound) and highest lane deviation for the least immersive one (2D; no sound). That could indicate that highly immersive driving conditions influence driving performance positively. Furthermore, second best performance regarding lane deviation was found for condition 2D with car sound. That could indicate that the factor sound exerts higher influence on driving performance than 3D view. Additionally, when investigating the relationship between driving performance and subjective impressions, a significant negative correlation was found for NASA-TLX and lane deviation within high immersive driving environment (3D; car sound). This reflects a decrease in perceived cognitive workload when driving within more immersive environments. Immersivity decreased lane deviation (not significant) and cognitive workload (significant) and therefore supports the implementation of 3D view and car sound regarding evaluations of human-vehicle-interaction. Although this study did not find significant differences between conditions and driving performance, future studies may find these results with more participants and a secondary task.

Unfortunately, due to recording failures and implementation issues eye-tracking data from 10 participants had to be removed later. For the remaining 10 participants no significant differences for eye-tracking data were found, though. Bugs and inaccuracies probably occurred, because of self-built shutter/eye-tracking glasses. Lack of eye-tracking recording was a limitation of this study that should be considered in future research with more professional equipment.

The results showed significant differences regarding PIT questionnaire between the four immersive conditions. Two of the three dimensions (spatial presence, quality of the interface, involvement) revealed a statistically significant difference among the four immersive conditions. *Spatial Presence* and *quality of the interface* were constantly higher rated for the most immersive condition and lowest for the least immersive one. Although, the dimension *involvement* did not show a significant difference the overall rating was the same, too. It corroborates our expectation, the more immersive the test environment, the higher the users' presence experience. Immersivity increased presence and immersive tendency and supports the implementation of 3D view and car sound regarding

evaluations of human-vehicle-interaction, too. In the end, this result is the most important finding of this experiment, because higher presence experience leads to more reliable and valid findings in the context of human-vehicle-interaction evaluations.

The data of the 20 participants showed a significant difference between the four immersive conditions regarding the NASA-TLX questionnaire in ratings of *temporal demands*. Highest scores were found for the least immersive condition (2D; no sound) and lowest for the most immersive one (3D; car sound). It indicates, the more immersive the driving environment, the less the temporal demand perceived while driving. No other significant differences regarding workload dimensions were present in the data. Indeed, the overall rating showed highest total workload scores for the low immersive condition and lowest workload scores for the high immersive condition. Further research is necessary to examine these effects while driving and interacting with in-car devices. Furthermore, a positive correlation was found between NASA-TLX and spatial presence (PIT) for each immersive condition, except the most immersive one. It can be assumed that there is a relationship between cognitive workload and increasing spatial presence. Although, it does not mean the higher spatial presence, the more demanding, because the most immersive condition did not reveal a positive correlation. To sum up, immersivity decreased subjective workload (temporal demand) and therefore supports the implementation of 3D view and car sound regarding evaluations of human-vehicle-interaction. Future studies may find additional significant results regarding other workload dimensions with more participants and a secondary task.

Indeed, no significant differences were found for the self-assessment questionnaire PDP, the overall rating of perceived driving performance showed highest scores for the most immersive condition (3D; car sound) and lowest scores for the least immersive condition (2D; no sound), too. Furthermore, investigating the relationship between NASA-TLX and perceived driving performance a negative correlation was found for each of the four immersive driving conditions. These correlations suggest that the better the own perceived driving performance the less demanding the driving task. Further research is necessary to examine these effects while driving and interacting with in-car devices. Future studies may find significant results with more participants and a secondary task.

Regarding SAGAT data no statistical differences were found, perhaps due to realization issues. Because of the current methodical procedure, without presenting further distracting stimuli or secondary tasks, the SAGAT questions were probably too easy to answer. Lack of significance in this experiment may have also been due to the low number of questions that the participants had to answer regarding their actual situational awareness. Future studies may find significant differences regarding situation awareness between different immersive conditions with more situational questions and a secondary task.

CONCLUSIONS

The outcomes lead us to conclude that high immersive driving environments (3D; car sound) increase impressions of reality for driving tasks compared to conventional driving simulator environments. Both objective and subjective data provide evidence to suggest a more aware and realistic perception of the driving situation. Therefore, higher immersive driving environments are suggested regarding prospective human-vehicle-interaction evaluations.

Future research will also investigate different secondary tasks and only two immersive conditions. The results of this study suggest that a *low immersive* condition is characterized by 2D view without car sound and a *high immersive* condition by 3D view with car sound. Although both immersive parameters are necessary to characterize a low or high immersive condition, it can be assumed that the factor sound exerts higher influence on perceived immersivity than 3D view. One possible explanation could be that even if 3D view is turned off participants are still able to *see* the driving environment, while if sound is turned off participants can´t *hear* anything. This means with the current experiment visual parameters are only improved whereas sound parameters vary effectively between on and off. Thus, with only two immersive conditions (low and high) a longer time could be spent driving with different secondary tasks and the effect of practice could be investigated.

Further research is also necessary due to other immersive influencing factors, such as haptic or tactile feedback, which may also affect presence experience and immersivity positively.

REFERENCES

Breuer, J., Bengler, K., Heinrich, C. & Reichelt, W. (2003). Development of advanced driver attention metrics (ADAM). In H. Strasser, K. Kluth, H. Rausch, & H. Bubb (Eds.). Quality of Work and Products in Enterprises of the Future. *Proceedings of the 50th – Anniversary Conference of the GfA and the XVII Annual ISOES Conference in Munich, 2003, pp. 37-40*. Stuttgart: Ergonomia.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37(1), 65-84*.

Hart, S. & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. Hancock & N. Meshkati (Eds.), *Human mental workload, pp. 139-183.* Amsterdam: North Holland.

ISO, (2008). *Road vehicles – Ergonomic aspects of transport information and control systems - Simulated lane change test to assess driver distraction.* ISO/TC 22/SC 13 N WG8 N579, DIS 26022.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire (SSQ): An Enhanced Method for Quantifying Simulator Sickness. *International Journal of Aviation Psychology, 3 (3), 203-220*.

Knappe, G., Keinath, A. & Meinecke, C. (2006). Empfehlungen für die Bestimmung der Spurhaltegüte im Kontext der Fahrsimulation. In M. R.K. Baumann, S. Leuchter, & L. Urbas (Hrsg.). *MMI-Interaktiv, Aufmerksamkeit und Situation Awareness im Fahrzeug, 11, 2006,* ISSN 1439-7854, 3-13.

Levy, J., & Pashler, H. (2001). Is dual-task slowing instruction dependent? *Journal of Experimental Psychology: Human Perception and Performance, 27, 4, 862-869.*

Mattes, S. (2003). *The lane change task as a tool for driver distraction evaluation*. In H. Strasser, H. Rausch & H. Bubb (Eds.), Quality of work and products in enterprises of the future. Stuttgart: Ergonomia Verlag.

Mayhew, D. J. (1999). *The Usability Engineering Lifecycle*. San Francisco: Morgan-Kaufmann.

Mehler, B., Reimer, B. & Zec, M. (2012). Defining workload in the context of driver state detection and hmi evaluation. *In Proc. AutomotiveUI`12, pages 187-191,* New York, NY, USA, 2012. ACM.

Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104, 3-65.*

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory period phenomena. *Psychological Review, 104, 749-791.*

Moreno, R. & Mayer, R.E. (2002). Learning Science in Virtual Reality Multimedia Environments: Role of Methods and Media. *Journal of Educational Psychology, 94(3), 598-610.*

Nechvatal, J. (2009). *Immersive Ideals / Critical Distances.*LAP Lambert Academic Publishing. 2009, pp. 48-60.

Salvucci, D. D. (2005) A multitaksing general executive for compound continuous tasks. *Cognitive Science, 29, 457-492.*

Salvucci, D. D. & Taatgen, N. A. (2008). Threaded Cognition: An Integrated Theory of Concurrent Multitasking. *Psychological Review, 2008, Vol. 115, No. 1, 101–130.*

Scheuchenpflug, R. (2001). Measuring presence in virtual environments. In M.J. Smith, G. Salvendy, & M. R. Kasdorf (Eds.). *HCI International 2001, pp. 56-58. New Orleans: HCI International.*

Schindler, J., Kelsch J., Heesen, M., Dziennus, M., Temme G. & Baumann, M. (2013). *A Collaborative Approach for the Preparation of Co-operative Multi-User Driving Scenarios*. 10. Berliner Werkstatt Mensch-Maschine-Systeme, 10.-12. Okt. 2013, Berlin, Germany.

Schneegass, S.; Pfleging, B.; Broy, N.; Heinrich, F.; Schmidt, A. (2013). A Data Set of Real World Driving to Assess Driver Workload. *In:Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 150-157*.

Shahrokhi, M. & Bernard, A. (2004). A fuzzy approach for definition of dangerous zone in industrial systems. *In: 2004 IEEE International Conference on Systems, Man, & Cybernetics, Vol. 7, 6318-6324.*

Taatgen, N.A. (2005). Modeling parallelization and speed improvement in skill acquisition: from dual tasks to complex dynamic skills. *Cognitive Science, 29, 421-455.*

Weber, L. & Wetzel, I. (2013). *Eignung unterschiedlicher Versuchsumgebungen für die Erfassung der wahrgenommenen Qualität von Flugzeugkabinen.* In E. Brandenburg, L. Doria, A. Gross, T. Günzler & H. Smieszek (Hrsg.). Grundlagen und Anwendungen der Mensch-Maschine-Interaktion. 10. Berliner Werkstatt Mensch-Maschine-Systeme, ZMMS Spektrum Band 34, 2013 (S. 259-263). VDI-Verlag.

Wittmer, B.G. & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7, 225-240.*

Section 12

Road and Rail:

Highly Automated Driving - Aspects of Driver Vehicle Interaction II

Hands and Feet Free Driving: Ready or Not?

Victoria A. Banks and Neville A. Stanton

Transportation Research Group
University of Southampton
Hampshire, SO17 1BJ, UK

ABSTRACT

"Almost-driverless" cars are coming with an aim to improve the safety of our roads. Full automation of longitudinal and lateral control will enable the driver to become "hands and feet free" but without active control, one question remains. What is the driver actually doing? This paper looks at how multi-system automation that enables the driver to become "hands and feet free" may affect the driving system and the role of the driver within it. Using Operator Sequence Diagrams to explore Distributed Cognition in the driving system, the authors explore the interaction that may occur between the driver and vehicle subsystems in a "hands and feet free" driving system and how this may change the drivers position within the control-feedback driving loops. Acknowledging the role of the driver in this way highlights the need for ongoing Human Factors research into the implications of highly automated vehicles on driver behavior.

Keywords: Autonomous vehicles, Control-Feedback Loops, Distributed Cognition, Systems Engineering

INTRODUCTION

Safety research suggests that driver inattentiveness and a lack of timely response to unpredictable or incomplete information are the most common driver errors that result in vehicular accidents (Amiditis et al. 2010; Cantin et al. 2009; Donmez, Boyle & Lee, 2007; Khan, Bacchus, Erwin, 2012). These external factors are typically random events that evolve to form complex interactions between driver and vehicle (Khan, Bacchus, Erwin, 2012). Without automated assistance, the driver may be underprepared or lack the training needed to respond to the situation accordingly. The vision is that highly automated vehicles will lead to accident-free driving in the future. This means that 100% of the active driving task will need to be completed by a combination of automated subsystems with the driver becoming a passive monitor of system operation (Flemisch et al. 2008). With both General Motors and Nissan reporting that driverless cars will be ready to market from 2020, it is clear that highly automated vehicles that combine multiple automated subsystems are coming whether we are ready for them or not. Whilst it is technically feasible to achieve full vehicle automation (Brookhuis et al. 2008), there is growing concern within the Ergonomics and Human Factors community that the role of the driver is not being fully recognized. Even if the vehicle is capable of controlling all of the physical and cognitive tasks associated with driving, it is unlikely that drivers will willingly and ever truly become disengaged completely from the task. Just like a pilot in a cockpit, the driver will assume a new supervisory role that will become more important as they must remain aware of multiple vehicle subsystems statuses simultaneously and respond accordingly in situations of malfunction or failure (Cuevas et al. 2012; Dehais et al. 2012; Walker et al. 2009). However, consideration of the drivers ability to actually undertake this new supervisory role is becoming increasingly important as the average motorist becomes less actively involved in traditional vehicle handling.

"Hands and feet free" driving has been on the horizon for a while with each facet of technology being a stepping stone to an increased level of autonomy. Since the 1990s, autonomous driving has been viable (for a comprehensive review see Dickmanns (2002). Key milestones are shown in Figure 1. First, there was Adaptive Cruise Control

(ACC) capable of controlling the speed and longitudinal headway between vehicles (de Waard et al. 1999). Then there was Autonomous Emergency Braking (AEB) built upon the existing architecture of ACC. AEB is capable of detecting and identifying imminent collision threats and applying the brakes if needed. Next came Lane Keep Assist (LKA) capable of maintaining lateral road position by applying gentle steering inputs to avoid lane deviation. It is clear to see how the combination of ACC (Stop and Go), an extension of LKA and automated braking, activated simultaneously may lead to a hands and feet free driving environment similar to that of autonomous vehicles.

This paper looks at how multi-system automation that enables the driver to become "hands and feet free" may affect the driving system and the role of the driver within it using the representational method of Operator Sequence Diagrams (OSD) (Kurke, 1961) and a discussion on control-feedback loops within the driving task. With "almost-driverless" vehicles on the horizon, it is imperative that more is known about the Human Factors implications of this shift in driver role by focussing upon the interaction that may take place in a fully automated driving system. This form of task analysis provides an insight into "who" owns "what" information and how this is communicated across the system as a whole (Kaber & Endsley, 2004; Hoc, Young & Blosseville, 2009).

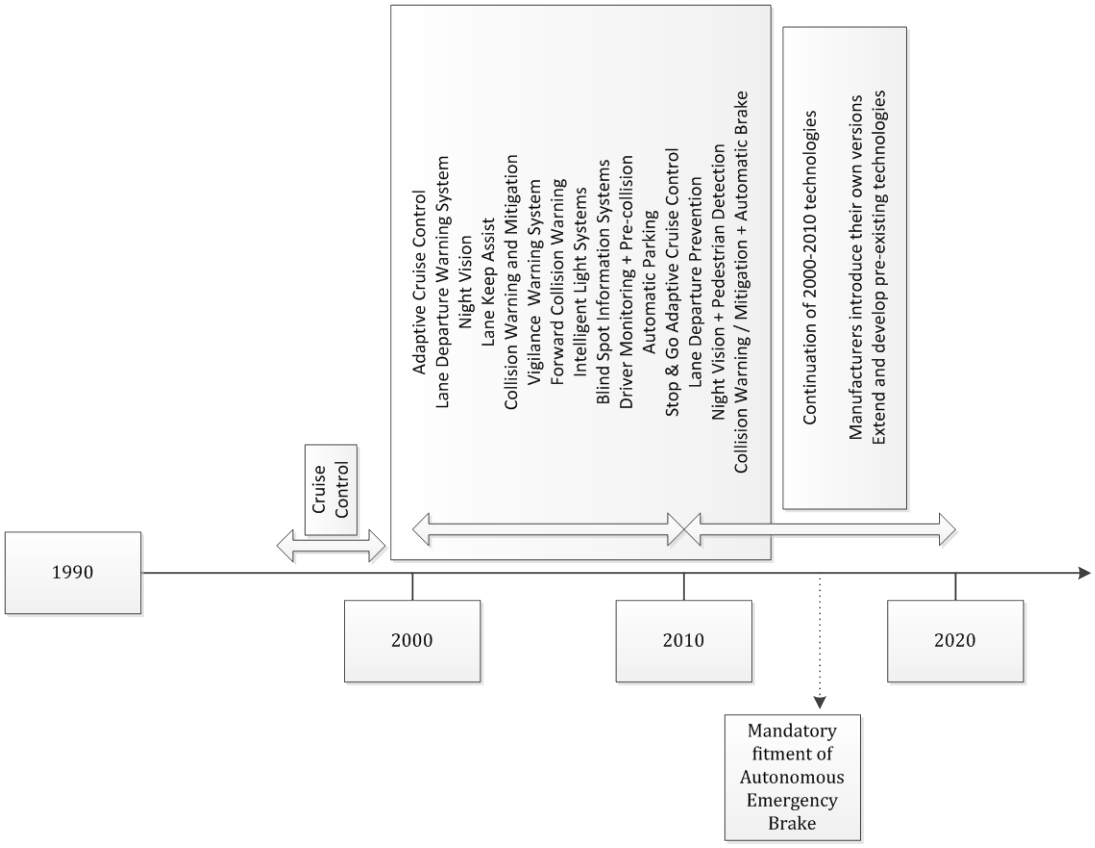


Figure 1. Milestones of vehicle automation since 1990

A "HANDS AND FEET FREE" DRIVING SYSTEM

With Systems Engineering integrating both technical and human-centered approaches, this theoretical approach enables us to understand the functioning and performance of joint cognitive systems such as driving (Salmon et al. 2008) and can be used in the design of adaptive automation (Hollnagel & Woods, 1983). An interdisciplinary approach such as this is extremely complex because the 'behavior' or interaction that occurs between system

components is not always well defined or understood. The aim of Systems Engineering is to better define and characterize subsystem behavior and the interaction that occurs between system agents. Distributed Cognition (Hutchins, 1995), unlike traditional theories, incorporates the idea that interactions can occur between humans, resources and materials over space and time (Hollan, Hutchins & Kirsh 2000; Hutchins, 1995; Walker, Stanton & Chowdhury, 2013). It recognizes that system agents can be both human and non-human (Stanton et al. 2006; Salmon et al. 2008) with Griffin, Young & Stanton (2010) suggesting that all system agents are vital to the flow of information in the control-feedback loops. With vast amounts of information being exchanged between multiple system agents, the ability to make sense of changes in the environment and respond accordingly implicates Situation Awareness (Endsley, 1995) and describes the essence of Distributed Cognition (Walker et al., 2010). However, many argue there is a need to move away from traditional notions of Situation Awareness to one that focuses on entire systems (Gorman, Cooke & Winner 2006; Salmon et al., 2008; Sorensen & Stanton, 2011; Walker et al., 2010). Distributed Situation Awareness (DSA; Stanton et al., 2006) assumes that Situation Awareness is a system level phenomenon (Salmon et al., 2008) and loosely holds systems together because one agent has the ability to compensate for degraded Situation Awareness in another (Stanton et al., 2006).

For driving automation, OSD's offer a means to explore distributed cognition within the driving system with an aim to provide a clearer understanding of how tasks can be shared between system agents and offers a unique opportunity to visualise how system agents may communicate (e.g. Banks, Stanton & Harvey, 2014). OSD's have previously been used to model interactions between system agents in a cockpit environment (Sorensen & Stanton, 2011) and driving emergencies at different levels of automation (Banks, Stanton & Harvey, 2014) as well as a number of other domains. Figure 2 offers a visual representation of Distributed Cognition in an "almost-driverless" system using standardised geometric features (see Table 1; Kurke, 1961) and the subsystems described earlier that allow for a near autonomous driving environment. Where Endsley & Kaber (1999) suggested that at higher levels of automation, human operators are completely removed from the control-feedback loops as intelligent subsystems become capable of completing all of the physical and cognitive responsibilities associated with a task, the authors argue here that far from being removed from the loop, drivers still remain passively involved via their monitoring of the wider environment and subsystem behaviour. Thus, although the driver is no longer required to perform any of the physical tasks associated with driving and in theory, can be relieved of any cognitive workload, it is clear that driver monitoring remains to some extent. Just like a pilot in a cockpit, the driver will assume a new supervisory role when automation is engaged and far from being removed from the control-feedback loop, drivers will find themselves to be an important link between these multiple subsystems. This is because they still accept ultimate responsibility for safe vehicle operation and have an innate desire to be in control (Parasuraman & Wickens, 2008).

It is also clear that the multi-subsystem approach on the road to full vehicle automation will bring with it a greater need for effective communication between system agents to ensure that safe driving practice is maintained. All system agents will need to have an awareness of how each subsystem works and have some level of intelligence relating to specific functional limitations of each subsystem. In essence, automated subsystems will need to monitor the behaviour of each intelligent counterpart as well as monitoring the wider environment and adapt accordingly signalling the development of Distributed Situation Awareness amongst automated subsystems (Stanton et al. 2006).

Table 1: OSD Key

Geometric Shape	***Meaning***
	Process
	Decision
	Terminator
	Manual input
	Delay
	Path of interaction

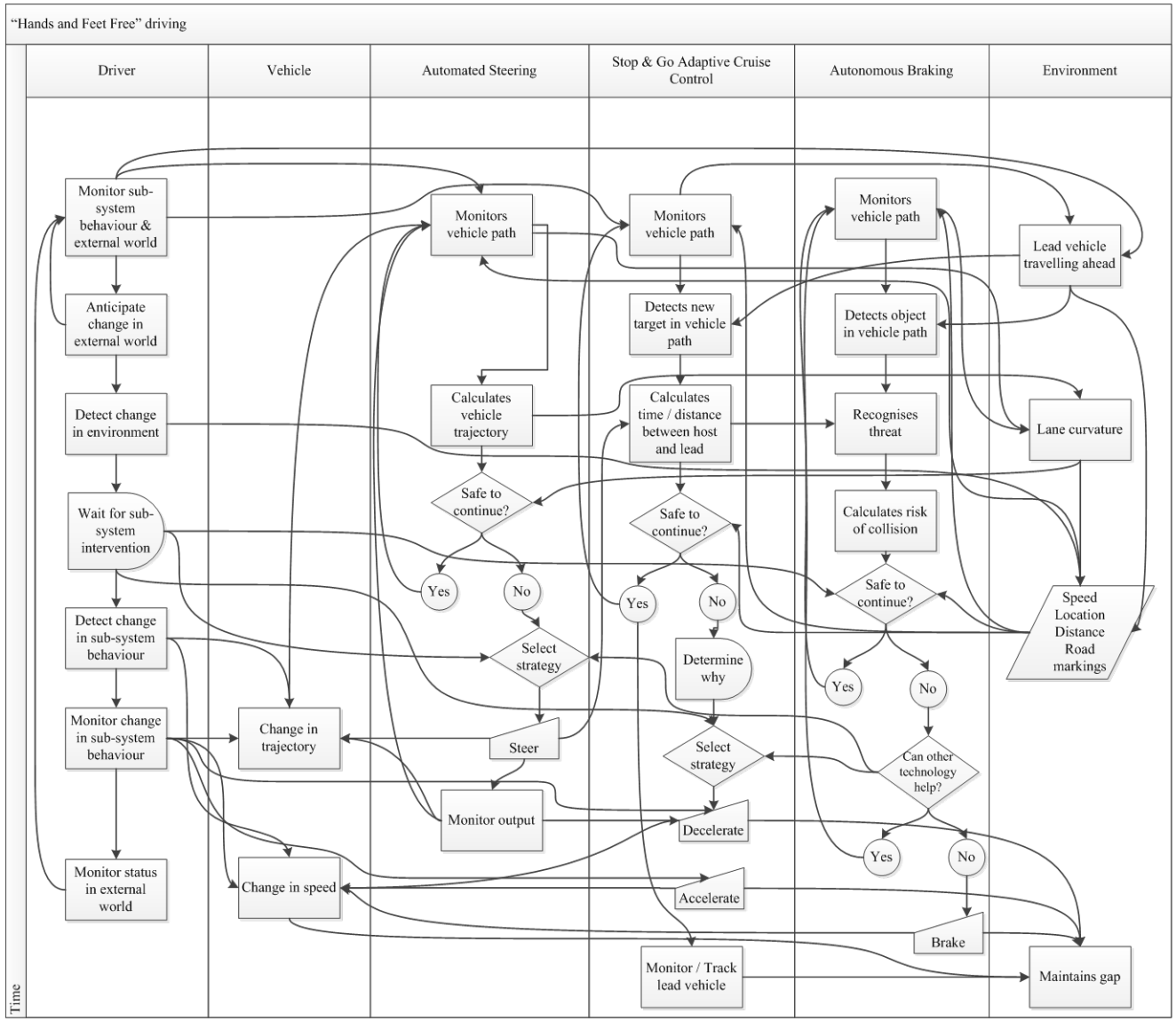


Figure 2.Division of labor in a "hands and feet free" driving system.

THE CHANGING DRIVER ROLE

At first glance, a "hands and feet free" driving system could lead to the traditional role of the driver becoming redundant. An "almost driverless" system such as this would see 100% of the active driving task being completed by automated subsystems (Flemisch et al. 2008) with the driver becoming a passive monitor of system operation. However, Figure 2 clearly indicates that driver monitoring will remain an important aspect of the task regardless of their 'location' within the driving control-feedback loops. Whether the driver wants to or not, they will continue to receive feedback from both the vehicle and environment in a highly automated driving system (Figure 3). They will also continue to have the ability to anticipate changes in the environment using feedforward information. The key difference between manual and highly automated driving is that the latter may see a disintegration or possible removal of links between the driver and vehicle subsystems (Stanton, Young & McCaulder, 1997; Stanton et al. 2007). This disintegration or removal of links could negatively impact upon the level of driver engagement (Table 2). Although it seems reasonable to suggest the driver will remain in-the-loop once automation is engaged for some time, prolonged durations of automated driving could lead to the driver becoming out-of-the-loop (e.g. Cuevas et al. 2007; Beckier, Molesworth & Williamson, 2012). A review of the literature indicates that automation can have both positive and negative impacts upon driver behavior. For example, Brookhuis et al. 2008 suggest that an individual's ability to monitor the visual scene efficiently may actually decrease under automated driving conditions since automation leads to changes in levels of vigilance and complacency. Although this is not necessarily a cause for concern if the level of automation within the driving system is failsafe, capable of sound judgement, analysis, decision making and learning, questions still remain over how we reengage the driver and ensure their levels of situation awareness would enable swift manual takeover if required. Worryingly, de Waard et al. (1999) reported that 50% of drivers failed to regain control following system malfunction on an automated highway system due to the belief that the system would intervene despite the system being compromised (signaling an active driver being out-of-loop). Although the reality of system failure is small in most cases due to an extensive testing phase, operational failings may leave drivers vulnerable to the need of intervention whether it is prompted by the system or not (Larsson, 2012) and to sudden increases to their workload (Parasuraman & Hancock, 2001). This suggests that automated subsystems may not be uniformly beneficial (Brookhuis et al. 2001; Lee & See, 2004).

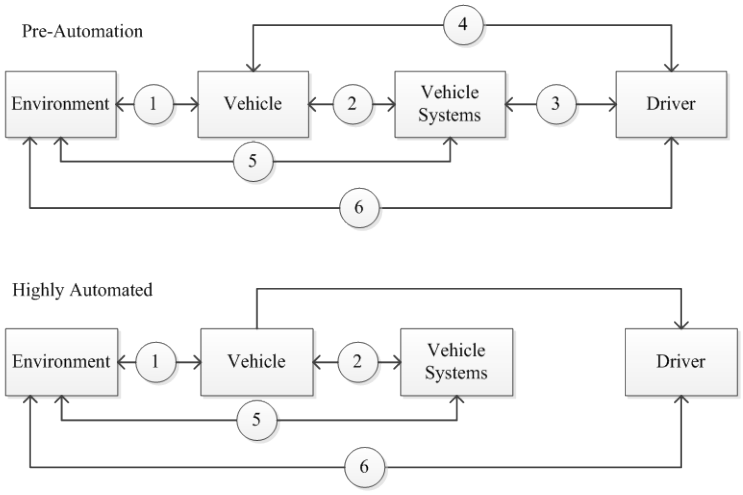


Figure 3. Changes to the Control-Feedback Loops in Driving.

Driver desensitization may mean that manual override in unanticipated and unexpected events will be difficult to manage, increase workload and stress as well as create surprise or startle effects (Sarter, Woods & Billings, 1997). However, if an element of command and control remains within the drivers grasp, these performance decrements could be reduced (Figure 4). One way of achieving this is to allow drivers to issue commands to the automation and in this way, act as an active supervisor rather than a passive monitor of the system. This would allow the driver to be maintained in the control-feedback loop and satisfy their 'need' or desire to remain in control of the vehicle. This

Table 2: Driver States of Engagement

	Loop	
	In	Out
Active	Driver in full control of the vehicle and is actively engaged in the driving task.	Driver in full control of the vehicle but show characteristics of out-of-loop driving e.g. driving without attention
Passive	Driver no longer in control of vehicle operation but remains actively engaged in the driving task.	Driver no longer in control of vehicle operation and becomes desensitized to the driving task.

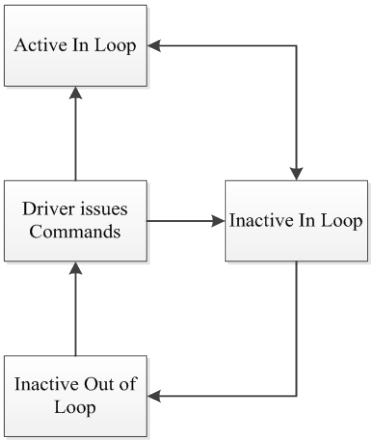


Figure 4. Keeping the driver engaged.

would go some way in addressing out-of-the-loop performance problems commonly reported in the literature (e.g., Kaber & Endsley, 2004; de Waard et al. 1999; Stanton & Young, 2000; Endsley & Kiris, 1995). In this way, the authors argue that the role of the driver could become analogous with the role of co-pilot within aviation. Although the status of the driver within the control-feedback loop has changed, an element of command and control could prevent the disintegration of driver-vehicle interaction and instead enable the driving system to become flexible and 'accommodate' for change. Furthermore, it is still important to note that the limitations of technology mean that at some point within the system of command and control, a human operator will be needed so having a mechanism to keep drivers engaged to some extent is important. In addition, it seems that until issues surrounding practicability, liability and individual preferences of the driver are addressed, control transitions will continue to be made between the driver and automated subsystems (SMART European Commission Study Report, 2010). This means that drivers remain a vital link within the system network and still accept ultimate responsibility for safe vehicle operation (Parasuraman & Wickens, 2008). Despite out-of-the-loop performance concerns, it is encouraging to see recent studies finding that although highly automated driving can lead to the driver becoming more involved with secondary driving tasks, drivers remain receptive to changing demands such as an increase in traffic density in the wider environment. For example Jamson et al. (2013) found that drivers were more focused on the roadway when traffic was heavy when driving a highly automated vehicle. This goes some way in supporting the representation

shown in Figure 2 and highlights that even when automated subsystems are capable of controlling the vehicle without any human input, the driver continues to engage in the control-feedback loops to some extent.

CONCLUSIONS

The complex web of interaction that an "almost-driverless" system is likely to introduce into the driving system is not something vehicle manufacturers should ignore. Although early research into automation seemed to focus most heavily upon autonomy, current research now focusses upon satisfying the requirements of joint activity, including human-machine teamwork (Klein et al. 2004). This is because questions still remain over appropriate levels of automation and whether or not fully automated driving enables the driver to become completely "mind free" as well as "hands and feet free". Although Ursom et al. (2008) suggest that perhaps we view the subject of autonomous vehicles too inflexibly and that drivers should instead learn to adapt to intelligent vehicles rather than the other way round, driving remains a social activity – one that incorporates subtle behaviours. For example, behavioural indicators displayed by other drivers such as varying the speed and gap between vehicles can indicate a willingness to allow other vehicles to join the flow of traffic. Additionally, eye contact and hand gestures demonstrated by other drivers can assist in driver decision-making. Groom and Nass (2007) argue that intelligent vehicles lack these human-like tendencies necessary to be fully integrated or accepted on our roads. Whatever the stance, highly automated vehicles are coming apace.

If as we suspect the future of vehicle automation points to "near driverless" cars, more research is needed to ensure that the driver is capable of manual takeover after prolonged exposure to a fully reliable system. It is important that that the principle of complementarity is adopted, with the allocation of tasks serving to maintain control whilst retaining human skill (Grote et al., 1995). As with Free Flight (Langan-Fox et al., 2009), driving automation poses many challenges with regards to the interaction between humans and automation including operational functionality and system management. There may be confusion over "who" (the driver or automated subsystem) has authority over 'which' vehicular controls as the level of automation increases. Presently, technologies such as Adaptive Cruise Control can be controlled by the driver. This means that the driver is able to manually switch the system on or off. This on the one hand supports driver preferences but prevents the technology reaching its full potential. Controlling technologies that enable the driver to become "hands and feet free" may actually desensitize the driver because the manual and cognitive skill that is required to drive a vehicle is not being used. This desensitization may mean that manual override in unanticipated and unexpected emergency events will be difficult for the driver to manage (Sarter, Woods & Billings, 1997). Future research should focus upon driver-vehicle interactions in catastrophic failure events. Although these failure events are likely to be rare occurrences, we need to ensure that the driver remains capable of manual takeover. In other words, we need to find out ways to keep drivers actively engaged in the driving task whilst automation is engaged. This is because manual override may continue to be an important aspect of emergency intervention. Only once these concerns are acknowledged and resolved can we say we are ready for "almost-driverless" vehicles.

ACKNOWLEDGEMENTS

This research was funded by the Engineering and Physical Sciences Research Council (EPSRC) and Jaguar Land Rover.

REFERENCES

Amditis, A., Pagle, K., Joshi, S., & Bekiaris, E. (2010), Driver-Vehicle-Environment monitoring for on-board driver support systems: lessons learned from design and implementation. *Applied Ergonomics*, *41*(2), 225-235.

Banks, V. A., Stanton, N. A., & Harvey, C. (2014). Sub-systems on the road to vehicle automation: Hands and feet free but not 'mind'free driving. *Safety science*, *62*, 505-514.

Bekier, M., Molesworth, B. R., & Williamson, A. (2012). Tipping point: The narrow path between automation acceptance and rejection in air traffic management. *Safety science*, *50*(2), 259-265.

Brookhuis, K. A., De Waard, D., & Janssen, W. H. (2001). Behavioural impacts of advanced driver assistance systems–an overview. *European Journal of Transport and Infrastructure Research*, *1*(3), 245-253.

Brookhuis, K. A., van Driel, C. J., Hof, T., van Arem, B., & Hoedemaeker, M. (2009). Driving with a congestion assistant; mental workload and acceptance. *Applied ergonomics*, *40*(6), 1019-1025.
Cantin, V., Lavallière, M., Simoneau, M., & Teasdale, N. (2009). Mental workload when driving in a simulator: Effects of age and driving complexity. *Accident Analysis & Prevention*, *41*(4), 763-771.
Cuevas, H. M., Fiore, S. M., Caldwell, B. S., & Strater, L. (2007). Augmenting team cognition in human-automation teams performing in complex operational environments. *Aviation, space, and environmental medicine*, *78*(Supplement 1), B63-B70.
Dehais, F., Causse, M., Vachon, F., & Tremblay, S. (2012). Cognitive conflict in human–automation interactions: a psychophysiological study. *Applied Ergonomics*, *43*(3), 588-595.
de Waard, D., van der Hulst, M., Hoedemaeker, M., & Brookhuis, K. A. (1999). Driver behavior in an emergency situation in the Automated Highway System. *Transportation Human Factors*, *1*(1), 67-82.
Dickmanns, E. D. (2002, June). The development of machine vision for road vehicles in the last decade. In *Intelligent Vehicle Symposium, 2002. IEEE* (Vol. 1, pp. 268-281). IEEE.
Donmez, B., Boyle, L. N., & Lee, J. D. (2007). Safety implications of providing real-time feedback to distracted drivers. *Accident Analysis & Prevention*, *39*(3), 581-590.
Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 65-84.
Endsley, M. R. (2006). Situation Awareness. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (3rd ed., pp. 528–542). New York: Wiley.
Endsley, M. R. & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462-492.
Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(2), 381-394.
European Commission Study Report. (2010). SMART 2010/0064: Definition of necessary vehicle and infrastructure systems for Automated Driving, 1-111. Version 1.2.
Flemisch, F., Kelsch, J., Löper, C., Schieben, A., & Schindler, J. (2008). Automation spectrum, inner/outer compatibility and other potentially useful human factors concepts for assistance and automation. *2008), Human Factors for assistance and automation*, 1-16.
Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, *49*(12-13), 1312-1325.
Griffin, T. G., Young, M. S., & Stanton, N. A. (2010). Investigating accident causation through information network modelling. *Ergonomics*, *53*(2), 198-210.
Groom, V., & Nass, C. (2007). Can robots be teammates?: Benchmarks in human–robot teams. *Interaction Studies*, *8*(3). 326–382.
Grote, G., Weik, S., Wäfler, T., & Zölch, M. (1995). Criteria for the complementary allocation of functions in automated work systems and their use in simultaneous engineering projects. *International Journal of Industrial Ergonomics*, *16*(4), 367-382.
Hoc, J. M., Young, M. S., & Blosseville, J. M. (2009). Cooperation between drivers and automation: implications for safety. *Theoretical Issues in Ergonomics Science*, *10*(2), 135-160.
Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *7*(2), 174-196.
Hollnagel, E., & Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies*, *18*(6), 583-600.
Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive science*, *19*(3), 265-288.
Jamson, A. H., Merat, N., Carsten, O. M., & Lai, F. C. (2013). Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. *Transportation Research Part C: Emerging Technologies*, *30*, 116-125.
Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, *5*(2), 113-153.
Khan, A. M., Bacchus, A., & Erwin, S. (2012). Policy challenges of increasing automation in driving. *IATSS research*, *35*(2), 79-89.
Klein, G., Hoffman, R. R., Feltovich, P. J., Woods, D. D., & Bradshaw, J. M. (2004). Ten challenges for making automation a" team player" in joint human-agent activity. *IEEE Intelligent Systems*, *19*(6), 91-95.
Kurke, M. I. (1961). Operational sequence diagrams in system design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *3*(1), 66-73.
Langan-Fox, J., Canty, J. M., & Sankey, M. J. (2009). Human–automation teams and adaptable control for future air traffic management. *International Journal of Industrial Ergonomics*, *39*(5), 894-903.
Larsson, A. F. (2012). Driver usage and understanding of adaptive cruise control. *Applied Ergonomics*, *43*(3), 501-506.
Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50-80.
Parasuraman, R., & Hancock, P. A. (2001). Adaptive control of workload. In P. A. Hancock & P. E. Desmond (Eds.), Stress, workload, and fatigue, pp. 305–320. . Mahwah, NJ: Erlbaum.
Parasuraman, R., & Wickens, C. D. (2008). Humans: Still vital after all these years of automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 511-520.
Salmon, P. M., Stanton, N. A., Walker, G. H., Baber, C., Jenkins, D. P., McMaster, R., & Young, M. S. (2008). What really is going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, *9*(4), 297-323.

Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. *Handbook of human factors and ergonomics*, *2*, 1926-1943.

Sorensen, L. J., Stanton, N. A., & Banks, A. P. (2011). Back to SA school: contrasting three approaches to situation awareness in the cockpit. *Theoretical Issues in Ergonomics Science*, *12*(6), 451-471.

Stanton, N. A., Stewart, R., Harris, D., Houghton, R. J., Baber, C., McMaster, R., ... & Green, D. (2006). Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. *Ergonomics*, *49*(12-13), 1288-1311.

Stanton, N. A., Walker, G. H., Young, M. S., Kazi, T., & Salmon, P. M. (2007). Changing drivers' minds: the evaluation of an advanced driver coaching system. *Ergonomics*, *50*(8), 1209-1234.

Stanton, N. A., & Young, M. S. (2000). A proposed psychological model of driving automation. *Theoretical Issues in Ergonomics Science*, *1*(4), 315-331.

Stanton, N. A., Young, M. S., & McCaulder, B. (1997). Drive-by-wire: the case of mental workload and the ability of the driver to reclaim control. *Safety Science*, *27*(2-3), 149-159.

Urmson, C., Anhalt, J., Bagnell, D., Baker, C., Bittner, R., Clark, M. N., ... & Ferguson, D. (2008). Autonomous driving in urban environments: Boss and the urban challenge. *Journal of Field Robotics*, *25*(8), 425-466.

Walker, G. H., Stanton, N. A., Baber, C., Wells, L., Gibson, H., Salmon, P., & Jenkins, D. (2010). From ethnography to the EAST method: A tractable approach for representing distributed cognition in Air Traffic Control. *Ergonomics*, *53*(2), 184-197.

Walker, G. H., Stanton, N. A., & Chowdhury, I. (2013). Self-Explaining Roads and situation awareness. *Safety Science*, *56*, 18-28.

Walker, G. H., Stanton, N. A., Kazi, T. A., Salmon, P. M., & Jenkins, D. P. (2009). Does advanced driver training improve situational awareness?.*Applied Ergonomics*, *40*(4), 678-687.

Taking Over Control from Highly Automated Vehicles

Christian Gold and Klaus Bengler

Institute of Ergonomics
Technische Universität München
Munich, Germany

ABSTRACT

Future cars will be able to execute the longitudinal and lateral control and other subtasks of driving. Automation effects, known in other domains like aviation, rail traffic or manufacturing, will emerge in road transportation with consequences hard to predict from the present point of view. This paper discusses the current state of automation research in road traffic, concerning the take-over at system limits. Measurements like the take-over time and the maximum accelerations are suggested and substantiated with data from different experiments and literature. Furthermore, the procedure of such take-over situations is defined in a generic way. Based on studies and experience, advice is given concerning methods and lessons learned in designing and conducting take-over studies in driving simulation. This includes the test and scenario design and which dependent variables to use as metrics. Detailed information is given on how to generate proper control conditions by driving manually without automation. Core themes like how to keep situation presentation constant even for manual drivers and which measures to use to compare a take-over to manual driving are addressed. Finally, a prospect is given on further needs for research and limitations of current known studies.

Keywords: take-over, automated driving, controllability, driving simulation

INTRODUCTION

The transferability of results from aviation for instance is very limited, because of many factors like variety of users, different object density, task complexity, time criticality and many more. That is why basic research in road transportation has to be done, caused by the fast progress of the development of automation in this specific area.

Automation can be divided into different stages with increasing levels of support by automation. With traffic-jam-assistance we have nowadays reached the level of partially automated driving. The task of the driver is changing from driving to monitoring the system and staying prepared to intervene at any point in time. A high driver benefit could be gained with the next level, the high automation. Following the definition of the BASt, the driver doesn't have to monitor the system any more, but has to be prepared to take over control of the vehicle within a certain time buffer (Gasser, 2012). Accordingly, the NHTSA (2013) defines the same stages of automation, using a different wording. The BASt "Partially automated driving" corresponds to NHTSA "Level 2 - Combined Function Automation" and "highly automated driving" and is also called "Level 3 - Limited Self-Driving Automation".

As soon as the driver doesn't have to monitor the system, he can deal with tasks other than driving and is able to completely withdraw himself from the driving task. While the automation deals with the driving within its system limits, the human and his performance is crucial in take-over situations that occur on system limits or at the end of the automation scenario. Concerning the aspects of driver vehicle interaction, ways have to be found to keep the driver as involved in the driving task as possible and to be aware of the current availability of the driver at any point in time while driving highly automated to ensure the safety of the overall system.

OVERVIEW ON CURRENT RESEARCH

Based on the conclusion of Reason (1990) that the cognitive structure of humans is not well suited for supervisory control and the "ironies of automation", introduced by Bainbridge (1983) warning that the higher the level of automation becomes the more important the role of the human within the overall system gets, the impact of automated driving has to be critically examined. As car manufacturers become more confident about the feasibility of such automated systems, this field of research grew in recent years and led to plenty of publications. Beside automation effects like overreliance, skill degradation or reduced situation awareness, the authority transition (Saffarian, Winter & Happee, 2012) is one key controllability aspect of automated driving. Based on a review of literature regarding adaptive cruise control and active lane assist systems, Stanton and Young (1998) stated that there are going to be "some problems with reclaiming control of the vehicle in failure scenarios". Except for those failure scenarios, transitions could likewise be caused by "road blockage, severe weather conditions, sudden maneuvers by another vehicle and operator preference" (Saffarian et al., 2012), wherefore detailed studies focusing transitions in automated vehicles are urgently required.

Several studies in recent years investigated impacts of different parameters on take-over situations, almost exclusively in driving simulator experiments. Automation effects of different manifestations have been found in nearly every known study. With higher automation levels, subjects tend to get more involved in non-driving related tasks (Carsten et al., 2012). Although one would expect, that distraction tends to worsen drivers' reaction in take-over situations, Neubauer, Matthews and Saxby (2012a) measured a shortening in brake reaction times of subjects making a phone call while driving highly automated compared to subjects driving automated without any non-driving related task. The author assumed a higher activation level to be the reason for this finding. The direction of the influence of non-driving related tasks seems to depend on the modality and/or type of the task. For example, Merat et al. (2012) measured an influence of the involvement in a questionnaire on reaction times in critical take-over scenarios. Also Radlmayr (2013) found differences in take-over quality between a visual and a cognitive task. Beller, Heesen and Vollrath (2013) figured out that the communication of systems uncertainty leads to greater time to collisions (TTC) at system limits. Even if the subjects themselves chose whether to activate the automation or not, they show increased reaction times compared to manual driving subjects (Neubauer et al., 2012b). Concerning the minimum time budget a driver needs to take over control, Damböck et al. (2012) and Gold et al. (2013a) suggested budgets greater than 7 seconds for visually distracted drivers in different take-over situations. Significant automation effects have been observed with 4, 5 and 6 seconds time budget. Petermann-Stock et al. (2013) measured take-over times up to 8.8 seconds independently of the age group of the subjects. Also other transitions than from highly automated to manual driving were addressed in a study by Petermann and Schlag (2010) in a wizard-of-oz vehicle. Similar to findings in Gold et al. (2013b) they found increased problems through mode confusion compared to transitions between highly automated and manual driving.

TAKING OVER VEHICLE CONTROL

As soon as the driver is out of the loop, he is likely to lose situation awareness by a shift of tasks, inattention, daydreaming, fatigue et cetera. This influences the shift back to the driving task, which is demanded within a limited time budget at system limits. A state where the automation and the driver both reach their limits has to be avoided as far as possible and therefore, as highly automated cars have system limits by definition, the take-over procedure has to be studied and the driver has to be granted as much time and support as possible. Taking over control from highly automated vehicles can be abstracted as shown in figure 1.

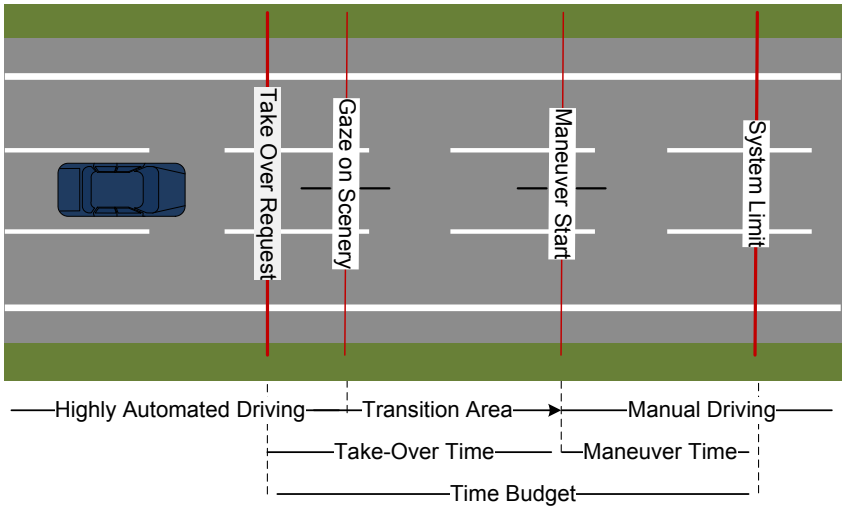


Figure 1. Generic take-over situation

The driver approaches a system limit in highly automated mode. With a certain time budget, the driver gets informed about approaching this system limit by a take-over request (TOR). Depending on the previous task of the driver, he has to interrupt his task and glance up to the scenery. With his eyes on road, the information perception and processing starts and a transition area is entered, where the level of automation is somewhere in between highly automated driving and manual driving. Depending on the system design, the vehicle continues the driving task until reaching the system limit, or, for example in case of a system failure, the driver has to perform the driving task immediately. In case of continued guidance by the automation, the driver has to indicate take-over readiness by inputs on a control element like the steering wheel. As soon as the driver starts his maneuver to respond to the system limit, the level of automation reaches “manual driving” and the take-over time ends.

There are more possibilities for transitions at system limits such as only delegating longitudinal or lateral guidance back to the driver. Those transitions had been considered in studies by Petermann and Schlag (2010) and Gold et al. (2013b). Whether those transitions help to improve overall performance of the human machine system could not be conclusively clarified yet, but problems with mode awareness seem to be more likely to occur.

HUMAN PERFORMANCE IN TAKE-OVER SITUATIONS

The measurements of drivers’ performance in take-over situations can be divided into two aspects: Take-over times and take-over quality. Although the driver might steer, brake or accelerate, it is not ensured that he assessed the situation correctly and came to an appropriate decision. It is therefore not only very important to provide a sufficient time budget and to not force the driver into premature decisions but also to assess drivers’ inputs and measure the take-over quality. With shorter time budgets, the take-over times get shorter but the take-over quality worsens (cf. Gold et al. 2013a). Both aspects therefore always have to be considered when assessing humans’ performance in take-over situations.

Take-over time and reaction times

Considering the take-over procedure, different time metrics can be measured. First of all, the take-over time - when the manual driving starts - is important. It includes the movement time to the driving position, situation perception and assessment and the decision process. Besides, a variety of times can be measured by methods like Eye-Tracking or based on drivers input on control elements. Depending on the situation and the required driver input, the end of the take-over time is not easy to define. Gold et al (2013a) suggests for a lane change a steering wheel angle of more than 2 degrees and a braking pedal position of more than 10 percent. Below those boundary values any driver input is supposed to be not related to a conscious maneuver but for the reason of vehicle stabilization which is not a reaction to the system limit. If the vehicle stabilization itself is the reaction to the system limit, the definition of the take-over time would have to be defined differently correspondingly.

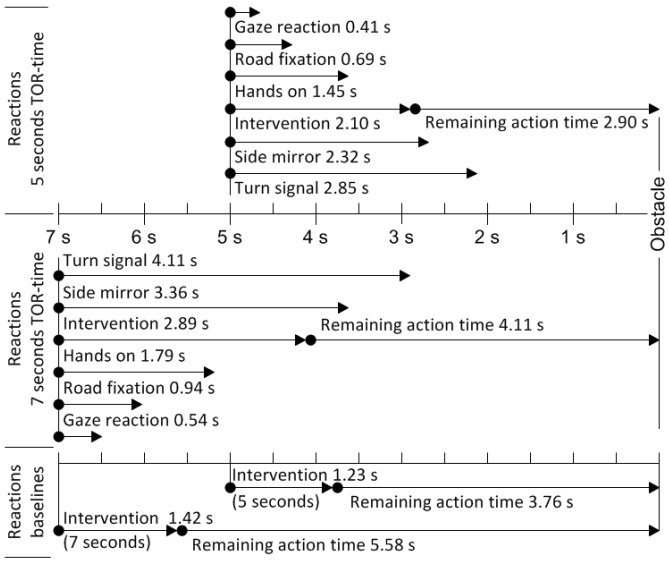


Figure 2: Reaction times in Gold et al. 2013a

Figure 2 shows results regarding different reaction times in two take-over scenarios with 5 and 7 seconds time budget. In this experiment, next to the take-over time (intervention) the gaze reaction (the time until the first reaction after the TOR is measured), the road fixation (time until first gaze on scenery), the hands on time (time when hands touched the steering wheel), the first gaze in the side mirror and the activation of the turn signal (indicators) were recorded. Depending on the situation and experimental question, the measurement of several different times like brake-reaction time can be reasonable.

Take-over quality

Measuring the drivers' performance during and after taking over control is essential to assess the overall system composed of driver, vehicle and automation. There are several ways to assess drivers' performance, depending on situation and necessary maneuver. If a stabilization of the vehicle is sufficient, common measures like standard deviation of lateral position or standard deviation of velocity can be measured. If the required maneuver is more complex, the measuring of take-over becomes elaborate. For measuring the quality of a lane change for example, lateral accelerations or the minimal observed time to collision can be recorded. Furthermore, the securing of the lane change by gazes in the mirrors or shoulder check can be examined.

Finally, it can be assessed whether the subjects chose the right maneuver as a response to the system limit. For instance, inappropriate braking or choosing the wrong lane can indicate insufficient take-over quality. As higher accelerations or smaller TTC clearly indicate a worse take-over quality, the determination as to whether a maneuver is appropriate is more complex as for example ethical aspects can influence this decision.

DESIGNING TAKE-OVER EXPERIMENTS

The following suggestions are based on the experience of the author in several take-over experiments. As in other fields, take-over experiments examine the influence of an independent variable on a dependent variable. An imaginable scenario is that in a within subjects design the situation is kept constant and with repeated measures the independent variable (e.g. different non-driving related tasks) varies. For this purpose every subject has to experience the same situation at least two times. Results show, that even with different visual representation of one situation, take-over times and quality improve on the second situation. Figure 3 shows additional data of the experiment of Gold et al. (2013b). Situation "Road Construction" and "Situation Person" were presented twice to each subject with a slightly different appearance. In both situations, the subjects had to react to an object which was entering the current lane from the right. The influence of the learning effect is obvious. Several dependent variables are showing significance tested with a t-test and a significance level of 0.05.

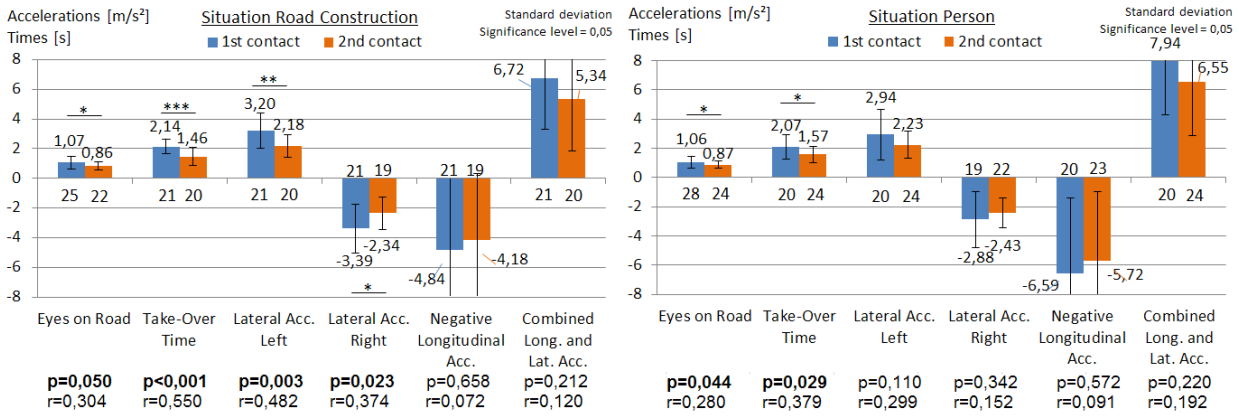


Figure 3: Example for Learning Effects

Those learning effects complicate conclusions regarding the influence of the independent variable. A between subjects design, where subjects experience only one take-over, would avoid influences of learning effects, unfavorable on costs of bigger samples and therefore a higher experimental effort. Nevertheless, when planning the experimental design, the question has to be raised as to whether repeated measures are usable or if big samples with only one take-over situation have to be preferred.

To judge not only whether certain aspects have an influence on the take-over time and quality but also whether a certain take-over is sufficient from a controllability point of view, the take-over performance has to be compared to manual driving. It is, and has to be further discussed, not only how much the performance of the overall system has to improve, but also if such take-overs have to be compared to manual driving without any assistance, or to the state of the art with driver support by advanced driver assistance systems. Furthermore, if a baseline consisting of manual driving subjects shall be recorded, it has to be ensured that the representation of the situation matches the automated condition. Here, the most critical value is the point in time when the system limit becomes visible. As automated driving subjects only have a certain time budget to react on the system limit, this limit should become visible to the manual driving subjects at the exact same point in time. This is possible by hiding the system limit behind any object, like a leading vehicle. The authors suggest to preferably letting the limit emerge suddenly in a simulator by displaying the limit in the very moment of the tor. Accepting small losses of sense of reality of subjects detecting this sudden appearance, automated and manual driving subjects have the same preconditions for solving the situation and artefacts are prevented as much as possible.

CONCLUSION AND OUTLOOK

The paper summarized questions and studies regarding automated driving and the take-over from automated vehicles. Directions were given and dependent measures suggested for designing such studies. By performing the experiments in real vehicles, future studies have to examine whether the transferability of driving simulator studies in this special research field is permissible. The impact of many kinds of influences on take-over time and quality also have to be further investigated and additional studies are needed regarding all kinds of automation effects probably emerging in automated vehicles, such as behavioral adaptation or skill degradation.

REFERENCES

Bainbridge, L. (1983). *Ironies of Automation*. Automatica, 19(6), 775–779.

Beller, J., Heesen, M., & Vollrath, M. (2013). *Improving the Driver-Automation Interaction: An Approach Using Automation Uncertainty*. Human Factors: The Journal of the Human Factors and Ergonomics Society, 55(6), 1130–1141. doi:10.1177/0018720813482327

Carsten, O., Lai, F. C. H., Barnard, Y., Jamson, A. H., & Merat, N. (2012). *Control Task Substitution in Semiautomated Driving: Does It Matter What Aspects Are Automated?* Human Factors: The Journal of the Human Factors and Ergonomics Society, 54(5), 747–761. doi:10.1177/0018720812460246

Damböck, D., Farid, M., Tönert, L., & Bengler, K. (2012). *Übernahmezeiten beim hochautomatisierten Fahren*. 5. Tagung Fahrerassistenz, 5.

Gasser. (2012). *Rechtsfolgen zunehmender Fahrzeugautomatisierung: Gemeinsamer Schlussbericht der Projektgruppe*. Bremerhaven.

Gold, C., Dambock, D., Lorenz, L., & Bengler, K. (2013a). *"Take over!" How long does it take to get the driver back into the loop?* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57(1), 1938–1942. doi:10.1177/1541931213571433

Gold, C., Lorenz, L., Damböck, D., & Bengler, K. (2013b). *Partially Automated Driving as a Fallback Level of High Automation*. 6. Tagung Fahrerassistenzsysteme, 28.-29. November, München, 6.

Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. (2012). *Highly Automated Driving, Secondary Task Performance, and Driver State.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 54(5), 762–771. doi:10.1177/0018720812442087

Neubauer, C., Matthews, G., & Saxby, D. (2012a). *The Effects of Cell Phone Use and Automation on Driver Performance and Subjective State in Simulated Driving*. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56(1), 1987–1991. doi:10.1177/1071181312561415

Neubauer, C., Matthews, G., Langheim, L., & Saxby, D. (2012b). *Fatigue and Voluntary Utilization of Automation in Simulated Driving.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 54(5), 734–746. doi:10.1177/0018720811423261

NHTSA. (2013). *Preliminary Statement of Policy Concerning Automated Vehicles*. Retrieved from www.nhtsa.gov/ staticfiles/ rulemaking/ pdf/ Automated_Vehicles_Policy.pdf

Petermann, I., & Schlag, B. (2010). *Auswirkungen der Synthese von Assistenz und Automation auf das Fahrer-Fahrzeug System*. AAET 2010 – Automatisierungssysteme, Assistenzsysteme und eingebettete Systeme für Transportmittel, 383–403.

Petermann-Stock, I., Hackenberg, L., Muhr, T., & Mergl, C. *Wie lange braucht der Fahrer?: Eine Analyse zu Übernahmezeiten aus verschiedenen Nebentätigkeiten während einer hochautomatisierten Staufahrt*. In 6. Tagung Fahrerassistenzsysteme, München, 2013 .

Radlmayr, J. (2013). *How varying traffic situations and secondary tasks affect the quality of the take-over process during highly automated driving.* Technische Universität München, München.

Reason, J. T. (1990). *Human error*. Cambridge: Cambridge University Press.

Saffarian, M., Winter, J. C. F. de, & Happee, R. (2012). *Automated Driving: Human-Factors Issues and Design Solutions.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56(1), 2296–2300. doi:10.1177/1071181312561483

Stanton, N. A., & Young, M. (1998). *Vehicle automation and driving performance.* Ergonomics, 41(7), 1014–1028. doi:10.1080/001401398186568

The 'Methods for Designing Future Autonomous Systems' (MODAS) Project: Developing the Cab for a Highly Autonomous Truck

Stas Krupenia[a], Anna Selmarker[a], Johan Fagerlönn[b], Katarina Delsing[b], Anders Jansson[c], Bengt Sandblad[c] and Camilla Grane[d]

[a]Cab Development, Driver Vehicle Interaction
Scania CV AB
15187 Södertälje, Sweden

[b]Interactive Institute Swedish ICT
Acusticum 4
94128 Piteå, Sweden

[c]Department of Information Technology
Uppsala University
75105 Uppsala, Sweden

[d]Department of Business Administration, Technology and Social Sciences
Luleå University of Technology
97189 Luleå, Sweden

ABSTRACT

Modern technologies have the potential to create a paradigm shift in the vehicle-driver relationship with advanced automation changing the driver role from "driving" to "supervising". To design new driver environments that caters for these emerging technologies, traditional approaches identify current human and technical constraints to system efficiency and create solutions accordingly. However, there are two reasons why such approaches are limited within the technologically-evolving automotive domain. First, despite significant progress in the development of socio-technical systems theory and methods, the application of these methods is largely constrained to the existence of a current system. Second, there are few structured approaches for using the analysis results to support design. In MODAS, an attempt is made to overcome these challenges by developing and implementing a method for analyzing and designing a non-existent sociotechnical system—a highly autonomous truck. The starting point for MODAS is the Goals, Method, Observability, Controllability (GMOC) model (Sandblad, Andersson, Kauppi & Wikström, 2007). In MODAS we also consider safety in human-automation system interaction and identify suitable assessment methods via a systematic analyze of estimated situations, goals, actions and behaviors that are of high importance from a safety perspective. A summary of the project is provided.

Keywords: Autonomous vehicles, Cognitive Systems Engineering, Human Automation Interaction, GMOC

INTRODUCTION

There are four goals of the MODAS project. First, to create a method for future (non-existent) systems design. Se-

cond, to apply the method to develop a driver environment that facilitates the safe, efficient, and pleasurable supervision and control of a highly autonomous commercial heavy vehicle. Third, to develop a method for assessing the user acceptance of the Auditory Display components of this multimodal driver environment. Forth, to develop a method for assessing secondary task performance in highly automated commercial trucks. These four goals are addressed within some broadly defined 'future world' constraints that are described below, starting with the question of why develop autonomous vehicles.

Why Develop Autonomous Vehicles?

Before addressing this overarching question, it is first necessary to clarify terms. Consistent with Wood, Chang, Healy, & Wood (2012), in the current manuscript we have chosen to refer to "autonomous" vehicles, instead of "automated" vehicles, primarily because the former term is used more widely. In the current manuscript we seek to avoid a discussion on the meaning of automation, automatic, autonomous, and so on. Instead, consider the practical case of developing a driver environment for the supervision and control of commercial truck vehicle that can, as a minimum, take operational and tactical driving decisions and actions in the absence of any human control.

There are several immediate responses that come to mind when considering why invest in the development of autonomous vehicles, most notably, safety, fuel efficiency, emission reduction, and supporting the mobility of persons with disabilities Indeed, it is possible to consider autonomous vehicles as reducing much of the 'marginal social costs' associated with travel (user costs, infrastructure provider costs, infrastructure user costs, society in general costs [air pollution, noise pollution…]; Jansson, 1997). Regardless of the issue of cost, a relatively agreeable perspective is that autonomous vehicles is one way to increase safety, pleasure, and efficiency of transport in an increasingly complex transport system. Consider too that the increasingly complexity of modern transport systems is inevitable. Transport Systems are regularly considered as a typical example of a Complex System. A Complex System is an entity composed of multiple interacting parts where the interaction of these parts (or agents) lead to large scale behaviors that cannot be predicted on the basis of any specific agent (Mitchell & Newman, 2002). Some of the key characteristics of Complex Systems are that they are Adaptive, include Emergent behaviors (novel behaviors cannot be predicted on the basis of the systems' fundamental properties), Self-organizing, include Attractors (i.e. involves a recognizable dynamic state that may appear continuously), Non-linear (i.e. small changes in one part of the system can have disproportionate effects in other parts of the system), and includes Phase transitions (the system's behavior may change radically once a certain threshold or phase has been passed) (adapted from OECD, 2009). According to Arthur (Arthur, 1993), a complex system will increase in complexity as functions and modifications are added to that system to break through current limitations or to handle exceptional circumstances, and in general to adapt to a more complex world. Recent national, regional, and global initiatives, for example, aimed at reducing pollution (Euro 6 emission standards) or personal injury (zero fatalities in road transport by 2050; European Commission, 2011) are some of the modifications (c.f. Arthur, 1993) that will lead to increased complexity within the Transport domain. Although it is possible to put forward a strong theoretical argument for why transport systems complexity will increase, in the current manuscript we are more interested in the implications of this growing complexity on commercial truck drivers. To this end, it is possible to take a more practical examination of transport and technology trends and make conclusions about the effect of this evolving complexity on drivers. To understand the developing complexity from the driver perspective, three factors are introduced; (1) Increasing traffic growth, (2) Novel Intelligent Transport Systems (ITS) solutions, and (3) Developing vehicle technologies.

First, consider the expected increase in road transport. Within the EU 25, the number of passenger kms by car is expected to increase by 20% from 2005 to 2020 and by 34% from 2005 to 2030. Even greater increases are expected within freight transport—truck ton-kms will increases by 27% from 2005 to 2020 and by 40% from 2005 to 2030 (Hansen, 2009; estimates based on a baseline model of growth). A key challenge for the transport industry as a whole, is how to cater for this increased traffic density and at the same time maintain reasonable levels of throughput. With new road highway construction costing about 10M€ per kilometer (Bentzrød, & Hultgren, 2009), national governments have a vested interest in avoiding the construction of new roads and instead need to find alternative means of catering for this expected increase in traffic density. Solutions to this challenge are being addressed within the transport sector, that is, via intelligent transport systems. Second, consider specifically the developments within Intelligent Transport Systems. Intelligent Transport Systems (ITS) are the advanced applications and integration of telecommunications, electronics, and information technologies together with transport engineering which aim to provide innovative services relating to different modes and transport and traffic management. These applications should support the better informed, safer, more coordinated, and more intelligent use of transport networks (European Parliament, 2010). ITS solutions of various levels of sophistication (or complexity) can be found within transport

systems, or under investigation. Some relatively simple ITS solutions that are found in most larger cities include dynamic highway speed control—where the maximum highway speed varies to support better traffic throughput. A slightly more complex solution is dynamic lane control (e.g. three lanes into the city in the morning, and one lane out which reverses to three lanes out and one in during the afternoon). More complex again is the notion of dynamic routing, here, it is not just lanes that may or may not be available during certain times, but whole roads (Taniguchi & Shimamoto, 2004). The potential for ITS solutions will continue to grow as the technology and methods for interacting with 'big data' develop. Third, consider the technological development occurring within vehicles (a factor sometimes included within the broader view of ITS). For example, current sensors coupled with automatic emergency braking supports safer driving at reduced vehicle separation (both laterally and longitudinally). Indeed, a recent European Regulation (EC Regulation No. 661/2009; European Parliament, 2009) dictates that from the 1st November, 2013 all new vehicles in the EU must be fitted with Advanced Emergency Braking Systems—systems that can "automatically detect and emergency situation and activate the vehicle braking system to decelerate the vehicle with the purpose of avoiding or mitigating a collision" (EC Regulation No. 661/2009, p. 8). As sensor sophistication increases, coupled with the capabilities afforded by vehicle to vehicle/infrastructure communication (and vice versa) it is likely (or inevitable?) that lateral and longitudinal separation can be further decreased.

Taken together, the implication of the three factors described above on transport complexity from a driver perspective is that 'future' truck drivers will find themselves working in an environment with significantly more traffic than is currently the case. Additionally, on the basis of ITS solutions, from day to day a driver may not know which roads would be available for use, nor which speeds to travel at, nor which lanes to use. As a result of vehicle technology developments, drivers will encounter reduced lateral and longitudinal separation between vehicles while maintaining current (or potentially higher) speed limits. In this kind of environment, it is possible to envisage that full manual driving is extremely difficult, or impossible. Furthermore, full manual vehicle control may be not allowed. Thus we take the point of view that automation is a survival tool, without which driving may not be possible or allowed.

The role of the driver

Given this future context, we must next consider the (expected) role of the driver. On this topic, we can use some existing theory on vehicle control and couple this with emerging technology trends to make predictions about the role of the driver. First, consider Michon's (1985) theory on level of vehicle control (more recently developed further by, amongst others, Lee & Strayer, 2004). According to Michon's model, there are three levels of vehicle control; operational, tactical, and strategic. The operational level of control includes actions such as braking and acceleration, behaviors occurring between 0.5 and 5 seconds. On the tactical level, we have things such as changing lanes, negotiating traffic solutions, taking turns, etc. The tactical level includes behaviors occurring between 5 and 60 seconds. Finally, on the Strategic level, we have all the planning activities—these actions occur over minutes or days. A review of current technological trends suggests that technology development is currently very much towards autonomous control at the operational and tactical level. Furthermore, despite very significant advances in cognitive computing over the last several years, developing autonomous planning tools will require some very significant developments on data fusion using information from very diverse areas (e.g. weather, traffic, organizational, and other databases).

Thus, as illustrated in the Figure 1, within the MODAS project, we take the perspective that the role of the driver in the future environment becomes one primarily of strategic control. However, given the future environment described above, coupled with the inevitability of the driver remaining the final litigious controller of the vehicle (at least in the medium term), we expect that the driver will remain an actively involved in supervising their autonomous systems, the environment, and taking supervisory actions. Thus an additional role of the Driver will be to Monitor and Supervise their systems and the environment (and take corrective actions as required). Finally, in addition to Planning, Monitoring, and Supervising, we expect a third role of the driver—Resilience (c.f. Hollnagel, Woods, & Leveson, 2006). Following (an inevitable) but wholly unexpected event, it becomes the drivers' responsibility to bring their system back to some kind of normal state. This is particularly important when considering the future world as described (increased traffic density travelling faster) where the potential for loss of life following an accident is arguably greater than is currently the case.

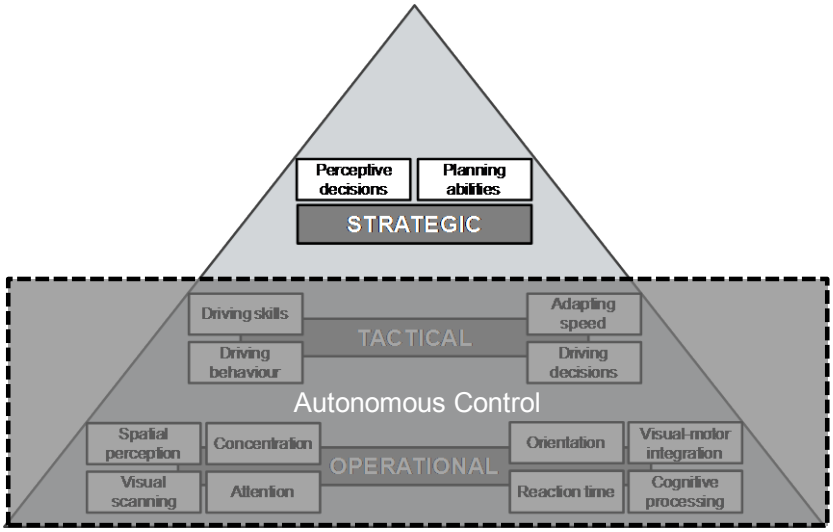


Figure 1. Adaptation of Michon (1985) showing levels of control (Operational, Tactical, Strategic) with Operational and Tactical actions under Autonomous (Vehicle) Control.

In summary, we envisage three key roles of the driver in a future scenario, Strategic Planning, Monitoring and Supervision, and supporting system Resilience. The implication of this is significant. If we consider these three components as the key parts of 'driving' in the future, then it is clear that a system that affords this kind of control can be significantly different to current vehicles. As will be elaborated upon further in the manuscript, we suggest that a system that affords this kind of control does not currently exist and consequently, current methods for iterative systems development are invalid. The fact that no system exists that could be said to afford the kind of control required in a future context is a main reason why we argue that a new method of systems development is required.

The nature of the future system

The next question to be addressed is how do we start developing a system that caters for the future traffic context as well as the wholly novel (expected) role of the driver. Here we make two arguments. First, we argue that that in moving forward with systems designs for highly autonomous vehicles we need to take a holistic approach that considers the changing nature of the driver-vehicle relationship (as explained above) as well as the driver and the vehicle as two parts of a joint team. Without making any reference to any specific theory regarding human-machine interaction (especially with respect to 'joint' socio-technical systems), we merely want to make the point that in the future system, where the vehicle itself posses a relatively high degree of 'intelligence' it is important to consider that both the driver and the vehicle must work together as a team to achieve the system goals. Finally, in the same way that large aircraft manufacturers are known to have clearly defined philosophies regarding the role of automation in their systems (Young, Stanton, & Harris, 2007), a clearly defined philosophy is needed to ensure a consistent application of autonomous technology (thus reducing driver confusion). In the text that follows we describe in detail three key theoretical parts of the MODAS project; the design method, the multimodal driver environment, and the challenges associated with assessing autonomous driving. Following this is a brief description of how the MODAS project is being implement. We conclude we a summary of the current state of the project.

SYSTEMS DEVELOPMENT

Current methods for complex system design, typically consider the user needs in the context of the current system. A standard user-centred approach to designing a new iteration of the current system normally involves some form of (user-oriented) system analysis, either descriptive (how users perform the task) or normative (how users *should* complete the task). These approaches involve interacting with end users to understand the current constraints to system efficiency and on the basis of the qualitative and /or quantitative data gathered, revisions to the system are proposed and implemented. Thus the new system improves system efficiency by removing the problems observed in the (now) old system. The automotive industry is currently experiencing a significant growth in the implementation of advanced automated driving technologies. When creating the control (driver) environment for interacting with these technologies, standard design methods are used. This means that the methods for interacting (accessing infor-

mation and manipulating system states) with the novel technologies are fitted into the legacy (existing) system.

There are two problems with this approach. First, predicting system wide effects when modifying some aspects of the system can be notoriously difficult (Parasuraman, & Riley, 1997). Furthermore, and as mentioned previously, when introducing automated technologies—which have the potential to significantly change the driver-vehicle relationship—these difficulties are exacerbated. Second, and given the point above, it is possible to consider system design of the future system, as wholly unique (thus removing the dependence on the legacy system). However, developing a new system in the absence of an existing system becomes more difficult (as descriptive analysis are impossible, and normative analyses questionable; Loukopoulos, Dismukes, & Barshi, 2001). What is needed is a structured method for designing future systems that removes dependence on the current system, and can take into account the unknown behaviours (human and technology) that can arise from the future system.

Currently, the preferred method for systems analysis in the absence of an existing system are formative analyses. As opposed to explaining how the system currently works (descriptive) or should work (normative), a formative approach describes how the system *could* work. Although some formative systems analysis approaches (e.g. Cognitive Work Analysis; Rasmussen, 1986; Vicente, 1999) do exist, the application of such approaches for wholly novel systems is difficult to validate. Furthermore, when the formative systems analysis is complete, there are few, if any, reliable and valid approaches for using the system analysis results for system design (Read, Salmon, & Lenné, 2012). Clearly, in the absence of an existing system, as can be the case when future technologies are considered, not only is the system analysis impossible, but system design becomes more of an art than a science.

One developing approach for future systems design is the Goal, Model, Observability, Controllability approach (GMOC; Sandblad, Andersson, Kauppi & Wikström, 2007; Jansson, Stensson, Bodin, Axelsson & Tschirner, 2014). The GMOC model defines the human control aspects associated with controlling dynamic systems. GMOC is used to describe and analyze the control tasks, design new control principles, design the user interface and decision support systems, and to support work organization and work place design. The new control strategy developed as part of the GMOC approach, "control-by-awareness", has shown to significantly improve the operational traffic planning and control processes. A key feature of GMOC as compared to other forms of system analysis is that the approach has an additional emphasis on system *design*. To support system design, GMOC, includes the following:

- Vision seminars, where the goal of the organization and promises from new technology are driving forces;
- Analyses of current work contexts and on-the-edge technology in driver environments, in order to separate current technology from intrinsic work constraints;
- Workshops around current and future automation as well as future interaction technologies;
- Design suggestions in terms of design principles and design concepts representing hypotheses about the relationship between the artefacts and human cognition and collaboration;
- Evaluation of prototypes in order to subject the design concepts to empirical testing and revision as long as the mutual shaping goes on between the artefacts and the actors in a specific field of practice.

In the current project we acknowledge that designing for the evolving supervisory control modes change the required interaction principles. However, it is a mistake to believe that higher levels of automation can be introduced as a simple substitution of machines for people (Dekker & Woods, 2002). Supervisory control of a future vehicle means the driving task changes from being a primarily operational to a tactical and even a strategic task. In MODAS we acknowledge that designing for cognitive control introduces new challenges that are not present to the same degree in the existing driver environment. One way to approach and understand the evolving driver task is to study and interpret it as a Dynamic Decision-Making task (DDM; Brehmer, 1992). GMOC can be seen as the applied version of the DDM-approach since it is based on the same requirements for control as those defined by DDM. It thus makes sense to investigate the drivers' actions using the following general problem formulation:

- How can the drivers' mental model development and goal formulation processes be supported by enhanced observability and augmented control functions?

One consequence of this general problem statement is that it is necessary to study, not only *what* drivers do when they perform judgments and make decisions, but also *how* they do it. This means identifying the drivers' strategies in different situations and contexts, and supporting the development of mental models in relation to these different strategies. One way to conduct such strategy analyses via the method developed by Hassall and Sanderson (2012).

MULTIMODAL DISPLAY DEVELOPMENT

A key goal of the future driver environment to be developed within MODAS is that the interface supports the driver acquire an accurate understanding of the state of their vehicle and of the external environment. It is presumed that in order to safely operate the vehicle some understand how the vehicle works is necessary. At a very superficial level it is necessary to understand that the accelerator is used to control acceleration, that the gear subsystem is used to control direction (forward/backward) , and the braking subsystem reduces motion. During operation, drivers receive information about the various subsystems (via the speed display, the gear display, the sound of the engine...) and use this to obtain an overall understanding of the vehicle. Thus in order to drive a vehicle, it is necessary to have some understanding of how the vehicle works as well as the current state of the vehicle, thus a "mental model" of the vehicle is required. More formally, a mental model can be defined as a representation of the system (or task) that is formed by the user, that is based on both previous experiences and current observations, which provides most (or all) of their subsequent system understanding and consequently dictates the overall level of task performance (Wilson & Rutherford, 1989). In contrast to private drivers, for professional heavy vehicle drivers, where overall vehicle efficiency is a key task requirement and the negative effects of safety mishaps are exaggerated, a more complete vehicle mental model is required.

A more concrete example from the driving is: based on a driver's previous experiences (the last time engine temperature increased, the cause was a split hose from the radiator) and current observations (my engine temperature is increasing now) an understanding is formed about the system (the cooling system has been comprised, probably due to a split hose) which dictates task performance (it is best to stop and attempt to fix the damaged hose). In this example, the driver has a formed a mental model of the vehicle mechanical subsystems (the coolant sub-system) that is formed by former and current observations. In this example, however, if the driver had received additional information, such as coolant pressure (which may have been normal), the driver would have formed an alternative mental model of the system state and implemented a different solution. Thus for safe and efficient control of (current) vehicles, drivers need an accurate mental model of the (mechanical) vehicular subsystems (e.g. engine subsystem, cooling subsystem, braking subsystem...). With comparatively simple mechanical subsystems the options for presenting this information are somewhat limited. However, with increasingly complex systems, the options for information presentation grow. Indeed, the key departure from traditional systems technology and display formats, to future vehicle systems is that with older vehicles, information was presented about the vehicle's mechanical subsystems, however, with future systems, drivers will require an accurate mental model of the vehicle 'cognitive' subsystems. It is one thing to understand, from a mechanical perspective, why the vehicle is behaving in a certain way, but it is vastly different to understand why the vehicle automation is behaving in a certain way.

From a theoretical perspective, there is much work describing how to identify the information requirements for supporting operator mental models, as well as accompanying models and theories of display design. As Moray et al. (1993) stated, "Even experts can only exercise their skills and expertise optimally if the pattern in which information is displayed matches their models of the dynamics of the problem". Whereas Moray's comments were directed primarily towards a design approach, we consider the statement from a more applied perspective. To accompany the overall design methods development, in the current research we will investigate the affordances of multimodal technologies to support the driver acquire an accurate mental model of the vehicle 'cognition'. Two complimentary lines of work will be completed; how to use sound and visual information to assist the driver acquire an accurate mental model of the vehicle cognitive subsystems. In other words, whereas the mental models requirements will be defined under the scope of systems development and user needs, the iterative development of design solutions will be investigated within the scope of multimodal display development.

Although supporting driver understanding of vehicle automation is essential to the present work, there are additional challenges for display design that will be considered as means of supporting efficient, safe, and pleasurable driver performance. As the vehicle will have operational and tactical control in many driving situations, its interface should be designed to assist drivers in making appropriate strategic decisions, and should enable them to override automation when necessary. Drivers must be informed not only about the state and relevant intentions of the vehicle, but also about events and conditions in the environment that are relevant for making decisions. For instance, displaying road conditions or weather on available routes to a destination may assist drivers in selecting the most appropriate route. Thus, work will be conducted to investigate how to design displays that support drivers in making strategic decisions. In addition, the project will focus on how to design a multimodal interface that assists drivers to easily asses and prioritize information, and presents the information in an appealing (high user acceptance) fashion.

Furthermore, while the present work focuses primarily on fully automated driving, autonomous commercial vehicles will likely require manual driving in some situations. In demanding contexts, such as extreme driving environments (e.g., forests and mountain roads) and circumstances (e.g., snowstorms), reliable fully autonomous driving support may simply not be available. Thus, technological solutions should not only support fully autonomous driving, but also safe and efficient manual (less automated) driving when necessary. By considering situations in which the driver has operational and tactical control (longitudinal and/or lateral) of the vehicle, solutions will be developed that support the drivers' decision-making process at these control levels. The vehicle may then provide perceptual, analytical and even decision support. For instance, during poor visibility conditions, the interface might display distances to other road users and obstacles, and indicate when safety margins are optimal from a fuel-efficiency perspective or too low from a safety perspective. In the present work, we will examine how display technology and solutions for highly autonomous vehicles can be utilized to support manual driving when necessary.

ASSESSING AUTONOMOUS DRIVING

As mentioned previously, new automation technologies will change the role of the driver resulting in new driver behaviors and errors. As Sarter, Woods, and Billings (1997) pointed out, automation can surprise operators via unexpected problems and consequences. Earlier examples from aviation and industry, highlight the difficulties in predicting the effects of large technology shifts. However, even an inaccurate prediction is at least one step towards building knowledge. In this project we aim to build a model for assessing automated driving.

Today, driving could be primarily described at a manual task, at least for most drivers. Still, much automation exists and assists the driver. Some automation systems are so common that we forget to view them as automated. One example is the fuel meter. Here, the vehicle automatically assists the driver in measuring the fuel and provides a warning (e.g. a red light) when fuel level is low. Applying Parasuraman, Sheridan and Wickens' (2000) description of automation types and levels, consider the task of travelling from point A to B with sufficient fuel. For most current vehicles, we could say that this task has the following automation support; a medium level of information acquisition (the car provides automatic measuring but the driver reads the display), a medium level of information analysis (the driver analyzes the situation but the car informs the driver when fuel level is critical), and finally a low level of decision selection and action implementation (the driver decides when and where to re-fuel and does this manually). Hence, driving could already be viewed as a semi-automated task. Using the levels of automation described by Sheridan (2011), current driving could be viewed as various actions where the car either offers no assistance (level 1) or executes a suggestion if the human approves to (i.e., cruise control, level 5). At the upper end of the scale, level 10, the car takes all decisions "ignoring the human". At a high automation level (e.g. 10) there is no need to assess driver behavior because the drivers' role reduces to being a passenger. On the contrary, if there is any risk of automation failure that requires a co-driver ready to take over the driver behavior becomes of high importance—possibly of even greater importance than is currently the case. Today the driver is active (or at least should be) and should have a level of situation awareness, alertness and attendance that supports (at least) operational control on the sub-second level. If new automation systems reduce the alertness of driver, reaction times will increase, comprising safety. Our suggestion is that, as long as vehicles are equipped with a steering wheel and a brake, and that driver 'driving' is needed, standard measures of driving behavior remain relevant. According to our view, the issue is not which assessment measures must be removed, but rather which must be added.

In an attempt to identify the future driver behaviors that must be assessed, we will make a systematic analysis of estimated future situations, driver goals, and driver actions that build a foundation for predicting driver behavior (Figure 2). A description of a specific situation makes it possible to predict driver goals linked to that specific situation. The driver goals then acquire driver actions to be fulfilled. When the driver actions are defined, then acceptable behavior can be specified where acceptable behavior is defined as performance measures (e.g. time on task and error rate) as well as measures such as situation awareness, fatigue and mental workload. There are likely unexpected driver behaviors for which effective assessment methods are currently unknown. The need for new assessment methods has followed the development of new complex sub-systems in vehicles and could be expected to grow even more with automation (Figure 3). As complexity grows, the assessment methods must grow. New automation technologies might reduce distraction but pose new problems. These problems may be avoided with smart systems giving the right information to the driver at the right time. Automation comes with an expectation of reducing accidents and environmental influence. However, if the systems are not well designed there is a risk that the errors will be reduced in numbers but increase in severe outcomes. Therefore, this new area of assessing automation in vehicles will need extensive research. The model built in this project is believed to become an important starting point but

will not give the whole picture.

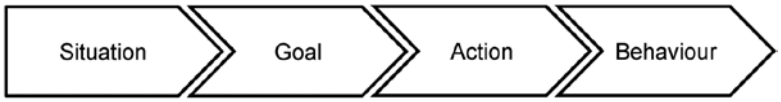


Figure 2. The planned method for predicting (future) desirable driver behavior.

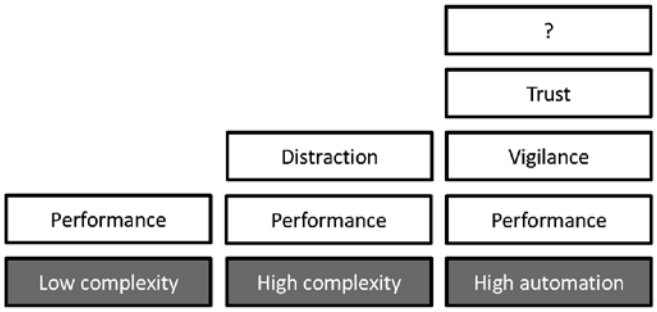


Figure 3. The scope of assessment measures grows with increasing system complexity.

ASSESSING AUTONOMOUS DRIVING WITH DRIVING SIMULATOR

Currently, the Lane Change Test (LCT) is the ISO standard for assessing the influence of secondary tasks on driving performance (Mattes, 2003). The LCT has the advantage of being relatively inexpensive and easy to use and has a high level of control. This makes the method suitable in early product development stages and in research when new areas are studied. The LCT is used to measure driving performance in terms of deviation from a normative driving route during the completion of other in-vehicle tasks (DaimlerChrysler AG, 2004). However, the value of the LCT for assessing future technologies is limited when automated driving features become more prevalent. For example, the LCT is useless for assessing secondary task performance during instances of autonomous lane changing. However, in a shift from automated to manual, control lane deviation measures become highly important. Driver distraction may affect driving in several ways, such as delayed reaction time (Alm & Nilsson, 1995; Treffner & Barrett, 2004) or decreased hazard detection (Strayer & Johnston, 2001; Reyes & Lee, 2008). These effects might increase when drivers are in supervisory rather than active control and suddenly need to take over the (manual) driving task. New technology solutions may also add effects that the current LCT method fails to capture (Engström & Markkula, 2007). Additional metrics measuring specific behaviors, and not only a general driving deviation, seems necessary (Rognin, Alidra, Val, & Lescaut, 2007), especially when the driver role changes due to automation. Previous research has been conducted that investigated what are the additional performance metrics can be obtained from the LCT with minimal methodological changes. Some performance measures identified include being the missed signs and erroneous lane changes (Bruyas et al., 2008; Young, Lenné, & Williamsson, 2011). However, given that the LCT was developed specifically to assess deviations in driving performance, the inclusion of these additional metrics must be tested. Indeed, recent work has been completed that attempted to specifically develop and test additional LCT metrics (Grane, 2012; Grane & Bengtsson, 2013). Whereas the development work on the LCT described by Grane (2012) moves towards validating some additional LCT performance measures that may be useful for assessing the impact of future technologies on 'driving' performance, the author noted that the measures investigated may reflect experimental design, rather than task demand. From a safety perspective, careful adjustments and further tests are required.

In the current project, we move forward from Grane (2012; Grane & Bengtsson, 2013) and aim to define a more generally useful set of performance measures that can be extracted from the LCT to assess the impact of future technologies on 'driving' performance. This work is necessary for assessing the impact of the to-be-developed system. Outcomes of this research will include recommendations for revisions of the LCT to remain valid for future technologies. This evolving methodology will be used to assess the driver environment underdevelopment in an iterative way with results used to revise both the driver environment and the assessment method.

PROJECT IMPLIMENTATION

Work will be completed within six key research and development workpackages The six key workpackages (two to seven) are described below and the relationship between workpackages is summarized in the PERT chart below.

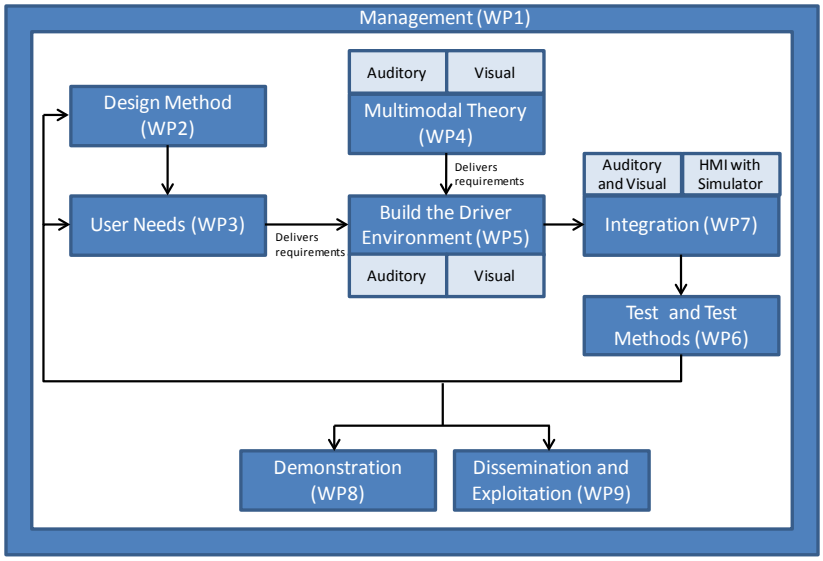


Figure 4. PERT chart showing relationship between Workpackages.

First, within the Design Method Workpackage (WP2), the overall design method will be developed. In WP2, a method for developing future systems is defined. The design method, a variant of the GMOC model will be adapted and refined for MODAS on the basis of the User Needs identified in the next workpackage. The key activity of the Design Methods workpackage will be to define the process of systems development that will be adhered to throughout the project. A challenge in developing this design method will be to identify a design process that considers future (currently unavailable) technologies. A good example of how a similar challenge was completed previously is Olsson and Jansson (2005) who successfully integrated future technologies into working practice. The methods of Olsson and Jansson (2005) will be extended in the current project.

Second, the Driver Needs workpackage (WP3) will occur closely with the Design Method Workpackage and aims to introduce the driver perspective to development. Whereas the Design Method Workpackage defines a general method for systems development, the adaptation and implementation of this methodology within the commercial truck domain occurs within the Driver Needs workpackage. This workpackage is thus dedicated to understanding (future and emerging) driver needs. The methods by which the driver needs are probed will be identified in Design Methods, but implementation (data collection and analysis) occurs within Driver Needs. Thus within this workpackage the information required to assist the driver develop an accurate mental model of the system state will be identified.

Third, in the Multimodal Affordances workpackage (WP4), the theoretical knowledge required for prototype design, integration and testing will be developed. The work here is divided into three main areas. First, we will investigate various possible technological solutions for a multimodal display. The work in this area will focus on the affordances of technologies and will occur within two modality directions: auditory and visual. The results will be used to support the process of decision-making on technological investments for the multimodal display. Secondly, literature studies will be conducted to support the multimodal display design work. The research will focus specifically on the vehicle and environment information identified in the Design Methods workpackage and the Driver Needs workpackage. The outcomes will be used to assess which driving-related information could be best expressed aurally or visually, as well as determine the most promising approaches for doing so. Finally, we will examine available assessment technologies, which will support decisions regarding investments in simulator test equipment.

Fourth, in the Driver Environment Development workpackage (WP5), the driver environment will be constructed using the methodology developed in the Design Methods work package, on the basis of the Driver Needs, and using the theoretical knowledge from WP4. The information display will be developed in parallel with the visual displays. Included within this WP will be the initial design, building and evaluating of the auditory and visual information

displays. During the design process, both display types will be evaluated and refined in an iterative way. An iterative approach is necessary given that the theoretical knowledge identified in WP4 is more general, and may not have been confirmed within a relevant user context and with appropriate end users. Carelessly implemented audio and visuals displays are known to cause annoyance, which can significantly impact the acceptance of the entire interface. The iterative process will contribute to a more context-specific and precise understanding of how the sounds and graphics should be designed and presented within the multimodal environment, in order to work effectively and be accepted by the users. Furthermore, as part of the iterative testing process, a new tool for audio display evaluations will be developed. Established methods for display evaluations already exist within the visual domain. However, while auditory cues are common components in various types of driver assistance systems, methods are lacking for how to evaluate the sounds in order to improve their design. Developing and applying a useful evaluation method for the auditory information display is crucial for efficient testing within the iterative design process.

Fifth, in the Integration workpackage (WP7) the auditory and visual displays developed in the Driver Environment Development workpackage will be integrated into Scania's high fidelity truck driving simulator. This workpackage includes two components, first, the visual and auditory displays will be integrated into a single driving environment, and second, the environment will be integrated into Scania's truck driving simulator. Key to this work is the development of a Multimodal Display Engine. Additionally, within this work package, the simulator scenarios necessary for implementing the Design Method will be developed, as well as those needed for Test Methods (WP6).

Sixth, in the Test and Test Methods workpackage (WP6) we systematically investigate the estimated situations, goals, actions and behaviors that might be of importance for future autonomous vehicles and safety. The results will give input to how the current ISO standard for assessing in-vehicle technologies (the LCT) could be extended to remain relevant for future technologies. In this WP we develop the test methodology for assessing the impact of the to-be-developed system. Outcomes here will include recommendations for revisions of the LCT to remain valid for future technologies. This evolving methodology will be used to assess the driver environment under development in an iterative way with the results influencing both system design and test method. A key difference between the work in this WP and that within WP2 is that developing the test method is conducted primarily with the involvement of domain experts (whereas system development is conducted with the primary involvement of potential system users).

It is important to note that within each WP, where relevant, an internal iterative process is applied. For example, within WP5, multiple iterations of display design occur with ongoing user assessment. In addition to regular deliverables from each Workpackage, three key Milestones have been defined. Milestone 1 is the 1st User Testing – where initial, conceptual auditory and visual (not integrated with each other or the simulator) displays can be presented to the user population. Essentially, these displays are to be developed on the basis of the first iteration of the Design Method. Milestone 2 is the 2nd User Testing, where the integrated Auditory and Visual displays are tested in the driving simulator and finally, Milestone 3 is the Final User Testing.

CURRENT STATE OF MODAS

The MODAS project has recently passed Milestone 1 (1st User Testing using non-integrated Auditory and Visual Displays). A brief summary of the work completed up to Milestone 1 is described. An initial proposal regarding the GMOC method was provided (by WP2) that including a first plan on how to implement the GMOC method within MODAS. Data collection with professional truck drivers (interviews and observations) and expert engineers (interviews) were then conducted (within WP3) to complete a Work Domain Analysis and a Control Task Analysis (Vicente, 1999). These analysis were used to identify the various levels (abstractions) of system constraints as well as critical situations for which further technology development was needed. A series of Hierarchical Task Analyses were conducted (also within WP3) for around fifty common driving activities. These activities were selected on the basis of the Control Task Analysis which identified these activities as being important to the overall goal. A series of GMOC interviews (a joint WP2/WP3 activity) with professional drivers was conducted using synthesized observation data to probe for the driver mental models. Follow up interviews were conducted to better understand what information is used (and where it is obtained) to build these mental models. Next, a series of interviews were conducted to identify driver preferences regarding automation and how it can be used to reach the high level system goals (identified from the WDA). In addition to identifying driver automation preferences, a literature review was conducted to identify the expected advantages of automation from a theoretical perspective (WP4). During this data collection, a series of literature and state-of-the-art reviews (WP4) were conducted to identify software and hardware requirements for building a prototype system of the type being investigated, as well as the requirements for assessing autonomous 'driving' performance. All this data was used as input for creating a set of visual and auditory

displays (within WP5) that were then used for a comprehensive week of user testing. The key topics investigated during this week of user testing were to what extent the visual displays can (1) support operational, tactical, and strategic vehicular control, (2) support the development of the driver mental models, (3) support current, manual driving, and (4) are liked/accepted by the drivers. Regarding the auditory displays, drivers reviewed the semantic connection between the auditory displays and the intended object of representation. Current activities include the iterative development of the auditory and visual displays via more detailed user involvement, as well as a Strategies Analysis. Additionally, the planned method for identifying future driver behaviors (Figure 2) is being implemented.

CONCLUSIONS

The MODAS project has positioned itself at the cusp of the burgeoning field of autonomous vehicles. Several novel approaches to the problem of 'human-automation interaction within future vehicles' are taken within the project. First, the project takes a comprehensive approach to systems development, investigating work constraints, technology and environmental trends, and a systematic approach for predicting future driver behaviors. Second, the project has significant user involvement throughout, from concept generation to final testing. Third, the project includes multiple rounds of iterative development, with user participation tailored for each stage of iterative testing. The key challenges associated with the project are eloquently summarized in Woods (1998) where he discusses the interplay between experimenter-as-designer and designer-as-experimenter—at what point does the project investigate new designs, compared to new artifacts that influence cognition and collaboration? Indeed, the MODAS project, to some extent, is an embodied attempt to solve the four questions considered by Woods (1998) to be pertinent to the 'designer-as-experimenter; (1) how to close the analysis-design gap, (2) how to use knowledge from other [application] domains, (3) what is the role of innovation, and (4) balancing short term [project] deadlines, with long term [product development] advantages. To date, the project has generated a number of novel visual and auditory displays for supporting safety, efficiency, and pleasure when controlling a highly autonomous vehicle. The final prototype is due for completion at the end of 2014 and the project is on track to meet all stated objectives.

REFERENCES

Arthur, W.B. (1993). Why do things become more complex? *Scientific American*, May, 1993.

Alm, H., & Nilsson, L. (1995). The Effects of a Mobile Telephone Task on Driver Behavior in a Car Following Situation. *Accidental Analysis and Prevention, 27*(5), 707-715.

Bentzrød, S.B., & Hultgren, J. (2009, Mar 19). Så mye motorvei får vi. Retrieved June 1st from http://www.aftenposten.no/nyheter/iriks/article2970807.ece#.UvtuBfldV8G

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica, 81*(3), 211-241.

Bruyas, M.-P., Brusque, C., Tattegrain, H., Auriault, A., Aillerie, I., & Duraz, M. (2008). Consistency and sensitivity of lane change test according to driving simulator characteristics. *IET Intelligent Transport Systems, 2*, 306-314.

DaimlerChrystler AG, Research and Technology. (2004). Lane Change Test – User Guide 1.2. Retrieved October 12, 2012, from: http://people.usd.edu/~schieber/pdf/LCT-UserGuide.pdf

Dekker, S.W.A., Woods, D.D. (2002). MABA-MABA or abracadabra? Progress on human automation co-ordination. *Cognition, Technology & Work, 4*, 240–244.

European Commission. (2011). Roadmap to a single European transport area – Towards a competitive and resource efficient transport system [White Paper, COM(2011) 144 final]. European Commission, Brussels.

European Parliament. (2009). Regulation (EC) No 661/2009 of the European Parliament and of The Council Concerning type-approval requirements for the general safety of motor vehicles, their trailers and systems, components and separate technical units intended therefore.

European Parliament (2010). Directive 2012/40/EU of the European Parliament and of The Council. On the Framework for deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport.

Grane, C. (2012). Methodological considerations when adding new measures to the Lane Change Test. In *Proceedings of the 44th Annual International Conference of the Nordic Ergonomics Society* (NES2012), Stockholm, Sweden.

Grane, C., & Bengtsson, P. (2013). Driving performance during visual and haptic menu selection with in-vehicle rotary device. *Transportation Research Part F, 18*,123-135.

Hansen, C.O. (2009): Report on Scenario, Traffic Forecast and Analysis of Traffic on the TEN-T, taking into Consideration the External Dimension of the Union – TRANS-TOOLS model version 2: Calibration and forecasts 2020 and 2030, Funded by DG TREN, Copenhagen, Denmark.

Hassall, M.E., & Sanderson, P.E. (2012). A formative approach to the strategies analysis phase of cognitive work analysis. *Theoretical Issues in Ergonomics Science*, *x*, 1-47.

Hollnagel, E., Woods, D.D., & Leveson, N. (Eds.). (2006). *Resilience Engineering: Concepts and Precepts*. Aldershot, UK: Ashgate.

Jansson, J.O., (1997). Theory and practice of transport infrastructure and public transport pricing. In de Rus, G. & Nash, C.,

(Eds.). *Recent Developments in Transport Economics*, Ashgate, Aldershot, UK.
Jansson, A., Stensson, P., Bodin, I., Axelsson, A. & Tschirner, S. Authority and level of automation: Lessons to be learned in design of in-vehicle assistance systems. In *Proceedings from HCI International Conference*, *Special session on In Vehicle Assistance Systems: Standalone to Social*, Heraklion, Crete, Greece, June 22nd-27th, 2014.
Lee, J.D., & Strayer, D.L. (2004). Preface to the special section on Driver Distraction. *Human Factors, 46*(4), 583-586.
Loukopoulos, L.D., Dismukes, R.K., & Barshi, I. (2001). Cockpit interruptions and distractions: A line observation study. In *Proceedings of the 11th Symposium on Aviation Psychology*, Columbus, OH, March.
Mattes, S. (2003). The Lane-Change-Task as a tool for driver distraction evaluation. In H. Strasser, K. Kluth, H. Rausch & H. Bubb (Eds.), *Quality of Work and Products in Enterprises of the Future* (pp. 57-60). Stuttgart, Germany: Ergonomia Verlag.
Michon, J. A. (1985). A critical view of driver behavior models: What do we know, what should we do? In L. Evans & R. C. Schwing (Eds.), *Human behavior and traffic safety* (pp.485–520). New York: Plenum.
Mitchell, M., & Newman, M. (2002). Complex Systems Theory and Evolution. In M. Pagel (ed.). *Encyclopedia of Evolution*, New York: Oxford University Press.
Moray, N., Jones, B.J., Rasmussen, J., Lee, J.D., Vicente, K.J., Brock, R.. et al. (1993). *A performance indicator of the effectiveness of human–machine interfaces for nuclear power plants* (NUREG/CR-5977). Urbana, IL, USA: University of Illinois, Urbana-Champaign, Department of Mechanical and Industrial Engineering; 1993.
OECD. (2009). Applications of Complexity Science for Public Policy: New Tools for Finding Unanticipated Consequences and Unrealized Opportunities. Paris OECD.
Olsson, E., & Jansson, A. (2006). Work on the bridge – studies of officers on high-speed ferries. *Behavior & Information Technology, 25*(1), 37-64.
Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors, 39*(2), 230-253.
Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 30*(3), 286-297.
Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction*. Amsterdam: North Holland.
Read, G.J.M., Salmon, P.M., & Lenné, M.G. (2012). From work analysis to work design: A review of Cognitive Work Analysis design applications. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*(1), 368-372, Boston, MA, USA,
Reyes, M. L., & Lee, J. D. (2008). Effects of cognitive load presence and duration on driver eye movements and event detection performance. *Transportation Research Part F, 11*, 391-402.
Rognin, L., Alidra, S., Val, C., & Lescaut, A. (2007). Occurrence of secondary tasks and quality of lane changes. In D. Harris (Ed.). *Engineering Psychology and Cognitive Ergonomics*, HCII 2007, LNAI 4562, 397-406. Berlin Heidelberg: Springer-Verlag.
Sandblad, B., Andersson, A.W., Kauppi, A. & Wikström, J. (2007). Implementation of a Test System for Evaluation of New Concepts in Rail Traffic Planning and Control. In J. Wilson., B. Norris., T. Clarke., & A. Mills. (Eds). *People and Rail Systems. Human Factors at the Heart of the Railways*. Hampshire, UK: Ashgate Publishing.
Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation Surprises. In G. Salvendy (Ed.), *Handbook of Human Factors & Ergonomics, second edition* (pp. 1-25). Wiley.
Sheridan, T. B. (2011). Adaptive automation, level of automation, allocation authority, supervisory control, and adaptive control: Distinctions and modes of adaptation. *IEEE Transactions on Systems, Man, an Cybernetics – Part A: Systems and Humans, 41*(4), 662-667.
Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. *Psychological Science, 12,* 462-466.
Taniguchi, E., & Shimamoto, H. (2004). Intelligent transportation system based dynamic vehicle routing and scheduling with variable travel times. *Transportation Research Part C: Emerging Technologies, 12*(3-4), 235-250.
Treffner, P.J., & Barrett, R. (2004). Hands-free mobile phone speech while driving degrades coordination and control. *Transportation Research Part F: Traffic Psychology and Behaviour, 7*, 229-246.
Vicente, K. J. (1999). *Cognitive Work Analysis: Towards Safe, Productive, and Healthy Computer-based Work*, Mahwah, NJ: Erlbaum.
Wilson, J.R., & Rutherford, A. (1989). Mental Models: Theory and application in Human Factors. *Human Factors, 31*(6), 617-634.
Woods, D.D. (1998). Commentary Designs are hypotheses about how artifacts shape cognition and collaboration. *Ergonomics, 41*(2), 168-173.
Wood, S.P., Chang, J., Healy, T. & Wood, J. (2012). The potential regulatory challenges of increasingly autonomous motor vehicles. 52nd Santa Clara Law Review, 4(9), 1423–1502.
Young, K.L., Lenné, M. G., & Williamsson, A.R. (2011). Sensitivity of the lane change test as a measure of in-vehicle system demand. *Applied Ergonomics, 42*, 611-618.
Young, M. S., Stanton, N.A., & Harris, D. (2007). Driving automation: learning from aviation about design philosophies. *International Journal of Vehicle Design, 45*(3), 323-338.

End-users' Acceptance and Use of Adaptive Cruise Control Systems

Niklas Strand[a,b], I.C. MariAnne Karlsson[a,c] and Lena Nilsson[a,b]

[a]SAFER, Vehicle and Traffic Safety Centre at Chalmers, Sweden

[b]VTI, Swedish National Road and Transport Research Institute

[c]Design & Human Factors, Chalmers University of Technology

ABSTRACT

An online survey was distributed to end-users of adaptive cruise control (ACC). In total 90 drivers answered the questionnaire, which covered e.g. ACC usage and how such use affects driver behaviour. According the responses: the ACC is used primarily on roads with higher speed limits; the end-users trust the system even though it has some functional limitations; they have a very positive attitude towards the system; and positive effects on comfort and safety are observed including reduced inclination to overtake and increased compliance with speed regulations.

Keywords: Adaptive Cruise Control, Acceptance, Assessment, Driver behaviour

INTRODUCTION

Advanced driver assistance systems (ADAS) in vehicles are associated with safety and comfort benefits for the driver (Brookhuis, de Waard, and Janssen, 2001). Adaptive Cruise Control (ACC) is an ADAS that controls speed and distance to vehicles ahead based on speed and time-gap settings (for a description of ACC systems see e.g. Naranjo, González, Reviejo, Garzia, and de Pedro, 2003, or Stanton and Young, 2005). Several studies concerning ACC systems have been carried out. A majority of these studies has however not focused actual users, instead car drivers in general have been appointed as participants. Such a recruiting strategy has advantages when addressing certain research questions, while for other questions knowledge and experience of a specific system are essential. In comparison with studies not including actual users studies including them are scarce. There are nevertheless a few exceptions including studies by Strand, Nilsson, Karlsson, & Nilsson (2010); Larsson (2012); Piccinini, Simões, and Rodrigues (2012), and Sanchez, et al. (2012) which all focus actual end-users and their assessment of and behavioural adaptation to the systems.

Behavioural adaptation (BA) is a psychological issue that explains behaviours that occur in response to change in a road-vehicle-user system (OECD, 1990). (For a full account of BA see the review by Saad, et al. (2004).) Strand et al. (2010) explored end-user experiences of ACC by means of focus group interviews and the study by Piccinini et al. (2012) used a similar approach. Both studies showed that, overall, end-users were satisfied with their ACC systems and that there are positive effects in terms of convenience as well as safety. Nevertheless, some negative effects were also identified. Strand et al. (2010) revealed that there are users carrying mental models, which to some extent can be considered rudimentary. Mental models can be viewed as our understanding of objects and processes that guide our interaction with them (Bernstein, et al. 2003) and proper metal model of system functionality are

important to cope with for instance mode errors (Sarter and Woods, 1995). Another finding was that the drivers had experienced situations, such as roundabouts, in which ACC functionality could be improved. Negative aspects related to ACC use concluded in the study by Piccinini et al. (2012) were that end-users were not fully aware of potential critical situations when driving with ACC and that improper usage of ACC occurred. They also found that tasks distracting driving were more frequent when driving with ACC, compared to manual driving. The study by Larsson (2012) reported on a survey to 130 ACC users. In this case the results indicated that use experience is crucial for the understanding of functional limitations of the ACC.

The report by Sanchez, et al. (2012) addresses user acceptance of ACC. The results are based on a survey to 227 participants in the EU-project euroFOT (www.eurofot-ip.eu). The majority was drivers of passenger cars equipped with ACC, but the study also included those who drove trucks equipped with the same system. The results suggest that acceptance (defined as perceived usefulness and driver satisfaction) was very high and, in addition, stable over time. The ACC was rated most useful in normal traffic on motorways. Furthermore, ACC increased perceived comfort (stated by 80 %) and safety (stated by 94 %).

The main purpose of this study was to follow up on earlier studies and investigate end-users' experiences of ACC with regards to drivers' use and acceptance of ACC; usability assessment; and perceived influence on driving behaviour. An additional aim was to investigate if these aspects were affected by ACC use experience (self-rated).

METHOD

Data collection

Data was collected by means of an on-line questionnaire, set up with SPSS DimensionNet software, version 5.5, patch level 3 (SPSS Ltd., 2008), and distributed by e-mail. The questionnaire included slightly more than 70 questions covering the topics: experience, usage pattern, acceptance and perceived usability. For acceptance, a new acceptance scale was used, SKAS (Strömberg & Karlsson, 2012). The scale includes 20 items covering four areas:

- *Trust & Control* attempts to capture how secure the user feels with the system. It consists of three items capturing aspects of the perceived technical reliability of the system, whether the user thinks that the information that the system gives or the action it takes can be trusted, and whether the user feels in control of the system.
- *Perceived benefit* aims to find whether the user perceives the system as something useful and as something that provides benefits (in terms of convenience, joy, efficiency etc.) to them in the task they are trying to perform.
- *Perceived effort* tries to identify the effort the users feel they have to put into gaining the benefits. It draws on classical usability in terms of ease of use, logic and coherence, etc.
- *Compliance* tests more abstract level aspects of acceptance. The dispensability item aims to capture whether the users feel that the problem that the product is trying to solve actually is a problem that needs solving, and the appropriateness item aims to capture whether the product is a suitable way of solving that problem.

All but one question in the questionnaire were closed-ended. The major part of questions was of a Likert type (Likert, 1932) with five response categories, including a neutral category. There were also a few questions with rating scales, from low to high (1-5). In SKAS the response scales were instead of a semantic differential type. The semantic differential scales used bipolar categories, with one positive and one negative pole. (See for instance McQueen and Knussen (2006) for a brief overview of rating scales, Likert scales, and semantic differential scales.)

Analysis

Statistical analyses were conducted with IBM SPSS statistics software, version 21 (IBM Corp., 2012). First a descriptive data analysis was conducted. This was followed by a Mann-Whitney test in which driving context (mainly within urban areas and mainly outside urban areas) was used as grouping variable. Only significant tests are presented in the results section.

Participants

A list of 414 end-users' of ACC was handed over by Volvo Cars Corporation. They were approached by e-mail and asked to answer the on-line survey: 90 end-users of ACC completed it. However, three of them stated that their cars were either not equipped with ACC. These were excluded from the analysis.

The respondents' ages varied between 31 and 76 years (M = 53.18, SD = 10.69, n=87). A majority had completed higher education (higher than upper secondary school). All of them had a valid driver's license and had held them for between 13 and 58 years (M = 34.75, SD = 10.69, n=87). The majority drove more than 25,000 kilometres per year. (See table 1 for an overview of participants' characteristics.)

Table 1: Participants Characteristics (n= 87)

Characteristic	***n***	**%**
Gender		
Men	78	89.7
Woman	9	10.3
Highest education level completed		
Compulsory school	3	3.4
Upper secondary school	29	33.3
Higher education: university college; university	49	56.3
Other education after upper secondary school	6	6.9
Annual mileage (kilometres)		
5,001–10,000	1	1.1
10,001–15,000	9	10.3
15,001–20,000	17	19.5
20,001–25,000	15	17.2
>25,000	45	51.7

About half of the participants were primarily engaged in non-professional driving, and one third drove equal amounts professionally and non-professionally. The remaining drove mainly in a professional role. Furthermore, roughly equal amounts of the drivers drove mainly within urban areas and outside of urban areas respectively. (See table 2 for overview of responses regarding context of driving.)

Table 2: Context of Driving (n= 87)

Context	***n***	**%**
Type of driving		
Professional	18	20.7
Non-professional	42	48.3
Equal amount of professional and non-professional	27	31
Area of driving		
Within urban area	40	46
Outside of urban area	47	54

The participants were all owners of cars from the S, V, or XC ranges manufactured by Volvo Cars. At the time of the survey (i.e. 2013) all cars were fairly recent year models with the oldest being of year model 2009 and the latest 2014. The majority (65,5%) drove a car with automatic transmission.

A majority rated themselves as experienced or very experienced ACC users and only a few considered themselves to be beginners why any analysis of differences between experienced less experienced ACC users was not feasible.

RESULTS

One of the questions posed to the respondents concerned their general attitude towards ACC. Overall the respondents had a very positive attitude towards the ACC. As many as 94.3 per cent was "very positive", 4.6 per cent was "somewhat positive" and 1.1% neither negative nor positive. None rated their attitude towards ACC as negative.

Earlier studies (Strand et al., 2010) have indicated that ACC users experience that other drivers react to the car driven with the support of an ACC. The respondents were therefore asked to answer the question: "Do you worry about how your driving with adaptive cruise control is perceived by your fellow commuters?" According to the responses to the survey, the major part of the respondents had no such worries. Instead it seems as if they are signifying a positive influence on other commuters (Table 3). Nevertheless, 14,9% notes that other road users have a negative perception of the car following distance. Approximately 22% provided answers suggesting that they had not received any reactions to their driving or that they had not considered the matter.

Table 3: "How Do You Think Your Fellow Commuters Have Perceived Your Driving with Adaptive Cruise Control Regarding the Following Aspects?" (n = 87). Percentages given are the valid per cent.

	Answer										
	N/A	**very negative**		**somewhat negative**		**neither nor**		**somewhat positive**		**very positive**	
Aspect	**N**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
Your car following distance	20	10	14.9	-	-	23	34.3	11	16.4	23	34.3
How you adapt your car following distance	21	-	-	7	10.6	23	34.8	9	13.6	27	40.9
How you adapt your speed	22	-	-	3	4.6	20	30.8	12	18.5	30	46.2
Your traffic rhythm	23	-	-	5	7.8	15	23.4	17	26.6	27	42.2

A particular topic addressed was the drivers' usage of the ACC (Table 4). When asked how often they activate the ACC under different circumstances the respondents answered that the system is frequently used under low as well as high traffic intensity. The ACC is also used when there is a queue, when it is raining, and when driving in the dark. However, there are, as indicated in earlier studies, also situations where drivers do not activate the ACC as often. These include snowfall and slippery road conditions.

Table 4: "How Often Do You Choose to Activate the Adaptive Cruise Control in the Following Circumstances?" (1 = never, 5 = always), (n = 87).

Circumstance	***M***	***Mdn***	**Mode**	***SD***
Low traffic intensity	4.49	5	5	.680
High traffic intensity	3.99	4	4	.946
Night, dark	3.89	4	4	.933
Rainfall	3.67	4	4	.923
Queues	3.53	4	4	1.302
Fog	3.11	3	4	1.342
Snowfall	2.80	3	3	1.109
Slippery road condition	2.69	3	3	1.194

The responses also show that the higher the allowed speed limit, the more frequently the ACC is used (Table 5). The drivers use the function the most on roads with speed limits of *100/110/120 km/h.* and the least in *living streets*.

Table 5: "How Often Do You Choose to Activate the Adaptive Cruise Control When the Following Speed Regulations Apply?" (1 = never, 5 = always), (n = 87).

Speed regulation	***M***	***Mdn***	**Mode**	***SD***
Living street	1.25	1	1	.766
30/40/50 km/h	2.44	2	2	1.291
60/70 km/h	3.38	3	3	1.123
80/90 km/h	4.09	4	4	.871
100/110/120 km/h	4.56	5	5	.659

The respondents' answers to how frequently they use ACC depending on the speed regulation reflect in the answers they provide on how satisfied they are with the function during the same speed regulations (Table 6): the higher the speed the more satisfied the user and the lower the speed the less responses are provided.

Table 6: "How Do You Perceive the Adaptive Cruise Control When the Following Speed Regulations Apply?" (n = 87) Percentages given are the valid per cent.

		Answer									
	N/A	**very dissatisfied**		**somewhat dissatisfied**		**neither nor**		**somewhat satisfied**		**very satisfied**	
Factor	***n***	***N***	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
Living street	60	2	7.4	3	11.1	8	29.6	3	11.1	11	40.7
30/40/50 km/h	19	-	-	9	13.2	11	16.2	12	17.6	36	52.9
60/70 km/h	6	-	-	1	1.2	7	8.6	17	21	56	69.1
80/90 km/h	2	-	-	-	-	1	1.2	16	18.8	68	80
100/110/120 km/h	1	-	-	-	-	1	1.2	9	10.5	76	88.4

The respondents' general attitude towards the ACC is positive, 94% are very positive and no one negative. The respondents are particular satisfied the ACC during l*onger trips* and when driving on roads without possibilities to overtake. The respondents' ratings of satisfaction when being overtaken and when overtaking was higher for those who drove mainly outside urban areas (mode = 5;5) than it was for those who drove mainly within urban areas (mode = 5;4): (U = 701.5, p = .021, and U = 555.5, p = .002). There are nevertheless specific situations, such as driving in *roundabouts* and *curves*, where satisfaction drops. For instance 35% of the respondents answers that they are somewhat or very dissatisfied with the function when driving in a roundabout. There are also a considerable number of responses missing which could indicate that the drivers do not use the function in this condition.

Table 7: "How Do You Perceive the Adaptive Cruise Control during the following circumstances?" (n = 87). Percentages given are the valid per cent.

		Answer									
	N/A	**very dissatisfied**		**somewhat dissatisfied**		**Neither nor**		**somewhat satisfied**		**very satisfied**	
Circumstance	***n***	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
When you are being overtaken	1	-	-	2	2.3	11	12.8	14	16.3	59	68.6
When you overtake	3	1	1.2	6	7.1	12	14.3	28	33.3	37	44
Access roads	18	-	-	9	13	17	24.6	19	27.5	24	34.8
Curves	1	3	3.5	19	22.1	11	12.8	28	32.6	25	29.1
Longer trips	-	-	-	-	-	2	2.3	3	3.4	82	94.3
Roads without overtake possibilities	1	-	-	-	-	3	3.5	10	11.6	73	84.9
Traffic roundabouts	39	4	8.3	13	27.1	18	37.5	11	22.9	2	4.2

The participants were also asked to assess if and in what way their access to ACC had influenced their driving and their experience of driving (Table 8). According the responses the most common effect concerns *comfort*. A majority of the drivers stated that access to ACC has increased their comfort. It is also apparent that the ACC has

perceived *safety* effects. Approximately half of the respondents indicated a reduced inclination to overtake other vehicles, a response which is in line with earlier findings (Strand et al. 2010) and almost three out of four answered that ACC has meant that their compliance with speed regulations has increased.

Table 8: "How Would You Assess the Influence of Using the Adaptive Cruise Control on the Following Factors?" (n = 87). Percentages given are the valid per cent.

	N/A	Answer									
		drastically reduced		somewhat reduced		neither nor		somewhat increased		drastically increased	
Factor	***n***	***n***	***%***	***n***	***%***	***n***	***%***	***n***	***%***	***n***	***%***
Safety	-	-	-	1	1.1	4	4.6	49	56.3	33	37.9
Stress level	-	15	17.2	40	46	15	17.2	10	11.5	7	8
Distances of travel	8	-	-	-	-	64	81	12	15.2	3	3.8
Fuel consumption	17	-	-	31	44.3	32	45.7	6	8.6	1	1.4
Compliance with speed regulation	-	-	-	2	2.3	20	23	53	60.9	12	13.8
Inclination to overtake	-	1	1.1	44	50.6	40	46	2	2.3	-	-
Annual mileage	4	-	-	-	-	79	95.2	4	4.8	-	-
Attention to other traffic	2	-	-	9	10.6	43	50.6	30	35.3	3	3.5
Comfort	-	-	-	-	-	2	2.3	42	48.3	43	49.4
Ability to judge following distance	-	-	-	-	-	40	46	36	41.4	11	12.6
How rested you are at the end of journey	1	-	-	-	-	24	27.9	46	53.5	16	18.6

Somewhat incongruent with the response regarding comfort, approximately 20% of the respondents' answer that their stress level had increased rather than decreased as a consequence of their access to ACC. Less or no effects were indicated for distances travelled, annual mileage, or fuel consumption. Negative effects were only reported by 10% of the drivers and concerned their attention to other traffic. Almost 40% thought, on the other hand, that the same function had resulted in increasing their attention. A rated increase in "*attention towards other traffic*" was more common for those who drove mainly outside urban areas (mode = 4) than it was for those who drove mainly within urban areas (mode = 3), (U = 671.0, p = .028).

A number of studies have shown that trust and control is a key component for acceptance. The answers to the question "Does it happen that the ACC is behaving in a way you did not expect?" show that the drivers find ACC to be a fairly predictable system (1 = always, 5 = never: N = 87, M = 4.07, Mdn = 4, mode = 4, SD = .938). Those who drive mainly outside urban areas (mode = 5) consider the system more predictable than those who drove mainly within urban areas (mode = 4): (U = 679.0, p = .018). More than half or 51.7% of the drivers feel very safe when handing over the control to the ACC, 46% feels safe and only 4% feels neither safe nor unsafe. No one feels unsafe.

Finally the respondents were asked to answer a number of questions that together made up a new acceptance scale (Strömberg & Karlsson 2012). The scale includes 20 items addressing the topics trust and control, perceived benefit, perceived effort, and compliance. Overall, the ratings indicate that acceptance of the ACC was very high: according to the ratings the ACC is reliable, usable, and driving becomes easier and more convenient. Driving does not however necessarily become more fun or more effective. The drivers assess the ACC not to completely "take control away" and ACC is not by all considered "absolutely necessary"

Table 9: Responses to the 20-item SKAS Acceptance Scale (n = 87)

	Item	Scale	N/A	*M*	*Mdn*	Mode	*SD*
Trust and control	ACC is…	1 = operationally reliable, 7 = prone to fuss	3	1.69	1	1	1.006
	ACC is…	1 = reliable, 7 = arbitrary	3	1.52	1	1	.719
	ACC…	1 = leaves all control for me, 7 = takes over control from me	3	3.62	4	4	1.536
Perceived benefit	For driving the ACC is…	1 = usable, 7 = unusable	3	1.14	1	1	.352
	If I use ACC driving becomes…	1 = easier, 7 = more challenging	3	1.57	1	1	.765
	If I use ACC driving becomes more…	1 = convenient, 7 = inconvenient	3	1.35	1	1	.526
	If I use ACC driving becomes more…	1 = safe, 7 = dangerous	3	1.61	1	1	.728
	If I use ACC driving becomes more…	1 = fun, 7 = boring	3	2.49	2	1	1.367
	If I use ACC during driving I perform driving more…	1 = effective, 7 = ineffective	3	2.05	2	2	.877
	If I use ACC the driving I perform will result in …	1 = less environmental impact, 7 = greater environmental impact	3	2.62	3	2	1.211
Perceived effort	To use the ACC is…	1 = easy, 7 = difficult	4	1.27	1	1	.607
	To orientate in the ACC is…	1 = easy, 7 = difficult	4	1.34	1	1	.720
	The design of the ACC is …	1 = consistent, 7 = inconsistent	4	1.43	1	1	.666
	To understand the information provided by ACC is…	1 = easy, 7 = difficult	4	1.33	1	1	.543
	To understand how I should act based on information provided by ACC is…	1 = easy, 7 = difficult	4	1.37	1	1	.657
	To understand what I should do to get the ACC to do what I want it to do is…	1 = easy, 7 = difficult	4	1.34	1	1	.547
	To learn ACC is…	1 = easy, 7 = difficult	4	1.31	1	1	.562
	To remember how ACC is used from time to time is…	1 = easy, 7 = difficult	4	1.22	1	1	.470
Compliance	ACC is…	1 = necessary, 7 = unnecessary	4	2.48	2	2	1.075
	ACC is…	1 = expedient, 7 = inexpedient	4	1.28	1	1	.502

When answering the questionnaire the participants were also given the opportunity to provide answers as to how they would like to improve the ACC. Even though the respondents were very positive towards and considered the ACC to work well in different situations, there were (as already indicated) some limitations and there were also quite a few suggestions for modifications. Three recurring themes could be distinguished from the provided answers (Table 10). These themes were connected to i) the functionality of ACC in relation to specific traffic infrastructure, (ii) a specific task relevant for driving, and (iii) the user interface.

Table 10: Recurring Themes Concerning ACC Improvements

Description	Illustrative quotes
1. Functionality in specific traffic infrastructure (roundabouts and steep curves): Aggressive accelerations; target loss; traffic infrastructure	"…today the ACC tend to accelerate to the max when the car in front turns"; "…the car should understand that it is still behind the vehicle in front, without increasing any speed"; "…sometimes it loses the car in front"; "The only negative is when driving in curves…"; "The feeling or experience when driving in a roundabout. The speed increase feels unpleasant"
2. Functionality in specific tasks relevant for driving (overtaking); harsh decelerations; incorrect target	"The only negative is /…/ and when overtaking it can target the wrong vehicle"; "The discomfort when it sometimes activates braking in a curve when overtaking for example a truck"
3. User interface and specific settings: usability; adaptability	"…I would like a warning when the radar sees the vehicle ahead but before the speed drops. This is relevant for taking the decision to overtake, or not, before my speed drops"; "…sometimes I would like to be able to increase the distance to the car in front even more"; "…I often have to look at the buttons"

Note. Some quotes are relevant for more than one theme.

DISCUSSION

In large, the results are in line with previous studies on the subject (e.g. Strand et al., 2010; Larsson, 2012; Piccinini et al., 2012, Sanchez, et al. 2012). It shows for example that end-users, in general, are satisfied with their ACC. More in particular, it confirmed many of the findings by Strand et al. (2010) regarding safety and comfort. The results also show that use frequency increase as do how positive the ratings of the performance of the ACC when speed limits increase and road types are more adapted towards higher speeds, a result which is in accordance with Sanchez, et al. (2012). However, some earlier findings could not be supported by results of the present study. For example the results do not show that worries about other road-users are a frequent matter.

The study by Piccinini et al. (2012) presented results indicating that end-users were not fully aware of the limitations of the ACC function. This study indicates that ACC is perceived as a very predictable system. These results could be interpreted as though end-users really are not aware of ACC limitations as was evident in the Piccinini et al. (2012) study, or that drivers are aware of the limitations and therefore view ACC as a predictable system. In this particular study the respondents were experienced ACC users and their use patterns indicate that they use it under some conditions and not under other. Hence, it is feasible that the end-users in this study are aware of the limitations and has adapted to it.

The study by Piccinini et al. (2012) suggested that drivers engage in tasks that distract them from driving when the ACC is activated. In the study by Sanchez, et al. (2012) 13% of the participants stated that they use ACC in order to free more time to perform other tasks, such as changing radio channel or eating. This study present results on a related question, namely if the drivers were less attentive to other traffic (as could be a consequence of a distractive task). Half of the answers were that ACC does not affect how attentive they are to other traffic. Of the remaining answers, the majority provided answers suggesting that they are even more attentive to other traffic. On the other hand about 10% provided answers suggesting that ACC had a negative effect on how attentive they are to other traffic (worth noting is that there were no answers stating that it had drastically reduced). The difference in responses between overall experienced ACC users is interesting but not altogether easy to explain. It is possible that the drivers do not want to admit to becoming distracted from driving, it is also possible that they are not aware that they are distracted. How drivers make use of the handing over control to an assistance system is a topic that needs further investigations.

Overall the drivers are satisfied with the ACC, a result with complies with the findings of Sanchez, et al. (2012), and their acceptance of the system is high. Nevertheless the drivers' responses to the different items in the acceptance scale and their comments indicate a potential for improvement. One concerns control. It is possible that a goal should be to design the function so that the drivers feel more in charge, or control, even when the ACC is performing its duties. Such reasoning would be in line with for instance Norman's (2007) proposition that one should "make people think they are in control" (p. 193). However, it can be debated whether this is a desired strategy when designing ADAS. Instead, the strategy could be to design systems which are more transparent in order to contribute to the drivers developing a correct and meaningful mental model of the system, developing enough trust in order to hand over the control, but at the same time be aware of the limitations of the system so that they will be able to master takeovers when so is required.

Another item concerns effectiveness of driving where the ratings suggest that drivers experience to become more effective in their driving with access to ACC. However the earlier study by Strand et al. (2010) highlighted a potential to improve ACC in situations where drivers overtake and the answers in the present study also display some improvement potential regarding such situations.

The drivers were very satisfied with the ACC but even though it was considered to contribute to comfort as well as safety the function was not considered absolutely necessary. The respondents in the survey are experienced drivers who have driven a considerable number of kilometres without the support of ACC or other ADAS. Safety and comfort benefits of ACC are by no means insignificant, but they are most probably viewed as a bonus to a higher order need, rather than crucial for it. An investigation some ten or twenty years from now may well tell another story. Future generation of drivers, given the present automation trend, may never have first-hand experience of driving without support systems. Rather than questioning the drivers' trust in handing over control to the system, the challenge may be to design vehicles in which the drivers are comfortable with taking over control.

The differences between the present study and earlier ones could be partly attributed the differences in methods used, but perhaps more so the sampling differences. The participants who took part of this study were very positive towards their ACC as well as rated themselves to be very experienced ACC users. At the same time the response rate was 21,7 %, which means that a substantial number of the population did not answer the questionnaire. The mean age of the respondents is fairly high and very few women answered the survey. It is possible that this profile reflects the end-users of ACC systems, but it could also be that the non-respondents represent another user group with other experiences of their ACC systems than the group that provided answers to this study. For instance, could the responses fail to reflect the views of users who have found the ACC less useful and who choose not to activate it because of its limitations in certain conditions. It is feasible that less experienced drivers are less prone to hand over control to the system as well as less aware of the limitations of the ACC. If so, the results are a display of the experiences of a particular user group, and not the whole ACC end-user population.

At the same time a major part of the results is consistent with earlier studies why it is feasible to assume that a considerable amount of drivers appreciate the ACC function and adapt to its limitations. There are benefits in terms of increased perceived comfort and safety. Reported changes in driving behaviour in terms of less overtaking and increased compliance with speed regulations support this assumption. At the same time some drivers report feeling more stressed when driving with ACC than without and there are contradictory results regarding drivers' attention to traffic. These factors need further investigations.

ACKNOWLEDGMENTS

This work has been carried out at SAFER – Vehicle and Traffic Safety Centre at Chalmers, Sweden. The authors would like to acknowledge valuable input on the development of the questionnaire from SHADES project members, co-workers at VTI and CHALMERS. The authors would also like to acknowledge Mr. Mats Petersson (Volvo Cars) for contributing to the selection of participants.

REFERENCES

Berstein, D. A., Penner, L. A., Clarke-Stewart, A., & Roy, E. J. (2003). *Psychology* (6th ed.). Boston, MA: Houghton Mifflin.

Bianchi Piccinini, G. F., Simões, A., & Rodrigues, C. M. (2012). Effects on driving task and road safety impact induced by the usage of Adaptive Cruise Control (ACC): a focus groups study. *International Journal of Human Factors and Ergonomics, 1*(3), 234-253.

Brookhuis, K. A., de Waard, D., & Janssen, W. H. (2001). Behavioural impacts of Advanced Driver Assistance Systems – an overview. *European Journal of Transport and Infrastructure Research, 1*(3), 245-253.

IBM Corp. (2012). IBM SPSS Statistics (Version 21.0) [Computer software]. Armonk, NY: IBM Corp.

Larsson, A. F. L. (2012). Driver usage and understanding of adaptive cruise control. *Applied Ergonomics, 43*(3), 501-506.

Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology, 22*(140), 1–55.

McQueen, R. A., & Knussen, C. (2006). *Introduction to Research Methods and Statistics in Psychology*. Harlow, Essex, UK: Pearson Education.

Naranjo, J. E., González, C., Reviejo, J., Garcia, R., & de Pedro, T. (2003). Adaptive fuzzy control for inter-vehicle gap keeping. *IEEE Transactions on Intelligent Transportation Systems, 4*(3), 132-142.

Norman, D. A. (2007). *The design of future things.* New York: Basic Books.

OECD. (1990). *Behavioral Adaptations to Changes in the Road Transport System.* Chicago: University of Chicago Press).

Saad, F., Hjälmdahl, M., Cañas, J., Alonso, P., Garayo, L., Macchi, F., Bekiaris, E. (2004). Literature review of behavioural effects (Deliverable No. D1_2_1). Retrieved from EU-project AIDE – Adaptive Integrated Driver-vehicle InterfacE website: http://www.aide-eu.org/pdf/sp1_deliv_new/aide_d1_2_1.pdf

Sanchez, D., Garcia, E., Saez, M., Benmimoun, M., Puts, A., Ljung Austm M., … Obojski, M-A. (2012). Final results: User acceptance and user-related aspects (Deliverable No. D6.3). Retrieved from EU-project euro FOT website: http://www.eurofot-ip.eu/download/library/deliverables/eurofotsp620121119v11dld63_user_acceptance_and_userrelated_aspects.pdf

Sarter, N., & Woods, D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors, 37,* 5-19.

SPSS Ltd. (2008). SPSS Dimension Net (Version 5.5) [Computer software].

Stanton, N. A., & Young, M. S. (2005). Driver behavior with adaptive cruise control. *Ergonomics, 48*(10), 1294-1313.

Strand, N., Nilsson, J., Karlsson, I.C.M., & Nilsson, L. (2011) Exploring end-user experiences: self-perceived notions on use of adaptive cruise control systems. In J. F. Krems (Ed.), Selected papers from the 2nd European Conference on Human Centred Design in ITS [special issue]. *IET Intelligent Transport Systems, 5*(2), 134-140.

Strömberg H. & Karlsson, I.C.M. (2012) A new scale for addressing acceptance. Working paper. Design & Human Factors, Chalmers University of Technology, Gothenburg

Section 13

Road and Rail:

Human Mobility Increase as a Key Element in Designing Innovative Transportation Systems

Ergonomic Issues in the Design of Innovative Means of Transport and Transportation Systems

Iwona Grabarek and Wlodzimierz Choromanski

Faculty of Transport
Warsaw University of Technology
Poland

ABSTRACT

This article focuses on the design of 'a driver-passenger' interface for two types of vehicles (Eco-car and PRT system). However, this solution has different purposes in each case. The Eco-car is driven by a human, so the issues under consideration are as follows: driving of the vehicle by upper limbs only and adaptation of key elements of the interface to the individual features of the driver. The PRT system itself combines some features of individual and public transportation. Despite the fact that the vehicle is not driven directly by the passenger, the passenger's influence on the vehicle operations is much stronger than in conventional means of public transportation in urban areas. This article deals with the innovative design of such an interface. This research has been carried out as part of the 'Eco-Mobility' project, 85% co-funded by the European Regional Development Fund under the Operational Program-Innovative Economy.

Keywords: innovative means of transport, ergonomic design, driver's and passenger's interface

INTRODUCTION

Degradation of quality of life in urban areas results from i.e. increasing transport congestion, alarming air pollution levels or degradation of 'acoustic climate' in urban areas. These factors imply new challenges for designers of means of transportation and transportation systems. The scientific research is focused on a number of issues:

- search for design solutions that would allow minimum energy consumption, minimum emissions and the use of new material technologies including recycled materials
- other organizational methods of vehicle use including vehicle sharing
- adaptation of vehicles for senior individuals with limited physical capabilities as well as for disabled individuals in wheelchairs
- designing new means of transportation and transportation systems that allow human-free operations (APM systems / Automated People Mover, ATS / Automated Transit System, ATN / Automated Transit Network). These systems strive for combining the advantages of public and individual transportation.

NEW DESIGN PHILOSOPHY IN MEANS OF TRANSPORTATION AND TRANSPORTATION SYSTEMS

Ageing societies in industrialized countries are a common and permanent phenomenon particularly visible in Europe. 100 years ago, the population of Europe constituted 15% of the world´s total. By 2050, this share is forecasted to be three times less. (Leśnikowska-Matusiak, I., Wnuk, A, 2011). The cyclical European Demographic Report published by the European Commission describes demographic facts and trends in the E.U. member countries. The number of people aged over 65 among the total population is set to increase from 17, 1% in 2008 (84, 6 million) to 30% in 2060 (151 million). The number of people aged over 80 is forecasted to triple from 21,8 million in 2008 to 61,4 million people in 2060 (2009 Ageing Report, The European Parliament).
The percentage of people with limited mobility in various age groups can be determined based on the 2000 GUS (Główny Urząd statystyczny – (Polish Central Statistical Office) data (Figure 1).

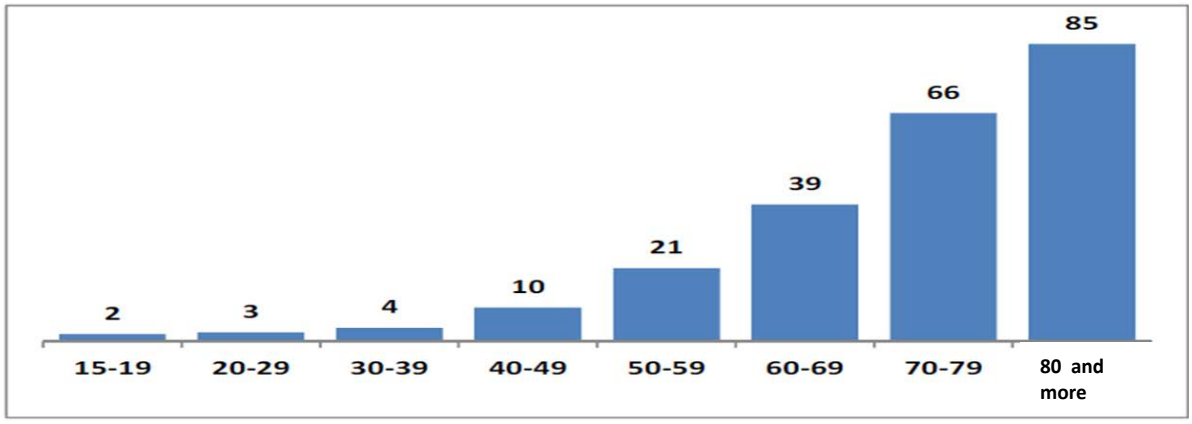


Figure1. Percentage of people with limited mobility % /GUS 2010/

Disabled people are part of the group of people with limited mobility. There are two types of disability that are recognized in Poland: legal disability and biological disability. The legal disability is being evaluated by authorized institutions. The biological disability is one of the most important health aspects. This issue becomes more and more fundamental as life expectancy increases. The assessment for biological disability is based on criterion of limited ability to perform day-to-day physical activities that persists for six months or longer.

Table 1: Disabled people by disability level between 2002 and 2011 (GUS, Raport z wyników, Narodowy Spis Powszechny Ludności i Mieszkań 2011)

	Total		Urban areas		Rural areas	
	2002	2011	2002	2011	2002	2011
	In thousands					
Disabled people	5 456,7	4 697,5	3 213,1	3 018,4	2 243,6	1 679,1
Legal disability	4 450,1	3 131,9	2 650,6	2 089,8	1 799,6	1 042,1
Biological disability only	1 006,6	1 565,6	562,5	928,6	444,0	637,1

The figures above can help the designers of transportation systems realize how challenging their tasks will be in the nearest future. This process should be intensified due to some other, equally important, challenges of today's civilization such as: increasing transport congestion, alarming air pollution levels or strong degradation of ´acoustic climate´ in urban areas. Designing innovative transportation systems, in particular for urban areas and for people of various physical fitness levels, requires a different approach and a different design philosophy. Mobility, i.e. ability of people, machines, goods or ideas to move, is one of the fundamental features of the modern world. Mobility allows:

- exploring the world
- individual development
- living social life.

The mobility of population increases in areas where transportation systems are made available. Accessibility primarily concerns people with reduced mobility, disabled people, elderly people, families with young children and young children themselves: they should have easy access to urban transport infrastructure. The social aspects of mobility in towns and cities present a challenge. Urban transport needs to be affordable also to people on low income. Citizens with reduced mobility and senior citizens expect increased and higher quality of mobility, as personal mobility is crucial to independence and freedom. Two design approaches (user-centered design and universal design) can potentially help meet expectations of people that use transportation systems in urban areas and increase their mobility.

What is the user-centered design? The key differentiating factor from other product-design philosophies is that the user-centered design tries to adapt the product to how users can, want or need to use it rather than forcing them to change their behavior to adjust to the product.

The International Usability Standard, ISO 13407, specifies the principles and activities that underlie the user-centered design:

- The design is based on an explicit understanding of users, tasks and environments
- Users are involved throughout design and development
- The design is driven and refined by user-centered evaluation
- The process is iterative
- The design addresses the entire user experience
- The design team includes multidisciplinary skills and perspectives.

The user-oriented design is always an iterative process that involves the testing of each introduced change. UCD can be seen as a specific philosophy of an organization, but also as an element of product life-cycle management.
The ergonomic design approach is summarized in the principle of user-centered design. Both groups of rules have the same goal - to design equipment that will achieve the best possible fit between the users (drivers, passengers) and the equipment (vehicle) such as the users' safety (freedom from harm, injury and loss) comfort, convenience, performance and efficiency are therefore improved.

The Universal Design is the design of products and environments that are to be:

- Usable by all people to the greatest extent possible
- Without the need for adaptation or specialized design.

The universal design principle stipulates that the primary solution must be designed to anticipate the needs of all users. One of the key goals of the universal design strategy is to promote equality and ensure full participation in social life to people with limited functions by eliminating the existing barriers and preventing new ones from arising. The universal design concept sets new ways of thinking. It is based on the equality principle to a larger extend than the concept of general availability for individuals with limited functions. Whereas general accessibility for disabled individuals can be achieved with sophisticated solutions, the concept of universal design is based on the principle that the basic functionalities and solutions will be designed to match the needs and capabilities of all potential users. Accessibility for all people should be seen as a starting point for the designer. Products and environment need to be designed to serve people of all ages, various capabilities and disability level. The key aspects that need to be considered here include ability to move, see, hear and perceive as well as environmental vulnerability. The key aspect of the universal design strategy is to always strive for better solutions.

Universal design strategy is innovative and can be defined with seven key rules:

- Equitable use
- Flexibility in use
- Perceptible information
- Simple, intuitive use
- Low physical effort
- Tolerance for errors
- Size and space for approach and use.

The Eco-Mobility project is aimed at designing innovative means of transportation. One goal of the project was to design the vehicles (the Eco-Car and PRT vehicle) in line with the two above mentioned design concepts, so that the vehicles are usable by users with various levels of physical ability.

The process of vehicle design begins with a discussion on the size and type of the vehicle and the number of occupants that the vehicle should accommodate. The very first step in designing a vehicle is to determine the user population(s) and their anthropometric and biomechanical characteristics. The anthropometric data of the user population will help determine various basic dimensions of the vehicle. The biomechanical data helps design the vehicle so that the users will not be required to exert or be subjected to forces that are above their tolerance or comfort levels (Bhise, 2012).

SPECIFIC FEATURES OF THE VEHICLE DESIGN

The object of ergonomic design is always a system consisting of two key sub-systems: human and technical, i.e. system human-machine (Tytyk, 2001). Designing means of transport in line with the human's psychophysiological requirements is crucial to their ergonomic quality. For most means of transportation the design is a two-way process. The designer needs to meet on the one hand, the requirements regarding the workplace of an operator (i.e. driver, pilot) and the passenger space on the other. Those two sets of requirements are not always in line with each other, but one area that is always common is ´transport safety´. Therefore, two systems are subject to analysis and design: the operator-means of transport system and the passenger – means of transport (Grabarek, 2009).

The defining ergonomic requirements for the newly-designed means of transport calls for a brief description. (Grabarek and Choromanski, 2012). The Personal Rapid Transit (PRT) – is a zero-emission, intelligent transportation system, combining features of individual and public urban transport (´point to point´ or ´door to door´) that involves small, four-passenger, automated vehicles operating on a network of specially built guide ways – usually an overhead monorail track. A non-stop, point-to-point travel bypassing all intermediate stations is fully available for people with various levels of disabilities. In the system, humans are passengers and the automated vehicle requires some very simple actions from them including selecting the destination station on a touch screen. One innovative feature of the urban, electric Eco-car is that it can be used by both able-bodied and disabled people in wheelchairs. The special interior design, including the driver's seat and control panel, allows a person with physical dysfunction of lower limbs to drive the car without the need to leave his or her wheelchair. In the Eco car, a human can play a role of the driver or a passenger.

Table 2: Initial assumptions on functionality of the PRT vehicle and the eco-car

PRT	**Eco-car**
Max 4 persons	Max 3 persons
Remote control /no operator/	Able-bodied driver in a driver's seat or a person with physical dysfunction of lower limbs in a wheelchair
One driving direction	Any driving direction
Travelling in seats	
Two persons facing forward, two backwards	All persons facing forwards
Door on one side	Doors on three sides

THE ROLE OF INTERFACE IN INNOVATIVE MEANS OF TRANSPORT

As mentioned above, both the Eco-car and PRT meet users' expectations. Both systems are designed as a support to urban public transport systems and are to be available to the public. The authors of this article focus on solutions that allow a simple and intuitive use of both types of vehicles. In case of the Eco-car the following processes are subject to analysis: entering the interior, adjusting the control panel to the driver and the driving of the car. In the PRT the analysis of vehicle/human interaction is much more limited and the following systems are not subject to analysis: CAN network information systems aimed at supporting all power supply, vehicle propulsion system as well as vehicle management systems. The user's interface (driver's or passenger's) channels communication between the

user and the vehicle. For each of the considered cases, the scope of communication is different, but in each case the user was able to accomplish the following actions unaided:

- Entering and exiting the cabin
- Placing a wheelchair in the right position and locking it
- Providing back support to a person in a wheelchair
- Accessibility of the interface / control panel
- Using of the interface / control panel.

Designing functional and ergonomic vehicles requires numerous analyses and extensive research i.e. surveys among disabled people acting as a group of experts. Building an intuitive interface that is optimal from the user's perspective requires some proven solutions, including:

- Effective communication of the system's functionality
- Informing the user about the status and all possible changes of the system
- Care for user's safety and error prevention
- Supporting user's autonomy and sense of control
- Interface use - easy to learn and remember
- Efficient system without delays

The key rules for designing user's interface include:

- Layout – the interface should be divided into different areas responsible for specific functions
- Keeping the user informed – the interface should inform the user on his/her current location and communicate additional key information
- Esthetics – the interface should strike a balance between the amount of the information displayed and its visual attractiveness
- User's experience – the interface should be easy to learn for beginners and easy to use for experienced users
- Consistency – the interface should be consistent in order to help the user anticipate the results of his/her actions
- Optimum effort – the interface should allow to achieve the desired goal with as few steps as possible.

Eco-Car

The innovative character of the electric Eco-car makes it usable by both able-bodied persons and disabled people in wheelchairs. Cabin design, and in particular the design of the driving seat and the control panel, makes it possible for a person with the physical dysfunction of the lower limbs to drive the car without the need to leave his or her wheelchair. People with higher levels of disability require an autonomous, ´driverless ´vehicle. Various ages and levels of disability determine the channels of communication between the user and the Eco-car. In the Eco-car, humans can play the role of a driver or a passenger. Whereas able-bodied people do not experience any serious barriers in driving a car, a disabled person in a wheelchair can only do it without assistance by having a remote control over actions that are necessary to set up the car before the driving starts. What is understood as a ´driver's interface´ in the Eco-car is the system that consists of three key elements: remote control, touch screen mounted in the control panel and a multi-function driving wheel. The control panel is mobile and can be easily adjusted, which makes it possible for every driver to feel comfortable (see Figure 2).

a) b)

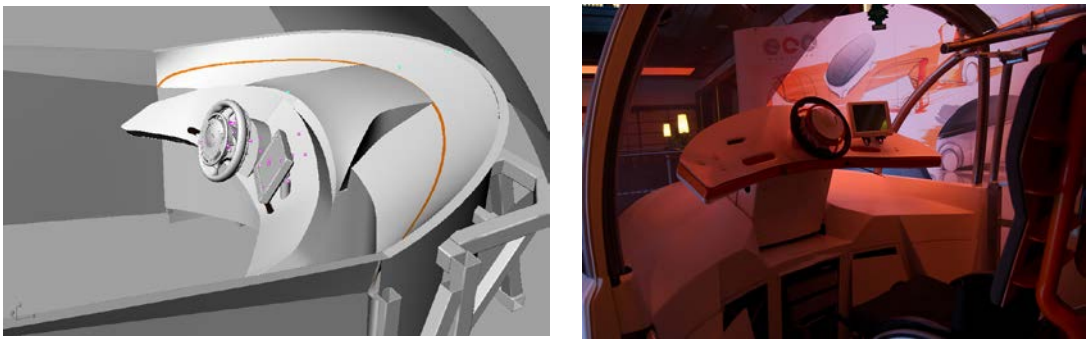


Figure 2. a) Design of the control panel area of the Eco-car b) Mock mobile desktop of the eco-car

The remote control is the first device used by the driver to select a driving mode (specific configuration of the Eco-car). There are four different driving modes in the Eco-car:

- Able-bodied driver (and possibly able-bodied passengers)
- Disabled driver (additionally – disabled passenger)
- Disabled driver with able –bodied passengers
- Disable driver with no passengers
-

Each mode requires different sequence of activities and commands prior to the starting of the car. While seated in its seat (or in a locked wheelchair), the driver can also use another element of the interface i.e. the touch screen mounted in the control panel.

The selection of any mode by a disabled driver in a wheelchair requires a sequence of actions, including the reconfiguration of the interior, allowing to locate the wheelchair in the right position and lock it in. The support for the drivers´ back is provided by the standard drivers' seat which has folded in the first place to allow space in front of the steering wheel for the driver's wheelchair. All these actions, with the exception of entering the interior by the driver in a wheelchair, take place automatically after a certain command is selected on the remote control (see Figure 3).

Figure 3. Sequence of screens showing selection of the user type /*driver - kierowca; passenger – pasażer*/

Mode selection: driver in a wheelchair – initiates the process of lowering the car's floor to the ground level as well as folding and moving the standard driver's seat inside the cabin. After the sequence is complete, one selected door opens. The figure below depicts the location of a locked wheelchair inside the cabin. The mode selection is available on the remote control or on the touch screen.

Figure 4. Location of the wheelchair in the cab's axis */ position your wheelchair as shown on the diagram – zajmij wyznaczone miejsce według schematu/*

Figure 5. Desktop adjustment manual /*set the desktop in the original position- ustaw pulpit w pozycji wyjściowej/*

a) b)

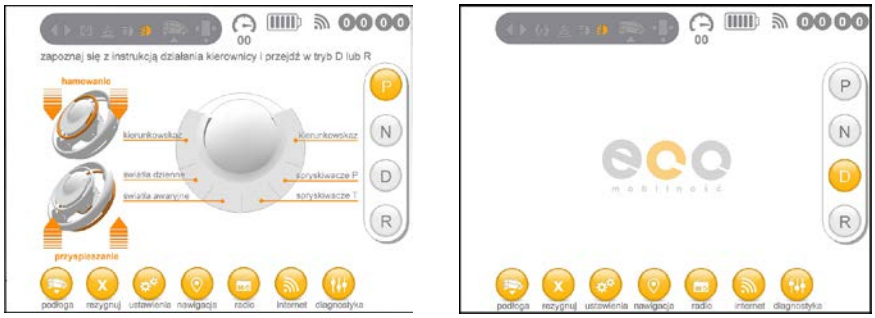


Figure 6. Consecutive screenshots from the touch screen: a) multifunctional steering wheel manual, b) screen appearing in the driving mode

Legend:
Refer to the service manual of the steering wheel and select mode D or R – zapoznaj się z instrukcją kierownicy I przejdź w tryb D lub R
Turnsignal - kierunkowskaz
Daytime / emergency lights - Światła dzienne /awaryjne
Sprinklers - Spryskiwacz
Breaking – hamowanie
Accelerating - przyspieszanie

Another element of the eco-car driver's interface is the circular multifunctional steering wheel that allows the driving of the car with the upper limbs only /Figure 7/.

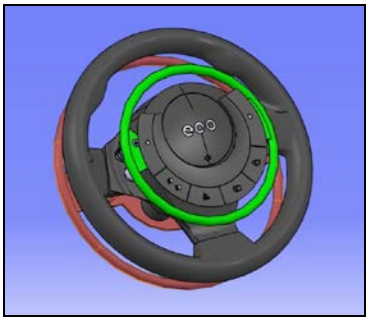


Figure 7. Circular multifunctional steering wheel

The steering wheel is one of the elements of the ´drive by wire´ system that was used in the Eco-car. This means that the traditional mechanical connections (between the steering wheel and the wheel turning mechanism or between the brake pedal and the wheel braking mechanism) were replaced by using electrical or electro-mechanical systems for performing various vehicle functions. This solution is a fundamental part of the design for the following reasons:

- allows a significant increase of the usable space through the elimination of many mechanical systems
- allows an easy connection of other steering mechanisms (e.g. a joystick)
- allows the parameterization of the steering system (e.g. an analysis of ´steering system sensitivity´ – i.e. the variability in the ratio of the steering wheel rotation to the corresponding turning angle of the turning wheels
- makes is easy to customize the steering system.

These advantages are however offset by some significant challenges, e.g.

- reliability and safety issues
- issue of synthesis of complex feedback loops: from the steering wheel to the turning wheels and from the turning wheels to the steering wheel.

All the issues mentioned above were subject to research within the ECO-Mobility project. The research focused in particular on the issue of the steering wheel coupled with the ´drive-by-wire´ system for drivers with disabilities. The in-depth analysis was presented in (Choromanski et al 2013). ´Steer-by-wire´ oversees turning wheels' angles, enhances driving comfort and allows the intervention of the vehicle's supervisory system irrespective of the driver. ´Brake-by-wire´ monitors the vehicle's braking system and also allows the intervention of the supervisory system, irrespective of the driver, in order to adjust the braking forces on each wheel of the vehicle to the values that are appropriate given the current road conditions. The performed simulation results show that the vehicle is able to pass the standardized test often called the ´moose test´. It is assumed that the driver may, to some degree, have difficulties holding the steering wheel. Thus, the analysis of the turning velocity of the wheels should provide an answer to the question of how to adjust the maximum turning range of the steering wheel in the vehicle.

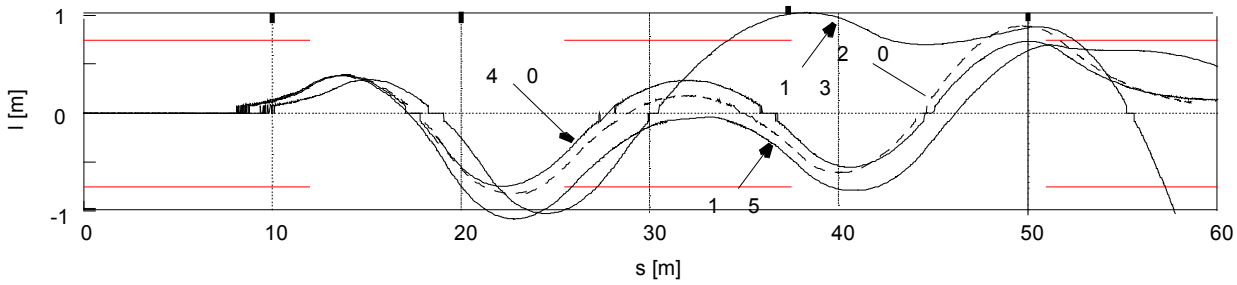


Figure 8. the results of driving simulation in the ´moose test´. Distance of the gravity center point of the car model from the road axis, for different preset values of the maximum speed of turning the steering wheel

Figure 8 shows the vehicle's trajectory simulation results of the dynamics model. The "13" line indicates the distance of the vehicle's gravity center point from the road axis, in the case of limitation to the maximum velocity of turning the front wheels to the value v_s = 13 ^{0}/s. In this case, it failed to keep the vehicle within the required distance range from the road axis. The "15" line shows the velocity v_s = 15 ^{0}/s, line "20" – velocity v_s = 25 ^{0}/s and line "40" – velocity v_s = 40 ^{0}/s. The results show that a steering velocity of at least v_s = 40 ^{0}/s is required for the positive test result. For a maximum turning angle of the front wheels α_{max}= 35 ^{0}/s, to be achieved through a 1.25 rotation of the steering wheel, the wheel would have to be turned at a velocity of at least v_k = 514 ^{0}/s, and for a 0.5 rotation, just 206 ^{0}/s. The result of a test is considered positive if the solution is contained between the two red lines (Figure 8). Only the trajectory corresponding to40 ^{0}/s has the positive test result. Reducing this velocity requires increased sensitivity of the steering wheel e.g. adoption of the maximum front wheels turn corresponding to only one wheel rotation (and not 1.25, as assumed in the simulations performed).

Personal Rapid Transit

Operations of a PRT vehicle are remotely controlled, and this process does not require human intervention. The main controlling device is a passenger interface, i.e. a touch screen. The passenger selects the final destination prior to departure. He or she can also change the selected destination during the trip, provided that the time needed for this task is shorter than the time to reach the destination. This restriction is implemented in the software. Moreover, in the interface case, four buttons are displayed: control center call- in order to obtain information, ALARM call, having priority among all signals reaching the center, as well as buttons to open and close the door.

The opening and closing of the vehicle's door is done manually during the stop at a station. On selecting the destination and pressing the START button, the door closes and the 'door buttons' in the interface casing become inactive. After coming to the stop at the destination, a signal is generated that opens the door. The functionality of the interface depends on its layout within the cabin space, which makes PRT usable by people in various physical conditions. During the trip, the passenger and his/her wheelchair is secured in the vehicle axis (facing forwards or rearwards). From this position, he or she shall have good access to the interface. Ensuring access to the monitor/interface for smallest person sitting in a wheelchair, or C5 ♀, is in this case a necessary and sufficient condition. By installing two monitors on the opposite walls, operations using the right hand are allowed, depending on the location of the wheelchair in relation to the direction of driving. The intuitive interface depends on the graphic design, enabling:

- the ability to detect information
- the ability to distinguish information
- the ability to recognize information content or meaning of the signal.

The ability to detect and differentiate information depends on the readability of the displayed content. Catering for information recognition leads to high intelligibility and an intuitive interface. Communication with the interface starts with the reading of information or questions transmitted by the device. In order to make it convenient for the user, appropriate parameters should be ensured:

- typeface and font size
- text layout
- color, shape and size of display graphics
- monitor brightness.

The Polish standard specifies the format for presenting the information (ISO 13407: 1999). The recommendations are aimed at assuring readability and legibility. It is, however, the users' task to comprehend the displayed information. The intelligibility of the displayed image depends on:

- the language in which text is displayed
- proficiency in the use of interactive or multimedia devices.

It is a valuable and widespread practice to allow the user to select the language of communication with the machine. Often elderly people may have a major problem handling multimedia devices. Accustomed to obtaining information in the form of text, they are not always able to digest information transmitted in other forms, such as graphic characters. The PRT interface uses both text and graphical elements (icons and symbols). One measure of the interface intuitiveness is the user's ability to predict the system's response to his or her actions. This results in a quick and convenient operation of the program and therefore, there is no need for further instructions or training.
The interface design allows the passenger to choose his or her destination in several ways. A person who does not know the area in which the PRT system operates may be in the situation that he or she only knows the name of the destination station. By contrast, for people using the system on a daily basis, it may be the fastest way to choose the destination by pointing it on a map. Groups of the same elements are present on many screens and they are located in the same spots. Consistency regarding shape, color and position enhances the intuitiveness of the interface as a whole. A user seeing such an element expects that the next screen will display similar objects in similar positions and playing similar roles. These elements may include, for instance: header, commands UNDO, NEXT.

The interface consists of a number of screens: the welcome screen allowing the selection of the target stop in several ways, the screen to confirm the selection and the screen displayed while driving.

Figure 9. shows the localization of functional areas on the screen used in the PRT cabin. The darker area on the right hand side of the screen is easily available only for able-bodied passengers. This area can be used to display information. However, the basic commands should appear on the main touch panel, which is accessible to all users

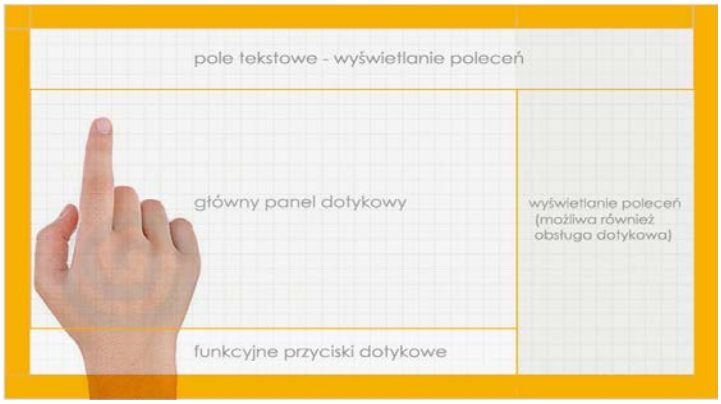


Figure 9. Functional chart of the monitor's screen

Legend:
Text box – displaying commands – pole tekstowe – wyświetlanie poleceń
Main touch screen – główny panel dotykowy
Displaying commands (touch screen functionality also available) – wyświetlanie poleceń (możliwa również obsługa dotykowa)
Functional touch buttons – funkcyjne przyciski dotykowe

Figure 10 shows the screens that appear when selecting the destination.

a) b)

Figure 10. a) Welcome screen; b) Screen allowing the selection of destination

During the transit a passenger can observe the vehicle move on another screen (Figure 11). He or she can also obtain information on service points located in the vicinity of the stops or use online services. The option to change the route, i.e. the destination, is doable until the vehicle reaches a critical distance. If such decision comes at a time when the vehicle approaches the destination station, the passenger is being informed that it is not possible to make any changes at this stage.

Figure 11. Screen displayed en-route to the destination station

In the interface design process, research was also conducted on increasing its accessibility for people with various disabilities, not just physical. These have included vision impairment, color perception impairment or motor disorders. The results of this research will be addressed in upcoming solution proposals.

CONCLUSIONS

This article focuses on the design of a 'driver-passenger' interface for two types of vehicles (Eco-car and PRT system). However, this solution has different purposes in each case. In case of the Eco-car, the vehicle is driven by a human, so the issues under consideration are as follows:

- driving using upper limbs only (disability of lower limbs was assumed) which requires a proper cooperation of a specially designed steering wheel and a touch screen
- adaptation of key elements of the interface to the individual features of the driver (e.g. regulated interface that allows a comfortable set up).

The 'steer by wire' solution adapter in the design of the vehicle has a fundamental impact on the interface design. This solution allows an additional adaptation of the vehicle steering performance to the individual features of a specific driver. It includes a variable responsiveness of the vehicle's wheels to the steering wheel movement. In the case of the PRT system, the passenger interface has a different role. It allows selecting and changing the route at any moment, getting on and off and emergency evacuation. PRT vehicles can be considered a specific type of driverless

car. The PRT system itself combines some features of individual and public transportation. Despite the fact that the vehicle is not driven directly by the passenger, the passenger's influence on the vehicle operations is much stronger than in conventional means of public transportation in urban areas. The proposed solutions are intended for users with varying physical abilities. The new design philosophy, i.e.: the user-centered and universal designs are to meet the expectations of users of transport systems, in particular of public transport, and significantly increase the mobility of this group. By observing trends in car design, it can be argued that such solutions as proposed in this paper will be implemented in the near future.

ACKNOWLEDGMENTS

The authors would like to acknowledge that the above mentioned research was carried out as part of the ECO-Mobility project, co-funded by the European Regional Development Fund under the Operational Program Innovative Economy (UND-POIG.01.03.01-14-154) – Project Co-coordinator Prof. W. Choromanski

REFERENCES

Bhise, D. Vivek (2012),"*Ergonomics in the Automotive Process",* CRC PRES, Taylor & Francis Group.

Choromański W., Grabarek I., Kozłowski M., Ufnalski B., Barwicki M. (2013), *"A new concept of electric eco-car"*. Selected proceedings of International Conference WCTR 2013, Rio de Janeiro, (http://www2.wctr2013rio.com/proceedings).

Communication from the Commission to The European Parliament, The Council, The European Economic And Social Committee and the Committee of the Regions (Dealing with the impact of an ageing population in the E.U. (2009 Ageing Report), COM (2009) 180, Commission of the European Communities, Brussels.

Grabarek, I.: *´Ergonomia środków transportu - rzemiosło, nauka, sztuka´*. Prace Naukowe Transport, z.71, Warszawa 2009, s.57-73.

Grabarek, I., Choromański, W. (2012), *´Innovative environmental design in means and systems of transport with particular emphasis on the human factor´.,* in Advances in Human Aspects of Road and Rail Transportation, edited by Neville A. Stanton, CRS Press Taylor &Francis Group, pp.273-282, ISBN 978-1-4398-7123-2.

ISO 13407: (1999), "*Human- Centered Design Process for Interactive Systems"*.

Leśnikowska-Matusiak, I., Wnuk, A.(2011), *"Kierowanie pojazdem a starzejące się społeczeństwa",* Transport samochodowy, No.1, pp.73-86.

Tytyk E. (2001), "*Projektowanie ergonomiczne*", PWN, Warszawa.

Lever Wheelchair for Disabled People: an Answer to the Increasing Demand for Wheelchair Mobility

Krzysztof Fiok and Włodzimierz Choromański

Faculty of Transport
Warsaw University of Technology
00-662 Warsaw, Poland

ABSTRACT

Young, active disabled people with impairment of lower extremities tend to be as active, as possible considering their disabilities. For purposes of leisure and sport many wheelchair solutions have been already proposed, yet the struggle to provide efficient and economically rational means of transport that increase their mobility is still ongoing. According to various authors lever wheelchairs have certain advantages in the given field: ergonomic work conditions and high human work efficiency resulting from the lever drive concept. In our works we created a pre-prototype of such wheelchair. Our projects goal was to try to implement many different solutions which were to provide rich functionality: adjustment of levers' length, adjustment of levers' axis of rotation position, changeable gears including reverse gear, innovative brake mechanism. As a result we could assess which of the functional assumptions were reasonable and easy to implement and which could be useful however make the design very complicated, heavy and thus irrational. Our main conclusion was that if the lever driven wheelchair concept is to become popular, one of the key-issues in their design must be: simplicity. Additional functionality seems attractive during theoretical studies, but it's implementation very often causes too much complication and weight gain.

Keywords: Wheelchair, lever wheelchair, disabled people mobility, prototype

INTRODUCTION

Lever wheelchair concept is not a new idea. Already in 1970s in Germany this type of wheelchairs was not only known, but popular. According to a study (Blohmke F. et al, 1975) 52,3 % out of 3482 interviewed disabled people used lever wheelchairs. Later on however, the idea of using such wheelchairs was overthrown in favor of classic push-rim wheelchairs. What was the reason for this change? Obviously it had to do with strong and weak sides of both concepts and how they changed along with technical, material and social development of both wheelchairs and their users. As a result the present situation is, that 90% of all wheelchairs are classic push-rim wheelchairs (van der Woude LHV et al, 2001).

However, at the turn of centuries various people worldwide start again to bring up the idea of lever wheelchairs. Further changes in our communities, increasing demand for mobility of people with impairment of lower limbs, made the concept live again. And so we can observe various proposals for design of lever wheelchairs depending on their destination: from simple and cheap terrain-adapted wheelchairs dedicated for African societies (Massachusetts Institute of Technology MIT, 2014), through proposals of wheelchairs with complicated power transmission systems (Taylor W. G., 1999), to compact, lightweight designs of wheels with levers that could be mounted on various wheelchair frames instead of classic wheels (Innovations Health Devices, 2012).

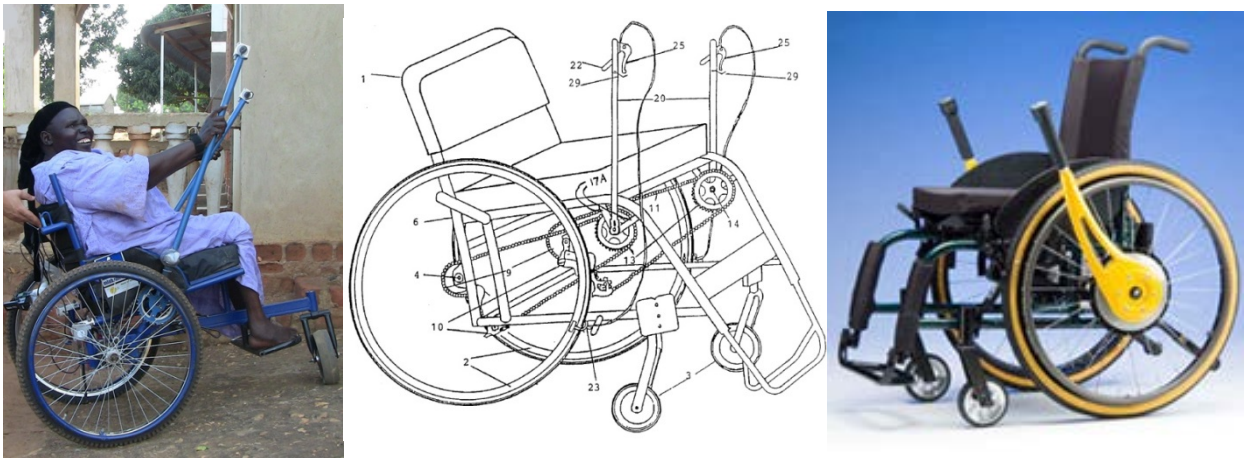


Figure 1. Various lever wheelchairs. From left to right: MIT design, Taylor W. G. design and Wijit wheelchair.

Analyzing figure 1 allows us to form a conclusion, that designs of lever wheelchairs can vary strongly depending on functional assumptions that were made in the beginning of their design process. While observing these differences we posed a question: which of these assumptions led to easily implementable solutions and which forced their designers to create sophisticated mechanisms that turned out to be too heavy, complicated or too expensive? Or maybe some of these assumptions are just, but haven't been yet properly addressed? In order to try answering some of these questions we have built our own lever wheelchair pre-prototype.

FUNCTIONAL ASSUMPTIONS FOR OUR PRE-PROTOTYPE

The functional assumptions for our lever wheelchair were derived from the purpose of constructing our pre-prototype: testing various ideas found in world-wide designs. Therefore the result-wheelchair was from the beginning condemned to become a test unit and not a real life wheelchair. As a result in our study we decided, that our lever-wheelchair will be:

1. Oriented for outdoor, terrain usage;
2. Dedicated for 50 percentile man (Polish);
3. Allowing ride in both front and rear direction achieved by the same move of pushing the levers;
4. Providing 3 changeable gears during ride in front direction;
5. Equipped with a brake that would work in 4 conditions: disallow rolling down a slope while allowing (at the same time) moving forward; disallow any movement (parking brake); allow movement backward while disallowing movement forward; allowing movement in both directions;
6. Allowing the user to control all brake and gear changes without moving hand from the grip;
7. Designed to allow slight modification in lever axis of rotation position (performed by a technician);
8. Designed to allow modification of levers' length (performed easily, also by wheelchair user);
9. Providing the possibility of easy back wheel detaching and backrest folding for transport purposes.

DESIGNING THE LEVER WHEELCHAIR PRE-PROTOTYPE

In order to bring to life sophisticated functional assumptions overcoming various design obstacles was necessary. Some issues regarding wheelchair design were solved with use of market-available parts (for instance gear mechanisms were bought from Shimano) whereas the wheelchair frame was designed and manufactured by Polish wheelchair company GTM-Mobil. However most of unique wheelchair mechanisms were of our original design. Because of the goal of the whole experiment, we didn't focus on the wheelchairs appearance. Neither did we try to optimize shapes and precision of not important elements. During the design phase for us the key issue was if a given problem can be easily solved at a glance, requires longer consideration, or if even after considerable work spent on solving the problem it was hard to achieve success. For this reason some mechanisms were developed and successfully implemented, some were carried out, but with mediocre success and some were abandoned during manufacturing because of excessive design problems and complication.

Lever wheelchair pre-prototype overview

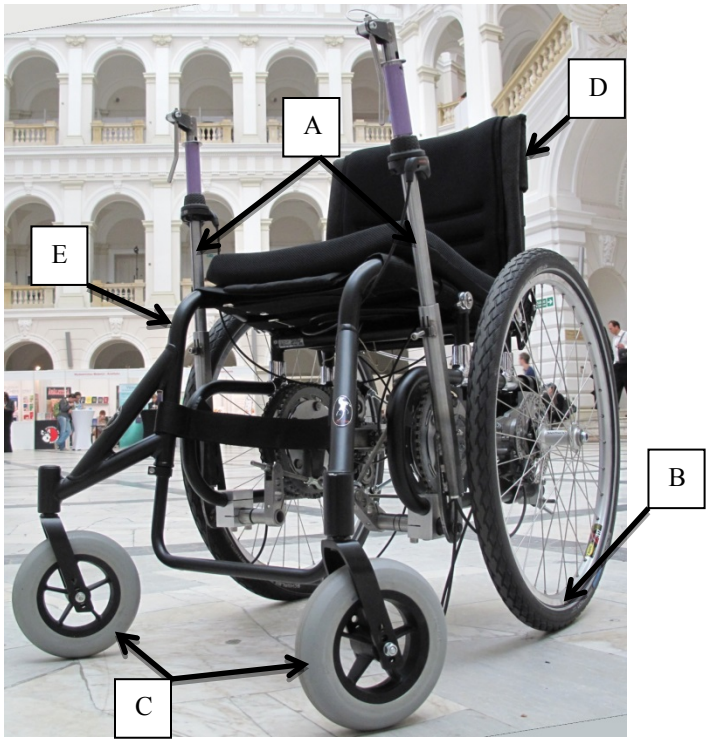


Figure 2. Overall view of the pre prototype.

The overall appearance of our wheelchair is dominated by main elements shown on figure 2: levers (A), big back (B) and front (C) wheels, typical wheelchair seat (D), frame (E) adapted to the wheels' dimensions and to carry noticeable propulsion mechanisms.

Wheelchair mechanisms

A more precise look on the prototype allows inspection of various mechanisms implemented due to established functional assumptions. The mechanisms are presented in figures 3-5. Since the levers axis of rotation position was required by the functional assumptions to be moveable, it was placed in front of the back wheels axis. Such decision implied the necessity of placing the lever in between the back wheel and the seat (when looking from above on the wheelchair). Resulting geometrical dependencies are presented in figure 6.

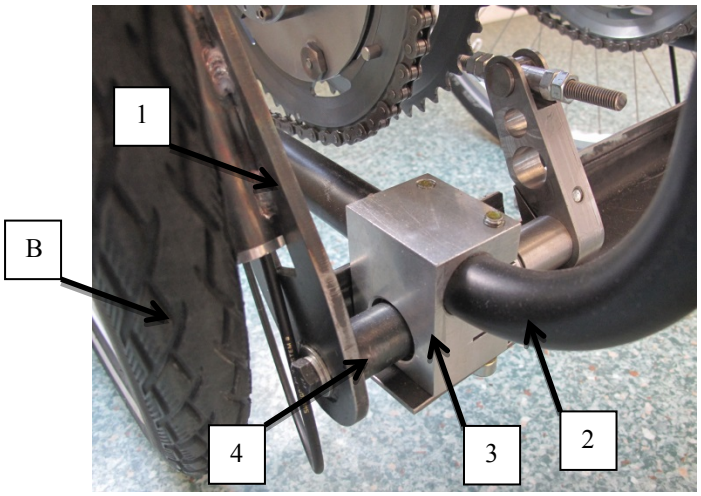


Figure 3. Moveable mounting of the levers axis of rotation: lever (1); wheelchair frame (2); levers axis rotation fastening moveable along the wheelchair frame (3); back wheel (B); levers axis of rotation (4).

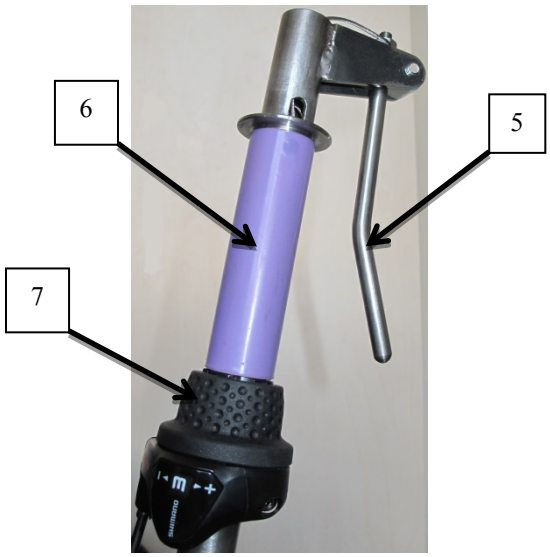


Figure 4. Grip with manipulation devices: service brake handle (5); grip with ride direction change (6); Shimano gear shifter (7).

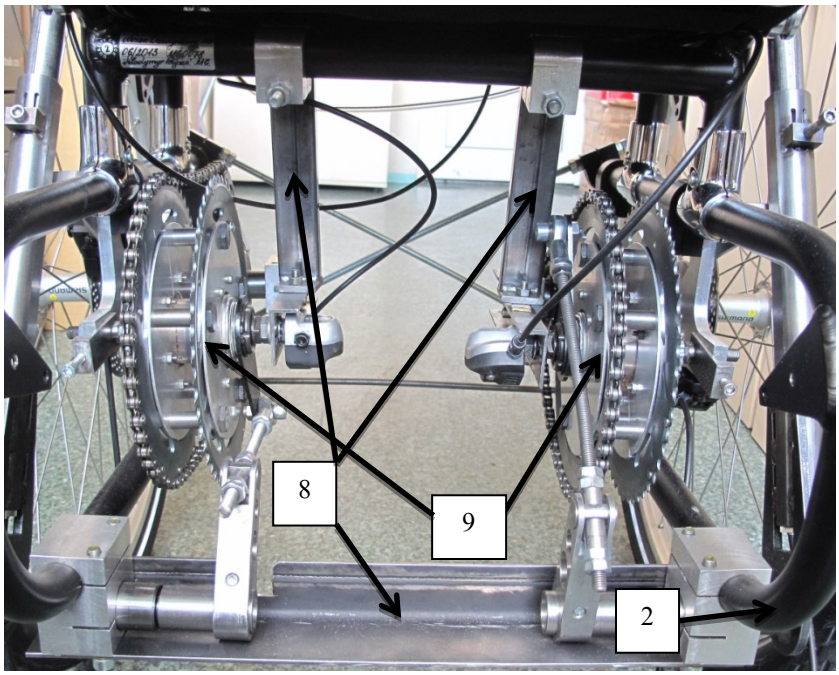


Figure 5. Stiffness increasing elements (8), wheelchair frame (2), modified Shimano 3 level gears (9).

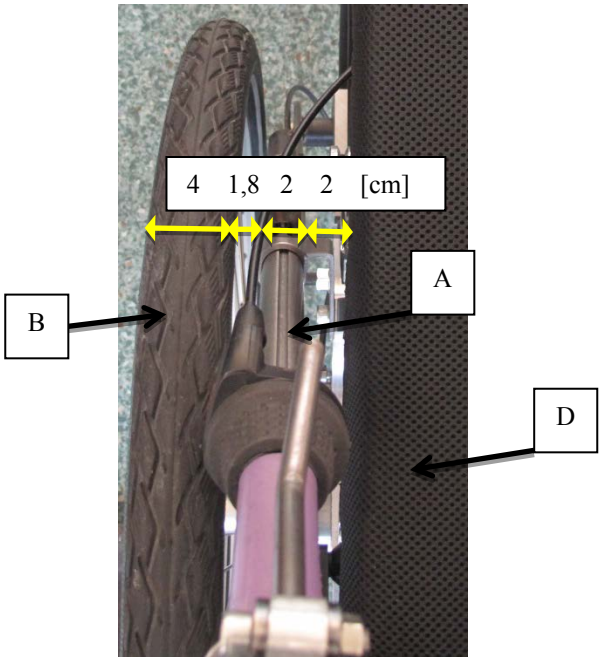


Figure 6. View of the wheelchair from the top: geometrical dependencies between back wheel (B), lever (A) and wheelchair seat (D).The distances between lever and other elements are indicated on the figure.

DISCUSSION

Implementation of mechanisms realizing sophisticated functional assumptions was a hard task and was not always carried out with 100% success. In the studied case the setbacks were:

- Implementation of innovative braking system that could realize functional assumption nr 5. The braking system had to be invented specially for fulfilling this assumption. A special spring-based solution was

proposed, which is now a subject of patent application. However the struggle to bring the invention to life demonstrated a serious of issues regarding control of the new mechanism which finally led to abandoning the concept in the discussed pre-prototype. It seems that even after overcoming encountered problems realizing this sophisticated braking system would require installing a significant number of small parts that on one hand would require precision in manufacturing and assembling and on the other hand would be probably characterized by low reliability. From our point of view at present even the positive impact of successful realization of functional assumption nr 5 would not pay for the design complication and loss of reliability. We recommend forming less demanding assumptions regarding lever wheelchair braking system;

- System for transmitting human push force from levers to driven wheels. Allowing adjustment of lever length and lever axis of rotation position required creating moveable parts and solutions which affected stiffness of the whole transmission system. As a result 2 iterations of adding elements increasing stiffness of the transmission system where carried out, however the results were still not fully satisfactory. Authors' opinion on the issue is, that this problem must be solved and the solution should be optimized from the design complication and weight point of view. Stiffness of this system is way more important than providing the possibility of lever length and lever axis of rotation adjustments. This additional functionality should only be added in case if it doesn't hinder the basic functionality. From our point of view it seems either very hard or impossible to maintain comparable systems stiffness without considerable weight growth while implementing assumption nr 7, therefore we recommend not to allow any adjustment of lever axis of rotation position;

- Implementation of functional assumptions nr 6 and 8. The possibility of manipulating gears and service brakes collided with possibility of easy changing lever length. The problem resulted from cords used to mechanically connect grip and manipulators with effectors (gears and service brakes). As a result changing lever length in the pre-prototype can be done with ease, however after performing such change the service brake and gear cords need to be adjusted in order to function properly. This cannot be done easily by the wheelchair user. Regarding changeable lever length: adding this functionality to a lever wheelchair design seems attractive and possible, however it requires clever design solutions not to hamper the more important functional aspect – manipulating gears and brake system without taking hand off the lever by the wheelchair user;

- System for forward-rear gear changing that would allow forward and back movement to be performed by pushing the levers (functional assumption nr 3). Authors of the pre-prototype tried to approach this issue in several ways, finally choosing, what seemed at the time, the best solution. In the end realization of this functional assumption caused major complication of propelling mechanisms and weight increase. Manufacturing and assembling parts without sufficient precision led to unsatisfactory performance of the mechanism. Namely one side of the wheelchair propulsion mechanisms worked almost correctly while second never managed to achieve the same performance. As a result during the alteration phase the mechanisms were uninstalled and substituted with a simpler version allowing only forward movement. In this case authors strongly recommend erasing this assumption. In case of lever wheelchairs designed for outdoor use the forward movement is dominating in 99% of the time. Moving backwards is only useful for maneuvers which do not require high speed or strength. Therefore in our opinion it is worth allowing the wheelchair user to pull levers during backward movement and pushing during forward movement: as a result the lever wheelchair gains simplicity, reliability and loses weight;

- Grip with ride direction change (6) shown on figure 4 didn't work properly. The authors design was assuming that moving the grip up and down should change the ride direction. During tests it happened quite often that the user unwillingly changed the ride direction of one of the wheels because of the force exerted by the human arm that was aimed to push the lever. Unfortunately it appeared that it also moved the grip undesirably. Probably this issue can be solved and the gear changing mechanism can be integrated with ride direction change;

- Figure 6 demonstrates geometrical dependencies between wheelchair main elements shown in the view from the top. The distances between lever, back wheels and seat shown in the figure are small. Our tests shown that, in some cases necessary, exerting high force values (200-300 N) by the human arm and transferring it onto the lever caused slight "shaking" of the lever to the sides. This shows that the stiffness of the lever and the joints in the system for transmitting human push force was too low. Our conclusion in

this case is that it might be hard to increase this stiffness in the given case. Increasing it in another design is possible, but will definitely cause weight gain.

Aside from negative conclusions, our experiment with constructing a lever wheelchair pre-prototype also led us to positive ones regarding some assumed functionalities:

- In the pre-prototype easily folded backrest was achieved by using a pre-designed market available solution (GTM-Mobil) and the possibility of detaching back wheels easily was achieved by our original design solution and worked in 100% as expected. Adding this facts led us to conclusion, that functional assumption nr 9 can be easily implemented without excessive complication of the design. Adding the fact that this functionality makes transport of a lever wheelchair much easier we highly recommend implementing this assumption;

- Designing the lever wheelchair as an outdoor transport mode seems just since its advantages are high human work efficiency and possibility to achieve high speeds (in comparison to classic push rim wheelchairs). As demonstrated in our pre-prototype, in a lever wheelchair it is not a problem to increase additionally front and back wheel sizes in order to make overcoming land obstacles easier. Performing tests on grass terrain showed how important these changes in design are. We strongly recommend adapting wheelchair tires and wheel sizes to the predicted terrain type of usage, as this does not hinder any other functionalities and fully compensates small increase of weight;

- Providing 3 changeable gears in the front ride direction required installing additional mechanisms in our wheelchair and mounting them onto the frame was a challenge. It complicated slightly the design and also added considerable weight (around 2,5 kg total). Nevertheless during test rides it proved to be worth installing. Building a lever wheelchair without changeable gears would either make it impossible to achieve any noticeable speed or make it very hard to start riding. Therefore we recommend implementing functional assumption nr 4 with the remark that it is positive to have wide spectrum of gear values (for example wider than found in some bicycles ranging from ~0,5 to ~2).

Our last and summing up conclusion was that probably designing lever wheelchair with levers axis of rotation position in the back wheels axis of rotation might be the right choice. Such solution would allow not to install many supporting, stiffness increasing elements. This would considerably decrease weight of the whole design. Also, implementing such idea would solve the problems shown on figure 6 – the lever can be placed simply "outside" the wheelchair, not in between of the back wheel and seat. In order to fulfill this proposal a designer willing to have changeable gears in the wheelchair will need to propose original gear mechanisms. Also, this approach would require original braking mechanism. Nevertheless, the work dedicated to realize this idea should pay out in a lightweight, compact design. From our point of view the Wijit design shown in figure 1 seems to be a proof that realizing this concept is possible.

ACKONWLEGMENT

The described research was financed from the ECO-Mobility project WND-POIG.01.03.01-14-154/09. The project was co-financed from the European Regional Development Fund within the framework of Operational Programme Innovative Economy

REFERENCES

Blohmke F, Boenick U, Grazianski T, Nietert M. (1975), "*Result of a user inquiry for the improvement of hand propelled wheelchairs*" (author's transl). Zietschrift fuer Orthopaedie und ihre Grenzgebiete. pp. 264-270.

Innovations Health Devices (2012), Wijit project website http://www.wijit.com

Massachusetts Institute of Technology (2014) website http://web.mit.edu

Taylor W. G. Patent USA. 5865455. (1999)

van der Woude LHV, Dallmeijer AJ, Janssen TWJ, Veeger D. (2001), "*Alternative Modes of Manual Wheelchair Ambulation: An Overview."* American Journal of Physical Medicine & Rehabilitation; Volume 10 No. 80 pp.:765-777.

Analysis of Ride Comfort on the Stairs Climbing Wheelchair

Grzegorz Dobrzyński , Włodzimierz Choromański and Iwona Grabarek

Faculty of Transport
Warsaw University of Technology
00-662 Warsaw, Poland

ABSTRACT

The paper presents an issue from the field of disabled people transport - analysis of ride comfort on the stairs climbing wheelchair. These studies were carried out in a simulation environment, and on a model prototype in the laboratory. The study was conducted for the author's design of a stairs climbing wheelchair and a wheelchair which is available in the commercial offer.

Keywords: Wheelchair, Ride Comfort, Simulation Research, Multi-Body Simulation (MBS), CAD

INTRODUCTION

Contemporary systems of computer stimulation allow to conduct experiments with some models of vehicles at the construction stage. The possible influence on people who would use such a model of vehicle is analyzed at this stage. There are some research works, which provide wheelchairs and tools research experience from motor vehicle traffic issues (Quaglia et al., 2009), (Suyang Yu et al., 2010). Restricting the area of research to the publications that concern only simulation models of wheelchairs that overcome barriers(obstructions) one can only find few R&D centers in the world of interest in the subject matter. The authors of this thesis in their previous publications (Choromanski et al., 2009), (Choromanski et al., 2007), (Dobrzynski, 2010), (Dobrzynski and Choromanski, 2010), (Choromanski et al., 2010), (Choromanski et al., 2012) have presented their own innovative design solutions and the results of the follow-up work were the basis for the development of the patent application.

In other research centers/scientific centres some authors from the USA in their thesis (Ding and Cooper, 2005) regularize the issue of new wheelchair design with have some special capabilities. In addition to defining the current situation, they suggest new directions of inquiry. An example of a computer analysis using simulation tools and optimization of design parameters to the real model are the development of the University of Castilla-La Mancha Campus, Ciudad Real, Spain. This is an example of a complex approach towards solving a scientific problem in terms of building/constructing a new means of transport. Research works (Morales et al., 2006), (Gonzalez et al., 2008), (Morales et al., 2007) present further steps of analytical considerations for static and dynamic models. The effect of various parameters of the control system and its impact on a person using the wheelchair have also been analyzed. A similar problem analysis methodology was adopted in the research center in Italy. The concept of a wheelchair overcoming the obstacles/barriers presented in the research work (Quaglia et al., 2011) is significantly different but the scope and problem analysis tools are very close.

The thesis (Ito, 2009), (Sugahara et al., 2011) present different design solutions of wheelchairs overcoming the obstacles (stairs climbing wheelchair) with an emphasis on analytical and experimental considerations. The authors from Japan have presented a very unique solution of a vehicle driven by the power of hand muscles. The Taiwan authors in their thesis (Lawn and Ishimatsu, 2003) presented a mechanically complex design where they limited the analysis of overcoming the obstacles only to conditions of keeping the static stability. A similar approach was also

presented in the thesis (Chun-Ta et al., 2008).

It was the subject of research that differentiated all the above mentioned works. Each of the research centers offered their design concepts and led the research for them. Despite the work on different structures, researchers pursue their projects in a similar range, using similar methods of analysis and simulation tools.

The purpose of the simulative research presented in thesis was an analysis of the features of a model with the regard to its dynamic characteristics in terms of a ride comfort. The subject of simulation was the implementation of the task of the entry stairs and analysis of the ride comfort. This work was carried out for the simulation model *EXPLORER* wheelchair which is in the commercial offer.

THE SIMULATIVE INVESTIGATION OF PROTOTYPE DYNAMICS

While building a nominal prototype model for the purpose of analyzing the dynamics simulation nominal in terms of overcoming the obstacles such as stairs climbing, the following simplifications were made. The wheelchair prototype is defined in 3 dimensions and has a symmetrical construction in relation to the vertical of the XY plane. The location of the model will be determined in terms of inertial frame of reference OXYZ associated with ground like in the Fig. 1.

The nominal model structure consists of one-box three-dimensional rigid elements, with the bulk parameters. The human model reflects the solid lump of defined parameters, named on the dummy/manikin. The model of the frame and the seat of the wheelchair represents one item that reflects also the bulk parameters of other parts connected to the frame, such as batteries, actuators, electronic module, front wheels, gears, etc. The leveling skid is connected on one end by a pair of rotational kinematic frame of the trolley. The other end of the bars shall be based on the ground. Windmill rolling is connected, in the middle of its length by a pair of rotational kinematic frame of the trolley. The ends of the arms of a windmill will alternately be based on ground. The rear wheels in its axis are connected with the frame by a pair of rotational kinematic. In rotating pairs listed above there is no friction. The structure of the model together with indications was presented on Fig. 1.

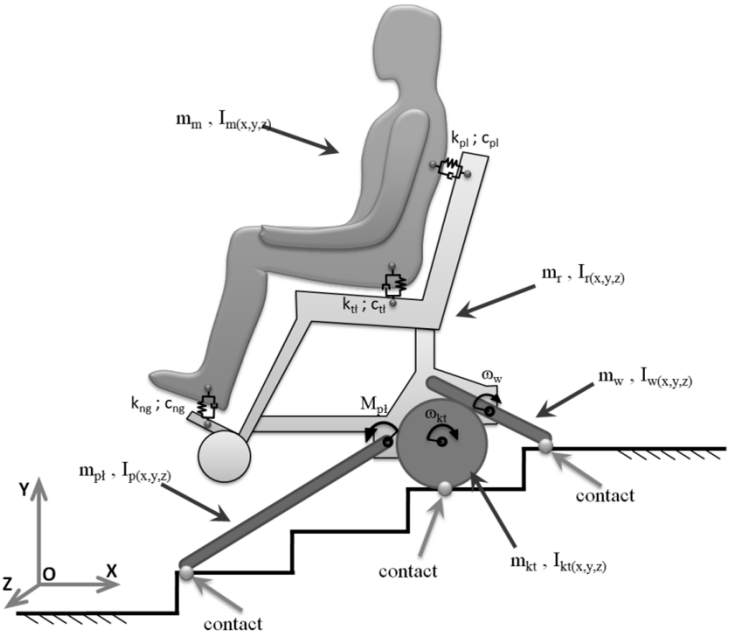


Fig. 1. The nominal model to study the dynamics of the overcoming obstacles

Due to the fact that in the process of overcoming obstacles like stairs, front wheels do not come into contact with the ground, they were not treated as a separate part, and mass parameters were attached to the frame.

Model of Contact Wheels, Skids and a Windmill with the Ground

Contact phenomena in dynamic passes decide, inter alia, the stability of motion, the intensity of use of the material-support, ability to change the direction of movement, etc. (Blundell and Harty, 2004). In this chapter, these phenomena will have a significant impact on the simulation results of overcoming obstacles. Look at the modeling issues contact phenomena have been assuming that the wheelchair is modeled as a mechanical discreet. This limits the scope of the applicability of the model to a relatively low frequency. At work examined the scope of the: 0.44-89,09 Hz.

A simplified model of the ground was adopted– it is rigid and consists of flat elements reflecting the floors and surfaces. During normal overcoming of the obstacles there are: hitting, rubbing and slippage. All of these phenomena must be reflected in the modeled system to properly put the actual mileage. Contact model shown below will be used to describe this phenomenon between the ground and the:

- rear wheels (right and left),
- rolling the Windmill (Windmill arm right and left),
- leveling skid

To describe the normal forces occurring in the contact with the ground such equation was taken:

$$F(y,\dot{y}) = \begin{cases} k(R-y)^{e} - c\dot{y} & dla \quad y \le R \\ 0 & dla \quad y > R \end{cases}$$

where:

y – is the distance between the axis of the wheels and the ground,

y – current speed changes the distance between the axis of the wheels and the ground,

R – is radius of the wheels,

k – modulus of elasticity,

e – exponent of elasticity,

c – the damping factor.

The normal force is potęgowo-squrely proportional to values and proportional to the speed of this strain. In the present analysis the deformation must be understood as deflection through the item when in contact with the ground. The nature of the component of normal force elasticity depends on the value of the exponent power and for the value of e = 1 its increment is proportional to the deformation, for e<1 it reflects its course < so called soft elasticity and for e>1 is tough elasticity (Documentation - MD.Adams)

Process of attenuation depends on the values of the deflection d with a value which is the damping coefficient changes course (c) during a transitional period (Fig. 2). In addition to the transition period, the item will come in contact, normal suppression component is null and the value of the damping for the deformation of the larger value (d) retains a fixed maximum value.

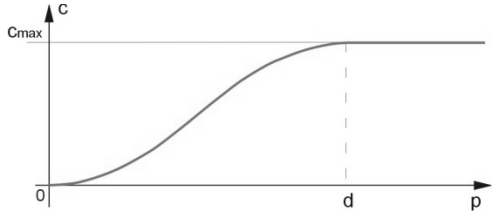
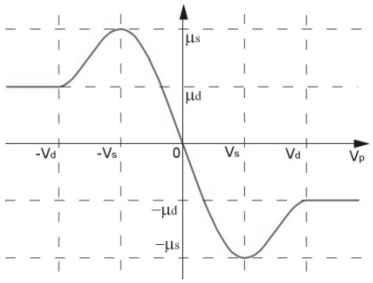


Fig. 2. Dependence of the Suppression of deformation (deflection) p where p = R-y (Documentation - MD.Adams)

In this, the model to reflect the phenomena of friction coefficient of friction course is variable characteristics Fig. 3.

The characteristics of this it is included both a static coefficient of friction and dynamic.



Rys. 1 Fig. 3. Process according to the coefficient of friction of sliding velocity Vp (Documentation - MD.Adams)

where:

- μ_s – static friction coefficient,
- μ_d – the dynamic friction coefficient,
- v_s – the threshold speed for μ_s,
- v_d – the threshold speed for μ_d,
- v_p – sliding velocity at the point of contact.

The Program has a contact model indicated Adams and it was used in the analysis (Documentation - MD.Adams). Contact Model will be used to describe the phenomena in the simulation model, for which parameters were nominal and they are presented in previeous work (Dobrzynski, 2013). That characteristics have been the subject of analysis of variants in which the simulation results were compared for their different values. More broadly they were described at thesis (Dobrzynski, 2013).

The Simulation Model of the Human Body and the Model of the Wheelchair-Person Configuration

The simulation model used in the research was simplified to a rigid body. Similar simplification were used in the thesis (choromanski, 2009). In addition, the guidelines used in the standard PN-ISO 7176-11 wheelchairs: the dummies dedicated to research (PN-ISO 7176-11:1998) and anthropometry of adult population for the purpose of Polish design (Batogowska and Słowikowski, 1994) were used in these measurements. The works mentioned above contain a description of the geometry of the dummy and their weight, but does not provide information about the position of the Centre of gravity and moments of inertia. This information is necessary for the proper construction of a simulation model of the human body.

To build a solid multisegmental model the dummy representing the human figure and give its parts uniformly distributed mass, corresponding to the human body the Software CATIA V5R20 and in particular the Ergonomics Design & Analysis were used. This tool has an in-built anthropometrical parameters library with the ability to alter them.The Model generated by this tool can be used to determine the spatial structure analyses for ergonomic workplace.

Portions of this dummy such as thighs, drumsticks and shoulders do not have independent parameters. All the dummy has assigned to the mass and the position of the center of gravity, but does not have information about the moments of inertia.

To specify them additional geometric models that correspond to the indicated portions have been built. The guidelines were taken from standards (PN-ISO 7176-11:1998). It indicates the value of the dummies' mass used in research and how to place the dummy in the wheelchair. There are three components of the construction of the model and their individual weight: body, thighs, drumsticks. There have been chosen three variants of the model with the masses: 50 kg, 100 kg and women's and the selected variants correspond to: 5 centile for women (48 kg), 50 centile for men (76 kg) and 95 centile for men (98 kg) according to (Batogowska and Słowikowski, 1994). These values are within the tolerances specified by the standard.

The dummy Model of mass parameters in line with variations, was introduced in the form of one section of the

mechanism, which has been given a model binding.

The standard (PN-ISO 7176-11:1998) provides general guidelines for the placement of dummies on the wheelchair, in order to carry out the research. The body block must be located as close as possible to the backrest of the seat, in an equal distance from each of its sides. The rear edge of the legs has a overlap with the rear edge of the footrest. According to those guidelines, each of the models was placed on his seat of the wheelchair according to the indicated Fig. 4.

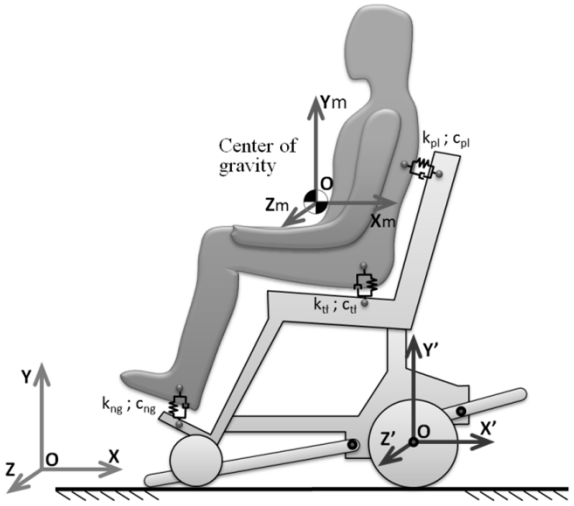


Fig. 4. The user coordinate system for the model definition dummy for his seat

While building a connection to the elastic-suppressant of the body block with the wheelchair seat the thesis (Gągorowski, 2009) was used. The same parameter values were used for all variants of the mass of the dummy. In the software used here, there is a possibility to build more complex human models, as well as different seat connections models. Because of the complexity of such activities and the fact that in this analysis the first and foremost goal was to recognize the problem of overcoming the obstacles the identified models had no complex details. The use of a one-block model, prevents also the possible assessment of vibration propagation in different parts of the body.

SIMULATION STUDIES

The purpose of the study on the simulation studies concerning the task of climbing the stairs, trolley-man was the analysis of dynamic phenomena. On its basis there was an attempt to assess the applications of the proposed concept as the means of transport. Simulation Model built in MD. Adams (Fig. 5.). In accordance with the assumption adopted for the nominal model, five parts of a related relevant constraints have been built.

Fig. 5. Simulation Model prepared in MD. Adams

The equations of motion integration GSTIF method-SI2 developed by C. W. Gear have been used. It uses the Backward Differentiation Formula (reverse method differentiation). This is the method for integrating with variable row, variable step and is dedicated to solving stiff differential equations systems (Documentation - MD.Adams).

Determination of Acceleration Model Elements

One of the essential features of the simulation model is the ability to determine the acceleration of vibrations, which are its elements. As the most important, from the point of view of effects on humans, the contact area of the trunk with the seat wheelchair (Grabarek, 2009), (Grabarek, 2008). The Fig. 6 shows the process changes the weighted vibration acceleration in X, Y, Z directions which affect the dummy model.

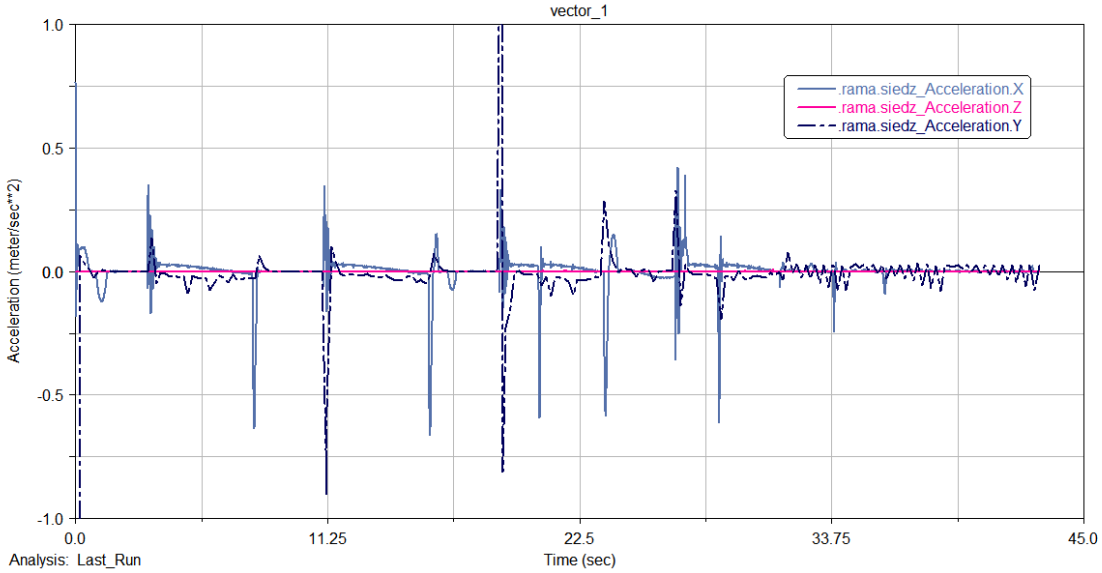


Fig. 6. Process changes, vibration acceleration in X, Y, Z directions which affect the dummy

Presented characteristics allows you to assess the nature and parameters of the Dunnage between the seat and the stroller frame. The results obtained have been exported accelerations to the DASYlab to analyze vibrations affecting human model in the course of the entry staircase, which will assess the comfort (Kardas-Cinal, 2010), in accordance with ISO 2631-1: 1997.

Analysis of Vibrations Affecting Humans in the Course of the Entry Stairs

Tests carried out at this stage of the development of the concept of design allow you to assess the impact of driving conditions on a moving truck. This is the use of simulation for virtual prototype testing. Analysis of vibrations affecting humans in the course of the entry stairs made it possible to assess comfort, in accordance with ISO 2631-1: 1997 Mechanical Vibration and Shock-Evaluation of Human Exposure to Whole-Body Vibration-General Requirements (ISO 2631-1:1997).

This standard applies to the measurements of vibration and transient, contemplated in 3 aspects: the harm, nuisance and comfort. It includes the following parameters of vibration: vibration acceleration value impact (equivalent to speed up or adjusted), the period of oscillation frequency in the series tripples or octave and the direction of impact. Accelerations are analysed in Cartesian coordinates XYZ, which determines the starting place of penetration of vibrations to the body. The Y direction is determined by the spine. In this thesis only the method of evaluation of ride comfort has been used. The analysis was limited to a single measuring point, located on his seat at the transmission of vibrations at the trunk of the dummy.

Method involves registering fixed time passes accelerations in three directions, in the accepted measurement point. The entry must contain data from a representative period, but its length is not strictly defined. Then the nature of the assessment should be carried out on an interim basis on the basis of the model vibration on the crest factor.

$$k = \frac{a_{wpeak}}{a_{wRMS}}$$

where:

a_{wpeak} – peak value of the weighted vibration acceleration in m/s2

$a_{w\,RMS}$ – effective value weighted vibration acceleration in m/s2

Depending on the result obtained, other values of the weighting factors are applied. In case it is from 1 to 9, the coefficients are used. When it is 9 or more some additional factors are used. Signal for each axle is then divided into frequency bands one-third. Process each band is multiplied by the corresponding weighting factor specified in the standard. For each band the value of the effective (RMS) is calculated . Then the value of successful bands for the three directions are summed up in vector, and the result is the basis for the comparison with the threshold values.

The next step was to calculate the value of the coefficient of the Summit. Its value amounted to 1.34, so in the next step the basic values of the weighting factors were used. A tool was built to calculate the weighted RMS value of the frequency.

In accordance with the guidelines measuring point corresponding to the site of penetration of vibrations in the human body was set. The point is associated with the domain seat wheelchair and it is located at the base of the dummy's torso. In signal analysis the phenomena occurring in 1 second was omitted, because they are associated with a model stabilization in the early stage of simulation.

Value weighted RMS acceleration in the tripple waves have been presented at the Fig. 7.

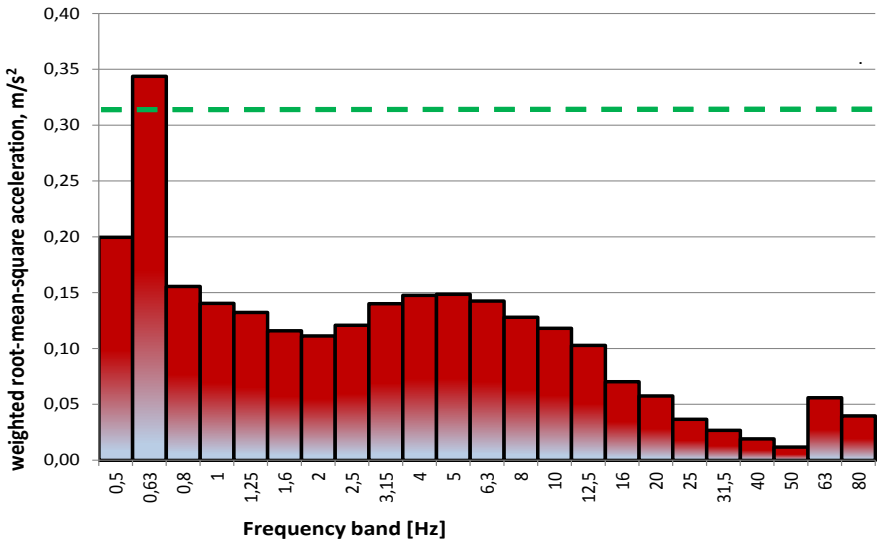


Fig. 7. Effective value of vibration acceleration recorded at the measuring points

The calculated value of weighted acceleration effect, accident frequency, have been referred to the comparative value of the threshold (Fig. 7.). The accelerates in 0,63Hz have the greatest value and in reference to the comparative value of the threshold they are with the border ranges conditions comfortable and slightly uncomfortable. The results for the remaining bands are at the limit of comfort.

The results of these analyses allow to conclude that the proposed working method of the wheelchair overcoming the stairs does not carry negative consequences in terms of ride comfort. Simplification applied at this stage of the simulation understood as the recognition of the model as a solid, rigid, could have caused an increase in the maximum values of the vibration acceleration.

EXPERIMENTAL RESEARCH OF THE PROTOTYPE WHEELCHAIR "WEKTOR"

In the following part of the work analyses the comfort analysis is presented for the prototype wheelchair. This design was developed with the participation of the authors of this paper, and is part of the patent application (Choromanski et al., 2010). In laboratory studies there has been an attempt to evaluate the ride comfort of the person moving on wheelchair prototype named "vector". The results of this analysis were compared to the results obtained for a tracked *EXPLORER*. This is a truck manufactured by Italian company TGR. This vehicle has two settings. To move around the level in the system used three wheels, two of which are driven by the rear and the front is responsible for broadcasting. The stairs and passes through tracks with independent drive to any of them, allowing you to beat any type of stairway.

The ride comfort analysis carried out for the presented prototype of a wheelchair is presented further in the paper. This design was developed with the participation of the authors of this paper, and is a part of the patent application (Choromanski et al., 2010). In laboratory studies there has been an attempt to evaluate the ride comfort of the person moving on wheelchair prototype called "WEKTOR". The results of this analysis were compared to the results obtained for a caterpillar track EXPLORER. This is a truck manufactured by Italian company TGR. This vehicle has two settings. To move around the flat surface it uses the three wheels system, two rear wheels are driven and the front one is responsible for giving directions. However, in order to go on stairs it moves by independently driven caterpillars, allowing to overcome any type of stairs.

In order to register vibrations three-axis accelerometers were used that were also linked to the registration system. A three-axis sensor was placed in the middle of the seat, in a place assumed as resultant reaction of a seat on a

passenger. The location of the accelerometer with directions of action is shown in Fig. 8.

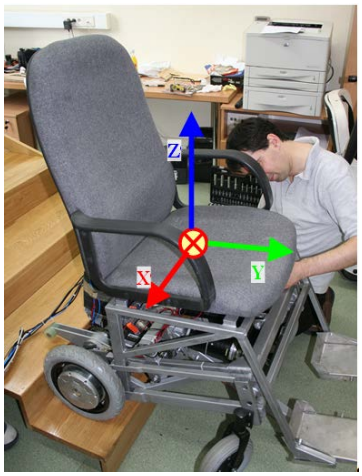


Fig. 8. Location and directions of triaxial acceleration sensor action line

Due to the nature of the direction of movement for the further analysis signals recorded in Y, and Z directions were use, where the greatest energy values were observed in Z direction. After having placed the sensor in the selected measurement point (Fig 8) wheelchairs were loaded with a mass of 75 kg. The distribution of weights at the approximate way allowed to replace the human model.

For each of the wheelchair the stage of entry and exit were registered separately. The tested wheelchairs have a different method of surmounting stairs and time of this task depends on many factors. For the purposes of this research work , it was based on the assumption that the operating range is more important which is understood as a job to perform, that is, the ability to overcome the 4 steps and not the time of analysis. The measuring signal was recorded since the moment of switching on the drive on the stairs mode until the moment of reaching the final part of the stairs and switching into the drive on horizontal surfaces mode.

The Fig. 9. below shows the timing of vibration during the entry on stairs. The *Explorer* wheelchair needed 42 seconds to overcome these stairs while the *Wektor* 54 sec. the comparison of vibration acceleration time passes have shown differences of maximum values for both.

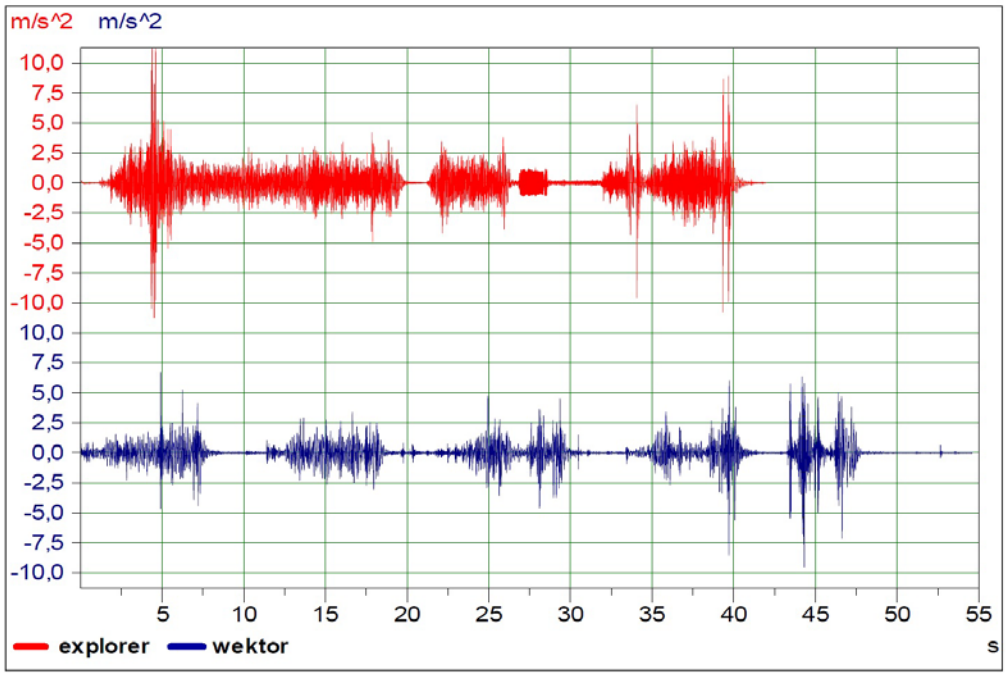


Fig.9. Timing of vibration during the entry on the stairs

The carried out analysis of vibrations affecting humans in the course of the entry stairs has allowed an assessment of comfort that is compatible with ISO 2631-1: 1997 Mechanical Vibration and Shock-Evaluation of Human Exposure to Whole-Body Vibration-General Requirements (ISO 2631-1:1997). To analyze the results the algorithm described earlier have been used. These results are described in the chart Fig. 10.

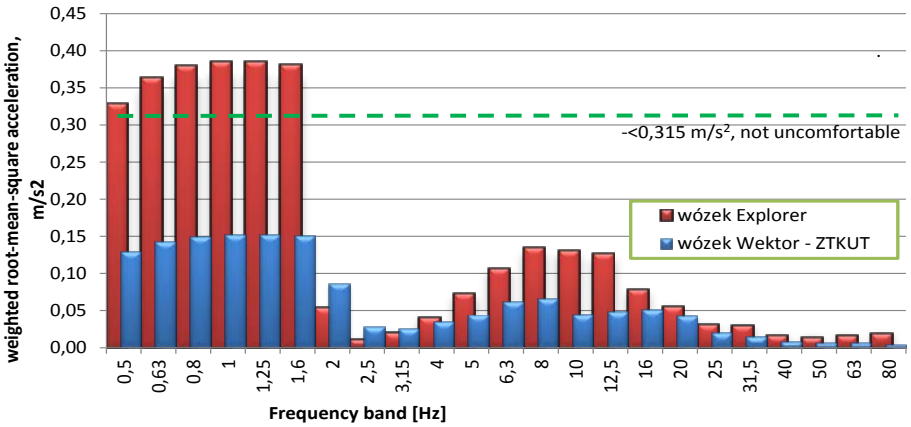


Fig. 10. RMS-adjusted values of the vibration acceleration chart for Explorer and Wektor wheelchairs

The chart shows the value of the weighted vibration acceleration RMS for Explorer and Wektor wheelchairs. The comparison of the calculated values from the comparative values of the threshold acceleration showed that the author's concept of getting results in terms of comfort and a stroller Explorer in the low frequency reaches values of slightly uncomfortable. The results indicate the validity of the assumptions of the method of overcoming stairs used in the construction of Wektor. This copyright method gives better results than commercial and recognized construction of Explorer wheelchair.

SUMMARY AND CONCLUSIONS

The thesis presented the issues related to the analysis of the model's dynamics of a wheelchair overcoming obstacles. Simulative studies has been shown together with experimental vibrational impacts on a person moving on the wheelchair. The results obtained have enabled the verification of simulation studies in terms of ride comfort. The use of simulation allows to shorten the development time and to reduce costs associated with prototyping. Another important issue of simulative studies was the simulation model validation studies, which helped to evaluate the compliance of the simulated processes with in reality. Simulation results were compared with the results of the tests on the prototype by obtaining quality compliance in terms of ride comfort.

The conclusions related to the structure of the wheelchair concern on the one hand, ride comfort and on the other hand, the way how to steer it. The completed research/studies and the evaluation of their results have shown that the simulation model design provides sufficient comfort when driving in conditions that may occur during the actual operation. It is worth underlining that the applied simplification of the framework of the cart, as perfectly solid rigid, as well as the way the dummy was designed may have an impact on the change in the value of the transferred acceleration of the dummy.

ACKONWLEGMENT

The described research was financed from the ECO-Mobility project WND-POIG.01.03.01-14-154/09. The project was co-financed from the European Regional Development Fund within the framework of Operational Programme Innovative Economy

REFERENCES

Batogowska A., Słowikowski J., (1994), "*Anthropometric atlas for the adult poulation of Poland for design purposes*", Institute of Industrial Design, Warsaw.

Blundell M., Harty A., (2004), „*The Multibody Approach to Vehicle Dynamics*", Elsevier Butterworth-Heinemann, Burlington-Oxford, UK.

Choromanski W., G. Dobrzynski, J. Kowara, (2009), *The dynamics of a mechatronic wheelchair with the function of overcoming physical obstacles*; Proceedings of IAVSD 2009 - 21st International Symposium on Dynamics of Vehicles on Roads and Tracks, 17 – 21 August 2009, KTH, Stockholm, Sweden, CD - volume I, p.143-146.

Choromański W., Dobrzyński G., Kowara J., (2007); "*Simulation Studies of New Solutions of Wheelchairs"*. Archives of Transport, vol. 19, issue 3, Warszawa 2007, str. 31-37.

Choromański W., Dobrzyński G., Potyński A., (2010), zgloszenie patentowe pt.: „*Środek transortu*", zgłoszenie patentowe nr P.393063; z dn. 26.11.2010r. BIULETYN URZĘDU PATENTOWEGO Nr 12 (1003) z dnia 4.06.2012 r.

Choromański W., Fiok K., Dobrzyński G., (2012), "*Optimizing the lever propelling system for manual wheelchairs*", Bulletin of the Polish Academy of Scienses – Technical Sciences, December 2012, pp. 793-800. DOI: 10.2478/v10175-012-0092-8

Chun-Ta Chen; Hoang-Vuong Pham, (2008); "*Enhanced development and stability analysis of a new stair-climbing robotic wheelchair*," Advanced robotics and Its Social Impacts, 2008. ARSO 2008. IEEE Workshop on , vol., no., pp.1-6, 23-25 Aug. 2008

Ding, D.; Cooper, R.A.; (2005) "*Electric powered wheelchairs,*" Control Systems, IEEE , vol.25, no.2, pp. 22- 34, April 2005, doi: 10.1109/MCS.2005.1411382

Dobrzyński G., Choromański W.; (2010), *Wózek inwalidzki z opcją jazdy po schodach – metodyka konstruowania*, Acta Mechanica Et Automatica, rok: 2010, Vol. 4, no. 2, s. 37-42.

Dobrzyński G., (2010); Badania symulacyjne i eksperymentalne wózka inwalidzkiego z opcją jazdy po schodach jako elemetu systemu transportu osób niepełnosprawnych. Materiały konferencyjne „Transport XXI wieku", Białowieża 21-24.09.2010r.

Dobrzyński G., (2013), Doctoral thesis: "*Synthesis of the mechatronic wheelchair properties as a part of the transport systems for people with disabilities*", Faculty of Transport of Warsaw University of Technology, defence of doctoral thesis: 25.01.2013, [access on 10.04.2013 at: http://bcpw.bg.pw.edu.pl/publication/4062]

Documentation „Online Help" - MD.Adams

Gągorowski A.(2009), *Metoda syntezy mechatronicznych siedzisk pojazdów z punktu widzenia minimalizacji oddziaływania drgań na człowieka*. Rozprawa Doktorska, Politechnika Warszawska 2009 rok.

Gonzalez A., E. Ottaviano, M. Ceccarelli; (2008), *On the kinematic functionality of a four-bar based mechanism for guiding wheels in climbing steps and obstacles*, Mechanism and Machine Theory, Volume 44, Issue 8, August 2009, Pages 1507-1523, ISSN 0094-114X, 10.1016/j.mechmachtheory.2008.12.004.

Grabarek I. (2009), „*Ergonomia – technika i technologia – zarządzanie*" : Metodyczne aspekty diagnozowania ergonomicznego układu operator – pojazd szynowy – otoczenie. /pod red. Marka Fertscha, Wydawnictwa Politechniki Poznańskiej, Poznań 2009, s. 59-68.

Grabarek I., (2008) „*Zasady ergonomiczne w projektowaniu układów operator-pojazd*". Wydawnictwo BEL Studio Sp. z o.o. ISBN 83-89968-28.

ISO 2631-1:1997 - Mechanical Vibration and Shock - Evaluation of Human Exposure to Whole-Body Vibration -- General Requirements.

Ito, T.:(2009) "*Simulation-Based Study Using a Stair Climbing Wheelchair*," Modelling & Simulation, 2009. AMS '09. Third Asia International Conference on , vol., no., pp.537-542, 25-29 May 2009; doi: 10.1109/AMS.2009.89; URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5072044&isnumber=5071935.

Kardas-Cinal Ewa: (2010), *Ride comfort for various passenger positions in a railway vehicle - simulation study*, Archives of Transport, Volume 22, Number 2/2010, p. 189-199, Versita, Warsaw, 2010. URL: http://versita.metapress.com/content/L5574X15L782G225

Lawn, M.J.; Ishimatsu, T.: (2003) "*Modeling of a stair-climbing wheelchair mechanism with high single-step capability*," Neural Systems and Rehabilitation Engineering, IEEE Transactions on , vol.11, no.3, pp.323-332, Sept. 2003. doi: 10.1109/TNSRE.2003.816875. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1231243&isnumber=27583

Morales R., A. González and V. Feliu: (2007) *Mechanical and Kinematics Design Methodology of a New Wheelchair with Additional Capabilities*, Bioinspiration and Robotics: Walking and Climbing Robots, Book edited by: Maki K. Habib, ISBN 978-3-902613-15-8, pp. 544, I-Tech, Vienna, Austria, EU, September 2007.

Morales R., V. Feliu, A. González, P. Pintado: Coordinated Motion of a New Staircase Climbing Wheelchair with Increased Passenger Comfort, Proceedings of the 2006 IEEE International Conference on Robotics and Automation, Orlando, Florida - May 2006, pp 3995-4001.

PN-ISO 7176-11:1998 Wózki inwalidzkie - Manekiny do badań.

Work by multiple authors edited by Włodzimierz Choromański, (2009) "*Transport of the disabled – innovative wheelchairs for persons with motor disabilities*", Publishing House of Warsaw University of Technology, Warsaw 2009.

Quaglia G., Walter Franco, Riccardo Oderio: (2009), *Wheelchair.q, a mechanical concept for a stair climbing wheelchair,* Proceedings of the 2009 IEEE, International Conference on Robotics and Biomimetics, pp. 800-805, December 19 -23, 2009, Guilin, China.

Quaglia G., Walter Franco, Riccardo Oderio: Wheelchair.q, a motorized wheelchair with stair climbing ability, Mechanism and Machine Theory, Volume 46, Issue 11, November 2011, Pages 1601-1609, ISSN 0094-114X, 10.1016/j.mechmachtheory.2011.07.005. (http://www.sciencedirect.com/science/article/pii/S0094114X1100139X)

Sugahara, Y.; Yonezawa, N.; Kosuge, K.: "A novel stair-climbing wheelchair with transformable wheeled four-bar linkages," Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on , vol., no., pp.3333-3339, 18-22 Oct. 2010. doi: 10.1109/IROS.2010.5648906; URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5648906&isnumber=5648787

Suyang Yu, Ting Wang, Xiaofan Li, Chen Yao, Zhong Wang and Di Zhi: Configuration and Tip-Over Stability Analysis for Stair-Climbing of a New-Style Wheelchair Robot, Proceedings of the 2010 IEEE, International Conference on Mechatronics and Automation, pp. 1387-1392, August 4-7, 2010, Xi'an, China.

Traffic Fatalities and Serious Injuries in Europe – A study of association with cultural, demographic and income variables

Isilda G. Barata [a] *and Denis A. Coelho* [b]

[a] *unaffiliated*
Covilhã, Portugal

[b] *Human Technology Group, Dept. of Electromechanical Engineering*
Universidade da Beira Interior
6201-001 Covilhã, Portugal

ABSTRACT

This contribution reviews key data on road traffic fatalities across Europe and seeks associative links between the afore-mentioned data and per capita income, new car sales and national cultural dimensions of European countries. The rate of population shift from poverty to affluence has been exponentially growing in many countries, with access to products and systems having increased. Understanding the association between cultural dimensions and road fatalities can point, at the European level, to the countries, where national dimensions of culture represent a predisposition for greater or less risk adversity and safe or unsafe road behaviors. This study may hence assist decision makers, nationally and at European level to tailor prevention campaigns and safety programs and measures to each national reality and culture. The regression model showed a single independent variable within the model of traffic fatalities per capita, which was power distance.

Keywords: Power Distance, Uncertainty Avoidance, Traffic Safety, Risk Taking in Traffic

INTRODUCTION

Previous studies on car seat comfort (Coelho & Dahlman, 1999, 2012) and car design and infrastructure paradigms (Camboa & Coelho, 2010; Coelho & Camboa, 2010) touched upon safety concerns. In the current paper, an association approach (Coelho et al., 2013; Coelho, 2011) is used to identify cultural predispositions for behaviors conducive to traffic fatalities. In the EU, for every person killed on the road, an estimated ten are seriously injured. Some 250 000 people are estimated to be seriously injured in road accidents every year - compared to the 28 000 road fatalities in 2012 (europa.eu/rapid/press-release_MEMO-13-232_en.htm). An appropriate mix of legislation, awareness-raising, enforcement, engineering, cooperation and knowledge transfer among relevant stakeholders, for example through the European Road Safety Observatory, plus research support will be needed.

Per capita income, new car sales and national cultural dimensions (Hoefstede, 1980; Smith-Jackson & Essuman-Johnson, 2014) are used as independent variables against which association is tested with regard to fatalities and serious injuries, taking each EU country as a case. Demographic pyramid information, focusing on the percentage of citizens over 64 year of age (http://ec.europa.eu/transport/road_safety/specialist/knowledge/old/index.htm) is also used as independent variable. Additionally, the variance in national income is also characterized and used in the

association study.

TRAFFIC RELATED DATA

Road Deaths

Road fatalities across the EU have decreased by 9% in 2012. In the following EU-27 countries are considered, in data for 2012 (or as close to this year as possible and available) (Table 1).

Table 1 - Country by country statistics on road deaths for 2012 (source: europa.eu/rapid/press-release_IP-13-236_en.htm)

	Fatalities per million inhabitants				Evolution of fatalities		
E. U. Member State	**1965**	**2010**	**2011**	**2012**	**Average annual decrease 2000-2010**	**2010-2011**	**2011-2012**
Belgium	147	77	78	73	-6%	2%	-12%
Bulgaria	91	103	89	82	-3%	-15%	-8%
Check Republic	150	76	74	71	-5%	-4%	-4%
Denmark	212	46	40	32	-6%	-14%	-18%
Germany	234	45	49	44	-7%	10%	-10%
Estonia	178	58	75	65	-10%	29%	-14%
Ireland	124	47	41	36	-7%	-12%	-12%
Greece	89	111	101	92	-4%	-9%	-10%
Spain	114	54	45	41	-9%	-17%	-9%
France	249	62	61	56	-8%	-1%	-8%
Italy	186	68	64	62	-6%	-6%	-2%
Cyprus	162	73	85	59	-5%	18%	-28%
Latvia	290	97	86	86	-10%	-18%	-2%
Lithuania	250	90	97	100	-9%	-1%	2%
Luxembourg	250	64	64	65	-8%	3%	3%
Hungary	86	74	64	60	-6%	-14%	-5%
Malta	36	36	51	26	-1%	40%	-48%
Netherlands	202	32	33	32	-7%	2%	-1%
Austria	252	66	62	64	-6%	-5%	4%
Poland	79	102	109	93	-4%	7%	-15%
Portugal	117	79	84	71	-6%	-7%	-16%
Romania	98	111	94	96	0%	-15%	1%
Slovenia	327	67	69	59	-7%	2%	-13%
Slovakia	128	68	60	55	-5%	-13%	-9%
Finland	230	51	54	48	-5%	7%	-11%
Sweden	170	28	34	31	-8%	20%	-7%
United Kingdom	146	31	31	28	-7%	3%	-12%
EU	171	62	60	55	-6%	-2%	-9%

In 2011, some 1.5 million people were reported injured on EU roads. Of these, around 250,000 were serious injuries (European Commission, 2014-a). But different EU countries use different reporting criteria, so the real figure – based on consistent medical criteria – could actually be significantly different. Comparable statistics across the European Union should be available, starting from the end of the year 2014. Hence, the following analysis considers comparable data across the European Union countries.

Older drivers have a relatively high fatality rate. Older drivers have the second highest fatality rate. Only the youngest group of drivers (18- and 19-year olds) has a higher fatality rate. Driver's age and fatality rate shows a well-known U-shape: fatality rates of the young drivers are high, after which the rate declines to a minimum for drivers of the age of 40-60 years. Then it increases again, to a maximum for those aged 75 and older. The injury rates, plotted against driver age also show a U-shape, but this one levels off on the right-hand side. Whereas young drivers (18-24 years old) have a relatively high injury rate as well as a high fatality rate, the injury rate for older drivers is much lower than their fatality rate (http://ec.europa.eu/transport/road_safety/specialist/knowledge/old/older_drivers_risky_or_at_risk/high_fatality_rate_more_crashes_or_more_severe_injuries.htm).
Individuals who have a low social status are more frequently involved in road accidents than individuals who have a high social status (ETSC, 2007). Research has been called for to study the association between social status and road accident risk, at the individual level. Based on a review made by Sælensminde (2001) and the review of de Blaeij et al. (2004), Table 2 shows the official monetary valuation of a road accident death in a number of countries according to ETSC (2007).

Table 2 – Official monetary valuation of a road accident death in selected European countries (ETSC, 2007)

country	**Monetary valuation (Euro in 2002 prices)**
Portugal	55,812
Spain	150,253
Greece	206,087
Poland	221,392
Belgium	462,717
France	589,177
Denmark	679,737
Italy	791,748
Austria	899,014
Ireland	1,170,695
Germany	1,266,000
Finland	1,273,372
Netherlands	1,741,000
Sweden	1,954,000
Switzerland	2,010,000
Great Britain	2,107,000
Norway	2,707,000

Passenger Vehicle Age, Density and Renewal

In 2010 cars in the EU were on average 8.3 years old (source ANFAC, ACEA, 2012), average car age for 12 European Union countries is shown in Table 3. Moreover, passenger car density per 1000 inhabitants in the European Union is also shown in Table 3. New passenger car registrations per 100 inhabitants in units for 2012 and 2012 GDP per capita are shown in Table 4.

Table 3 — Average passenger car age by country in 2010 (source: ANFAC, ACEA, 2012) for vehicles in use and passenger car density per 1000 inhabitants in the European Union (2011) (source, ACEA 2012)

country	average car age (2010)	passenger car density per 1000 inhabitants (2011)
Luxembourg	-	658
Italy	-	610
Malta	-	589
Lithuania	-	570
Finland	11.9	551
Cyprus	-	545
Austria	7.5	535
Germany	8.3	525
Slovenia	-	519
France	8.2	502
Belgium	8.0	490
Spain	-	482
Poland	-	470
Netherlands	-	470
United Kingdom	7.3	466
Sweden	9.8	464
Greece	10.7	461
Portugal	10.1	447
Czech Republic	-	436
Estonia	12.0	428
Ireland	6.3	417
Denmark	-	394
Bulgaria	-	368
Slovakia	11.5	324
Latvia	-	300
Hungary	-	298
Romania	-	203

Table 4 – New passenger car registrations per 100 inhabitants in units for 2012 (source ACEA, 2014) and 2012 GDP per capita in PPS (Index EU28=100; Eurostat, 2014)

country	New passenger car registrations per 100 inhabitants (2012)	GDP per capita in PPS (Index EU28=100; 2012)
Luxembourg	9.8	263
Switzerland	4.9	158
Belgium	4.4	120
Austria	4.0	130
Germany	3.8	123
Norway	3.7	195
Iceland	3.5	115
United Kingdom	3.3	106
France	3.1	109
Denmark	3.1	126
Netherlands	3.0	128
Sweden	3.0	126
Slovenia	2.4	84
Italy	2.3	101
Finland	2.1	115
Ireland	1.7	129
Czech Republic	1.7	81
Spain	1.5	96
Cyprus	1.3	92
Estonia	1.3	71
Slovakia	1.3	76
Portugal	0.9	76
Poland	0.7	67
Hungary	0.5	67
Greece	0.5	75
Latvia	0.5	64
Lithuania	0.4	72
Romania	0.3	50
Bulgaria	0.3	47

Traffic Fatalities reported by Age Group

Fatalities reported by country (different years from 2009 to 2012) have been broken down according to age groups and shown in percentage according to age structure in Table 5.

Table 5 – Age group percentage of reported fatalities in EU countries (European Commission, 2014) and age structure for selected age groups for 2011 (Eurostat, 2014)

country	year	<15	15-17	18-24	25-49	50-64	65+
Belgium	2011	4% (17%)	2%	17%	41%	15%	22% (17%)
Bulgaria	2009	3% (13%)	3%	18%	41%	17%	18% (18%)
Czech Republic	2012	2% (14%)	2%	13%	42%	19%	21% (16%)
Denmark	2011	4% (18%)	4%	16%	29%	19%	29% (17%)
Germany	2011	2% (13%)	3%	18%	32%	18%	26% (21%)
Estonia	2009	4% (15%)	2%	22%	40%	13%	19% (17%)
Ireland	2010	3% (21%)	3%	27%	40%	13%	14% (12%)
Greece	2011	2% (14%)	3%	15%	41%	17%	23% (19%)
Spain	2011	2% (15%)	1%	11%	43%	18%	24% (17%)
France	2012	3% (18%)	4%	21%	37%	15%	20% (17%)
Italy	2010	2% (14%)	3%	14%	39%	16%	26% (20%)
Cyprus	2012	0% (17%)	4%	20%	45%	14%	18% (13%)
Latvia	2012	4% (14%)	3%	10%	40%	24%	20% (18%)
Luxembourg	2012	3% (18%)	6%	18%	41%	6%	26% (14%)
Hungary	2012	3% (15%)	2%	7%	39%	28%	20% (17%)
Malta	2010	8% (15%)	0%	31%	54%	8%	0% (16%)
Netherlands	2012	4% (18%)	2%	13%	30%	16%	33% (16%)
Austria	2011	2% (15%)	5%	15%	31%	18%	29% (18%)
Poland	2012	3% (15%)	2%	16%	37%	23%	18% 14%)
Portugal	2011	2% (15%)	1%	12%	34%	21%	30% (19%)
Romania	2012	4% (15%)	3%	11%	36%	23%	22% (15%)
Slovenia	2011	4% (14%)	3%	12%	44%	21%	16% (16%)
Slovakia	2010	3% (15%)	2%	19%	40%	21%	15% (13%)
Finland	2011	3% (16%)	5%	17%	32%	15%	28% (18%)
Sweden	2010	3% (17%)	2%	18%	31%	20%	26% (18%)
United Kingdom	2012	3% (18%)	4%	19%	35%	16%	23% (17%)

Official road accident statistics are incomplete and inaccurate in all countries. The level of reporting for injuries treated in hospitals is, on the average, less than 50%. Injuries are not always correctly classified by severity in police accident reports (ETSC, 2007). The long-term consequences of transport-related injuries within the EU are to a large extent unknown. Mortality rates are fairly well known in the different member states. Statistics on survivors are much less reliable, especially for slight injuries. These patients are usually only to a small extent included in the trauma registries or police records, even though the long-term consequences of injury might be severe (ETSC, 2007).

NATIONAL MEASURES OF CULTURE

The argument according to which culture exerts a profound influence on the innovative capacity of a society has been largely supported by empirical research. Barnett (1953) postulated a positive correlation between the individualism of a society and its innovative potential: the greater the freedom of the individual to explore and express opinions, the greater the likelihood of new ideas coming into being. Hofstede (1980) indicated that societies that score high on individualism and low on power distance tend to display higher growth and innovation rates. Shane (1993) found that individualistic societies that accept uncertainty and exhibit a low level of power distance are those who attain better innovation performance. Hussler (2004) introduced a culture-based taxonomy of innovation performance. Societies that succeed by innovating on their own are those that possess a "culture of endogenous innovation". Vice versa, those countries with high uncertainty avoidance and high power distance can be defined as "cultures of imitation". The four measures of national cultures, initially identified by Hofstede (1980, 2001), are summarized in the following paragraphs.

Power distance is the extent to which less powerful members of organizations expect power to be equally distributed (Hofstede, 1980). In low power distance countries there is limited dependence of subordinates on their bosses. Power is very decentralized as well as decision-making. In contrast, in high power distance countries, hierarchy is the fundamental principle on which all relationships are based. Power is centralized as well as decision-making, leading to more emphasis on formal methods for gathering and analyzing external information (Flynn and Saladin, 2006).

Individualism is the degree to which people are oriented towards acting as individuals as opposed to acting as a group (Hofstede, 1980). In individualist countries people tend to value individual success and achievement. Members of individualist countries are autonomous and confident, tending to rely primarily on their own ideas (Snell and Hui, 2000). In collectivist countries, people are bound in groups such as the extended family or the village and are more likely to rely on information provided by others in formulating their opinions (Snell and Hui, 2000).

Masculinity is the extent to which success and aggressiveness are valued (Hofstede, 1980). In high masculinity countries, high earnings, advancement through opportunities and challenging work are mostly emphasized. The use of information to support decision-making is dependent on its expected effectiveness in gaining advantage over competitors (Flynn and Saladin, 2006). In contrast, in high femininity countries, relationships, concern for the others, inclusiveness and society's best interest are valued. Cooperation is often a visible feature. The use of information to support decision-making is very typical of a feminine national culture (Wacker and Sprague, 1998).

Uncertainty avoidance is the degree to which people feel confident about the future (Hofstede, 2001). National cultures that score high in uncertainty avoidance have an emotional need for rules. Vice versa, national cultures that score low in uncertainty avoidance dislike formal rules, setting them only when it is necessary (Flynn and Saladin, 2006).

Long term orientation stands for the fostering of virtues oriented towards future rewards, in particular perseverance and thrift. It's opposite pole, short term orientation, stands for the fostering of virtues related to the past and present, in particular, respect for tradition, preservation of 'face' and fulfilling social obligations.

In what concerns the current validity of Hofstede's cultural measures, criticisms addressed to the construct of national culture as a suitable variable for differentiation, apply directly to all four measures (Coelho, 2011). Different corporate, organizational, industrial and/or sector specific cultures may co-exist within the same firm and might as well conflict and counterbalance the national one (Vecchi and Brennan, 2009). Furthermore, in many countries, different ethnic or national cultures co-exist (Au, 2000), as result of people mobility around the globe.

Within the same country, different sub-cultures might persist, standing apart on religious, language or ethnicity grounds. As a consequence, the four measures of national cultures could be far from being reliable proxies for cultural homogeneity for a given national culture (Vecchi & Brennan, 2009).

The data extracted from Hofstede's measures of national culture, that was used in the correlation analyses is presented in Table 6. The data has been cross-validated in an empirical study by Van Oudenhoven (2001) for Belgium, Denmark, Germany, United Kingdom, Greece, Spain and the Netherlands.

Table 6 - National culture dimensions for European countries (Hoefstede, 2001)

Country	Power Distance	Uncertainty Avoidance	Individualism (vs. Collectivism)	Masculinity (vs. Femininity)	Long/Short-term Orientation
Belgium	65	94	75	54	38
Bulgaria	70	85	30	40	-
Czech Republic	57	74	58	57	13
Denmark	18	23	74	16	46
Germany	35	65	67	66	31
Estonia	40	60	60	30	-
Ireland	28	35	70	68	43
Greece	60	112	35	57	-
Spain	57	86	51	42	19
France	68	86	71	43	39
Italy	50	75	76	70	34
Luxembourg	40	70	60	50	-
Hungary	46	82	80	88	50
Malta	56	96	59	47	-
Netherlands	38	53	80	14	44
Austria	11	70	55	79	31
Poland	68	93	60	64	32
Portugal	63	104	27	31	30
Romania	90	90	30	42	-
Finland	33	59	63	26	41
Sweden	31	29	71	5	33
United Kingdom	35	35	89	66	25

RESULTS OF ANALYSIS OF ASSOCIATION

The percentage of fatalities for under 15 years of age is strongly correlated with individualism (rho=0.619; p=0.011) and long-term orientation (rho=0.693; p=0.003). The percentage of fatalities in the age group from 15 to 17 years old is negatively correlated with uncertainty avoidance (rho=-0.546; p=0.029) and power distance (rho=-0.541; p=0.031). Interestingly, opposite results are obtained for the 25 to 49 year old group, showing strong and positive correlations with power distance (rho=0.557; p= 0.025) and uncertainty avoidance (rho= 0.560; p= 0.024).

A composite variable obtained by calculating the ratio between the two variables obtained in the two last columns of Table 6, which is a variable that provides a measure of the degree to which more people in the 65 plus age group die in traffic accidents than the demographic structure would support, showed a correlation with masculinity. It is a negative correlation (rho=-0.590; p=0.016). Additionally, the GDP per capita shown in Table 4 is negatively correlated with power distance (rho=-0.598; p=0.014).

Total fatalities per country in the period from 2010 to 2012 are very much associated with two dimensions of national culture: power distance and uncertainty avoidance. The Spearman rank order correlation coefficients and significance are shown in Table 7.

Table 7 – Association between fatalities per capita per country in the period from 2010 to 2012 are very much associated with two dimensions of national culture obtained from Spearman rank order correlation coefficient analysis (significance shown as p-values in parentheses).

Reference year	Power Distance	Uncertainty Avoidance
2010	0.654 (0.006)	0.853 (0.000)
2011	0.627 (0.009)	0.831 (0.000)
2012	0.600 (0.014)	0.817 (0.000)
2010 to 2012 (average)	0.651 (0.006)	0.857 (0.000)

Moreover regression analysis using the backward entry method (Hastie, 2009) was used as a means to test the possibility of modeling the phenomenon. The regression model showed a single independent variable for fatalities as dependent value, this was power distance.

DISCUSSION

Deaths of people under 15 years of age are associated with high individualism and high long term orientation. Individualism is the degree to which people are oriented towards acting as individuals as opposed to acting as a group. As people under the age of 15 may not be driving automobiles, but can be driving bicycles and in some countries mopeds, it would be necessary to expand the data to understand if these underage deaths result from being a passenger or a two-wheeled vehicle driver. Hence individualism can harm passengers of this age or the driver her or himself in the latter case. Long term orientation stands for the fostering of virtues oriented towards future rewards, in particular perseverance and thrift. Hence the association between this cultural dimension and fatalities in this age group is somewhat unexpected.

Fatalities in the age group from 15 to 17 years old are negatively correlated with uncertainty avoidance and power distance. This may be explained by the association of low uncertainty avoidance and low power distance with more affluent countries, where people in this age group have access to motorized vehicles (e.g. scooters). The biggest age group (25 to 49), shows positive correlation for fatalities and uncertainty avoidance as well as for fatalities and

power distance. This may be explained by the fact that societies with high uncertainty avoidance and high power distance have a tendency to 'bend' rules and have decision-making centralized.

More feminine countries have a higher rate of fatalities than more masculine growth for the age range over 65, because in the former, more women have been driving for longer and they outlive men, and keep driving well past 65. Hence, more older drivers on the road and greater longevity of women explains this association.

Interestingly, economic indicators such as GDP per capita, included in the analysis, did not show any association with fatality data across the European union countries, but in this case GDP per capita is negatively associated with power distance. This is a reminder that any many of the less affluent European countries, the difference between rich and poor has been growing.-

The results are extremely clear in what concerns power distance and uncertainty avoidance, as these are strongly and significantly associated with fatalities per capita. In high power distance countries, hierarchy is the fundamental principle on which all relationships are based. Power is centralized as well as decision-making, which strongly associates with increased traffic fatalities. Uncertainty avoidance is the degree to which people feel confident about the future. National cultures that score high in uncertainty avoidance have an emotional need for rules. Apparently traffic rules are not totally fulfilling that need, at least not in every occasion, given the high number of per capita traffic fatalities registered.

These countries with high traffic fatalities per capita also have inherent high uncertainty avoidance besides high power distance can be defined as cultures of imitation, as opposed to countries of innovation.

REFERENCES

ACEA – European Automobiles Manufacturers Association (2014). URL: http://www.acea.be/statistics/tag/category/passenger-car-fleet-per-capita. Accessed February 24th, 2014.

Au, K. Y. (2000). Intra-cultural variation as another construct of international management: a study based on secondary data of 42 countries. J. Int. Manage., 6, 217–238.

Barnett, H. G. (1953). Innovation: The Basis of Cultural Change. Mc Graw Hill, New York.

Blaeij, A. de., Koetse, M., Tseng, Y-Y., Rietveld, P., & Verhoef, E. (2004) Valuation of safety, time, air pollution, climate change and noise; methods and estimates for various countries. Report prepared for ROSEBUD. Department of Spatial Economics, Vrije Universiteit, Amsterdam.

Camboa, A. S., & Coelho, D. A. (2010). SharE--An Engineering System Concept Proposal for Sustainable Personal Mobility. International Review of Mechanical Engineering, 4(1), 106-111.

Coelho, D. A. (2011). A study on the relation between manufacturing strategy, company size, country culture and product and process innovation in Europe.International Journal of Business and Globalisation, 7(2), 152-165.

Coelho, D. A., & Camboa, A. S. (2010). Batteries, vehicle and infrastructure: interlocking elements of a new engineering system concept for personal mobility. International Journal of Electric and Hybrid Vehicles, 2(4), 298-307.

Coelho, D. A., & Dahlman, S. (1999). A pilot evaluation of car seat side support: Leading to a redefinition of the problem. International journal of industrial ergonomics, 24(2), 201-210.

Coelho, D. A., & Dahlman, S. (2012). Articulation at shoulder level–A pilot experimental study on car seat comfort. Applied ergonomics, 43(1), 27-37.

Coelho, D. A., Harris-Adamson, C., Lima, T. M., Janowitz, I., & Rempel, D. M. (2013). Correlation between Different Hand Force Assessment Methods from an Epidemiological Study. Human Factors and Ergonomics in Manufacturing & Service Industries, 23(2), 128-139.

European Commission (2014-a). Serious Injuries. Mobility and Transport. URL: http://ec.europa.eu/transport/road_safety/topics/serious_injuries/index_en.htm. Accessed February 24th 2014

European Commission (2014). Reported fatalities in EU countries ,Mobility and Transport Statistics. URL: http://ec.europa.eu/transport/road_safety/specialist/statistics/index_en.htm, accessed February 24 2014.

ETSC – European Transport Safety Council (2007). Social and Economic Consequences of road traffic injury in Europe. Brussels, Belgium: ETSC.

Eurostat (2014).Population age structure by major age groups, 1991 and 2011 (% of the total population). URL: http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Population_structure_and_ageing, accessed February 24th 2014.

Flynn, B. B., & Saladin, B. (2006). Relevance of Baldridge constructs in an international context: a study of national culture. J. Oper. Manage., 2006 24, 583–603.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). Linear Methods for Regression (pp. 43-99). Springer New York.

Hofstede, G. (1983). National cultures in four dimensions: A research-based theory of cultural differences among nations. International Studies of Management & Organization, 13(1/2), 46-74.
Hofstede, G. (1980). Culture's Consequences: International Differences in Work-related Values. Sage, London.
Hofstede, G. (2001). Culture's Consequences: Comparing Values, Behaviours, Institutions and Organizations Across Countries. Sage, Thousand Oaks, CA.
Hussler, C. (2004). Culture and knowledge spillovers in Europe: new perspectives for innovation and convergence policies? Econ. Innov. New Technol., 13 (6), 523–541.
Sælensminde, K. (2001) Verdsetting av trafikksikkerhet i ulike lands nyttekostnadsanalyser. Arbeidsdokument SM/1352/2001, Oslo, Transportøkonomisk institutt.
Shane, S. (1993). Cultural influences on national rates of innovation. J. Bus. Venturing 8, 59–73.
Smith-Jackson, T., & Essuman-Johnson, A. (2014). Cultural Ergonomics.- Overview and Methodologies. In T. L. smith-Jackson, M. L. Resnick and K. T. Johnson (Eds.) Cultural Ergonomics – theory, methods and applications. Boca Raton, FL: CRC Press.
Snell, R. S., & Hui, S. K. K. (2000). Towards the Hong Kong learning organization: an exploratory case study. J. Appl. Manage. Stud., 9 (2), 150–175.
Van Oudenhoven, J. P. (2001). Do organizations reflect national cultures? A 10-nation study, International Journal of Intercultural Relations, 25, 89–107.
Vecchi, A., & Brennan, L. (2009). A cultural perspective on innovation in international manufacturing, Research in International Business and Finance, 23: 181–192.
Wacker, J. G., & Sprague, L. G. (1998). Forecasting accuracy: comparing the relative effectiveness of practices between seven developed countries. J. Oper. Manage., 16 (2), 271–290.

Simulation Study of Wheelchair Manoeuvrability

Grzegorz Dobrzyński , Włodzimierz Choromański and Krzysztof Fiok

Faculty of Transport
Warsaw University of Technology
00-662 Warsaw, Poland

ABSTRACT

The article presents simulation study results of maintaining a pre-set trajectory by wheelchairs. The objects of the study were the models of electric and manual wheelchairs. The aim of these simulations was to analyse the influence of various drive models on the capability of maintaining a pre-set trajectory.

Keywords: Wheelchair, Manoeuvrability, Simulation Research, Multi-Body Simulation (MBS), CAD.

INTRODUCTION

The issues of simulation researches of four-wheel vehicles, wheelchairs in particular, are dealt with by a few research centres in the world. A high level of construction diversity of these wheelchairs is a result of the demand among users in relation to the functionality of this means of transport, but first of all it is directly connected with requirements concerning adaptation to various types of disabilities. A wheelchair designed for a particular person is selected in such a way that during its use the person applies possibly the most extensive parts of healthy muscles. This allows for improving mental comfort of the disabled through ensuring them a relatively independent existence, but also enables them to keep fit. Therefore, the manual wheelchairs are most often used by people suffering from paraplegia or other impairments of walking ability, however, with properly functioning upper limbs. More complex motor impairments require assistance of an able-bodied person in moving this type of wheelchair, which limits self-reliance of the disabled. In such cases it is indispensable to use a wheelchair with autonomous drive e.g. electric.

The paper presents simulation study results of maintaining a pre-set trajectory by wheelchairs. The objects of the study were the models of electric and manual wheelchairs. The aim of these simulations was to analyse the influence of various drive models on the capability of following a pre-set trajectory. For the purpose of simulation the wheelchairs (fig. 1) designed within the framework of Eco-Mobility Project realized at the Faculty of Transport of Warsaw University of Technology have been selected.

THE OBJECT OF SIMULATION RESEARCHES

The object of the researches are the prototypical and continuously developed constructions, namely, an electric wheelchair with stair-driving option and a manual wheelchair driven by levers (fig.1). The present paper is part of activities aiming at designing a universal simulation environment of various construction wheelchair movement. This environment is to play the role of a tool for virtual prototyping and experiment. In the abovementioned

activities, the research problem is the assessment of influence of regulation parameters and dynamic features of a simulation model on the capability of maintaining a pre-set trajectory.

Fig. 1 Prototypical construction of an electric wheelchair and a wheelchair driven by levers

Scientific researches within the area of simulation studies of maintaining a pre-set trajectory concern mainly motor vehicles. The issues within the range of wheelchairs are dealt with only by a few research centres in the world. The basic difference in models used in these centres is the method of pre-setting a driving direction. In the case of wheelchairs, the most frequently used manoeuvre is turning through rotational speed diversification of wheels on a given axle. Swivel casters are placed on the other axle.

For simulation model construction Matlab – Simmechanics environment has been used. The structure of this model has been presented in Fig 2. The figure presents the division into the main modules of the model. The module constitutes a combination into a group of elements commonly realizing a given function.

Each of given modules of this model is defined in accordance with nominal model assumptions and represents respectively:

- **1.** – the module including mass parameters of the frame with the seat together with rigidly connected drives, gears and levers etc.
- **2. (RR, RL, FR, FL)** – module representing the wheels: RR and RL are respectively rear right and rear left wheels, FR and FL are respectively front right and front left wheels together with a swivel axle. The rear wheels together with the frame make a rotational kinematic pair without friction. The front wheels are connected with a swivel axle as a rotational kinematic pair without friction. The swivel axle is connected with the frame as a rotational kinematic pair with friction. Axis of rotation of the swivel axle is vertical.
- **3.** – module representing decisions of the person driving the wheelchair. The task of this module is to reflect the reaction of the person who is trying to maintain a pre-set trajectory at a given speed.
- **4.** – module representing the source of the driving torque. In the case of an electric wheelchair, it represents electronic control system of the wheelchair. For the sake of these simulations, the module structure has been limited to a speed control system and driving direction by means of pre-setting the driving torque. In the case of a manual wheelchair, this module represents upper limbs as the source of the driving torque whose value is pre-set in a periodic manner.

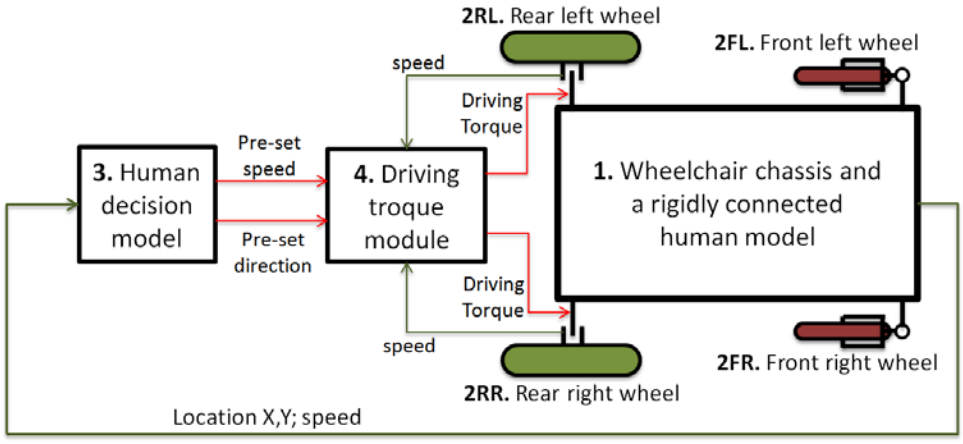


Fig. 2 Flowchart presenting the structure of simulation model prepared in Simmechanics

Module No. 1 Structure – Wheelchair Frame and a Human Being

The simulation objective was task realization of driving on a flat surface maintaining a pre-set driving direction. With such a task, the nominal model of module no.1 has been simplified to rigid bodies of the structure comprising:

- body of the frame with the seat together with rigidly connected drives, gears and levers etc,
- body representing a human being rigidly connected with body of the frame,
- bodies representing rear right and left wheels. The rear wheels together with the frame make a rotational kinematic pair without friction,
- the front wheels are connected with a swivel axle as a rotational kinematic pair without friction. The swivel axle is connected with the frame as a rotational kinematic pair with friction. Axis of rotation of the swivel axle is vertical.

This model is presented in 0, indicating the assumed coordinate systems. The origin of the coordinate system connected with the body of the wheelchair frame is situated in the middle of rear wheel axle.

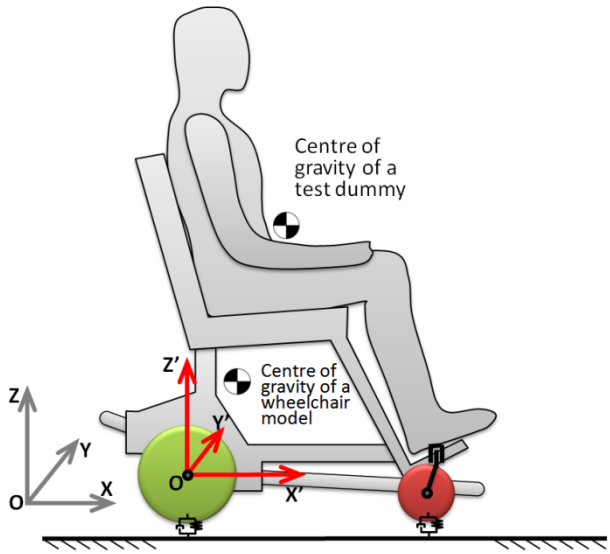


Fig. 3 Nominal model prepared for simulation of maintaining a pre-set trajectory

When modelling the frame, its mass parameters were described excluding from the whole structure the mass parameters of rotating wheels' elements. The values have been set in Table 1.

Table 1: Parameters defining a wheelchair model

Parameter Name	Indication	Parameters of electric wheelchair model	Parameters of manual wheelchair model	Unit
Mass	$m_{zr.r}$	60.5	23	kg
Location of the centre of gravity of the frame within OX'Y'Z' system	x'_{cg}	303	77	mm
	z'_{cg}	64.3	20.5	mm
	y'_{cg}	0	0	mm
Moments of inertia within OX'Y'Z' system	I_{XX}	3.4	8.2	kgm^2
	I_{yy}	7	3.6	kgm^2
	I_{zz}	7.8	8.3	kgm^2
Wheel base		514	620	mm
Rear track		640	600	mm
Front track		600	560	mm

When building a simulation model of a human body it was simplified to a rigid body. A similar simplification was applied in paper (Choromanski, 2009). In addition, the guidelines were taken into consideration that are included in the standard PN-ISO 7176-11 Wheelchairs: Test dummies (PN-ISO 7176-11:1998) and The anthropometric atlas of the Polish population - data for design (Batogowska and Słowikowski, 1994). Table 2 comprises the nominal values of values defining a test dummy model.

Table 2. Mass parameters of a test dummy model

Parameter Name	Indication	Mass parameters of test dummies of the mass of 75 kg:	Unit
Location of the centre of gravity of a test dummy within OX'Y'Z' system	x'_{cg}	217	mm
	z'_{cg}	553	mm
	y'_{cg}	0	mm
Moments of inertia within $OX_mY_mZ_m$ system (centre of gravity of test dummy)	I_{XX}	7.918	kgm^2
	I_{YY}	11.054	kgm^2
	I_{ZZ}	4.916	kgm^2
	$I_{XY}= I_{YZ}$	0	kgm^2
	I_{XZ}	3.8	kgm^2

Wheel Models

In a simulation model, two pairs of wheels have been defined that differ from each other, namely, front and rear. The MF-Tyre program has been used, which was designed by the university TU Delft (Web Page: http://home.tudelft.nl/) and TNO organization (Web Page: http://www.tno.nl). MF Delft Tyre provides a semi-empirical tyre model to be applied in various simulation environments, among others, in Matlab – Simulink – SimMechanics environment. In the description of grip, the equations of Pacejka's magic formula have been used (Pacejka Formula Magic). Tyre deformations have been described by the so called brush model [20]. The model is designed for calculations of forces affecting a tyre in a static state or in a slowly variable state of dynamic interactions the frequencies of which do not exceed 8Hz. The model is designed for steerability tests of a vehicle during cornering conducted in simulation researches of virtual pre-prototyping on non-deformed (flat) surfaces. In view of the above, the model meets the criteria of a wheelchair design. In the present paper the model in version 6.1.2 with a definition of contact option has been used: smooth road contact, circular cross section (motorcycle tyres).

Parameters have been experimentally indicated of radial stiffness and attenuation of tyre model. This stage of works was conducted within the framework of Eco-Mobility Project and has been extensively presented in paper (Dobrzynski and Barwicki, 2011). The tyres tested were pumped to the nominal pressure value stipulated by the producer, namely, 35 psi (241318 Pa).

Table 3. Indicated parameters with values.

Values	Parameter values of Wektor wheelchair tyres
Radial stiffness [N/m]	Front tyre: 32'880 Rear tyre: 74'600
Attenuation constant [kg/s]	Front tyre: 446,0 Rear tyre: 372,6

Centres of gravity of wheels are located in their geometrical centres and with reference to the system connected with this location inertia moments have been defined. Parameters defining the model of rear and front wheels have been presented in Table 4. Front wheel parameters were identical for both constructions. Parameters not mentioned here originate from TNO Delf-Tyre definitions of tyres used in motorcycles.

Table 4. Parameters defining the model of front and rear wheels

Parameter Name	Indication	Rear wheel – electric wheelchair	Rear wheel – manual wheelchair	Front wheels	Unit
Mass	m_{kT}	6.7	1.6	1	kg
Moments of inertia	$I_{XX} = I_{ZZ}$	0.048	0.058	0.002	kgm^2
	I_{yy}	0.09	0.116	0.005	kgm^2
Wheel radius	r	155	330.2	90	mm
Tyre radius	r	30	14	20	mm

Module Representing Decisions of the Wheelchair Driver

The objective of the module representing decisions of the wheelchair driver is to reflect the reaction of a person who tries to maintain a pre-set trajectory at a given speed. For description of behaviour of a person on a wheelchair, the so called anticipatory control block has been used (Augustynowicz, 2009), , (Bułka et al., 2006).

PID controller used describes the reaction of a person driving within the process of joystick angle inclination or decisions concerning diversification of manual power driving the wheelchair levers. The controller reflects psychomotor features of the driver. The segment „saturation" describes maximum value of inclination of joystick lever that can be realised by the driver. At the system output a signal is obtained which controls the difference of moments driving the wheels of the values maintaining the pre-set values accounting for typical reactions of the driver and technical limitations of the pre-setting system operation.

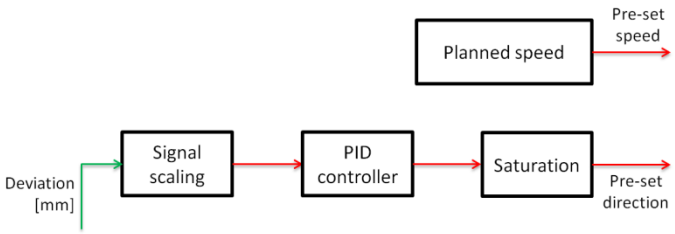


Fig.4 Diagram of module representing decisions of a wheelchair driver.

Parameters of the controller applied have been analysed further in the paper. The following values have been assumed as nominal ones: P=1; I=0.1; D=1;

The speed of model movement is pre-set through the defined characteristics and it gradually increased to 1.2 m/s. The maximum speed value (PN-EN 12184:2010), at which the wheelchair can move equals v=3.33m/s.

Module of Driving Torque Pre-Setting

The module representing the object of pre-setting the driving torque in an electric wheelchair plays the role of an electronic control system. Its operation is limited to speed and driving direction control systems (Grzesiak et al., 2011), (Kozłowski et al., 2011). The diagram of this module is presented in Fig. 5.

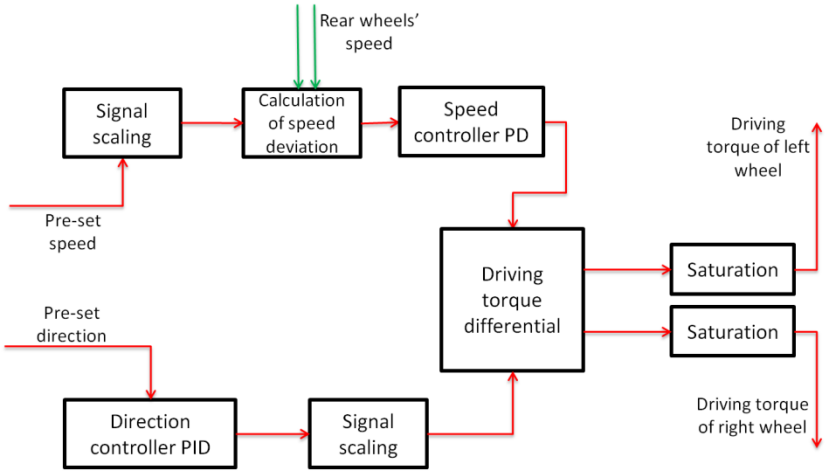


Fig. 5. Diagram of the module representing electronic control system of a wheelchair

This module operates on the basis of rotational speed of wheels feedback. This system has the structure with average speed controller (Kozłowski et al., 2011).

Parameters of controllers used have been analyzed further in the paper. The following values have been assumed as nominal ones:

- speed controller: P=1; D=-0.3 ,
- direction controller: P=1; I=0.02; D=3.

Simulation Researches

Simulation researches were conducted in Matlab-Simmechanics environment. The solver "ode15s stiff/NDF" was used for calculations. The studies conducted allowed for drawing up characteristics of the whole model response to coercion, which is the pre-set trajectory. The trajectory on which the wheelchair is moving is dependent on mass parameters of the model, contact parameters and elements of control system. What is also important are the psychomotor features of a person driving the wheelchair as well as a pre-set speed. Only the combination of the features of these elements allows for analysis of the whole model behaviour. Detailed descriptions of operation of these elements of the model are included in the paper (Dobrzynski, 2013).

Fig. 6 presents the characteristics of PID controller operation representing human properties for nominal parameters.

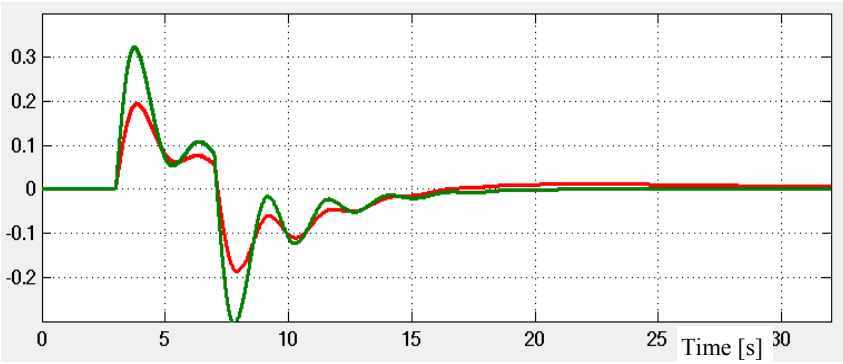


Fig. 6 PID controller operation – human being: red line - deviation, green line – registered value on controller output

One of the significant elements of researches on dynamic systems properties is the study of dependence of system model response on its parameter changes. The analysis results presented so far were prepared for nominal parameters. Further in the paper, an attempt has been made at assessment of parametric sensitivity in relation to human model features. Assuming various values of parameters of the model, a situation was simulated in which a wheelchair is driven by persons of different psychomotor predispositions.

Nominal parameters of a PID controller constituting a human model have been assumed on the basis of (Augustynowicz, 2009), (Bułka et al., 2006), (Grzesiak, 2011). Subject literature concerning the models of persons driving wheelchairs is not known to the authors, therefore parameters have been assumed describing vehicle drivers. Nominal values and parameter values of a PID controller for subsequent series of simulations have been set in Table 5.

Table 5. Parameters of a PID controller of a human model

Series	P	I	D
n - nominal	1	0.1	1
a	1	0.4	1
b	1	1	1
c	1	1.5	1
d	1	1.5	0.5
e	1	1.5	0.3

As a result of simulations conducted the influence of parameters of the human element model has been observed on the behaviour of simulation model. In order to picture the differences in operation of this controller, the graph (Fig. 7) has been prepared which shows deviation and regulation for simulation with nominal parameters and subsequent series (a-e) with altered parameter values.

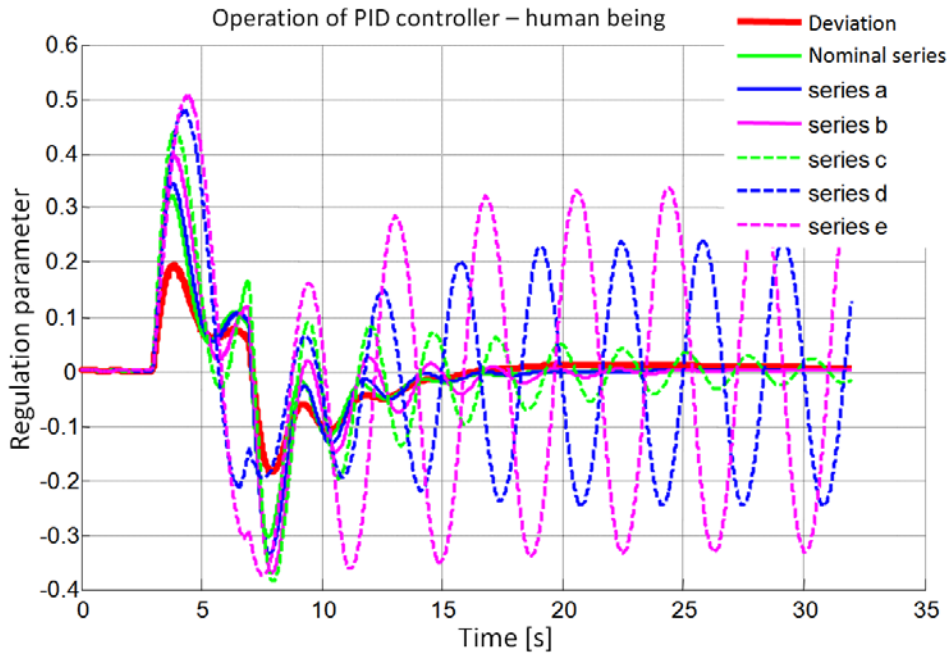


Fig .7 Operation of PID controller of a human model at various settings

The drawn up characteristics presenting trajectories of a wheelchair model movement for subsequent simulation series (0), presents parametric sensitivity of the whole system. When analysing the above graphs one can observe series for which the model after manoeuvres oscillates around a pre-set direction gradually decreasing the amplitude. These are series n, a, b, c. These cases correspond with proper operation of control system.

In series d, e maintenance of inclinations amplitude has been observed from a pre-set direction within the whole period of simulation, which means lack of stabilization of a wheelchair driving direction. At the assumed control system, the wheelchair behaves differently depending on the assumed parameters of a human model.

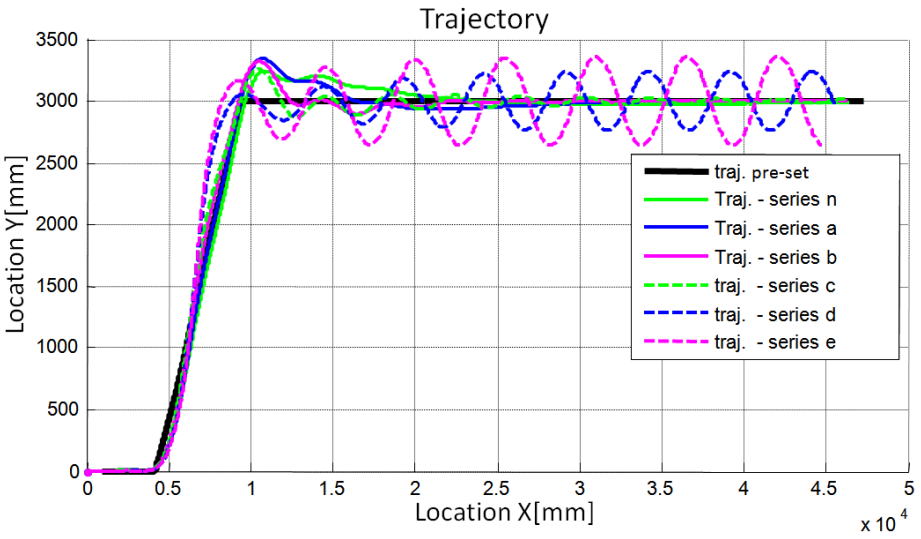


Fig. 8 Trajectories indicated by an electric wheelchair model obtained at various settings of a PID controller of a human model

At the next step of simulation the abovementioned series of parameters describing decision model of a human being, a manual wheelchair has been used for a describing model. The results of these simulations have been set in Fig.9.

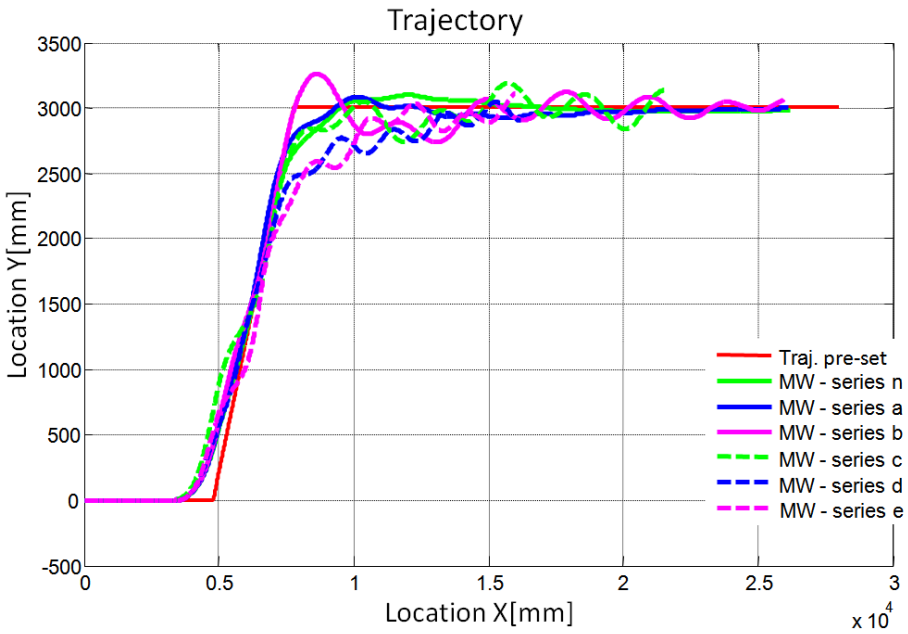


Fig. 9 Trajectories indicated by the a manual wheelchair model obtained at various settings of a PID controller of a human model

The above simulations allowed for omitting parameters, which when applied to the control system in the prototype, could pose a threat to safety of the persons driving the wheelchair during the tests. These activities also allowed for limiting the range of parameters tested on the prototypical construction discussed here.

CONCLUSIONS

The authors of the present paper have discussed simulation researches within the range of driving on a designated trajectory. The simulation results have been presented of a model designed within Matlab-Simmechanics

environment. These simulations were conducted in two stages. At the conclusion of stage I, results were obtained for the nominal model, which became the basis for comparison with simulation results for altered parameters in stage II.

The simulation results have indicated the need for constructing a control model of different parameters and also of a different structure (e.g. accounting for a learning model of a human). It seems reasonable to undertake works aiming at identification of parameters of a human model controlling a wheelchair. The results obtained should be treated as an initial analysis of the issue, because in the simulations, descriptions of models were used of a human driving a motor vehicle, in relation to whom different requirements are formulated due to higher driving speeds.

The simulation models prepared give the possibility of conducting further research work comprising the following:

- Analysis of parameter libraries of tyre models of a complex structure.
- Analysis extension of control structures of vehicles driver by two electric motors.

ACKONWLEGMENT

The described research was financed from the ECO-Mobility project WND-POIG.01.03.01-14-154/09. The project was co-financed from the European Regional Development Fund within the framework of Operational Programme Innovative Economy

REFERENCES

Augustynowicz A. (2009), "*Modelling of a car driver type*", Publishing House of Opole Univertsity of Technology, Opole**.** [URL: http://www.dbc.wroc.pl/Content/3706/SiM_232_scal_popr.pdf].

Batogowska A., Słowikowski J., (1994), "*Anthropometric atlas for the adult poulation of Poland for design purposes*", Institute of Industrial Design, Warsaw.

Bułka D., Walczak S., Wolak S., (2006), "*Preview predictor driver mode used ID computer program for the simulation of car accidents*", Reasearch Bulletin of Kielce University of Technology, Mechanics 84, p.147, Kielce.

Dobrzyński G, Barwicki M., (2011) "*Parameter identification for simulation model of a"Simple Tire" used on an electric wheelchair*", Report on work conducted within the framework of Eco-Mobility Project UD/0145/1160/2011.

Dobrzyński G., (2013), Doctoral thesis: "*Synthesis of the mechatronic wheelchair properties as a part of the transport systems for people with disabilities*", Faculty of Transport of Warsaw University of Technology, defence of doctoral thesis: 25.01.2013, [access on 10.04.2013 at: http://bcpw.bg.pw.edu.pl/publication/4062]

Grzesiak L., Choromański W., Grabarek I., Kamiński G., Ufnalski B., Sczypior J., Brach P., Choromański K., Ambrożak A., Nowak J., Kozłowski M., Tomczuk K., Barwicki M., Orłowski J., (2011) "*New Concept of an Electrical City Car and its Infrastructure"*, Polish-Chinese seminar materials, PAN, ECO-Mobility, Warsaw University of Technology, ITS, Warsaw 16 – 17 June 2011.

Kozłowski M., Tomczuk K., Szczypior J., (2011), "*Methodology of determining basic technical parameters of electric-drive car*", Electrical Review, issn 0033-2097, 87 no. 10/2011 pp. 299 – 304.

PN-EN 12184:2010 "*Electrically powered wheelchairs, scooters and their chargers - Requirements and test methods*".

PN-ISO 7176-11:1998 "*Wheelchairs – Test dummies*".

Work by multiple authors edited by Włodzimierz Choromański, (2009) "*Transport of the disabled – innovative wheelchairs for persons with motor disabilities*", Publishing House of Warsaw University of Technology, Warsaw 2009.

TNO: Netherlands Organisation for Applied Scientific Research TNO - Web Page: http://www.tno.nl .

TU Delft: Delft University of Technology – Web Page: http://home.tudelft.nl/.

Zegelaar P. W. A., (1998), "*The dynamic response of tyres to brake torque variations and road unevennesses*", Delft University of Technology, podręcznik ze strony web page: http://www.tno.nl .

Section 14

Road and Rail:

Perception and Control Issues in the Design of Advanced Driving Assistance Systems

Augmented Reality Head-Up-Display for Advanced Driver Assistance System: A Driving Simulation Study

Lynda Halit [a,b,c], Andras Kemeny [a,b], Hakim Mohellebi[a], Samir Garbaya[b], Frédéric Mérienne[b], Sylvain Michelin[c] and Valentin Albaut[d]

[a]RENAULT, Technical Center for Simulation
TCRAVA 013, 1 avenue du Golf 78288 GUYANCOURT

[b]Arts & Métiers ParisTech – Institut Image ENSAM
2 rue Thomas Dumorey 71100 Chalon sur Saône

[c]Université de la Rochelle, Laboratory of Informatics Image Interaction
Ave M. Crépeau 17042, La Rochelle, Cedex 01

[d]THEORIS, 101 Rue La Fayette, 75010 PARIS

ABSTRACT

Research and technological advance in the field of Augmented Reality (AR) is growing rapidly (Mas, 2011). One of the new application domains is the automobile industry, linked to the necessary men machine aspects of Advanced Driving Assistance Systems (ADAS). Relevant road traffic as well as useful navigation or path planning information may be displayed using partially or totally the windshield surface thanks to these emerging technologies. However, the way road traffic, signs or vehicle information is displayed impacts strongly driver's attention with increased mental workload and safety concerns. Research in perceptual and human factors assessment is needed for relevant and correct display of this information for maximal road traffic safety as well as optimal driver comfort. At Renault, research is carried out in a number of automotive AR domains: the used information type and visual grammar, visual perception for the displayed information (depth, localization), and real time mixed reality, that is matching virtual and real environment. The main goal of this presented experiment was to study whether head movement impacts AR depth perception and thus modifying displayed image quality and decreasing driver performance. For this purpose, a driving simulator experiment, carried out in the CAVE Immersive Integration Platform (P2I) Driving Simulator, we analyzed: driver head movements during different realistic situations, preferences scale in each specific situation, and thus with and without head tracked conditions. Primary data shows a strong preference for the tracked system condition and the statistical scale factor was very significant.

Keywords: Augmented Reality (AR), Head-Up-Display (HUD), Head Motion Parallax, Advanced Driver Assistance Systems (ADAS), Driving Simulation, Cognitive Psychology, Human Factors, Ergonomics

INTRODUCTION

Many clues are used in everyday life for depth perception and estimation. Some of them are binocular cues requiring combined information from both eyes, while others are monocular, requiring information from one eye alone. The research study done by Faubert in 2001, illustrates several approaches where motion parallax cue is used to better perceive objects depth. Movement parallax can be defined as the optical change of the visual field of an observer

which results from a change in his viewing position (Gibson et al., 1959). This optical motion, described by Gibson, can be divided into 2 different categories: the apparent motion due to the object motion and the movement of the observer (Gibson, 1968). When the observer is in motion and focuses on a point in the scene, objects that are in front of the focal point will appear to move in the opposite direction from the observer, and objects behind that point will appear to follow the same direction as the observer. The motion of each point on the retina depends on the relative position and relative motion of the object and the observer (Kemeny et al., 2003). Analyzing the movement parallax cue is even more complex in a driving environment; especially when we come to highlight some information in the scene using Augmented Reality (AR). In fact, the driving environment is usually rich with static and dynamic objects. Dynamic objects, for example cars and pedestrians, move at different speeds, and the vehicle moves at a variable speed. Adding to that, errors may be caused by driver's head movements.

A previous study carried out by Andersson Hultgren et al. in 2012, showed that subjects moved their head when positioned behind another vehicle trying to get a better view of the oncoming traffic. This study has been recently perused by Lisa et al. in 2013, and the results were much more promising when they adapted the experimental conditions. This later work was very inspiring and we find similar approaches in the presented work. Another case where drivers tend to tilt their head is during curve driving. Mestre et al. 2007, concluded in their study that subjects exhibit systematic head roll tilts toward the interior of curves, and the angle tilt is accentuated with the increase of road curvature.

Our research, therefore sought to analyze the impact of motion parallax, generated by the driver, while he moves his head horizontally and vertically from his initial frontal position, thus in a driving environment with a virtual Head-Up-Display (HUD) that serves as an Advanced Driving Assistance System (ADAS), during reduced visibility conditions on the road. The innovative method of Charissis et al. in 2007, revealed that focusing ability is less apparent in very low visibility situations, because the distance between the farthest visible object and the HUD plane was less pronounced. In another study, done by the same group the optimal operational distance of the HUD was investigated, participants preferred the longer focal distance. Thus because, it reduces refocusing strain on the eyes, but it was much less pronounced under very low visibility (Charissis et al., 2007). They also concluded that when the HUD was used under good visibility, it was distracting and it clutters the driving scene, and in this case all users preferred driving without the HUD.

Building on our initial work, the present document describes our success at enhancing the precision of an AR-HUD by adding the Head Tracking function so as to allow better perception of Bounding-Box alignment with the corresponding vehicle. Within this research we integrated different realistic road situations in the scenario (straight lines, curves, intersections, overtaking), and we introduced two secondary tasks (reaching the phone and hang out, opening and closing the glove box). This experiment was done by using virtual simulation, which is a cost-effective way for identifying design limitations before physical mock-ups (Charissis et al., 2007).

During all the diverse situations, we analyzed head displacement, and through a written questionnaire we have asked the subjects to evaluate the precision of the AR-HUD when Head-Motion-Parallax was taken in consideration and when it wasn't, and they therefore evaluated the difference scale and preferences. We found out a significant preference for the AR-HUD with Head Tracking conditions, and in comparison between the two systems those participants judged that there was a large difference between the two systems. For those participants, we also evaluated the degree of precision of each system they tested, grades are significantly better for the AR-HUD with Head Tracking conditions.

EXPERIMENTAL DESIGN OF AUGMENTED REALITY HUD FOR MOTION PARALLAX EVALUATION

When virtual information is superimposed to the real environment it is important to align them correctly. Accordingly, in the present problematic we are concerned on whether we should consider or not head movements during the driving task, while we display the Augmented Reality (AR) information on a Head-Up-Display (HUD) Layer. The main objective of this experiment, was to demonstrate the impact of motion parallax generated by driver's head movements, in the context of Advanced Driving Assistance System (ADAS) with an Augmented Reality (AR) Head-Up-Display (HUD). Results of this study could be exploited to investigate whether the driver

needs to be tracked in order to display AR information adapted to his perspective, when such a technology is used in the car.

Purpose of the Study

During bad weather conditions and reduced visibility, the driver has more difficulties to perceive information around his environment. The workload is more important when driving in reduced visibility conditions (McLean et Hoffmann, 1975). In this experiment, we wanted to better show one type of information for the driver using Augmented Reality (AR) technology, without cluttering his vision and saturating his mental workload. The augmented reality information is here to increase visual saliency and attract driver's attention to the proximity of the next and coming vehicle. When driving in reduced visibility conditions, we tend to leave less inter-distance with the next vehicle and the preceding vehicle serves as a landmark for trajectory control (Caro, 2008). In our case, our concern is to check if the fact that we highlight the next vehicle differently (with and without head motion parallax integration) will change our perception of the road and have an impact on the vehicle trajectory. We are assuming in this study that we know the respective positions of the different vehicles to highlight all the time, thus to display virtual information and overlay them in real-time. We are not concerned about predicting functions to estimate successive positions of objects.

A virtual Head-Up-Display is used in this study to project information directly in the driver's view. The experiment was implemented in order to compare two approaches for displaying the augmented reality information. The purpose is to compare driver's perception and performances, and quantify this difference, if there is any, with statistical study. This study emphasis on driver's behavior, trajectory control with respect to the experimental conditions in different situations. To measure the precision of the system we rely on the questionnaires completed by each participant after each session. The virtual plane was projected 3 meters in front of the driver, no color code is associated to a hazardous vehicle. In fact, a neutral color was chosen to not interfere with the driver's judgment and the distance he keeps with the next vehicle.

Driving Simulator and Virtual Environment

The methodological tool used in this experiment is one of Renault's CAVE, an Immersive Integration Platform (P2I), with 4 display surfaces (3m x 2.5m). It allows virtual stereoscopic vision from the tracked position (thanks to ART Tracking), with a resolution of 1024 x 768 per face. A static Driving Simulator was installed in the middle of the P2I CAVE to conduct the presented experiment. Virtual information were embedded in the driving scene through the driving simulation software SCANeR, it gave the impression that Augmented Reality Information was projected on a virtual plane between the vehicle cockpit and the driving scene. The next vehicle in the same lane, and the left lane (for cars coming on the opposite direction), were framed by an adaptive bounding box. Two bounding boxes in maximum are framed simultaneously.

Figure 1: P2I CAVE Driving Simulator

Participants

27 participants were invited to participate to this experiment, about half man and half women and their age ranges from 22 to 58 years old. Out of the 27 persons, 4 of them had to stop the experiment because of motion sickness. In order to avoid bias due to the learning effect the order of the two tested systems was altered each time. Subjects ignored totally the changing condition between the two tested systems. They were only ask to evaluate them relatively to their precision and indicate which system they preferred.

Calibration Phase

Calibration is required in accordance to driver's body and driving position, and this can be done manually or automatically (Charissis et al., 2007). In our study, the initial frontal position is recorded before taking off, then head displacement is measured all the way, but no feedback was present for the driver concerning his movements.

Driving Scenario and Instructions

The driving path is defined in an urban environment, where realistic traffic conditions are reproduced. The scenario was thought as it reproduces different driving perspectives. Considering that motion sickness is generally higher in static simulator, the time of the driving sessions was purposely limited. In fact the training session lasted about 3 minutes and the two driving sessions about 6 minutes each. Participants were asked to express any discomfort during the experiment. We had four cases of dizziness, and they were stopped as soon as they reported it to avoid further motion illness. The two driving sessions were identical with the same traffic conditions, same road events and instructions. This allowed us to have a maximum consistency for results analysis and comparison. However, in order to avoid bias due to the learning effect the two conditions order were counterbalanced. Participants were asked to follow the police vehicle as an itinerary indication. The police vehicle was chosen just because it was easy to recognize. It was indicated to the participants that if they happen to lose the vehicle to not rush to catch up, because the vehicle waits for them anyway. They were also asked to be careful during the sessions just like they would behave in real life.

Secondary Tasks during the Driving Sessions

The main goal during the experiment was to reproduce different driver's positions, we explicitly asked drivers to stress on their safety as they would act in their car, so they were asked to remain cautions and look at the road while doing the task. Participants were asked to do two secondary tasks at specific moments (defined through the scenario). They were told to take their time to accomplish it and especially to keep an eye on the road while doing it.
Task 1: participants were asked to perform this task when they hear the phone ring. The phone was initially placed on the copilot seat, and they were asked to reach up and grab the phone, while they keep looking at the road. Then they were supposed to hang up and hand it over to its initial position.
Task 2: participants were asked to perform this task when they hear the beep sound. Thanks to the stereovision it was possible to visualize the dashboard of the car and the glove box. The instruction for this task was to move and reach out to the glove box as if they would open it or close it.

Hypothesis

In the driving session with head tracking, the AR-HUD was updated relatively to driver's view point, so even when they moved considerably their head, the bounding-box followed the target car. This wasn't the case in the other driving session without head tracking, instead the AR-HUD was all the way displayed relatively to the measured initial frontal position. Consequently, if drivers tend to tilt their head from their initial positions they were expected to more likely observe misalignments of bounding-boxes with the corresponding cars. We expected a large difference in the perception of precision when head movements were important. For example, when drivers approach an intersection, measures of driver's head movement were expected to increase, thus modifying the perception of the none-Tracked HUD system. This led us to divide the session into 6 different intervals of interest: straight lines, curves, intersections, overtaking, and secondary tasks. In fact, this has allowed us to compare driver's perception in different situations relatively to his head movements. We didn't expect the AR-HUD to obscure other important information in both experimental conditions.

Protocol Validation Phase

This phase was achieved before convening participants to evaluate the difference between our two display prototypes. This first stage of the experiment was done by interacting with Engineers and Experts at Renault Technical Center for Simulation (CTS), to provide us with their expertise about some of the simulation aspects. The aim was to adapt the content of the experience to better answer the problematic. During this phase, it has been decided to project the virtual information 3 meters away from the driver. The projection distance is maintained constant during all the sessions. We also realized that if the driver is far behind the police vehicle he follows, the driver doesn't necessarily pay match attention to the bounding box alignment, even while doing the secondary tasks. Besides, if we happen to use this system, it will mostly find its interest when the front vehicle is somehow close to us. For those reasons, we decided to add a supplementary analysis hypothesis, which is that the difference will only be visible when the driver is not very far apart from the preceding vehicle.

Conducting the Experiment

Participants were first asked to read a presentation sheet, which contained the object of the study and the content. This was to guarantee that all subjects had the same quantity of information and not influence their judgment. They were told that the goal of the experiment was to compare two approaches for calculating the bounding boxes and that we wanted to evaluate the two prototypes relatively to their precisions. They were not informed about the changing condition between the two driving sessions. After that, the experimenter made an oral summary of the different stages of the experiment and clarified instructions. A training session was necessary: to experience the virtual environment, make sure the participant doesn't feel motion sickness, rehearse secondary tasks, and get a first idea of what will the bounding box look like so the participant is not surprised during the experimental sessions. This later consisted of 6 minutes drives, each integrated 6 important situations of interest: straight lines, curves, intersections, overtaking, and secondary tasks. After each the two driving session, participants answered a written questionnaire. A final questionnaire was field in to compare the two display prototypes and participants were asked to evaluate their preferences. The experiment lasted about 45 minutes including all stages. (Below is a summary of the various stages of the experiment)

1 - Presentation Sheet
2 - Oral Overview
3 - Training Session (about 3 minutes)
4 - Driving Session N °1 (about 6 minutes)
5 - Questionnaire « Part 1 »
6 - Driving Session N °2 (about 6 minutes)
7 - Questionnaire « Part 2 » & « Part 3 »

Drivers were asked to adapt a natural behavior, and drive as if they were in their car, obeying all road's rules. However, they weren't supposed to consider the clutch pedal but drive instead with an automatic gearbox. For each participant, measures of the frontal position were taken before starting the two driving sessions, which permitted to record head movements for both conditions (with and without head tracking).

Motion Sickness

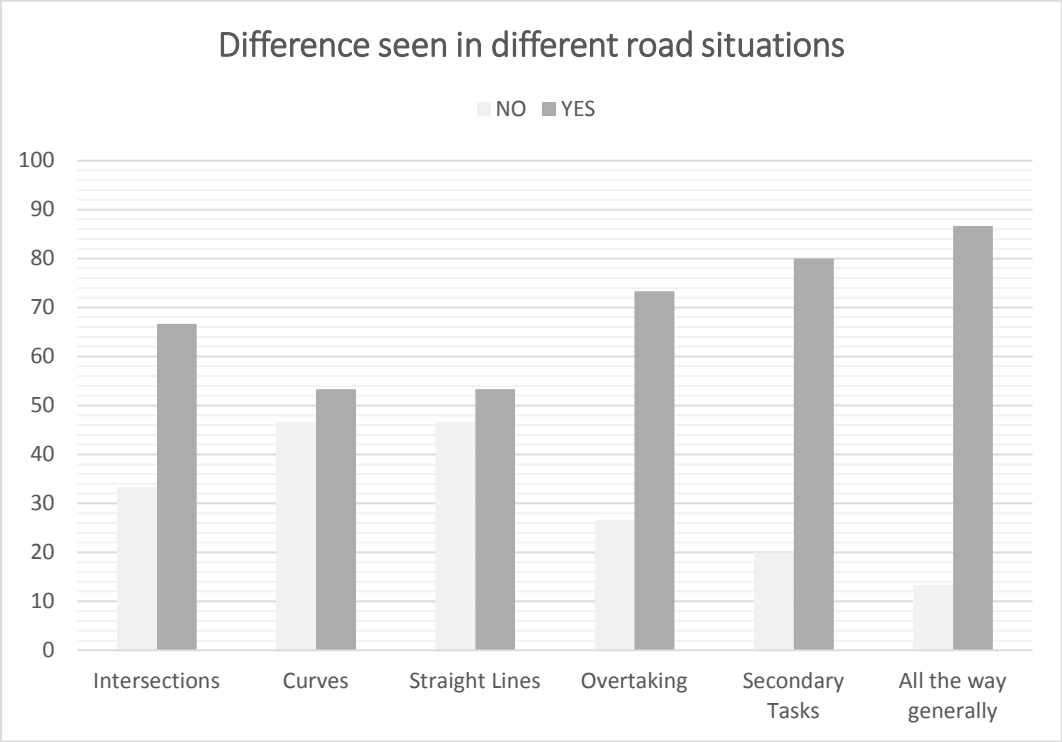
Participants were asked to inform the experimenter quickly if they felt any discomfort (nausea, headaches, heart palpitation, sweating, or any other symptom), and thus no matter at what stage of the experiment they were in. Four persons out of 27 participants had to stop the experiment because of discomfort apparition due to motion sickness.

RESULTS EVALUATION

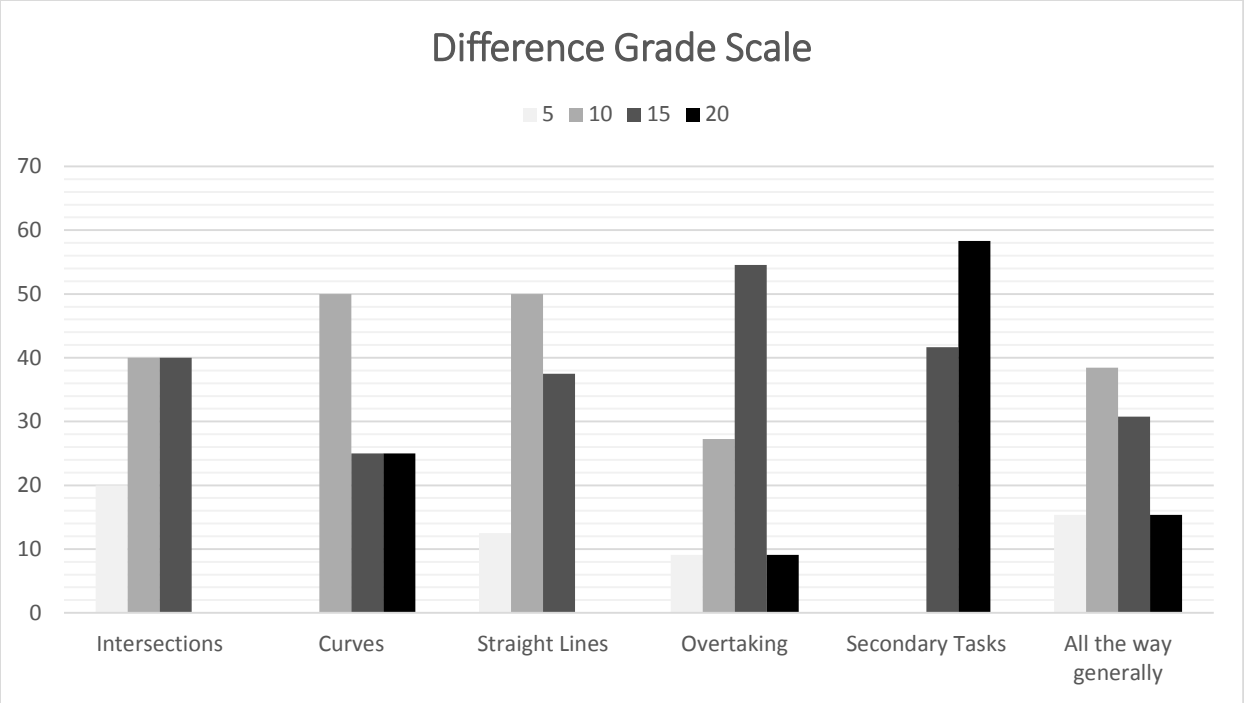
Questionnaire Analysis

Questionnaires were used to gather subjective ratings as how they perceived the AR-HUD and how they felt during the experiment. Out of 23 subjects in total, 15 persons had a preference for AR-HUD with tracking conditions (65.5%), 4 of them had a preference for the none-tracked system and the last 4 didn't have any preference. The AR-HUD with tracking conditions had 15 as median grade with a standard deviation of 6.5 (average = ~13), and the none-tracked system 10 median grade with a standard deviation of 2.5 (average= ~11).

Let's start by looking further into the 15 people that had a preference for the AR-HUD with Tracking Conditions. The first question that rises, is whether they have seen the difference all the way or only in specific situations. The histogram below shows the percentage of persons that have seen/ have not seen the difference between the two systems in different road situations.



After that, we wanted to look further into the scale factor in the case where the difference was seen between the two systems. In the graph below, we can see the number of persons with respective ranks for each specific road situation. The grades go from 5 (weak difference) to 20 (strong difference).



Statistical Test Results «Wilcoxon Ranksum»

For the 65.5% people that preferred the HUD with Tracking, our initial hypothesis was that the precision difference would be significant between the two systems and that imprecision ranks would be different in all of the following situations: intersections, curves, straight lines, overtaking manoeuvers, secondary tasks, and all the way in general . In order to measure this difference we used "Wilcoxon Statistical Test", which is a non-parametric statistical test. In our case, we used the matched data for two tasks (driving session with tracking and driving session without tracking). This test yielded the following results:

Intersections: p = 7.1901e-004, z = 3.3822, Curves: p = 0.0246, z = 2.2479, Straight Lines: p = 7.1901e-004, z = 3.3822, Overtaking: p = 0.0025, z = 3.0225, Secondary Tasks: p = 8.1464e-005, z = 3.9401, All the way in General: p = 0.0019, z = 3.1116.
As we can see *p* values are <<0.05, which means that the difference judged by the 15 participants is very significant. When taking all ranks for the same conditions (Tracking ON/ Tracking OFF) the global p value yielded by the same test is highly significant: p = 4.5667e-008 with z = 5.4674.

As mentioned above, a large majority preferred the AR-HUD with tracking conditions, and those participants evaluated the system to be more precise in all different road situations than the other system. It was surprising that some persons (8 of them) didn't manage to see the difference between the two systems. On closer examination, it was found that the majority of these drivers left a significant gap with the front vehicle, so they either had no preference or preferred the second configuration system. For the 4 persons that preferred the none-tracked system, 3 of them started the experiment with this system and only one person started with the tracked system. All of them reported a better adaptation for the system they preferred and less eyestrain. When a person leaves a large gap with the vehicle in front and the same vehicle is highlighted at 3 meters there is an adaptation problem due to eye confusion to make the focus. The order factor was significant because eyestrain was more important in those cases. In the debriefing phase, some of them admitted that they didn't pay much attention to system alignment since it was far, and all of them complained that it was hard to focus when the car was far away. Furthermore, head movement data has been explored and primary analysis show that while head movements varied significantly between test subjects, those that didn't see the difference were less likely moving their head during the driving session.

DISCUSSION

As we have seen, results were most significant when participants didn't leave too much gap with the preceding vehicle. At present, results for other participants who either showed no preference or preferred the other system remain to be fully assessed over time. We can, however, say that early observations show that most of these participants had more eyestrain problems due to the significant gap they left with the preceding vehicle and some of them didn't even pay attention to the Bounding-Box alignment with the corresponding car. Adding to that, only few participants showed a preference for the non-tracked AR-HUD, and this preference was weak. We thus find every reason to believe that this was due to the order effect since almost all of them showed a preference for the first system they started with.

It is generally accepted that using Augmented Reality to highlight some information might result in ignoring other important information. Even though it was not the purpose of our study, this idea was verified for many participants. This was noticed when the front vehicle was framed all the time, if a vehicle was present in an intersection and was not framed (in our system depends on the gap distance the driver keeps with the front vehicle, the virtual HUD is only provided on the front face of the CAVE for technical limitations), we have observed that driver's attention was most of the time diverted (considering a subject average). Even though the simulator offered a large stereoscopic view, it actually lacks other important cues, and the image resolution (**1024** x **768** par face) was not sufficient in the yield study and most likely weakens the results, as many participants reported it and others confirmed when they were asked about it at the debriefing phase. Many participants reported that the image resolution was low and it makes it hard to focus properly, and this was even more pronounced when the person wears glasses. Some participants indicated that they were confused in some situations because it was distracting to have the system activated all the time. Four participants withdraw due to motion sickness. Two persons reported that the noise was not realistic. Future work will be carried out using Renault's new CAVE like dynamic Driving Simulator CCARDS, which offers a better image resolution. It is equipped with a motion platform and it allows reproducing acceleration sensation and the steering wheel force-feedback, thus increasing the level of immersion and decreasing motion sickness.

CONCLUSIONS

Augmented Reality Technologies are the new tends in many domain. One of the new application field is the Automobile Industry, linked to the necessary men machine aspects of Advanced Driving Aid Systems (ADAS) is Augmented Reality (AR). The experiment aimed to compare the perception between two prototypes of Augmented Reality Head-Up-Display (Head-Up-Display) with and without Head Tracking.

Our research has shown that the majority of participants saw a precision difference between the two AR-HUD prototypes, and preferred the one with Head-Tracking function. Indeed, these participants noticed that this system was meaningfully more precise in all road situations, even if they didn't know the experimental conditions. As discussed above, comparing the statistical data in the two experimental conditions allows us to say that the large majority of drivers that preferred the AR-HUD with tracking condition saw a significant precision difference between the two AR-HUD prototypes. Subjects that had either no preference or preferred the other system left a significant gap with the leading vehicle. We are firmly convinced that this was due to this gap, which undoubtedly inhibited them from seeing the Bounding-Box over highlighted cars. As for the participants that preferred the non-Tracked system, we think that the order effect might have influenced their choice since the more they tried to focus the more eyestrain was important.

According to our experimental results, taking in consideration Head Motion Parallax improved the precision of the AR-HUD. Our future work will be carried out on performance parameters analysis related to trajectory control, lateral position, and steering wheel, relatively to head movements in different road situations using Renault's new CAVE like dynamic Driving Simulator CCARDS.

ACKNOWLEDGMENTS

Authors would like to thank Emmanuel Icart, Baptiste Laborie and Florent Colombet for their support and their advices to put in place this experiment. Special thanks to CTS Engineers and Experts for their suggestions and recommendations during the protocol validation phase.

REFERENCES

Caro, S. (2008). *"Perception et contrôle de la distance intervéhiculaire en condition de brouillard: Etude sur simulateur"*. (Doctoral dissertation, Université Pierre Mendès-France-Grenoble II).

Charissis, V., Naef, M., & Patera, M. (2007, January). *"Calibration requirements of an automotive HUD Interface using a Virtual Environment: Methodology and Implementation"*. In Proceedings of: International Conference in Graphics and Visualisation in Engineering,(GVE'07), Clearwater, Florida, USA.

Charissis, V., & Naef, M. (2007, June). "*Evaluation of prototype automotive head-up display interface: testing driver's focusing ability through a VR simulation"*. In Intelligent Vehicles Symposium, 2007 IEEE (pp. 560-565). IEEE.

Gibson, E. J., Gibson, J., Smith, O. W., Flock, H. (1959). *"Motional Parallax as a Determinant of Perceived Depth"*. Journal of Experimental Psychology. Vol. 58, pp. 40-51.

Gibson, J. J. (1968). *"What gives rise to the perception of motion?"*. Psychological Review, 75(4), 335.

Hultgren, J. A., Blissing, B., & Jansson, J. "Effects of motion parallax in driving simulators".

Kemeny, A., & Panerai, F. (2003). "*Evaluating perception in driving simulation experiments.* Trends in cognitive sciences, 7(1), 31-37.

Mas, A., Merienne, F., & Kemeny, A. (2011). *"Lateral control assistance and driver behavior in emergency situations"*. Advances in Transportation Studies RSS2011 Special Issue, 149-15

McLean, J.R. & Hoffman, E.R. (1975). "*Steering wheel reversals as a measure of driver performance and steering task difficulty*". Human Factors 17,248-256.

Mestre, D.R. & Lepecq, J.C. (2007). *"Assessments of head roll during simulated driving"*. In A.G. Gale (Ed.), Vision in Vehicles XI, Elsevier Science B.V. (North-Holland) (in press)

Palmqvist, L. (2013). *"DEPTH PERCEPTION IN DRIVING SIMULATORS: Observer-produced motion parallax, how it affects car drivers' position and perceived presence"*. Bachelor's thesis

Eco-Driving with Electric Cars: Effect of Continuous and On-Demand Feedback on Driving Behavior and Safety

Beloufa Sabrina[a,b,c], Vailleau Benjamin[a], Boucheix Jean-Michel[b]
Kemeny Andras[a,b] and Mérienne Frédéric[b]

[a]RENAULT
Center for Virtual Reality and Immersive Simulation, France

[b]Arts & Métiers Paristech
Institut image ENSAM, France

[c]LEAD
Université de Bourgogne, France

ABSTRACT

Analysis of driver behavior has been the purpose of many studies for several years. A new aspect is the introduction of electric vehicles on the car market and the fact drivers are more concerned with their consumptions in EV (Electric Vehicle) due to the lack of the appropriate infrastructures. However, only a few studies focusing on ecological behavior and consumption have been carried out. The primary goal of the present study was to assess empirically the impact of feedback on EV driving behavior involving autonomy management performance. We hypothesized that we could help the driver to manage his autonomy more efficiently if we'd give him feedback on his driving behavior and battery consumption. Our second hypothesis is that participants would allocate more attention on feedback information when they reach a critical battery autonomy situation, so the subjects had to drive a course with 100% and 15% autonomy left. Preliminary results indicate better driver performances and enhanced safety with Econometer gauge feedback, probably due to the increased driver control and better applied driver attention.

Keywords: Electric Vehicle, Human-Machine Interface, Range Anxiety

INTRODUCTION

More and more companies are concerned with eco-driving, in order to reduce fuel consumption and CO2 emissions (Anable & al, 2007), and reduce accidents risks. The interest in this field of driving studies and analyses of driving behavior is also linked to the expansion of driving assistance systems in modern vehicles. In the meantime, improvement in digital and immersive technologies has enabled more precise experimental study conditions for behavior analysis with the help of simulation technology for well controlled immersive driving situations.

In the frame of EV research field, the energy issue comes a different way due to the lack of charging infrastructures. People are more concerned with their consumptions. They fear that a vehicle has insufficient range to reach its destination and would thus strand the vehicle's occupants. This phenomenon is called "Range Anxiety" (Eberle & al, 2010; Backstrom & al, 2009; Schott & al, 2009; Rahim & al, 2010).

Graham-Rowe & al. (2011) suggest that the EV possibilities are in contradiction with drivers' expectations about their cars specified by Mann and Abraham (2006) like the need for autonomy, the positive experience produced by the feeling of being independent, comfort, physical effort to make a course, etc. Yet Graham-Rowe & al. (2011) and Wellings & al. (2011) also indicate that the anxiety is reduced when the driver is accustomed to EV. Those results are coherent with Taylor (2009) which shows that anxiety decreases after a few weeks of EV usage.

Several studies suggest that the autonomy available in the majority of EV actually matches the average daily courses majority of drivers do with traditional ICE (Internal Combustion Engine) cars (less than 50km per day), nevertheless it does not satisfy autonomy drivers are asking for (Sonnenschien, 2010). Thus, Sonnenschein (2010) confirms that the "Range Anxiety" problem is not only autonomy but how the driver deals with autonomy and the anxiety that results in it.

The primary goal of the present study was to assess empirically the impact of feedback on EV driving behavior involving autonomy management performance. Two specific interactive guidance systems based on ecologic and safety criteria were developed at this occasion. We hypothesized that we could help the driver to manage his autonomy more efficiently if we'd give him feedback on his driving behavior and battery consumption.

This feedback display consisted in a guidance system integrated in a motion based driving simulator (Kemeny 2003) that give to the driver information about his consumption depending on his driving behavior. Participants were randomly assigned to four different groups depending on the conditions. The first group participants were only told to drive normally. The second group had a continuous feedback display that gave them real-time information about their consumption in the form of a classical Econometer gauge. But this type of feedback, despite of his wide spread democratization, may add cognitive load during driving task since it's providing permanent information. So with the third group, we built a new type of feedback that provides more control to the driver. This asynchronous Econometer gauge appeared only when the drivers wanted by pushing on a button placed on the steering wheel and gave the driver his average consumption behavior. In the last group, participant had the two types of feedback available on their dashboard.

Our second hypothesis is that participants would allocate more attention on feedback information when they reach a critical battery autonomy situation, so the subjects had to drive a course with 100% and 15% autonomy left.

Another objective of this experiment is to enable driver to improve their eco-performance, in the respect of environmental and safety rules, with a string practical experience, in order to reinforce position and utility of driving simulation in practical learning

ECO-DRIVING SIMULATION

Driving simulators are mainly used for driving learning, experimental research on human behavior, and automotive engineering (Parodi-Keravec & al 2011; Azzi & al 2001; Kemeny, 2003). With evolutions and improvements of simulation techniques, as well as the increasing popularity, simulators have gradually become testing tools that could be provided turnkey. These more accessible solutions enabled to focus on contents (teaching, scripting, realism, interactivity). Fidelity of simulation provided by these tools is useful for application such as eco-driving learning. This experimental study required specific content in order to enable validation of simulator as a pertinent learning tool for eco-driving. New interactive guidance techniques based on ecologic and safety criteria were set up.

The CARDS simulator is a driving simulator developed by the Center for Virtual Reality and Immersive Simulation at RENAULT. The simulator is composed of a cabin including an adjustable seat, a steering wheel, gas and brake pedals. The simulator had been physically modified to fit like an EV, the clutch and gearstick had been dismounted for the experiment. The participant was immersed in the simulated environment by means of three screens, audio speakers located in front and behind the seat, and vibrators embedded in the seat. While driving, the scenarios are projected on the screens in front of the cabin, and in the wing mirrors and in the rear-view mirror (which is projected on the screen).

This dynamic simulator is built on 6 hydraulic cylinders, and reproduces the acceleration sensations, the swaying motion, switchback motion, and pitching motion. A microphone and 4 cameras are available in the cabin. The driver is in permanent audio and visual contact with the supervision room. The simulator is used for the development, the tune-up of the systems of interface man-machine, adaptive cruise control or characterization of driver's behavior in

accident situations. For further studies on driver behavior, the simulator is an ideal tool to create reproducible and safe experiments.

Figure 1: CARDS prototype at RENAULT

- **Software architecture**

Simulation was based on SCANeR™ Studio v1.2, developed by Renault and Oktal and SCANeR™. Software evolutions were made so that utilization of SCANeR™ tool could remain generic. Yet, some developments had been set up concerning the integration of new simulator cockpit specific acquisition module to Studio v1.2 and integration of new information from cockpit commands, and the set-up of the VEN interface (Virtual Extended Network), enabling communication and data exchange between distinct soft environments (vehicle data, driver evaluation variables, etc…).

Figure 2: CARDS virtual cockpit

- **Eye-tracking material**

The eye-tracking is a faceLAB apparatus and the eye-tracking data and the driving performance data are recorded at a frequency of 60 Hz.

- **Scenarios and event design**

The scenarios were developed in order to put the driver in various situations, including urban and rural context of traffic. The rural phase had the driver to drive fast and put the driver in situations where he was able to anticipate some expected events (e.g. a “stop” signal) or unexpected events (e.g. a road work zone). The urban phase had the driver to face low speed driving. Some of the events were announced by a road sign (e.g. modification of the speed limit), but some other hazardous events could not be anticipated (e.g. pedestrians). Drivers were guided through the scenarios by arrows seated on the floor.

In order to prevent from bias induced by difference in courses, or even habituation, a counterbalance system was used for initial and final courses between two half of subjects group. Courses were also equivalent in terms of curves, straight lines, specific events, etc.

Driving scenarios were designed, in order to evaluate different guidance techniques proposed. Each scenario lasted about 10 minutes of driving time.

- **Eco-driving analysis**

A driving style measurement and analysis method was implemented and improved in order to adapt former measurement system and to validate this tool for simulated and real driving. A technical demonstrator was built in order to test interactions between SCANeR™ and the feedback display tool.

This tool analyses driving style with data based on driver actions on vehicle. These data acquisition system is implemented via a CAN bus interface and takes into account all vehicle and engine data (pedals position, acceleration, speed, distance, etc…). While comparing several identical driving courses data it enables to determine global driving behavior (cumulated braking time, accelerations duration) and autonomy consumption.

- **Interactive guidance technique**

Innovative interactive guidance techniques for eco-driving on EV were built and implemented. A preliminary study was carried out, focusing on the guidance tools optimization, and the courses the subjects would have to drive on. We also wanted to be sure that the guidance tools did not disturb the driving behavior and activity.
These guidance metaphors was connected to vehicle data and acquisition module, they had to be interactive, easy understandable and not too heavy in terms of cognitive load. It consisted in visual information released on the dashboard screen, relative to driver's consumption and battery state of charge situation.
For all the courses, dashboard provided speedometer and gearbox indicator information.

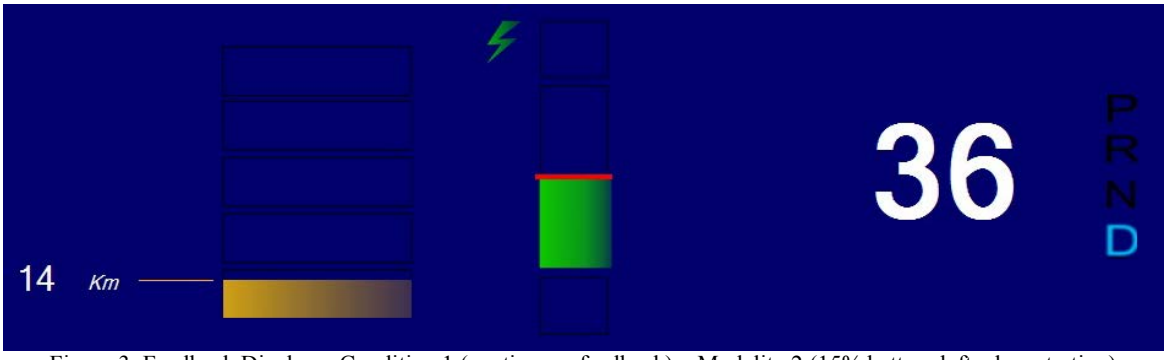


Figure 3: Feedback Display – Condition 1 (continuous feedback) – Modality 2 (15% battery left when starting)

ECO-DRIVING SIMULATOR EXPERIMENT

- **Participants**

55 participants were recruited at Technocentre Renault. All the participants who worked in a domain that could potentially be in touch with the field of EV and HMI had been excluded. Participants were specifically asked to have their driving license since at least a year and wear no glasses because of the eye-tracker material. Due to few technical problems and simulator sickness, some subjects had to stop the experiment and their record had not been used for the driving analysis. 44 participants were left for the study. The sample were composed of 29 males and 15 females, the average age was 37,4 y.o.

- **Procedure**

Subjects were randomly assigned to four different groups. All the participants started with a habituation phase while they learned about the simulator functioning and the safety instructions. The participant also had to answer a small survey about their driving experience, and knowledge about EV. Then the experiment started. The control group had

to drive with no particular condition. The two following groups had a guidance system diffusing driver's consumption information. They were told that they could possibly manage their autonomy consumption with the help of the Econometer gauge. They either had an Econometer gauge which gave continuous consumption's information, or they had to display this gauge by pushing a specific button on the screen whenever they wanted to. The Econometer gauges were filled with green when the driver's behavior lead to low energy consumption and orange when his behavior lead to high energy consumption. A specific section of the Econometer gauge was filled in blue when the driver's behavior leads to energy recovery (e.g. while braking). The last group had both continuous and on demand feedback available on the dashboard.

Medical matters wouldn't allow us to let the participants drive for hours on the simulator so we had to simulate the battery discharging performance. Each participant had to drive on 3 different battery state of charge context. First, a neutral context with no specific information about battery state of charge. They started the second course with 100% autonomy left. During the last course they had to experiment a critical situation where they started to drive with only 15% autonomy left.

RESULTS

First, it is important to point out that several indicators were analyzed to evaluate driver behavior in terms of road safety: time leaving the road, eyes position, violation of highway rules, steering wheel position, and ability to maintain the car in the middle of the lane. None of these indicators, showed a deficit in safety between the courses.

There is no significant difference in driving time to achieve courses between control group and other conditions group with different guidance systems displayed on the dashboard. Regardless of the groups the participants were in, they took approximately the same time to drive all the scenarios. There is yet a significant effect of the modality on driving time. Participants used to have longer driving time when they had 100% energy left, F (2,97)=3,397, p<.05. Conversely, when they have to start the course with only 15% energy left, their driving time match is more close to the control condition driving time. This could suggest that people are more likely to apply the Econometer gauge when they feel comfortable with their energy range. This is confirmed by the fact participants tend to spend more time looking at the dashboard when they have 100% energy left.

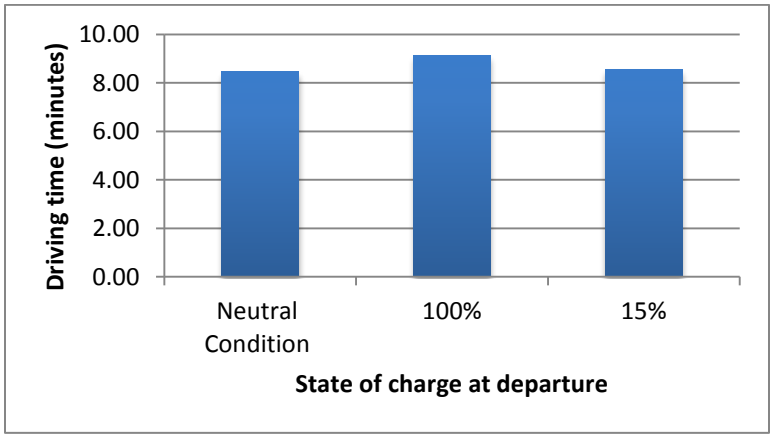


Table 1: Average driving time for each modality F (2,97)=3,397, p<.05

In the meantime, the analysis performed on the energy consumption with an ANOVA revealed a reliable decrease of energy consumption when the subjects are on the critical 15% energy left driving condition comparing to the two other starting state of charge condition, F(2,97)=6,794, p<.01. This result shows that it is possible to modify one's own driving behavior without losing any time in completing one's drive.

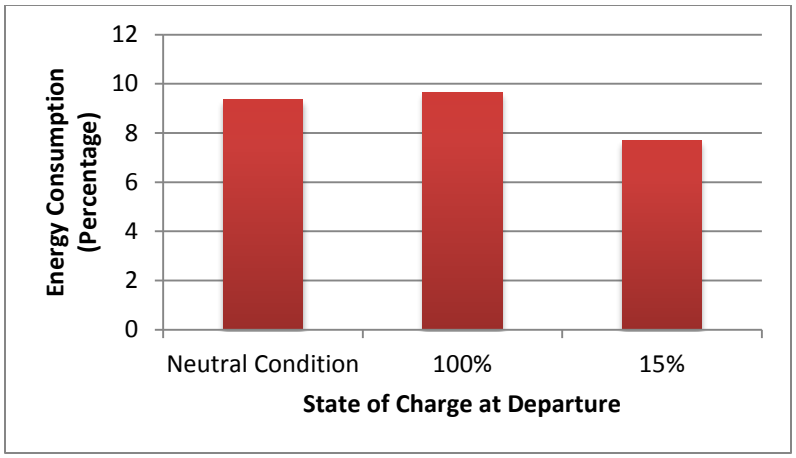


Table 2: Average energy consumption depending on the State of charge at Departure, F(2,97)=6,794, p<.01

This result could be explained by the measures on driving acceleration and braking behavior. The pressure on brake pedal, F(2,75)=3,584, p<.05, and acceleration pedal, F(2,75)=8,076, p<.01, is significantly more smooth on the critical context. The fact the subjects change their behavior toward brake pedal also enables energy recovery, which could explain the better consumption scores.

If we focus on different guidance system available for participants, we can see that drivers had a tendency to have a lower score in energy consumption with on-demand feedback display. The amount of pushing button in the conditions where it was available (condition 2, only the on-demand feedback display available, and condition 3, continuous and on-demand feedback display available) were significantly more important in the condition on-demand feedback were available alone, F(1,33)=8,956, p<.01, while the consumption were higher in the last condition with two guidance system simultaneously. Also, acceleration and braking behavior were also higher in terms of pedal pressure in this condition. Those results may suggest that people require the help of this feedback display more frequently, but the asynchronous aspect of this type of feedback makes it harder to understand and follow as an instruction, and participants in the last condition finally base their driving behavior on the continuous feedback.

CONCLUSION

Results first confirm that feedback display helps the driver to manage his autonomy and energy consumption comparing to a control condition. Participants feel more comfortable with continuous feedback when they feel like they have the time to achieve a driving course without having autonomy issues, in this situation they spend more time looking at the dashboard, their driving time increase, we can think they are more disposed to try to apply eco-driving in this context. On-demand feedback was conceived in order to enable the driver allocate more attention to the driving behavior, as a result, the asynchronous isn't helping and they couldn't change their acceleration/braking behavior on pedals pressure, and this scores are close to the control condition.

Using different type eye movement analysis (e.g. fixation length instead of fixation occurrences) the results might have been different. Another result is that no significant difference was noticed in safety measures, and driving time remains close to control condition. Further work may be carried out to gather additional to further analyse electric vehicle driver behaviour using man-machine on consumption autonomy.

REFERENCES

Anable J., Bristow A. L. "Transport and Climate Change", *Supporting document to Climate Change Working Group of the Commission for Integrated Transport*, Sep. 2007.

Azzi S, Reymond G, Kemeny A, Mérienne F. "Eco-Driving Performance Assessment With in-Car Visual and Haptic Feedback Assistance". *In Trends in driving simulation design and experiments, Proceedings of the Driving Simulation Conference* Europe 2010, pp. 181-190.

Backstrom, Michael (May 22, 2009). "Comments of Southern California Edison Company on the California Public Utilities Commission Staff's White Paper, Light-Duty V". p. 4. Retrieved 26 June 2010.
Eberle, Ulrich; von Helmolt, Rittmar (2010-05-14). "Sustainable transportation based on electric vehicle concepts: a brief overview". Royal Society of Chemistry. Retrieved 2010-06-08.
Graham-Rowe, E. & al (2012). Mainstream consumers driving plug-in battery-electric and plug-in hybrid electric cars: A qualitative analysis of responses and evaluations. Transportation Research Part A: Policy and Practice, 46(1), 140-153.
Kemeny A., Panerai F. « Evaluating perception in driver simulator experiments. TRENDS in Cognitive Sciences Vol.7 No.1 January 2003.
Kruger, J., D Dunning - Journal of personality and social psychology, 1999.
Mann E, C. Abraham (2006). The role of affect in UK commuters' travel mode choices: an interpretative phenomenological analysis. British Journal of Psychology, 97, pp. 155–176.
Ona Ebgue, Suzanna Long 2010 Barriers to widespread adoption of electric vehicles: an analysis of consumer attitudes and perceptions.
Parodi-Keravec A, Azzi S, Filliard N, Vailleau B, Icart E, Kemeny A, Mérienne F, Martinez JL. « Eco-driving performance assessment with in-car visual and haptic feedback assistance". *Proceedings of Virtual Reality International Conference (VRIC 2011),* 6-8 April 2011, Laval, France. Driving Simulation Conference 2012 Enhanced game mode for Eco-driving simulator Paper Number **-9-** DSC'12.
Rahim, Saqib (May 7, 2010). "Will Lithium-Air Battery Rescue Electric Car Drivers From 'Range Anxiety'?". *The New York Times*. Retrieved 26 June 2010.
Schott, Ben (January 15, 2009). "Range Anxiety". *The New York Times*. Retrieved 26 June 2010.
Sonnenschein, J. (2010) Preparing for the Roll Out of Electric Vehicles: Exploring how cities can become plug in ready. IIIEE Theses 2010:XX. Lund University, Sweden.
Taylor, D. (2009). The Differences and Similaritities between Plug in Hybrid EVs and Battery EVs. 24th International Batteries, Hybrid and Fuel Cell Electric Vehicle Symposium (EVS24), Stavanger.
Wellings, T & al (2011) Human machine interfaces in low carbon vehicles: Market trends and user issues. Dokument No. HMI 2.1. Low carbon vehicle technology project: workstream 13.

Effects of Interface Complexity of Head-up Display on Drivers' Driving Performance and Self-Perception

Shuling Li, Dunxing Wang and Wei Zhang

Department of Industrial Engineering
Tsinghua University
Beijing, 100084, China

ABSTRACT

This paper focused on the effects of interface complexity of HUDs on drivers' driving performance and self-perception. During the experiment, participants need to perform four tasks, namely car-following, lane changing, speed control, and response to an urgent event, all with either of five randomly chosen HUD interfaces with different levels of complexity shown to the drivers. The questionnaires to investigate drivers' driving easiness, satisfaction, efficiency, mental workload and task complexity were collected immediately after finishing the tasks. Results showed that both perceived effectiveness and satisfaction have reverse "U" shape relationship with HUD interface complexity. This paper made a contribution to interface design of HUDs on automobiles as well as made suggestions on information complexity design, which can improve drivers' driving performance and safety.

Keywords: Head-up display, Interface Complexity, Driving Performance, Driving Self-Perception

INTRODUCTION

In recent years, smart devices of automobiles have been developed rapidly. Head-up display (HUD) is one of them. HUD can present data without requiring users to look away from their usual viewpoints. It was originally developed to help pilots operate aircrafts, especially military aircrafts, more safely. The first use of HUD technology was in 1988, and it was introduced to the automobile industry by General Motor (GM) and Pontiac models (Smith and Fu, 2011). Nowadays, some automobiles have been equipped with an HUD in front of their windshields. HUD has become an essential device in most luxury vehicles, but it is not widely used in daily life. As for the design of HUD, there can be many different kinds of information on the display, such as speed, warnings, gas level, gear position, radio setting, temperature, position navigation and so on, which can help the driver operate safely and efficiently.

HUD on Driving Performance

In prior studies, some found that HUDs could make some positive effects on driving performance. The use of HUDs can improve driving control and driving safety (Charissis, 2006). Drivers' response time to an urgent event is faster with an HUD than that with a HDD (head-down display), and speed control is also more consistent with an HUD (Liu and Wen, 2004). In addition, mental workload can also be decreased, and it is good for first time users (Liu, 2003). Drivers feel safe and satisfied when they are driving with a HUD (Tonnis and Lange et al., 2007). With an HUD, drivers may have a clear viewpoint of the driving task, making driving more efficient. At the same time, it can

reduce complexity of driving. In 2003, Yung-Ching Liu investigated difference in driving performance between drivers' attention on HUDs under low and high road conditions (Liu, 2003). However, there are also some bad effects on using HUDs, such as focal distance may affect drivers' accommodation and perception of actual objects while driving. Meanwhile, HUD images may clutter or block drivers' view and affect visual attention (Tufano, 1997). As for the measurement of driving performance, Yung-Ching Liu used mental workload and speed control (Liu and Wen, 2004). Some researchers also use error rates and response time in their study.

Since HUDs have effects on driving, interface design of HUDs is very important. Watanabe and Yoo et al. studied the effect of HUD warning location on drivers' response and performance enhancement (Watanabe and Yoo et al., 1999). Other researchers developed a full-windshield HUD interface to improve drivers' spatial awareness and response time under low visibility conditions (Charissis and Papanastasiou, 2010). Shana Smith used Kansei engineering to produce a new HUD which is more acceptable based upon consumers' feelings and demands. Since driving is a heavy visual task, drivers need a visual interface to help them focus their attention on the road ahead. HUDs have been used to reduce drivers' visual and cognitive workload, without any physical interaction (Smith and Fu, 2011). The results of their paper can also be used to customize an HUD presentation image, which caters to the drivers' feelings and emotions. Tonnis et al. built a virtual bar and projected it in front of a car using an HUD, to assist drivers in longitudinal and lateral control (Tonnis and Lange et al., 2007).

Information Complexity on Driving Performance

Previous researches have shown that decisions are affected by uncertainty ambiguity and limited human capacity. It is hardly possible for anyone to be observant of everything at the same time. It is important to take into account what is known about limitations of human information processing (Svensson and Angelborg-Thanderez et al., 1997). Moray cited research where attempts were made to decide what is the optimal number of elements in decision making in uncertain situations. Normally it is not an advantage to have more than seven elements. The same conclusions are valid for judgment and estimation. The numbers even appear also in applied research on pilot performance (Moray, 1986). Miller found that humans cannot discriminate between more than half a dozen one-dimensional entities; nor can they handle more objects in their short term memory or control more content in their attention (Miller, 1956).

Eustase found that when more than seven simultaneous threats were presented to the pilot, his outcome, in terms of number of omissions and errors, radically deteriorated (Easterbrook, 1959). Way et al. stressed the importance of reducing display clutter and making critical information more evident (Way and Martin et al., 1987). Other researchers have done some analysis on the cognitive information complexity. Kushwaha found that all the nine Weyuker properties have been satisfied by cognitive information complexity measure and hence established cognitive information complexity measure based on information contained in the software as a robust and well-structured one (Kushwaha and Misra, 2006).

Preceding literatures showed that there were lots of researches on HUDs, especially for the information presentation and the effect on drivers. Driving performance and self-perception are way to measure whether the design of HUD is user-centered or effective for driving. However, the interface complexity design is widely used in aviation, which is not a consideration on the study of HUDs on vehicles before. There are many surveys on information complexity, but few on HUD designs. Thus this paper aims to study the effects of interface complexity of HUD on driving performance and self-perception, so that it can provide useful advices on HUD designs, which will improve the drivers' driving performance.

METHODOLOGY

In order to investigate the influence of interface complexity on driving performance and self-perfection and gather sufficient information to analyze, a questionnaire, which investigated the entire different hypotheses, statements and tendencies, is essential. Meanwhile, it is necessary to classify the questionnaire into three parts with each part researching on demographics, easiness, interface complexity and satisfaction on account of keeping the questionnaire's validity. Thus, the most frequent 7-point Likert-scale questionnaire from strongly disagree to strongly agree was used.

Firstly, in order to keep the inter consistency of different parts of this questionnaire, we use three independent questions with respect to each dependent variable. And all of these questions were either collected from the widely used CSUQ questionnaire (Computer System Usability Questionnaire) and SUS questionnaire (the System Usability

Scale) or revised according to our experiment request (Lewis, J. R.， 1995; Bangor, A. et al, 2008). Secondly, each participant would immediately receive one paper of questionnaire to evaluate the current HUD interface in case that they may forget their real feeling after the experiment. Thirdly, all the answers given by the participants would be checked whether this data satisfied the normality assumption. And for those data, which don't satisfy, a throwing-out or changing will be done to them. Then, we would check if there is good internal consistency among every three questions and whether they all examine the same dependent variable by checking the value of Chronch Alpha of the answer. Fourthly, one-way ANOVA with LSD ($\alpha = 0.01$) would be carried out to find out whether there is significant difference between the interface complexity and driving performance and self-perception. Finally, we will discuss the result of this study in detail and compare our findings with literatures. At the same time, we also explain every detail finding and limitation and point out the future research direction.

Participants

Twenty participants (all males, mean age=23.7, SD=3.81), who are students of a university, took part in the experiments. All of them have a driving license and driving experience. They also have experiences in a virtual environment and used a driving simulator before, such as video game car driving simulator.

Tasks

The experiment contains four tasks, which are car following, lane changing, speed control, and response to an urgent event. The participant will drive in a virtual city road. Firstly the participant will have a one-minute straight drive following a car, and then the participant will change his lane to left and then right. Secondly the participant will turn left on a cross, then go straight and stop before a sidewalk. During the experiment, the participant must follow a car once, change his lane five times, turn left once and stop before a sidewalk. When the participant finishes this task, he will fill out the questionnaire. The driver needs to repeat it five times. Every time the interface on the HUD will be changed, and the five images will be shown randomly to the driver.

Apparatus

This experiment was taken in a full cockpit simulator, which consists of a modified car and a projection system. The HUD we use is made of a projector and a film. We put this film on the windshield of the car, and let the projector projects the HUD images when the driver is driving. In this way, we can show the interface that designed to the participant (see Fig.1). The software we use is Vega Prime.

Figure 1. State Key Laboratory of Automobile Safety and Energy

As for the HUD interface, after we found out the dependent and independent variables, we used these items to design five HUD image samples. It represents five levels of the interface complexity of HUD from No.1 to No.5. The information on the interface is increasing from speed to road condition and navigation. The images are shown as below (see Fig.2).

Fig.2 Interfaces of five HUD designs

Procedure

Participants will spend about 40 minutes to finish the experiment, including an introduction, a driving training, a series of driving tasks, and filling out the questionnaires. Firstly, we will give a detailed introduction to the participant about the HUD and the participant will fill out the personal information questionnaire. Secondly, the participant will have driving simulation training. Thirdly, the participant will finish the driving tasks mentioned above and fill out the questionnaire. After all of this the participant will get his/her remuneration.

RESEARCH FRAMEWORK

The main objective of this study was to measure the influence of HUD interface complexity on drivers' driving performance and self-perception, which can be measured in terms of perceived complexity, subjective effectiveness, easiness, satisfaction, mental workload, and driving task complexity. Firstly, measuring perceived complexity of these five HUD interfaces can not only provide us with drivers' real feeling of HUD interface complexity, but it is also a good form of checking the consistency between designed interface complexity and perceived interface complexity (Liu, Y., & Wen, M., 2004). Secondly, measuring subjective effectiveness, easiness, and satisfaction is of great importance since these measurements demonstrate drivers' feelings and thoughts in the whole driving process (Koss, B., & Sieber, A., 2011). In addition, drivers are the ultimate users and purchasers of HUDs and hence, drivers' satisfaction and preference are the most vital standards of checking the effectiveness of HUD interfaces (Kiefer, R. J., 1998). Thirdly, it is also of great significance to measure drivers' mental workload and driving task complexity. For one thing, drivers' mental workload is a measurement which reflects drivers' mental requirement and occupation and will play an important role in influencing drivers' driving performance in the long run. For another, measuring driving task complexity can offer drivers' subjective efforts in the whole driving process and undoubtedly is a vital index (Wolffsohn, J. S. et al, 1998).

According to the existing researches and a short interview with experienced drivers, a total of 5 hypothesizes are proposed as follows.

Hypothesis 1a: Both the most complex and the simplest HUD interface will not bring drivers with the best perceived effectiveness, thus, the effectiveness curve should be reverse "U" shape (Chu, K. et al, 2008).

Hypothesis 1b: Similar to Hypothesis 1a, there should be a reverse "U" shape relationship between interface complexity and perceived easiness.

Hypothesis 1c: Similar to Hypothesis 1a and Hypothesis 1b, there should be a reverse "U" shape relationship between interface complexity and satisfaction.

Hypothesis 2a: With the increasing of HUD interface complexity, the amount of information that HUD provides is also increasing; thus, there should be a positive relationship between HUD interface complexity and drivers' mental workload (Frey, T. W., 2001).

Hypothesis 2b: Similar to hypothesis above, there should be a positive relationship between HUD interface complexity and driving task complexity.

RESULTS

Perceived complexity

One-way ANOVA with LSD (α=0.01) was conducted among these HUD interfaces. Results show that there is a significant difference between perceived complexity and HUD interface complexity: F (4, 95) =83.740, p<0.001.

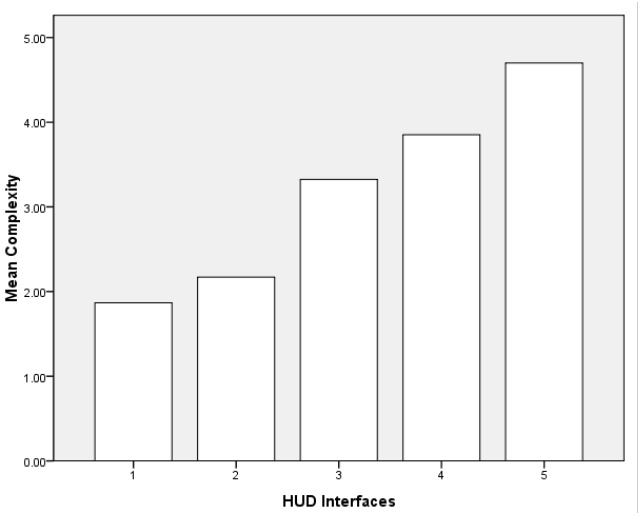


Fig.3 Comparison of mean perceived complexity between different interfaces

As for significance among these interfaces, except for the HUD interface No.1 and No.2, the left ones all have significance between each other (p<0.01). We could also see an increasing trend of interface complexity (see Fig.3).

Subjective Effectiveness

Similar method was also conducted to examine the relationship between subjective driving effectiveness and HUD interfaces, and there is a significant difference between them, F (4, 95) =3.983, p=0.005 (see Fig.4).

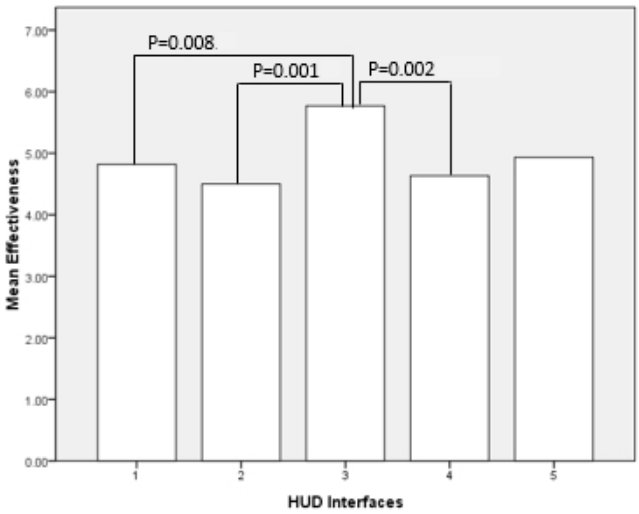


Fig.4 Comparison of mean effectiveness between different interfaces

There is a nonlinear relationship between subjective driving effectiveness and HUD interface complexity. It is obvious to see that drivers' subjective effectiveness increases when the interface complexity increases from interface No.1 to No.3. However, there is a significant decreasing trend from interface No.3 to No.4. To step further, three interfaces, namely No.1, No.2 and No.4, have significant differences with interface No.3 (p=0.008, p=0.001 and p=0.002 respectively).

Easiness

One-way ANOVA was conducted between easiness and HUD interfaces, and results indicate that there is a significant difference between them: F (4, 95) =4.643, p=0.002. Drivers' feeling of easiness decreased when the HUD interface complexity increased (see Fig.4), which was beyond our expectation. Three interfaces, namely No.1, No.2 and No.3, have significant differences with interface No.5 (p<0.01, p=0.001 and p=0.004 respectively). As for the interface No.4 and interface No.5, significance was found at the 0.05 level (p=0.011).

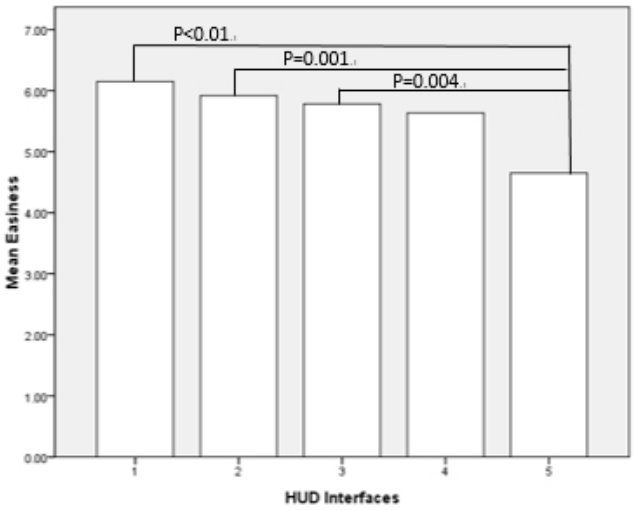


Fig.5 Comparison of mean easiness between different interfaces

Satisfaction

One-way ANOVA was conducted between satisfaction and HUD interfaces, and results indicate that there is a significant difference between them, F (4, 95) =6.515, p<0.001. As shown in Fig.6, there is an increase of drivers' satisfaction when the HUD interface complexity increases from interface No.2 to No.3 (p<0.01), which is the same as No.2 to No.5 (p=0.005). Nevertheless, we can also receive that drivers' satisfaction begin to decrease when it comes to HUD interface No.3 and interface No.4 (p=0.054).

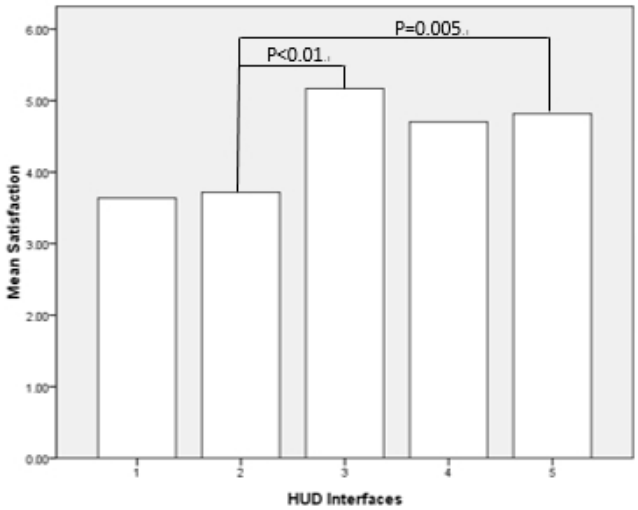


Fig.6 Comparison of mean satisfaction between different interfaces

Mental Workload

In addition to measure drivers' subjective feelings in the driving process, we also wonder the influence of HUD interface complexity on drivers' mental workload. This is of great importance since high mental workload in driving implies underlying dangers, such as tiredness, distraction, and etc. Significance was found between mental workload and interface complexity, F (4,95)=4.933, p=0.001.Results were shown in Fig.7.

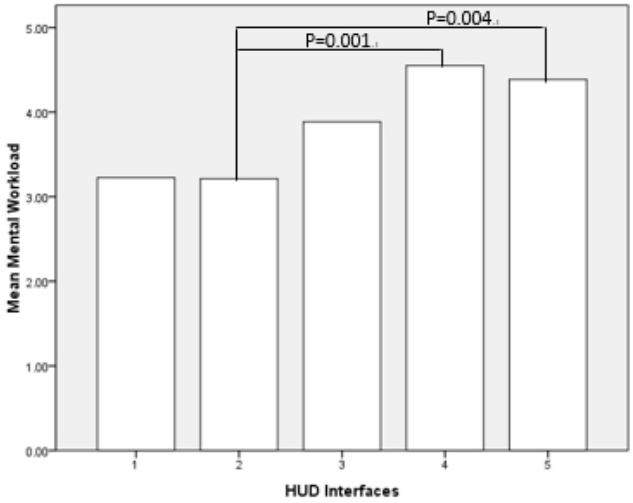


Fig.7 Comparison of mean mental workload between different interfaces

As shown in Fig.7, there is a significant increase of drivers' mental workload when the HUD interface changes from interface No.2 to No.4 (p=0.001) and No.5 (p=0.004), which is the same as No.1 to No.4 and No.5 (p<0.01). Besides, we can also see a marginally significant increase of drivers' mental workload from interface No.3 to No.4 (p=0.099) and a slightly decrease of mental workload from interface No.4 to No.5.

Driving Task Complexity

Though we have examined the perceived complexity of HUD interfaces, another interesting question is that whether the relationship between driving task complexity and HUD interfaces is the same as the relationship between perceived complexity and HUD interfaces. Significance was found between them, F (4,95)= 8.417, p<0.001. And results were shown in Fig.8.

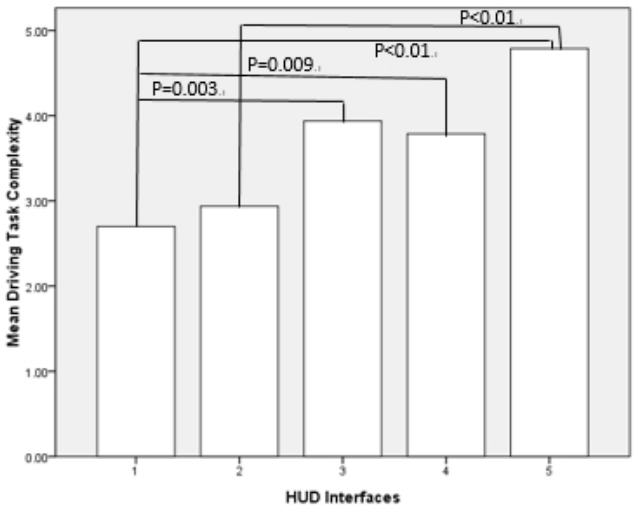


Fig.8 Comparison of mean driving task complexity between different interfaces

As demonstrated in Fig.8, in general, there is an increase of driving task complexity when HUD interfaces change from interface No.1 to No.5. In addition, we can also obtain that there is a significant difference in terms of mean driving task complexity between HUD interface No.1 and No.3, No.4, No.5 ($p=0.003$, $p=0.009$ and $p<0.01$) and interface No.2 and No.5 ($p<0.01$).

DISCUSSION

Perceived complexity. Perceived complexity of these five interfaces is dominantly in accord with the amount of information displayed on each interface. That is, the more complex the HUD interface was designed, the more complex the drivers would perceive. This result was also an illustration that the complexity of HUD interface would really determine the perceived complexity of drivers who used it. Among all these five HUD interfaces, interface No.4 and interface No.5 are the most complex ones for drivers and interface No.1 and interface No.2 are the simplest ones.

Subjective effectiveness. There is a significant increase of subjective driving effectiveness when the HUD interface complexity is growing and there also is a significant decrease when the HUD interface complexity is too high. This means the hypothesis that there should be a reverse "U" shape relationship between HUD interface complexity and effectiveness has been proved, that is, there is an optimal HUD interface complexity existing in terms of subjective effectiveness. However, in the existing researches, optimal interface complexity was found related to user characteristics, which means optimal interface complexity is different mutually (JH Carlisle, 1974; TJ Massaro, 1996).

Easiness. Since one of our hypotheses is that drivers will not feel easy in the driving process when the interface complexity is too low or too high, we expect to see a reverse "U" shape in relationship between easiness and interface complexity. Results were shown in Fig.5. Though this result is beyond our expectation, it is still important to pay attention to this finding. Simply speaking, one reason why drivers got less easiness when the amount of information is increasing may be the distraction, which is caused by information shown on the HUD interfaces, is also increasing (Olsson, S. et al., 2000;Nowakowski, C. et al., 2001). Hence, drivers under condition that HUD interface complexity is high may not be able to keep their attention on the road as easy as those under low complexity.

Satisfaction. The most important index of measuring drivers' performance in a subjective way is drivers' satisfaction under these five kinds of driving conditions. The result has verified our hypothesis that there should be a reverse "U" shape relationship between drivers' satisfaction and HUD interface complexity. This is a finding with significance since it reminds both the researchers and HUD designers that appropriate complexity of HUD interface is highly needed. This confirms the previous finding that there is an increase of satisfaction when interface complexity increases not too much and a decrease of satisfaction when interface complexity becomes too high

(Nielsen, J., 1993; Ahmed, S. Z., 2004).

Mental workload. There is a significant increase in terms of drivers' mental workload when the HUD interface changes from interface No.2 to No.4 and No.5 and marginally significant increase from interface No.3 to No.4. This is consistent with the existing studies that driving leads to a great mental workload for the drivers by the more complex driving context (Cantin, V. et al, 2009).

Driving task complexity. To have a comprehensive understanding of the increasing driving task complexity, we know that the increasing information of HUD interface from No.1 and No.2 to the others is road warning, road condition and navigation. Thus, this kind of information may occupy drivers' attention and hence, increase driving task complexity, which has been shown in other researchers' work (Wickens, C. D., 1994; Wang, W., 2003).

CONCLUSION

The objective of this study is to measure the influence of HUD interface complexity on drivers' driving performance and self-perception. The main limitation of this study is that drivers driving in simulated driving environment may not behave the same as when they are in the real world since simulated lane and pedestrian are somewhat different in view. Another underlying influence on the experiment results may situate at the sample characteristic of this study since all drivers recruited in this study are university students. Nevertheless, there are six broad conclusions that we can conclude from this study.

Firstly, drivers' perceived complexity has been proved to be consistent with designed complexity. Specifically, except for the HUD interface No.1 and No.2, and the left ones all have significance between each other ($p<0.01$). This is also an illustration of the effectiveness of our study. Secondly, the hypothesis that there should be a reverse "U" shape relationship between effectiveness and HUD interface complexity has been proved. To be specific, on one hand, three HUD interfaces (No.1, No.2 and No.3) have significant differences with HUD interface No.3. On the other hand, we can conclude from this research that HUD interface No.3 is the most welcomed and preferred one by drivers. Thirdly, the hypothesis that the easiness curve should be reverse "U" shape has not been proved. Specifically, drivers' feeling of easiness decreased step by step when the HUD interface complexity increased. Fourthly, the hypothesis that both the most complex and the simplest HUD interface will not bring drivers with the best satisfaction, and the satisfaction curve should also be reverse "U" shape has been proved. This is the most significant finding of this study. Specifically, there is an increase of drivers' satisfaction when the HUD interface complexity increases from interface No.1 to No.3. However, drivers' satisfaction begins to decrease when it comes to HUD interface No.4 and No.5. Fifthly, the hypothesis that the more complex of the HUD interface is, the higher drivers' mental workload will be has been proved. To be more detailed, there is nearly no increase of drivers' mental workload when the HUD interface changes from interface No.1 to No.2.and from No.4 to No.5.

Finally, in terms of driving task complexity, the hypothesis that the more complex of the HUD interface is, the higher driving task complexity will be has been proved. This indicates that the driving task complexity is increasing with the increasing of HUD interface complexity. Consequently, this study has started with research on the relationship between HUD interface complexity and driving performance and has found some significant findings. Though the research object is HUD interface, it can also be generalized to similar domain which research on interface complexity.

REFERENCES

Ahmed, S. Z., McKnight, C., & Oppenheim, C. (2004). A study of users' performance and satisfaction with the Web of Science IR interface. Journal of Information Science, 30(5), 459-468.

Alexander, L. (2005). Driver Assistive System Displays for Highway Vehicles. ADEAC, 5, 38-41.

Andersen, G. J. and R. Ni, et al. (2011). "Limits of spatial attention in three-dimensional space and dual-task driving performance." Accident Analysis & Prevention 43 (1): 381-390.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. Intl. Journal of Human–Computer Interaction, 24(6), 574-594

Boyer, C., Gauthier, F. H., & Gerbe, J. P. (1986). U.S. Patent No. 4,600,271. Washington, DC: U.S. Patent and Trademark Office.

Bozdogan, H. (2000). "Akaike's information criterion and recent developments in information complexity." Journal of mathematical psychology 44 (1): 62-91.
Buckley, E., & Stindt, D. (2008). Full colour holographic laser projector HUD. In Proc. SID Annual Symposium on Vehicle Displays (Vol. 15, pp. 131-135).
Cant, S. N., Jeffery, D. R., & Henderson-Sellers, B. (1995). A conceptual model of cognitive complexity of elements of the programming process. Information and Software Technology, 37(7), 351-362.
Cantin, V., Lavallière, M., Simoneau, M., & Teasdale, N. (2009). Mental workload when driving in a simulator: Effects of age and driving complexity. Accident Analysis & Prevention, 41(4), 763-771.
Carlisle, J. H. (1974). Man-Computer Interactive Problem Solving: Relationships Between User Characteristics and Interface Complexity. DTIC Document. (Reprinted).
Caudell, T. P. and D. W. Mizell (1992). Augmented reality: An application of heads-up display technology to manual manufacturing processes. System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on, IEEE.
Charissis, V. (2006). "Driving Simulator for Head-Up Display Evaluation: Driver's response time on accident simulation cases." DSC-A/P 2006.
Charissis, V. and S. Papanastasiou (2010). "Human–machine collaboration through vehicle head up display interface." Cognition, Technology & Work 12 (1): 41-50.
Chen, K., Chiu, S. C., & Lin, F. C. (2007). Kansei design with cross cultural perspectives. In Usability and Internationalization. HCI and Culture (pp. 47-56). Springer Berlin Heidelberg
Coburn, W. J. (2009). Attitudes in psychoanalytic complexity. Beyond postmodern: New dimensions in clinical process and theory. New York: Routledge.
Doshi, A. and S. Y. Cheng, et al. (2009). "A novel active heads-up display for driver assistance." Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 39 (1): 85-93.
Easterbrook, J. A. (1959). "The effect of emotion on cue utilization and the organization of behavior." Psychological review 66 (3): 183-201.
Fadden, S. and P. M. Ververs, et al. (1998). Costs and benefits of head-up display use: A meta-analytic approach. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications.
Freides, D. (1974). "Human information processing and sensory modality: cross-modal functions, information complexity, memory, and deficit." Psychological bulletin 81 (5): 284.
Horrey, W. J. and C. D. Wickens, et al. (2003). The effects of head-up display clutter and in-vehicle display separation on concurrent driving performance. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications.
Carlisle, J. H. (1974). Man-Computer Interactive Problem Solving: Relationships Between User Characteristics and Interface Complexity: DTIC Document. (Reprinted).
Kiefer, R. J. (1998). Quantifying head-up display (HUD) pedestrian detection benefits for older drivers. 16th International Technical Conference on Experimental Safety Vehicles. Windsor: NHTSA.
Kushwaha, D. S. and A. K. Misra (2006). "Improved cognitive information complexity measure: a metric that establishes program comprehension effort." ACM SIGSOFT Software Engineering Notes 31 (5): 1-7.
Lee, D. S., Yoon, W. C., & Choi, S. S. (1998). An entropy-based measure for evaluating the cognitive complexity of user interface. Korean Journal of The Science of Emotion and Sensibility, 1(1), 213-221.
Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. International Journal of Human‐Computer Interaction, 7(1), 57-78.
Liu, Y. (2003). "Effects of using head-up display in automobile context on attention demand and driving performance." Displays 24 (4): 157-165.
Liu, Y. and M. Wen (2004). "Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan." International Journal of Human-Computer Studies 61 (5): 679-697.
Liu, Y. and T. Wu (2009). "Fatigued driver's driving behavior and cognitive task performance: Effects of road environments and road environment changes." Safety Science 47 (8): 1083-1089.
Massaro, T. J., & Schmidt, D. J. (1996). Method and apparatus for variable complexity user interface in a data processing system: Google Patents. (Reprinted).
Miller, G. (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information." The psychological review 63: 81-97.
Moray, N. (1986). "Monitoring behavior and supervisory control." Handbook of perception and human performance. 2: 40-1.
Nielsen, J., & Hackos, J. T. (1993). Usability engineering (Vol. 125184069). Boston: Academic press.
Nowakowski, C., Friedman, D., & Green, P. (2001). Cell phone ring suppression and HUD Caller ID: Effectiveness in reducing momentary driver distraction under varying workload levels. NASA STI/Recon Technical Report N, 3, 13747.
Olsson, S., & Burns, P. C. (2000). Measuring driver visual distraction with a peripheral detection task. Obtained from August.
Seder, T. A. and J. F. Szczerba, et al. (2010). "FULL-WINDSHIELD HEAD-UP DISPLAY ENHANCEMENT: ANTI-REFLECTIVE GLASS HARD COAT.".
Smith, S. and S. Fu (2011). "The relationships between automobile head-up display presentation images and drivers' Kansei." Displays 32 (2): 58-68.
Svensson, E. and M. Angelborg-Thanderez, et al. (1997). "Information complexity-mental workload and performance in combat aircraft." Ergonomics 40 (3): 362-380.
Tonnis, M. and C. Lange, et al. (2007). Visual longitudinal and lateral driving assistance in the head-up display of cars. Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on, IEEE.

Tönnis, M., Klinker, G., & Plavšic, M. (2009). Survey and Classification of Head-Up Display Presentation Principles. Proceedings of the International Ergonomics Association (IEA).
Tufano, D. R. (1997). "Automotive HUDs: The overlooked safety issues." Human Factors: The Journal of the Human Factors and Ergonomics Society 39 (2): 303-311.
Wang, W., Chao, D., Kao, C., Wu, H., Liu, Y., Li, C., King, C. (2003). High levels of plasma dengue viral load during defervescence in patients with dengue hemorrhagic fever: implications for pathogenesis. Virology, 305(2), 330-338
Watanabe, H. and H. Yoo, et al. (1999). The effect of HUD warning location on driver responses. Sixth World congress on Intelligent Transport Systems, Toronto, Canada.
Way, T. C. and R. L. Martin, et al. (1987). "Multi-crew pictorial format display evaluation."
Wickens, C. D., & Long, J. (1994, October). Conformal symbology, attention shifts, and the head-up display. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 38, No. 1, pp. 6-10). SAGE Publications.

Influence of Information Flows on the Results of a Driver's Activity in the System «Driver – Car – Road – Environment»

Viktor Dolya, Ivetta Afanasieva, Urii Davidich, Iryna Englezi, Nizami Gyulyev, Oleksii Prasolenko

O. M. Beketov National University of Urban Economy in Kharkiv
Kharkiv, 61002, Ukraine

ABSTRACT

The article deals with the impact of information flows on results of the driver's activity in the system «driver – car – road – environment». Existing research on the impact of information flows on the characteristics of the driver are analyzed in this paper. It is determined that the previous studies do not study sufficiently the impact of information obtained from the additional sources (bill-boards on the side of roads, cell phones etc.) on the activity of the driver and do not research time of driver's distraction from performing main activities. As a result of laboratory studies the mathematical model of the impact of information flows based simultaneously on indicators of Electroencephalography (EEG) and EEG and ECG of time of driver's distraction from performing main activities is formed. Correlations of impact of information obtained from additional sources on time of driver's distraction from performing of main activity are received. Statistical evaluations of the obtained models demonstrate their adequacy and possible implementations.

Keywords: information flow, functional state, a driver, distraction.

INTRODUCTION

Any modern human faces with a problem of choice and processing of information coming to him/her in the age of information breakthrough. Human activity in systems "driver - car - road - environment" - is not an exception. The role and importance of information in all spheres of human activity increased significantly. The intensity of the flow of information that affects the person (a driver), is increasing constantly due to the rapid development of modern facilities of electronic telecommunications, increasing the number of roadside advertising, active use of cellular devices of information transmission. Increasing the flow of information both outside the car (excess unwanted information, the number of roadside advertising) and inside the car (using of the cellular and radio), affects the emotional state of the driver and reduces its reaction, increases the likelihood of traffic accidents (TA). Most accidents are caused by drivers, not by the car or by the road. The personal factor - the human psyche is the base of accidents. Taking into account this matter, the study of the impact of information flows on the driver is relevant.

METHODS USED IN RESEARCHES OF THE EFFECT OF INFORMATION FLOW ON THE DRIVER

Information flows

Directions for improving the organization of work of the drivers are based on the drivers' psychophysiological opportunities and their ability to receive information quickly and to act in stressful situations accurately, ie taking into account the driver's reaction time (Davidich, 2006). Furthermore, that directions are composed of control information transmitting to drivers tools. Effect of road conditions and information traffic in the human body causes fatigue and increases reaction time , decreased speed of information processing , ie, information overload the driver. Research on the influence of information flow on the characteristics of driver activities always attract attention

because it affects the security of the person. In this paper we consider the influence of information obtained from other sources (billboards on the roadside, cell phones, etc.). All this led to the goal and objectives of the study.

Methods

There are many methods for studying human factors in systems (Charlton, 2002). During the investigation the most informative parameters of the impact of information flows on the performance of the driver in the system were chosen. The parameters were decided to be determined using the following methods: the method of EEG (Gnezditskiy, 2002) with using the hardware-software complex "Neurocom" (CSE., 2008), used to determine changes in the electrical activity of the driver's brain when processing the input information, the method of ECG (Akselrod, 2010) with using the hardware-software complex "Cardiosens" used to determine the level of driver fatigue in research; tabular method "correction test" used to determine the execution time of a driver distraction core activities.

STUDY OF INFORMATION FLOWS

On the road, every driver is faced with a huge amount of information flows (Fig. 1), in the present work we investigated the influence of information obtained from other sources.

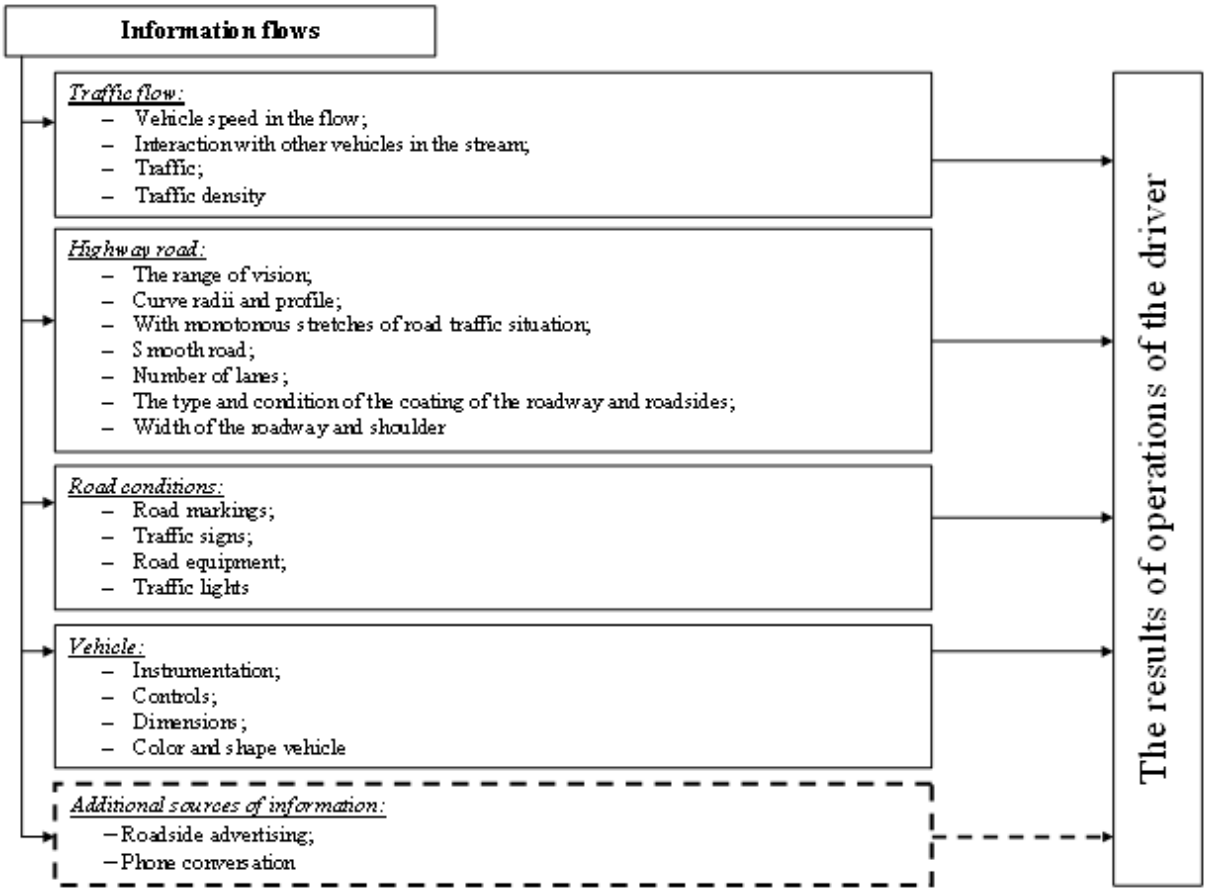


Fig. 1. Sources of information affecting performance of a driver

Study the impact of information flows on the performance of the driver consisted of two parts: laboratory studies and field studies. The first part (studding the impact of information flows on the performance of the operator) was carried out in the laboratory by the algorithm (Fig. 2).

The study was conducted in the laboratory under natural light. During the experiment the testee held a comfortable position. Small electrodes connected by wires to an electroencephalograph were attached with a special helmet to the testee's head. Before starting the experiment, the testee was informed about the procedure of EEG and its painless. EEG brain research carried out for 20-30 min to obtain reliable data.

During the first stage (Fig. 2) the conditions of the laboratory experiment using an electroencephalograph, when the quantity of artifacts (N_{apr}) tends to a minimum are defined. Recommendations to reduce artifacts during laboratory studies are shown in the Table. 1.

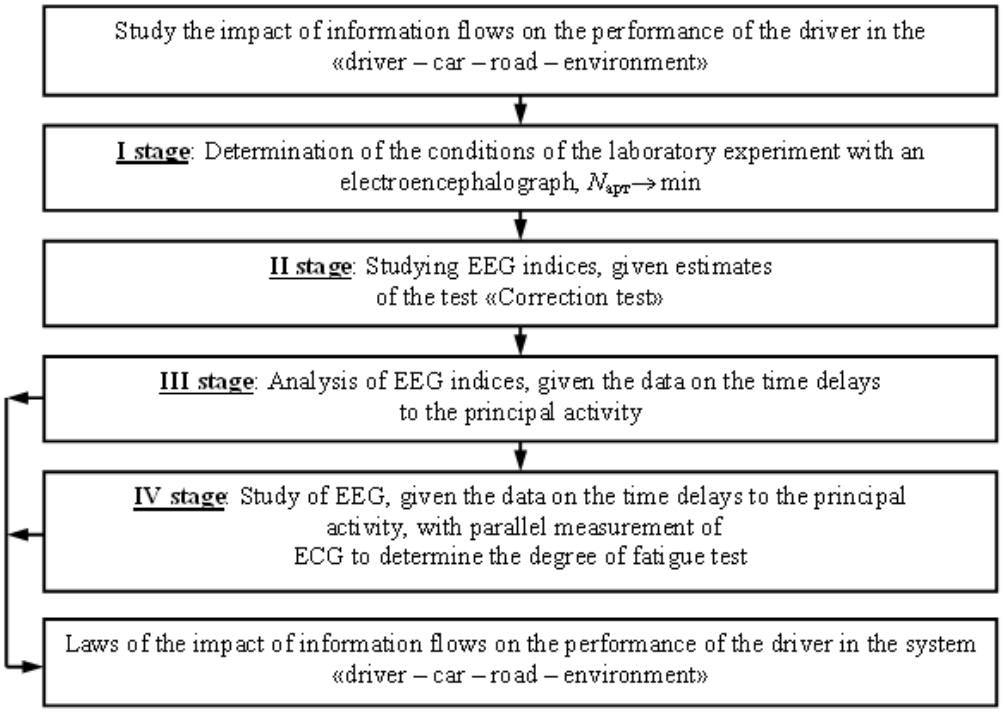


Fig. 2. Stages of the research

Table 1

Recommendations for reducing artifacts in laboratory studies using EEG

View artifact	The cause of	Recommendations to reduce its impact
Muscle artifact	Muscle activity (body motion test)	• Before starting the experiment, the subject must take a comfortable position; • During the experiment the test legs should be fixed, smooth head movement is possible in case of need under the experimental conditions; • Movement of the hands at the maximum possible limitation of motion of the forearm
Oculographical artifact	The rotation of the optic apple and winking test	• Before starting the experiment, the subject is able to use eye drops to moisten; • Using the occipital leads (O) for the analysis of EEG data; • Removal of visual artifacts using the technology ICA
Cardio artifact	ECG signal is close to the testee	• Conclusion of the experiment to the member whose registered this artifact
Pulsation artifact	Application of EEG electrode on a blood vessel of the testee	• Adjustment of the electrodes
Spiro artifact	Breath test	• During the experiment, the subject must breathe slowly;
Galvanic skin artifact	Sweating skin test	• Alcohol treatment places electrodes fixation
Voicekinetic artifact	Voice signals test	• During the experiment, the testee must speak without the use of abrupt facial movements
Extraphysiological artifact	Poor contact, AC, movement of objects around the subject, the sound wave directed at EEG amplifier	• Checking technical conditions before the start of the experiment; • Traffic technicians resolved within a 2 m, provided the use of rubber-soled shoes; • Do not direct sound wave from any source to the amplifier EEG

During the second stage the existence of the correlation between EEG and indicators of psychological test " correction test" was registered. The study involved 12 men aged 20-30 years. EEG recording of the test carried out in a state of quiet observation with open eyes. Monopolar recording was carried out in the 0-70 Hz bandwidth with a sampling frequency of 250 Hz with 19 active electrodes. Recording duration of the test was limited to complaints in the area of EEG and mounting hardware was 20 min. Estimates of test "correction test" indicate that an excess of information (additional questions) affects the development of fatigue during the experiment. However, these figures do not take into account the change in the amount of perceived information in real time. Taking into account the above mentioned items, the third phase of the study determined the correlation between EEG and delay processing of the information obtained during the test run its core activities due to the distraction of getting information from secondary sources.

During the third stage of the experiment, the testee must perform the test focused "correction test" , that was the main activities and main source of information obtained through the visual analyzer test . Simultaneously with this task subject received related to the main activity information from another source. One of the important tasks in the study - the right to answer the questions posed in the form of oral examples for easy mathematical calculation consists of two steps without being distracted from his main task - the test run " correction test." During this experiment was conducted surveillance of the work of the testee, by which time it was recorded distraction from performing basic activities.

Temporary sites were selected in each sample in order to analyze EEG data, corresponding to the time of distraction from the main activity of the test run . Time distraction was determined using a video of each sample and represents the time during which the subject ceased to perform its main function (passing the test " proof- test") , ignoring the questions.

Based on the analysis of quantitative characteristics of the intensity of the fast (β, γ) (Dolya et al., 2011) and slow (δ, θ) (Dolya et al., 2011) EEG rhythms constructed a general model of the impact of information flows that appears EEG , at the time of driver's distraction from performing its principal activity is as follows:

$$\Delta\tau_{i1} = 2{,}733 + 0{,}0423 \cdot \delta + 0{,}0295 \cdot \theta - 0{,}039 \cdot \beta - 0{,}018 \cdot \gamma, \quad (1)$$

$\Delta\tau_{i1}$ – while diverting drivers from performing their primary type of activity reflected EEG.

The results of calculations of the model parameters are given in Table 2 , the statistical evaluation of the model - in Table 3 . The resulting Fisher criterion (Table 3) shows the influence of informative model of information flows . Multiple correlation coefficient reflects a high degree of closeness of the connection. Assessment of the adequacy of the resulting model (1) was carried out in terms of average error of approximation, the value of which is equal to 10.86 %, which corresponds with the limits and indicates the adequacy of the model.

Table 2

Statistical evaluation of the general model of the information flows impact on the time of driver distraction from the main activity, considering the performance of EEG

Data	Values			
Factor	Percent delta-rate signal from the total power of the signal in the cerebral cortex	Percent theta-rate signal from the total power of the signal in the cerebral cortex	Percent beta-rate signal from the total power of the signal in the cerebral cortex	The percentage of gamma-rate signal from the total power of the signal in the cerebral cortex
Designation, dimension	δ, %	θ, %	β, %	γ, %
Coefficient	0,0423	0,0295	-0,039	-0,018
Standard error	0,329	0,0056	0,009	0,004
Student's *t*-test: calculated tabular	 8,296 1,98	 7,613 1,98	 3,184 1,98	 -9,639 1,98

In the fourth stage of the research organization of an experiment carried out in the same manner as in the third stage, with the additional use of hardware and software "Cardiosens." At this stage our task was to determine

the degree of operator fatigue, there are different methods for determining fatigue (Oven, 2009; Mocco, 2010). In our studies, we used the method allows the ECG in real time to determine the degree of fatigue.

Table 3

The results of assessment of the overall impact of the model information flows at the time of driver's distraction from the main activity, taking into account indicators

Data	Values
Multiple correlation coefficient, %	90,86
The average error of approximation, %	10,86
Fisher's exact test: settlement tabular	241,04 2,46

The mathematical model of the impact of information flows was built based on analysis of the data obtained during the fourth step, taking into account the performance of EEG and ECG at the time of the driver's distraction from performing basic activities:

$$\Delta\tau_{i\ 2} = \left(5{,}302 + 0{,}0899 \cdot \delta + 0{,}052 \cdot \theta - 0{,}0721 \cdot \beta - 0{,}055 \cdot \gamma\right)\left(\frac{P_r}{P_n}\right)P_r^{\left(1-\frac{P_r}{P_n}\right)}, \quad (2)$$

$\Delta\tau_{i\ 2}$ – while diverting drivers from doing their core business, reflected EEG and ECG;

P_r – value of the index activity of regulatory systems in the billing period;

P_n – value of the index activity of the regulatory systems of the background sample.

The results of calculations of the model parameters are given in Table 4, the statistical evaluation of the model - in Table 5.

Table 4

Statistical evaluation of the general model of the impact of information flows at the time of driver's distraction from the main activity, considering the performance of EEG and ECG

Data	Values				
Designation, dimension	$\left(\frac{P_r}{P_n}\right)P_r^{\left(1-\frac{P_r}{P_n}\right)}$	$\delta\cdot\left(\frac{P_r}{P_n}\right)P_r^{\left(1-\frac{P_r}{P_n}\right)}$	$\theta\cdot\left(\frac{P_r}{P_n}\right)P_r^{\left(1-\frac{P_r}{P_n}\right)}$	$\beta\cdot\left(\frac{P_r}{P_n}\right)P_r^{\left(1-\frac{P_r}{P_n}\right)}$	$\gamma\cdot\left(\frac{P_r}{P_n}\right)P_r^{\left(1-\frac{P_r}{P_n}\right)}$
Coefficient	5,302	0,0899	0,052	-0,0721	-0,055
Standard error	0,714	0,011	0,018	0,008	0,009
Student's t-test: calculated tabular	 7,42167 1,98	 8,34781 1,98	 2,8517 1,98	 -8,6942 1,98	 -6,4535 1,98

Table 5

The results of assessment of the overall impact of the model information flows at the time of driver distraction from the main activity, considering the performance of EEG and ECG

Data	Values
Multiple correlation coefficient, %	99,11
The average error of approximation, %	9,32
Fisher's exact test: settlement tabular	913,1 2,46

Based on the data shown in Table 5, its calculated Fisher criterion more table index, which indicates the informative model. Meaning of the multiple correlation coefficient is 0.99, which reflects the high quality features and the

coupling strength indicates that , through the use of a comprehensive registration and evaluation of ECG and EEG indicators was obtained more qualitative mathematical model to determine the influence of information flows at the time of distraction from performing drivers to their main activity . Average value of the approximation error for the model (2) is 9.32 %, which corresponds with the limits and indicates the adequacy of the model. Compared with the model (1), the value of the average error of approximation which is 10.86 % , we see that by incorporating ECG in the model (2) , managed to get a more accurate model of the impact of information flows at the time of driver distraction from performing its core activities. On the basis of the calculations can be concluded possibility of using the mathematical models (1, 2) in further studies.

The second part of the study was carried out in real conditions. The study involved 10 drivers aged 24 to 45 years with the experience from 6 to 12 years, excluding the type of temperament and gender of the driver.

In the first phase of this study was determined by the possibility of electrophysiological studies of EEG and ECG in natural conditions. Was chosen for two vehicles with different transmissions : a manual and automatic . The purpose of this phase of the study - to identify all possible artifacts in the equipment for recording EEG and ECG signals. For this experiment was chosen route traffic on the ring road of the city of Kharkiv .In the laboratory, the subject was fixed electrodes Holter "Cardiosens" electroencephalograph electrodes fixation " Neyrocom " (Fig. 3)carried out in the car, which was parked in an authorized location at the beginning portion of the selected route.

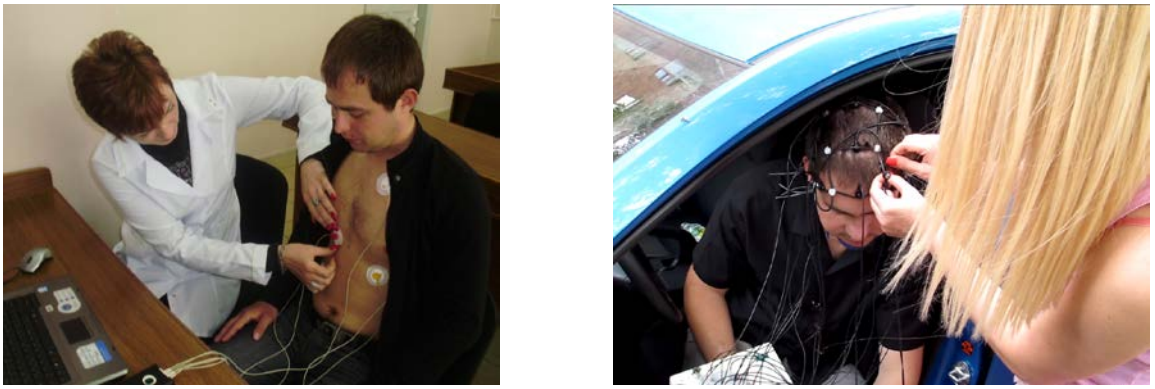


Fig. 3 Fixing electrodes

The experiment led to the following conclusions:

1) bumps in the road affect the quality of the recording EEG and ECG signals, creating strong interference in the record. In view of this route in the route selection for further experiments must take into account the quality of the pavement;
2) impact of the use of vehicles with different transmissions: a manual and automatic. While driving using a manual transmission driver performs a plurality of mobile functions, preventing the recording of EEG and ECG signals. In order to avoid additional sources of interference in further experiments will use cars with automatic transmission;
3) while driving in conditions of quality pavement to minimize muscle movements driver may write quality signal EEG and ECG.

Fig. 4 types of outlets video

During the second stage of the study conducted experiments in natural conditions in different parts of the roads within the Kharkiv city, and on the roads at the entrance to Kharkiv in areas with a high density of roadside advertising (billboards). Investigated the effects of additional sources of information (billboards , telephone calls using the system BLUETOOTH HANDS FREE, music and radio in the car) on the driver with the definition of the time delay in getting information from the primary source (road conditions). Subjects drivers before entering the route has been focused on research and on the experimental vehicle. The test drivers in this study controlled car equipped with two cameras to remove the review and removal of front face of the test (Fig. 4).

For analysis of data 45 studies 10 and portions with a total collection time of 15 hours were obtained. Total 90 videos were analyzed and identified 121 cases of identifying additional sources of influence on the driver.
Influence of information obtained from a telephone conversation on the location in the field of vision prominent advertising (additional sources) on driver distraction from the road conditions (mandatory information source) was determined taking into account the resulting model (2). Using the model, the impact of information flows at the time of distraction from performance drivers of their main activity (2), the dependences shown in Fig. 5.

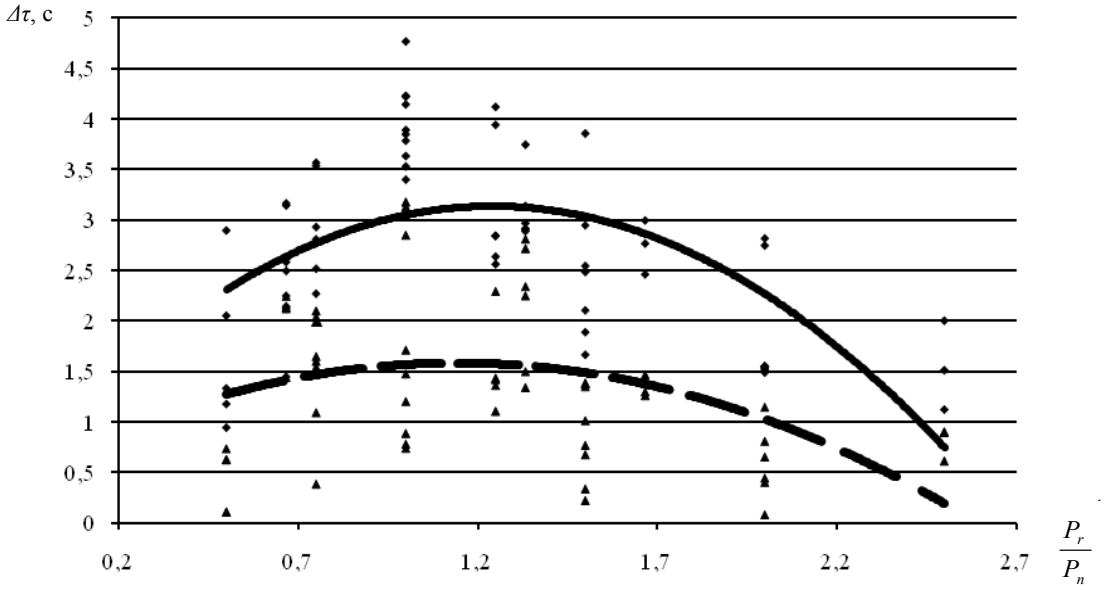


Fig. 5. Dependence on the time change of driver's distraction from performing basic activities in view of its functional state when the information obtained from the telephone conversation and roadside advertising:
- The impact of the telephone conversation;
- The impact of off-road advertising.

The resulting curve (Fig. 5) suggests that after the core business and increase driver's fatigue during the distraction of additional sources of information decreases. This is due to the fact that the driver is focused on the treatment of the basic information from the road and the environment. These dependences show that the additional information obtained from a telephone conversation when the driver core activities , distracting him twice as much as the information that the driver receives from roadside advertising.

CONCLUSIONS

As a result of studding the influence of additional sources of information on the time of driver's distraction from the main activity was determined that the information flow because of the driver's conversation by phone with the use of BLUETOOTH HANDS FREE distracts the driver from 0.94 s. to 4.77 s.

Information flows arising from the location in the field of driver's vision prominent advertising while driving distracted him/her from performing basic activities from 0.23 s. to 2.81 s. results of influence of the telephone conversation and roadside advertising on the driver were obtained in this study, but in general this method of research conducting lets to determine the impact of any sources of information, such as music , radio , passengers in the car , etc. on a driver or operator of any system .

REFERENCES

Давідіч Ю. О. (2006), *Проектування автотранспортних технологічних процесів з урахуванням психофізіології водія* – Х. : ХНАДУ. – 292 с.

Charlton S. G. *Handbook of human factors testing and evaluation* – Mahwah : LEA. – 2002. – p.536. – 2nd. Edition.

Гнездицкий В. В. (2002), *Обратная задача ЭЭГ и клиническая электроэнцефалография* – Таганрог: ТРГУ. – 625 с.

Компьютерная система электроэнцефалографии НейроКом : инструкция по эксплуатации. – Х.: НФУ «ХАИ», 2008. – 171 с.

Аксельрод А.С. (2010), *Холтеровское мониторирование ЭКГ: возможности, трудности, ошибки* – 2-е изд., исправ. и доп. – М.: МИА,. – 192 с.

Доля В. К., Енглезі І. П. , Афанасьєва І. А. (2011), *Вплив інформаційного навантаження на параметри основної діяльності водіїв (гальмівні процеси)* // Інформаційно-керуючі системи на залізничному транспорті. – № 3. – С. 32–35.

Доля В. К., Енглезі І. П. , Афанасьєва І. А. (2011), *Вплив інформаційного навантаження на параметри основної діяльності водіїв (збуджувальні процеси)* // Восточно-европейский журнал передовых технологий. – 1/2(49). – С. 65–68.

Oven S. *Survey and evaluation of fatigue assessment software [Електронний ресурс] / Project, Fall. − 2009. – p. 30. − Website: http://www.utmis.org/sv/aktuellt/Documents/Oven_Project (2).pdf*

Моссо А. *Усталость. Физиологические и психологические аспекты/ А. Моссо. / Пер. с итал. − Изд. 2 − М. : ЛКИ, 2010. – 328 с.*

Section 15

Road and Rail:

Transport Data and Analysis

Deconstructing the Whole Journey Traveler Experience

Andree Woodcock1, Jane Osmond1, Gabriela Rodica Hrin, Olympia Papadopoulou3, Oded Cats4, Yusak Susilo4, Tiago Pimentel5 and Owen O'Reilly6

1 Coventry School of Art and Design, Coventry University, Coventry, UK

2 INTEGRAL CONSULTING R&D, Bucharest, Romania
3TERO Ltd., Thessaloniki, Greece

4Department of Transport Science, KTH Royal Institute of Technology, Stockholm, Sweden
5VTM, Portugal
Interactions Ltd., Ireland

ABSTRACT

Understanding travelers' experiences across the whole journey lifecycle is becoming increasingly important for defining and optimizing mobility for all travelers (active travelers and those using public and/or private modes of transport), for transport operators, transport authorities, municipalities and the designers of transport technology. Integrated transport solutions will see greater amalgamation of journey planners, route navigation systems and advanced parking and payment applications, through to vehicle prioritization. The FP7 project, METPEX aims to create a suite of measurement instruments which will capture in real time or retrospectively the whole journey experience, from planning to arrival at destination, across different transport modes. The paper will firstly outline the human factors of each stage of the traveler's journey. Secondly, it will review the approaches currently used by transport operators to measuring the quality of the passenger experience. Thirdly, the aims and objectives of the METPEX project will be discussed together with work undertaken in the first year of the project, in relation to the development and validation of an extensive set of variables.

Keywords: Transport, mobility, user surveys

INTRODUCTION

The Lisbon Strategy set a goal for the European Union to become the most competitive and dynamic knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion. Public transport has a central role to play in ensuring equitable access to social, economic, educational and health services. Its effective use is also seen as key to reducing urban congestion and greenhouse emissions. However, public transport is regarded by many as an inferior form of transport, a perception which prevents it from realizing its full contribution to reducing road accidents, congestion, energy consumption and pollution. A greater understanding of mobility behavior and the expectations of users is required to enhance service provision, design and performance and to ensure the relevance of local, national and regional policies and strategies.

Although much research has been conducted on the integration of transport modes and travel information on the one hand, and travel behaviour and demand analysis on the other, these two streams of knowledge have not had adequate interaction hitherto. Furthermore, key stakeholders such as transport authorities, urban and regional local government and transport operators need to have access to standardised methods and results in ways that can inform policy and analysis, enabling the implementation of integrated transport systems and increased accessibility. Therefore there is a gap in the market to create a validated PanEuropean to measure and benchmark passenger experience which can be used by transport groups, operators and authorities.

OVERVIEW OF METPEX

METPEX (ww.metpex.eu) is a three year, FP7 funded project. Its focus is on the measurement of the quality of the passenger journey (door-to-door), including private or individual forms of transport such as walking, cycling and car sharing. METPEX will develop, validate and evaluate a standardised tool to systematically measure passenger experience across whole journeys, to develop best practice and benchmark services. Therefore deconstructing the journey into its different element is necessary in order to develop a systematic approach to consideration of human factors.

The information derived from the use of the METPEX measurement instruments will ultimately be used to inform policy makers in providing inclusive, passenger-oriented integrated transport systems that are accessible by all citizens. Specific objectives include:

1. Developing an integrated approach to the measurement of the whole journey passenger experience that takes into account human (physiological, perceptual, cognitive, sensory and affective) socio-economic, cultural, geographic and environmental factors.
2. Assessing the costs of 'inaccessible transport' for different sectors of society (such as those from low income groups, rural communities, the elderly, disabled and those with lower levels of literacy).
3. Assessing the extent to which the measurement of the passenger experience can be used to drive innovation and attention to transport quality from the customer's perspective in the transport industry.
4. Evaluating the passengers experience from different regions of Europe and to support the integration of regional transport networks into an European transport network
5. Facilitating the harmonization of travel behaviour research and analysis across European Union Member States.

DECONSTRUCTING THE JOURNEY (PART 1)

Woodcock (2012) stressed the need to take a systematic of the whole journey experience, in which each element should be optimized for each user, and that each part of the journey, including movement between transport modes and to and from transport gateways contributed in part to the overall experience. As such, the selection of transport mode is dependent on the sum of previous experiences, and each element of the journey may contribute negatively or positively to transport decision making. From this perspective the journey may be broken down into a number of elements (an example of these is shown in Table 1), each of which may have specific factors which need to be considered from the perspective of different user groups. This section focuses on some of these issues.

Table1: Example of a journey deconstruction

Journey stage	Description
1. Assessment of need for the journey	
2. Journey planning	Including assessment of mode and time of travel, on line ticket purchase, finding routes to the destination, collating information
3. Preparation for the journey	Including gathering primary journey artifacts (e.g. tickets, bus passes, passports,) and secondary journey related artifacts (e.g. entertainment, mobility aids, carriers and food)

4.	Movement from the origin to the transport gateway	Negotiating the route from the door to the first vehicle. Little attention is paid to this stage of the journey by transport operators, but it is of key importance to municipalities in planning transport provision
5.	Interaction with transport service	Including payment, ingress, travelling and egress from transport vehicle
6.	Travelling on the vehicle	Including vehicle design, service operation, quality of service, accessibility. Providing a high quality, friendly service is key to winning back support for public transport.
7.	Negotiating transport interchange (with iterative loops from 2 to 6).	This may require change of transport mode, finding the location of transport stops, travel information for the onward journey, purchase of new tickets. Of especial importance are the needs of vulnerable and disabled travelers
8.	Egress from the service to the destination	Although travel from the home to public transport may be considered, this stage is not so prioritised (Woodcock et al, 2014 and below)

Of the eight stages, the first two are the most crucial in terms of achieving EU transport targets as they relate to whether and how journeys should be undertaken. Integrating transport with smart city developments could significantly reduce the amount of journeys made, for example by supporting location independent working, e-commerce and penalties and rewards for certain types of behavior (such as higher parking charges in city centres, dedicated lanes for lift sharing). The impact of the destination has been studied in terms of tourism and vacation travel (e.g. Murphy et al, 2000), but it has only started to be addressed in terms of everyday travel. It may be argued that the attractors of a destination (Rajesh, 2013) in addition to the costs (e.g. physical, cognitive, temporal, financial and affective) need to be understood more fully. A starting point might be the attributes already developed in relation to tourist destinations. For example, Chi et al (2008) developed 8 categories – travel environment, natural attractions entertainment and events, historic attractions, infrastructure, accessibility relaxation, outdoor activities and value – some of which may be key elements in deciding whether to make a journey in to a city centre. The destination has to meet requirements, fulfill stated and unstated needs, especially if high costs are being incurred in the travel (e.g. in terms of journey preparation, see below)

The quality of travel information can contribute substantially to the overall satisfaction with public transport (e.g. Balcombe et al., 2004; Stradling et al., 2000) whether it is static, dynamic or real-time, provide in advance, wayside or station or en route (FTA,2003; Hine and Scott, 2000). Factors influencing the usability of the information include the content, design, condition and timeliness of the information. Recognition of the whole journey experience highlights the need to provide multimodal information, to enable planning, and ease of transfer between modes in the provision of integrated multimodal information services (Lyons and Harman, 2002) which ‘‘minimise the effort for the user in acquiring information on mode choice options and is able to expose the user to information on such options’’ (Kenyon and Lyons, 2003). Groetenhuis (2005) mapped out the effectiveness of different forms of information, provided at different stages of the journey in terms of physical, cognitive and affective savings.

The deconstruction of the journey (shown in Table 1) includes a stage related to ‘preparation for the journey’. This has been included based on observations and personal experience of ‘the difficulty of getting out of the house’. Logically, if planning phases are included in discussions about the journey, then this stage should also receive some consideration as well. For many, once the decision has been made to make a journey, the route and mode of transport decided upon, it is quite easy to depart. For others, such as the elderly and those who have responsibility for others, this may not be so simple. Travel related artifacts needed to be found, transferred to appropriate clothing and bags etc. Primary artifacts may include keys, tickets, travel passes/permits/identity cards: secondary items could include mobility aids (such as walking sticks, buggies), appropriate clothing for the journey, maps; and tertiary items (non essential, but useful in making the journey more enjoyable), could include entertainment (devices), food etc. Obviously each element may have design/ergonomics issues associated with it. However, what may be important is the extra stress and time this places on potential travelers. It may be hypothesized that the less important the journey/the less strong the attractors of the journey, and the more complex the pretrip arrangements , the less likelihood there will be of undertaking a journey. For the elderly/ people with young children, this may lead to greater isolation.

The design of the interchange is crucial to actual and perceived seamless travel (Hickman et al, 2012). Its design

effects the time spent waiting, the ease of transfer between vehicles and the attendant risks and inconvenience that are involved with this activity. Interchanges have been shown to foster uncertainty in the mind of the traveler (Chartered Institute of Transport, 1998; Colin Buchanan and Partners, 1998). Requirements relate to personal security, travel information, ticketing arrangements, predictability of service, low waiting times, and a reduction in organizational and institutional barriers (e.g. Faber, 1996; GMPTE, 1997).

Hine and Scott (2000) identified a number of factors associated with bad interchange design which included poor quality waiting environments, Payment for toilet facilities and inadequate toilet facilities, poor provision of timetable information (outdated, difficult to find),unreliable telephone information services, low levels of personal security, poorly lit facilities in the evening, confusing ticketing and pricing systems, carrying luggage long distance, no public telephones, poor signage within the interchange and for the bus or train services, lack of staff to ask for help, large distances between different interchanges, difficulty getting to the first vehicle and then finding the next service during the interchange activity, long boring waits for commuters. They found that interchanges represent a source of anxiety, uncertainty and powerlessness. This could be reduced if up to date information was given at key decision points, and with greater interoperability between service providers.

Walking forms a significant part of most journeys, yet it is only recently been given attention, through the links made between public health and active transport. However, walking routes are often impeded through poor maintenance such as cracked paving stones, puddles caused by blocked drains and broken lights. The London Councils recognized institutional barriers to promotion of walking (such as low priorities in terms of funding), lack of integrated funds for walking, safety concerns, misperceptions of walking speeds and distances, poor urban environments and habits. In terms of the whole journey experience, poorly maintained urban environments, or urban environments which prioritize motorized forms of transport lead to disadvantages for walkers including those using public transport. Mackett et al (n/d) listed examples of barriers to walking for those who are socially excluded as including changes of level, concerns about finding the way, difficulty in crossing roads, difficulties in moving along the pavements, difficulty in walking long distances and fear.

The objectives of METPEX are to measure the quality of the whole journey experience. The above review has been used to situate METPEX in terms of the wider factors involved in travelling. The measurement tools will enable us to capture (sometimes real time) information about most of the stages shown in Table 1, the interactions between passengers and the services, and to understand the relative importance of different factors (e.g. perceive quality of service). The toolset will not however gather information about Stages 1 and 3 as outlined in Table 1.The next sections of the paper will provide an overview of activities conducted in the first year of the project in support of the generation of the measurement instruments

UNDERSTANDING STAKEHOLDER REQUIREMENTS

Although guidelines and standards aimed to accommodate the different needs of different travelers have been established, there is still a lack of knowledge on what is really valued by different groups of travelers who use different travel modes, and the requirements of those who do not use public transport. Moreover, previous studies often ignore the impact of the access and egress legs on the overall travelers' journey satisfaction. Taking a holistic approach to the study of the passenger experience and journey satisfaction, not only from a users' perspective but also that of stakeholders', will provide an important bridge between action and intention to use more sustainable travel mode (Friman et al., 2011).

Prior to developing the METPEX tools (e.g. mobile app, on-line questionnaire, focus group and interview schedules), it was important to build on knowledge already gained in previous projects (such as those conducted under the CIVITAS programme) and to understand what information is currently recorded by transport planners, authorities and operators and where the gaps are. From these studies a comprehensive set of variables could be developed from which tools could be derived. To support this, a series of semi structured interviews were held with stakeholders across Europe to ascertain what journey information was currently being collected and the value placed on this. The following results provide an overview of this study, with especial reference to the situation in the UK.

The stakeholder interview survey was launched in April 2013 for 6 months across ten cities - Bucharest, Dublin, Grevena, Rome, Stockholm, Turin, Valencia, Coventry, Vilnius and Zurich - and one European body - the European Disability Forum (see http://www.edf-feph.org/). Stakeholders include local transport authorities, transport operators, government bodies, municipalities, passenger interest groups, those responsible for different aspects of

the network and major regional employers.

A standardised interview schedule was drawn up to enable control of data gathering across the different countries. Questions related to the extent to which the stakeholder was informed about the perceived quality of service (such as vehicle and station comfort, network comprehension); the importance of such information; their knowledge about the passenger journey (e.g. origin, final destination, travel and waiting times); the importance they placed on that information; accessibility of the service, infrastructure and network (including travel information and ticketing provision, stations, vehicles and interchanges) for those with reduced mobility; the likely impact a project such as METPEX might have in increasing customer trust and patronage, operating costs, policy and employee satisfaction; and interest in benchmarking.

The interview also gathered stakeholders' views on a pilot METPEX passenger survey questionnaire (developed from the initial literature review, Cats et al, 2014) on important travel experience factors, target user groups, common practices and policy priority areas. Stakeholders discussed which variables were important from their perspectives, identified variables that may be missed or unique from city to city and offered suggestions concerning questionnaire format and survey design.

In the UK, the interview was delivered in a semi structured manner in the stakeholder's office. Respondents were encouraged to read the questionnaire beforehand and provide supporting material (e.g. data collection tools).

Results

A total of 45 stakeholder interviews were conducted with planning authorities, public transport agencies and operators, non-governmental special interest groups and miscellaneous groups (e.g. national research institutes and passenger groups). An overview of the results is provided in Table 1 (Woodcock et al, 2014). Most the stakeholders appreciated and saw the benefit of METPEX. They believed that the identification of passengers' travel needs would be essential for future action plans and transport service modernization. Further, they identified that one of the most useful outcomes of the project would be to quantify and rank the contribution of each quality factor to the overall experience and use this to guide investment in service improvement. One of the stakeholders highlighted the potential of this project to provide a better and more (economically) efficient service during the period of economic crisis.

Table 2: Variables most valued by different stakeholder groups (Woodcock et al, 2014)

Operator	**Authorities**	**Specific Needs Groups**	**Other**
Subjective well being			Subjective well being
	Attitudes and opinions towards model specific preferences, social norm, transfer preference, traffic congestion and pollution, safety and security while travelling		
	The main purpose of the trip		
	Carrying heavy or bulky items when travelling		
	Familiarity with the trip		
	Trip arrival constraint		
	Access to public transport card		
The use of pre trip information			
			Satisfaction level of current choice
		The occurrence and impact of disruptions	
Detailed trip stages, including waiting and on-vehicle time and speed, travel time and punctuality		Detailed information on perception of time reliability	Detailed trip stages, including waiting and on-vehicle time and speed, travel time and punctuality
Information acquisition		Information acquisition	
Time utilization on board and at stops		Time utilization on board and at stops	
	Overall satisfaction in general, compared to expectations, towards other mode choices and travel modifications		
Parking price and easiness to find parking spot	Travel experiences among car travelers, which include the reliability, travel time, speed and information provision, parking provisions and fees.		

	Travel experiences among cyclists, which include the feeling of safe and being prioritized on the road, availability of the relevant information, route connectivity and the availability of bicycle parks at the destinations		
	Travel experiences among pedestrians, which include the quality and design of the pedestrian paths, feeling secure and safe while walking and the availability of relevant information		
Open Suggestions to improve travel experience			
			Gender, age, disability, household composition, income and education information
Special group needs including way-finding, accessibility, stress, travel information and lighting.			
Passenger satisfaction on: service availability (frequency and stop location), travel speed (both subjective and relative speeds), information at stations and on-board, information about ticketing, comfort (quality on on-board, fellow travelers, seat availability, seat comfort, easiness to buy ticket, crowding both at stops and on-board, station facilities), appeal (physical environment, vehicle quality, cleanliness both at stations and on-board), safety (at stops and on-board), overall reliability (including regularity and punctuality), personnel availability at stops and on-board, price (value-for-money and fairness), connectivity (network-wise and easy transfer), travel sickness, and environment issue.			

Different factors were valued differently by different classes of stakeholders. Operators were mostly interested and concerned about the impact of detailed level-of-service related variables on passenger experience (e.g. the use of travel information, time utilization whilst on-board, more detailed impacts of disruptions, detailed trip pattern), and less interested in the overall satisfaction of whole journey, questioning the value of quantifying the impact of past poor experience. Furthermore, they also showed limited interest in variables that could neither be used in understanding their customers' behavior, nor in detailed planning processes (e.g. the trip satisfaction can be improved in detailed level).

In contrast, planning authorities were more interested with wider general urban and public transport planning issues and the multi-modal travel patterns (e.g. different impacts of level-of-service for different travel modes and trip purposes). This is expected given their responsibility is to improve the transport service at the network and city (or even bigger) level. They were also interested in the impact of congestion and pollution in general and what particular conditions and locations public transport is needed from a planning perspective. The special interest groups were more interested in their constituents' interests and for them, the questionnaire lacked detail. Some argued that the questionnaire was not detailed enough to explore the disabled travelers' needs and considered that asking about personal information was insensitive.

The 'other' group of stakeholders was mainly government research institutes, many of whom were interested with more detailed trip patterns and behavioral variables that underlie the travelers' decision making processes in order to inform policy decisions. They were also interested in multidisciplinary issues such as the role of subjective well-being, stress and the impact of travelers' time constraints. Local stakeholders highlighted the benefits of having similar questions that are comparable with ones that they already use.

UK based results

Public transport users in the UK have the opportunity to comment on the quality of public transport regularly through market research and evaluation conducted by transport authorities (such as CENTRO and Transport for Greater Manchester), Passenger Focus (the independent passenger watchdog which conducts national surveys on buses, trains and trams), surveys conducted by transport operators such as Virgin, Arriva and National Express amongst others and major regional employers (such as Birmingham International Airport).

Although much information is collected on the passenger experience, there is still a shortage of information about why people do not use public transport and the quality of the whole journey itself. It is argued that the wider traveller experience, including intermodal travel, the experience of travelling to the transport node, and the journey from the transport to the final destination, or the next stage of the journey may influence and restrict use of public transport. Currently information tends not to be collected by the transport operator if it is about issues which are beyond their control about, for example interchanges, stations or bus stops. They may include questions on it (for completeness) and pass information on. Indeed, most transport operators were only interested in the perceived quality of their service. The municipalities and regional transport authorities were more concerned with the overall

quality of experience and the traveller's journey (for example the provision of way finding information, tactile pavements and dropped kerbs. Many respondents were starting to consider social media or apps as a way of collecting travel information, but were not sure how to use the information or how representative it was. Additionally certain groups may be underrepresented in the surveys such as women with young children and those who did not use public transport

Therefore there is a clear role for METPEX in investigating the quality of the whole journey experience as experienced by an individual traveller, paying particular attention to intermodality and those parts of the journey conducted by more active forms of transport. For the municipality and local transport authority, detailed local information is important in addressing need and planning new transport provision. Current surveys may not be able to capture the level of information needed to start planning for new models of mobility.

DEVELOPING THE TOOLSET

Deconstructing the journey (part 2)

For METPEX, the whole journey from inception to arrival is the main unit of interest. In order to measure this, an inclusive set of measurement instruments is being developed which can be used to assess a particular stage of a journey, or the whole journey, including modal transfers, for different types of users. This may include journeys made wholly or partly by active, public or private forms of transport. Providing inclusive measurement instruments, which can be used by a wide range of respondents (in particular those with disability classified as Vulnerable Road Users or those from lower socio economic groups), is key to evaluating transport service provision. Therefore the measurement instruments include CASIC (Computer Assisted Survey Information Collection (in the form of on line surveys using SurveyMonkey), PAPI (Pen and Paper) systems (e.g. semi structured interviews) and focus groups for targeted user groups. Additionally a game app and dynamic questionnaire linked to a satellite navigation system (sbNavi™) are being developed to enable the collection of real time data (Woodcock et al. in press).

In order to develop this suite of measurement instruments, the first year of the project was used to understand and consolidate previous research and research instruments, through desk based research and interviews outlined above. The subsequent analysis showed that over 400 variables would need to be considered if the whole journey was to be analysed. From this, an extensive pilot questionnaire was developed and administered to over 200 people from Lithuania, UK, Greece, Ireland, Portugal, Italy, Sweden and Romania in both on-line and pen and paper formats. This showed the operational complexity of such a survey and the burden completing such a question would place on travelers. The subsequent analysis showed the importance of some factors in determining overall satisfaction (e.g. the longest leg of the journey, overall satisfaction and mood) and the relationship between the variables.

Following the pilot study, the original categorization of variables into political, organizational, functional, environmental, technological, and social dimensions was removed (from the users perspective, although each item can be traced back to these dimensions). The variables were recategorised to make them more respondent and journey oriented, prioritized, filtered and mapped on to the proposed measurement instruments so that only the most useful and highly prioritized variables for each user and journey type will be presented to a potential respondent. In this way METPEX will meet one of its objectives, i.e. to support and encourage adjustability and adaptability according to the context of use, e.g. time period, targets' framework and resource limitations. Therefore, the full list of variables can be filtered so that only the most useful and preferred variables for each condition (e.g. user, mode of transport and journey stage) are included.

Initial filtering is being undertaken by colleagues across the project, and is focusing on the inclusion of variables that have been given a high priority by independent raters. At the end of the project, after the usefulness of the variables has been validated through our pilot study in 8 trial sites, guidelines will be provided to allow filtering by a researcher, who can use subsets of the 'full survey' to initiate targeted campaigns, or create a set of rules to generate appropriate research instruments.

CONCLUSIONS AND FUTURE WORK

The METPEX trials will be undertaken in 8 cities (Coventry, Dublin, Grevena, Vilnius, Stockholm, Valencia, Bucharest and Rome) in autumn 2014. The deconstruction of the journey has enabled an extensive set of variables to

be created from which a variety of measurement instruments can be generated, enabling real time and retrospective data gathering from different user groups, using different modes of transport. Prior to the trials each survey will be translated into the appropriate language, ethical and stakeholder approvals obtained, and Site Evaluation Assistants trained. The data from each site, and each measurement instrument will be collated on a central database enabling comparisons and benchmarking of services in different countries. This will also be used to establish which are the most reliable, valid and useful questions to include in the final toolset. The results will be presented back to the stakeholders in order to investigate how such information is used to inform policy and practice.

ACKNOWLEDGEMENTS

METPEX is funded under FP7-SST-2012-RTD-1, under the SST.2012.3.1-1. Research actions regarding the accessibility of transport systems work programme. The authors would like to acknowledge the contribution of all partners from Interactions (Ireland), SIGNOSIS (Belgium), ITENE (Spain), ZHAW (Switzerland), Euokleis s.r.l. (Italy), POLITO (Italy), ANGRE (Greece), KTH (Sweden), INTECO (Romania), FIA (Belgium), VTM (Portugal), SmartContinent (Lithuania), SBOING (Greece), TERO (Greece), RSM (Italy).

REFERENCES

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman. M., White, P. (2004), *The demand for public transport; A practical guide*. TRL Report, TRL 593, Crowthorne, UK

Cats, O., Susilo Y., Hrin, R., Woodcock, A., Diana, M., Speicyte, E., O'Connell, E., Dimajo, C., Tolio, V., Bellver, P., and Hoppe, M. (2014), An Integrated Approach to Measuring the Whole Journey Passenger Experience, *Transportation Research Arena*, Paris

Chartered Institute of Transport. 1998. *Passenger Interchanges - A Practical Way of Achieving Passenger Transport Integration*. Passenger Integration Working Party: CIT UK.

Chi, C. G.Qing, & Qu, H. (2008), Examining the structural relationships of destination image, tourist satisfaction and destination loyalty: An integrated approach. *Journal of Tourism Management*, 29: 624--636

Colin Buchanan and Partners (1998), *Transport Interchange - Best Practice Report*. Department of Environment, Transport and Regions: London.

Faber, O. (1996), *Ameliorating the Disbenefits of Rail Interchange*. A Report for Rail Operational Research.

Federal Transit Administration (FTA). 2003. Customer Preferences for Transit ATIS: Research Report. U.S. Department of Transportation, Battle Memorial Institute & Multisystems Inc., Washington Psychology, Bern, 4-7 September 2000.

Friman M., Fujii S., Ettema D., Gärling T. & Olsson L.E. (2011). Psychometric analysis of the satisfaction with travel scale. In: *Presentation at the 9th Biannual conference of Environmental Psychology in Eindhoven*, pp. 26-28, The Netherlands.

GMPTE (Greater Manchester Passenger Transport Executive). 1997. *Bus Station Audit*, GMPTE.

Grotenhuis, J.W.R., 2005. *Integrated multimodal travel information in public transport*. Final Thesis, Utrecht University, Utrecht

Hickman, R., Hamiduddin, I., Sellman, T. amd Hall, P. (2012), The seamless public transport journey: instrumental and affective, *RGS,* Edinburgh

Hine, J. and Scott, J. (2000), Seamless, accessible travel: users' views of the public transport journey and interchange. *Transport Policy* 7:217-226

London Councils, Living Streets and Walk London (n/d), *Breaking Down Barriers to Walking in London* http://www.londoncouncils.gov.uk/policylobbying/transport/cyclingwalking/breakingdownbarrierstowalking.htm

Lyons, G., Harman, R. (2002), The UK public transport industry and provision of multi-modal traveller information. *International Journal of Transport Management*, 1, 1–13.

Mackett, R.L., Achuthan, K. and Titheridge. H. (n/d), Overcoming the barriers to walking for those who are socially excluded, http://www.walk21.com/papers/Mackett,%20Roger-B-Overcoming%20the%20barriers%20to%20walking%20for%20peop.pdf

Murphy, P., Pritchard, M. P., & Smith, B. (2000), The destination product and its impact on traveller perceptions". *Tourism Management*, 21(1): 43-52

Rajesh, R. (2013), Impact of Tourist Perceptions, Destination Image and Tourist Satisfaction on Destination Loyalty: A Conceptual Model , *PASOS*, 11, 3, 67-78

Stradling, S.G., Hine, J. and Wardman, M. (2000) 'Physical, cognitive and affective effort in travel mode choices'. Symposium on Travel Mode Choice. *International Conference on Traffic and Transport*

Tovey, M (ed) (2012), *Design for Transport, A user centred approach to vehicle design and transport*, Gower, UK

Woodcock, A., Berkeley, N., Cats, O., Susilo, Y., Hrin, G.R., Reilly, O., Markucevičiūtė, I. and Pimentel, T. (2014), Measuring quality across the whole journey, *IEHF2014*, Southampton, UK, April 2014

Woodcock, A., Petridis, P., Liotopoulos, F., Georgiadis, A., Brady, L. (2014), The challenges of developing an on-line tool to measure the quality of the passenger experience in a PanEuropean context, *HCI 2014*, Cyprus, Greece

Woodcock, A. et al (2014) Measuring quality across the whole journey. *IEHF conference*, April 7-10, Southampton, UK

Comparative Analysis of Driving Maneuvers Using 3D Body Posture Data

Alexandra Kondyli [a], Virginia P. Sisiopiku [b], Liangke Zhao [c] and Angelos Barmpoutis [c]

[a] *Department of Civil and Coastal Engineering*
University of Florida
Gainesville, FL 32611-6580, USA

[b] *Department of Civil, Construction, and Environmental Engineering*
University of Alabama at Birmingham
Birmingham, AL 35294-4440, USA

[c] *Digital Worlds Institute*
University of Florida
Gainesville, FL 32611-5800, USA

ABSTRACT

Research has shown that driver inattention is the most prevalent cause of traffic collisions accounting for an estimated 25 to 56% of crashes in the US. Driver inattention may result from drivers engagement in secondary activities (such as texting or cell phone use), or lack of awareness of the surrounding environment. The main objective of this research is to investigate the relationship between potentially unsafe driving events and the actual driver body posture and movements when performing a driving maneuver under different traffic configurations. The paper presents results from a pilot study that captured the 3-D posture and activity of three drivers while performing both mandatory (merging) and discretionary (lane changing) maneuvers on freeway and arterial segments in Gainesville, Florida. The body posture of the drivers was captured through the use of a low-cost infrared depth sensor. A 7-point human skeletal model was fit to the captured depth frame sequences using our proposed framework. The comparative analysis of the participants' body movements while performing the maneuvers revealed differences between the participants' body activity when performing the same maneuvers. The findings of this research provide significant insights regarding which body movements may hide unsafe situations while performing a driving maneuver that requires the attention of the surrounding environment.

Keywords: Infrared depth sensors, driver behavior, merging process, lane changing, body posture

INTRODUCTION

Driver error constitutes a major cause of traffic crashes internationally (Peden et al., 2004). According to the World Health Organization (WHO), annually there are over 1.2 million fatalities and over 20 million serious injuries worldwide. In the US, the 100-car naturalistic study (Dingus et al., 2006) concluded that driver distraction is responsible for about 80 percent of crashes and 65 percent of near crashes. In an effort to assist drivers in the driving process and reduce the degree of uncertainty due to driver error, a number of advanced (or intelligent) driver assistance systems (ADAS) have been introduced (Shaout et al., 2011). Such systems are typically designed to

provide assistance or warnings to drivers by taking into account the position of the vehicle or other vehicle-related components. Examples of driver-assistance systems include lane departure warning, pedestrian protection, blind spot detection, collision avoidance, lane change assistance, and automatic parking.

Recognizing the importance of driver assistance with respect to traffic safety, the Federal Highway Administration (FHWA) and the US Department of Transportation (USDOT) introduced the connected-vehicle research program. This program offers a mixture of cutting edge technologies such as advanced wireless communications, on-board computer processing, advanced vehicle-sensors, GPS navigation, and smart infrastructure, to identify and warn the drivers on unsafe roadway conditions. The connected-vehicle program supports vehicle-to-vehicle (V2V) and vehicle-to-infrastructure communication research activities. The vehicle-to-vehicle communication is related to the exchange of data (e.g., speed, acceleration, heading angle, etc.) over wireless network that provide information on surrounding vehicles status and allows for performing calculations and issue driver warnings to avoid crashes. The communication option is based on Dedicated Short Range Communications (DSRC). Although the development of the communication component of this program is not complete to date, a number of crash avoidance systems (e.g., blind spot and lane changing warning, forward collision warning, etc.) have been established so far. In August 2012, DOT launched the Safety Pilot "model deployment" in collaboration with University of Michigan's Transportation Research Institute (UMTRI), where nearly 3,000 vehicles were deployed in the largest-ever road test of V2V technology. DOT testing is indicating interoperability of V2V technology among products from different vehicle manufacturers and suppliers and has demonstrated that they work in real-world environments (NHTSA, 2014).

It should be noted that currently available ADAS systems are designed to look at vehicle's external environment whereas systems that focus on the drivers' behaviors and actions inside the vehicle are still limited. However, the next generation of advanced driver assistance systems should also consider human factors issues and take into account interactions among drivers and vehicles, either in relation to their own vehicle or the surrounding vehicles. A multifaceted approach that looks at both the ergonomics aspect of human factors research as well as the psychological aspects of the engaged participants is expected to advance the state-of-the art of ADAS systems in the future.

Toward this direction, this paper investigated the relationship between potentially unsafe driving events and the actual driver body posture and movements when performing a driving maneuver (e.g., lane changing, merging) under different traffic configurations in a naturalistic setting. The ultimate objective of this research is to develop a framework for constructing an in-vehicle driver-assistance system that accounts for the driver's body posture and movements, rather than considering solely the vehicle position relative to other vehicles on the road.

LITERATURE REVIEW

A significant amount of research has focused on understanding driver intentions and actions (e.g., inattention or distraction states), using vision-based systems. Research studies primarily look into tracking of the head and the face of the driver. For example, Huang and Trivedi (2004) and Murphy-Chutorian and Trivedi (2010) developed a system that monitors and predicts drivers' head pose using video detection. Their head detection system was designed for identifying drivers' inattention and distraction, however, the authors did not specifically address how inattention was detected. Similarly, Braathen et al. (2001), developed an approach for identifying spontaneous facial expressions, such as blinking, to monitor alertness and anxiety. Huang et al. (2003) and Wu and Trivedi (2008) proposed a model that combines head pose detection with actual vehicle movement direction.

In addition, past research (e.g., Tijerina et al., 2005; Trivedi et al., 2007; McCall et al., 2005; Doshi et al., 2011) analyzed combination of head pose and gaze data for identifying and predicting driver's intent to change lanes and perform a maneuver. Research has also studied the position of hands and the grasp in conjunction with monitoring the head pose for lane change intent analysis and prediction (Cheng and Trivedi, 2010) or for driver distraction monitoring (Tran and Trivedi, 2009). A system that was developed to track the 3D body movement combined with head pose was also introduced in Tran and Trivedi (2010), where preliminary results of body posture and lane changing activity were collected in a simulation environment. Tran et al. (2012) used video-based analysis in conjunction with pedal sensor measurements and looked at drivers' foot behavior. They developed prediction models for braking and acceleration and concluded that the foot behavior depends greatly on the driver type.

A significant amount of research has dealt with predicting automobile driving posture for purposes of vehicle interior design. Reed et al. (2002) developed a model that predicts driver posture through a series of cascading regression functions. The final model is constructed to produce the best fit for the eye and the hip position using data from drivers with varying anthropometric characteristics. Fatollahzadeh (2006) proposed mathematical models that predicted truck drivers' comfortable sitting posture and position. Mathematical models using multiple regression analyses on selected body landmarks as well as anthropometrical measures were developed which proposed a linear correlation between parameters. Kovacevic et al. (2010) explored the impact of anthropometric measurements on ergonomic driver posture and concluded that anthropometric measurements of drivers and surrounding measurements of car controls affect traffic safety and driver's fatigue.

Overall, the literature review reveals that a significant amount of research has been involved with the development of advanced driver-assistance systems; however, most of these systems rely on the automobile position on the roadway and do not necessarily consider the drivers posture and actions. Apart from that, the lane trajectory and position of the vehicle could potentially differ from the driver's intent to change lanes. In addition safety research has focused on eye tracking as a means of capturing driver's attention, fatigue, or drowsiness; however, the entire body posture of drivers when performing a maneuver as well as different postures between various groups of drivers may also reveal behaviors that contribute to unsafe driving conditions.

METHODOLOGICAL FRAMEWORK

This research proposes a novel approach for studying the actual movements of drivers inside the vehicle, when performing specific maneuver types or while engaging to secondary tasks that require a certain body movement. With the use of a low-cost infrared depth sensor, the 3D shape of selected participants is being constructed, as they are performing various driving maneuvers and/or as engaged in secondary tasks while driving. A brief description of the methodology undertaken to collect and analyze the 3D data is presented here; however, more detailed information can be found in Kondyli et al. (2013). Each data frame captured by a digital depth sensor is a two dimensional array of depth values (i.e., distance between the sensor and objects). Similarly, a collection of frames is a three dimensional array that can be represented as $D \in R^{W \times H \times N}$, where N denotes the total number of recorded frames, and W and H denote the number of pixels across the width and height of the depth frame respectively. The depth value in a particular pixel with coordinates (i, j) on frame i is denoted by $D_{i,j,t} \in R^{+}$. In practice, each depth camera has a specific range of operation, which restricts accordingly the range of the recorded values (see depicted field of view in Figure 1). The depth frames can be equivalently expressed as quadratic meshes given by $X_{i,j,t} = (i - i_c) D_{i,j,t} f^{-1}$, $Y_{i,j,t} = (j - j_c) D_{i,j,t} f^{-1}$, $Z_{i,j,t} = D_{i,j,t}$, where (i_c, j_c) denote the coordinates of the central pixel in the depth frame, and *f* is the focal length of the depth camera. One of the advantages of the quadratic mesh representation of the depth frames is that they can be easily visualized using virtual lighting, shading, perspective and point of view using standard computer graphics techniques (Faugeras, 1993). An example of the quadratic mesh of a captured depth frame is shown in Figure 1 (left).

The 3D shape of the body of the driver and part of the vehicles' cabin are clearly captured in the depth frame. Optionally, the color information from a video frame can be applied as a texture to the quadratic mesh of the depth frame (Figure 2) and can also be used to enhance the detection of the body features.

The primary goal of our data processing method was to trace body features using the captured depth frame sequences. The body features of our interest included the *X, Y, Z* coordinates of the wrists, elbows, and shoulders as well as the orientation of the torso. The values of these quantities can be estimated by fitting a human skeletal model to each of the depth frames in our datasets. The main challenge in the skeletal fitting process is that the human body in our particular field of view is very close to other objects such as the driver's seat, the steering wheel and the driver's door. Any generic skeletal fitting algorithm performs better when the human body is clearly visible and at a distance from nearby objects, and therefore will fail in our in-cabin setting. For instance, the skeleton tracking algorithm included in the Microsoft Kinect Software Development Kit (SDK) fails in detecting the driver's body as it was previously reported in Kondyli et al. (2013).

Figure 1. Left: visualization of a depth frame. Right: The corresponding mask with enhanced boundaries between objects, computed using our framework

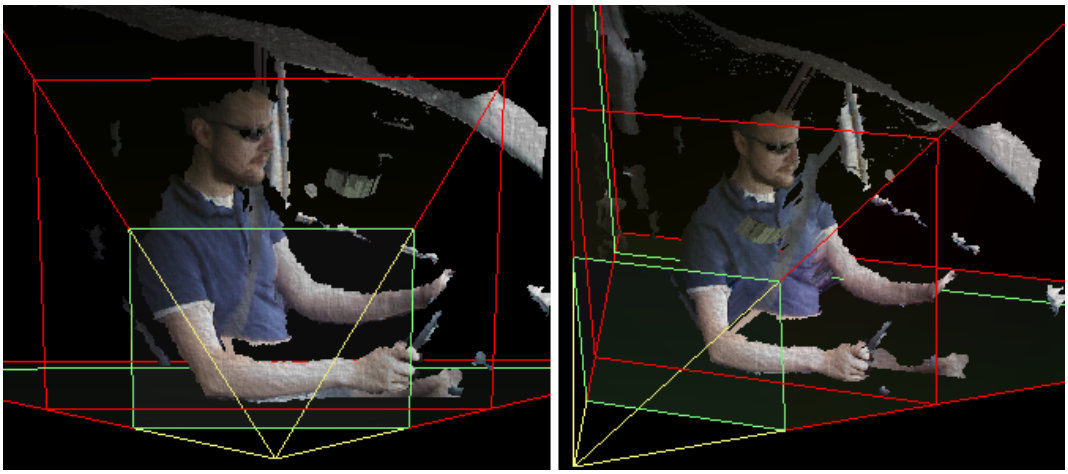


Figure 2. Two 3D views of the same frame from the recorded dataset. The video and depth frames are presented as a sequence of textured 3D frames. The field of view of the depth camera is also shown as a trapezoid.

In order to overcome the aforementioned skeleton fitting challenges we developed a novel algorithm that was designed to fit a 7-point skeletal model to the body of the driver using a sequence of depth frames. Our skeletal model included the line segments between the following joints: right wrist, right elbow, right shoulder, neck, left shoulder, left elbow, and left wrist. The skeletal model is visualized in Figures 5 and 6. In our visualization we also show the triangle formed by the left shoulder, the right shoulder and the neck, whose normal vector was used as an indicator of the torso orientation.

The proposed skeleton fitting algorithm scans the depth frames in a diagonal fashion from upper right to lower left, pixel stripe by pixel stripe until the entire image is covered. In each diagonal pixel stripe the medial points of the masked regions (see mask in Figure 1) are detected. The medial points that belong to the same region of the mask are connected so that they form medial line curves in all the masked regions. It should be noted that the medial points and curves are calculated in 3D and not in the 2D coordinates of the frames. After that the detected medial curves are filtered so that potential noise caused by the depth sensor is removed. Finally the curves that correspond to the arms are detected by using spatial constraints as well as geometrical constraints regarding the size, orientation and curvature of the arms. This process fits our 7-point skeletal model to the best matching medial curves. The proposed algorithm has linear complexity, which allow us to perform the fitting of the skeleton in real time in less than 15 milliseconds per depth frame in the computer configuration described in the next section.

DRIVER BEHAVIOR DATA COLLECTION

The field data obtained for this research were collected along a 2.6 mi stretch of Interstate 75 (I-75) in the southbound (SB) and northbound (NB) directions, and a 0.7 mile long arterial segment (Newberry Road eastbound and westbound approaches) in Gainesville, FL. The freeway segment along I-75 has three lanes per direction and the arterial segment has three through lanes per direction, several median openings, and includes a total of six signalized intersections. A schematic of the data collection sites is presented in Figure 3.

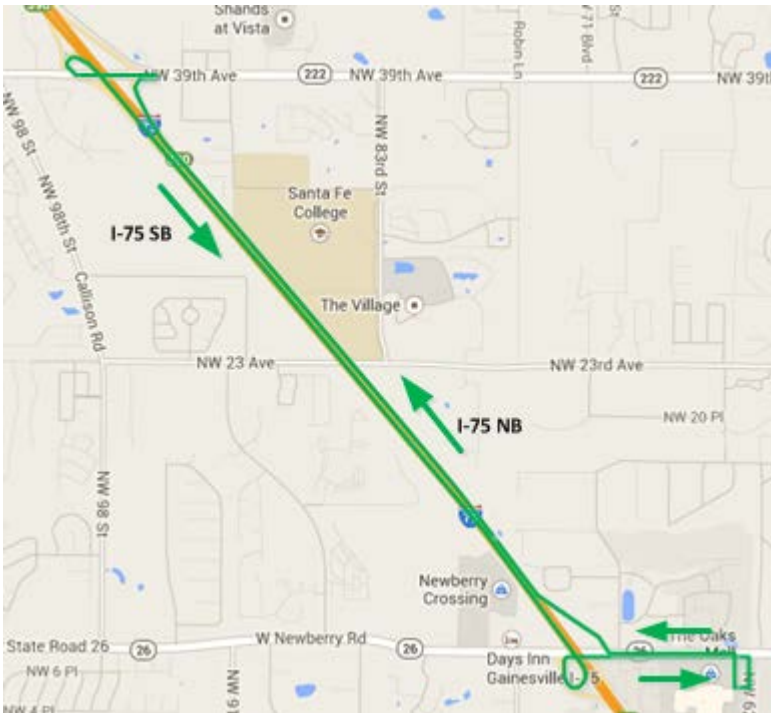


Figure 3. Map of the data collection route along I-75 and Newberry Road in Gainesville, FL.

The data collection effort took place on Sunday, September 1st 2013, between 10 am and noon. Traffic conditions were generally uncongested and free-flowing, especially on the freeway segment. Traffic on the arterial segment was light, although towards the end of the data collection effort the flows were considerably increased. For the purposes of this pilot study, three participants affiliated with the research team were asked to complete one route along the freeway and arterial segments. The participants performed two mandatory lane changes (i.e., merging onto the freeway) and several discretionary lane changes on the freeway and the arterial street. The entire duration of the experiment for each participant was approximately 20 minutes.

The real-time driver behavior data were acquired using the PrimeSenseTM depth sensor contained in the Microsoft Kinect™ sensor. The device was connected (via a USB 2.0 port) to a 64-bit computer with Intel Core i5 (quad core) CPU at 2.53GHz and 4GB RAM. The computer and the sensor were both powered using a 75 Watt car power inverter. The resolution of the depth camera was 320 × 240 pixels with horizontal field-of-view angle (FoV) angle of 57°. The resolution of the video camera was 640 × 480 pixels with horizontal FoV of 62°. The range of the camera was calibrated so that it records depth values in the range from 0.5m to 3.0m, which is suitable for the limited space of the cabin of a typical passenger vehicle. The sensor was fixed on the front passenger's door, so that the driver is within the field of view of the depth and video cameras. **Error! Reference source not found.** shows the field of view of the depth camera. The green rectangle depicts the closest plane of sensing, which is located 0.5m in front of the sensor (shown as the tip of the yellow pyramid in Figure 2).

DATA ANALYSIS

The video and depth sequences captured during this study were segmented into several fragments that correspond to the merging and lane changing maneuvers that were performed as participants were driving along the freeway and the arterial segment. Each of the fragments was analyzed independently using the framework presented earlier, and a comparative analysis was performed across the corresponding datasets from different participating drivers. The proposed framework was implemented in Java using the Java-for-Kinect library introduced by Barmpoutis (2013).

One characteristic of the Microsoft Kinect is that it has embedded an accelerometer that can be used to obtain the signature of the roadway segment that was used in this study. The output of the accelerometer is a unit vector, which typically indicates the direction of gravity in the case of a steady sensor. If the sensor moves due to the motion of the vehicle, the output of the accelerometer may be affected by the slope of the roadway as well as changes in the speed of the vehicle. Figure 4 shows the the X and Z coordinates of the accelerometer reading for the three drivers participated in this study. It should be noted that the Y dimension is perpendicular to the roadway and therefore is mainly affected by the vector of gravity and therefore it does not capture significant information about the signature of the road. Although the time stamp varies by driver due to differences in their driving speeds, the actual profiles of the accelerometer signature are very consistent among all three drivers. These signatures essentially describe changes in the elevation and slope of the roadway segment. For example, the first part of the graph clearly depicts the vertical alignment change due to vehicle driving from the arterial street on the on-ramp (upgrade) and onto the freeway (downgrade). Similarly, the last part of the graph shows the change in elevation due to the vehicle exiting the freeway through the off-ramp (downgrade). The information from the accelerometer is very useful for identifying the exact locations for obtaining the video and depth sequences of interest (e.g., merging onto the freeway).

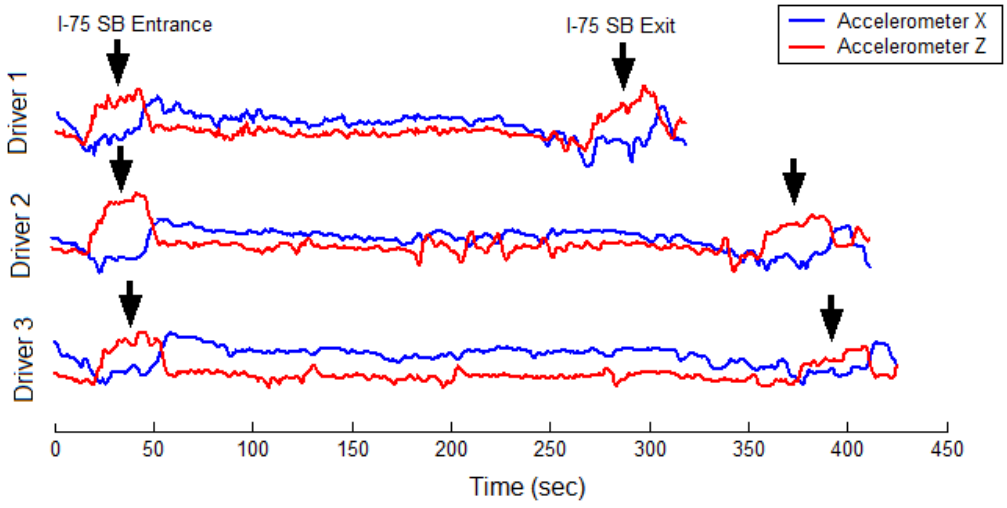


Figure 4. Microsoft Kinect accelerometer data obtained for three participants.

Using the framework presented earlier, this study examined differences in the body posture during a lane change maneuver for two of the three participants. Figure 5 shows the seven point skeleton model before and after a lane change maneuver for Driver 1 and Driver 3. The differences in the body posture between the two drivers are apparent from this figure. The torso of Driver 1 remains practically unaltered during the maneuver, whereas Driver 3 clearly shifts her body to the left in order to have a better visual of the traffic at the next lane. On the other hand, Driver 1 shifts only the head to identify potential conflict at the next lane through the rear mirror.

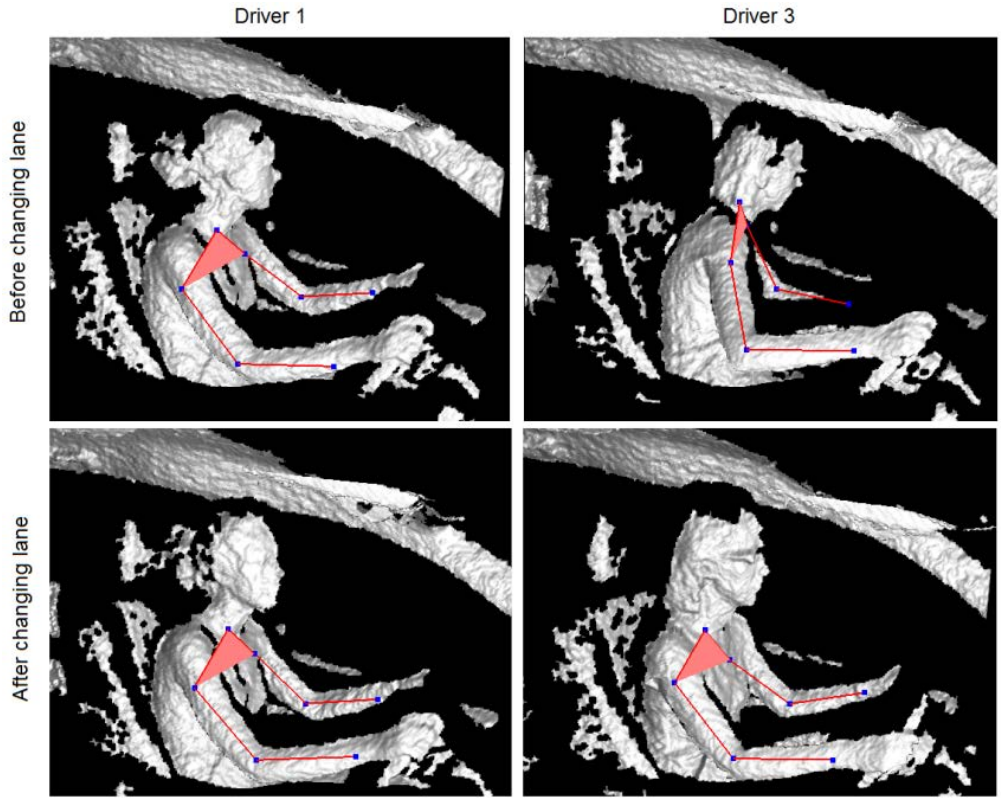


Figure 5. Change in body posture due to a lane change maneuver shown by the fitted skeletons in the depth frames.

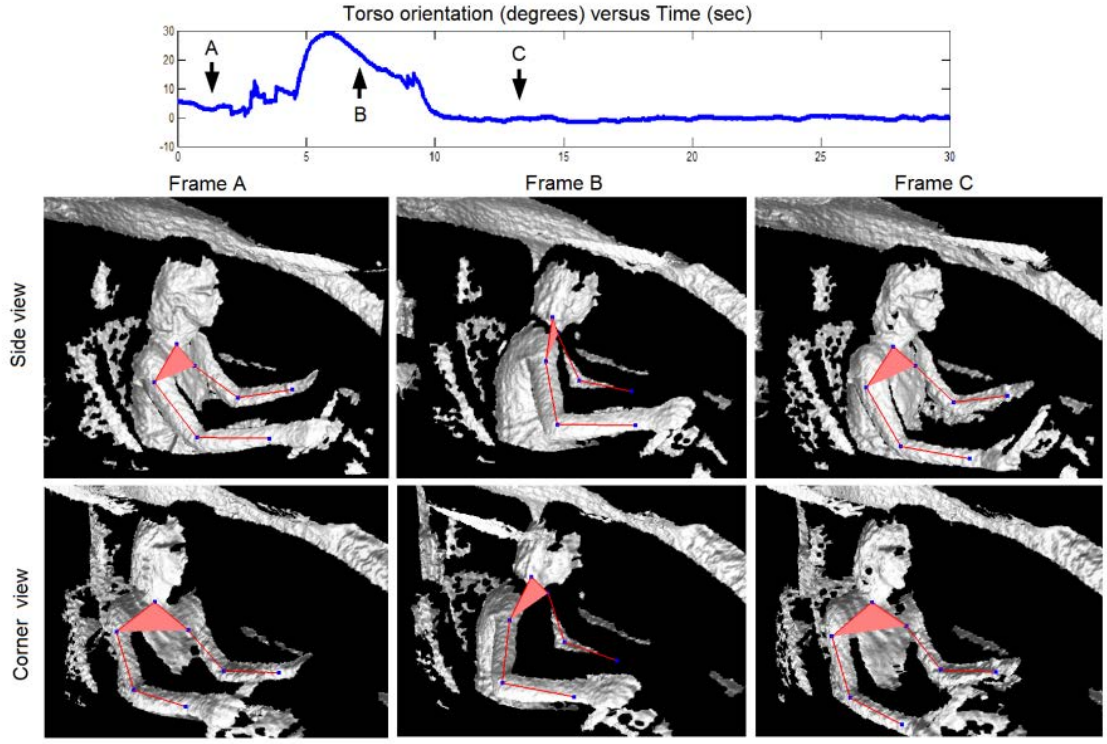


Figure 6. Change in body posture due to a merging maneuver for Driver 3. The plot shows the torso orientation during the merging maneuver. The depth frames and fitted skeletons of 3 frames are shown from two 3D perspectives.

In addition to the lane change maneuver, a comparative analysis of the body posture during a merging maneuver was also performed. Figure 6 presents the frame sequence during a merging maneuver for Driver 3, along with the corresponding time-series of the torso orientation. In this graph the torso orientation represents the rotation in degrees from the torso position perpendicular to the steering wheel. The orientation is positive for left-turn rotation and negative for right-turn torso rotation. Frames A, B and C are taken as before, during, and after the execution of the merging maneuver. From these graphs it is clear that the torso rotation of Driver 3 is considerably increased during the merging task. Driver 3 torso orientation during this merging maneuver is consistent with the lane changing example shown in Figure 5.

The same analysis was performed for the remaining two participants while merging on the freeway. The time-series plots of the torso orientation are presented in Figure *7*.

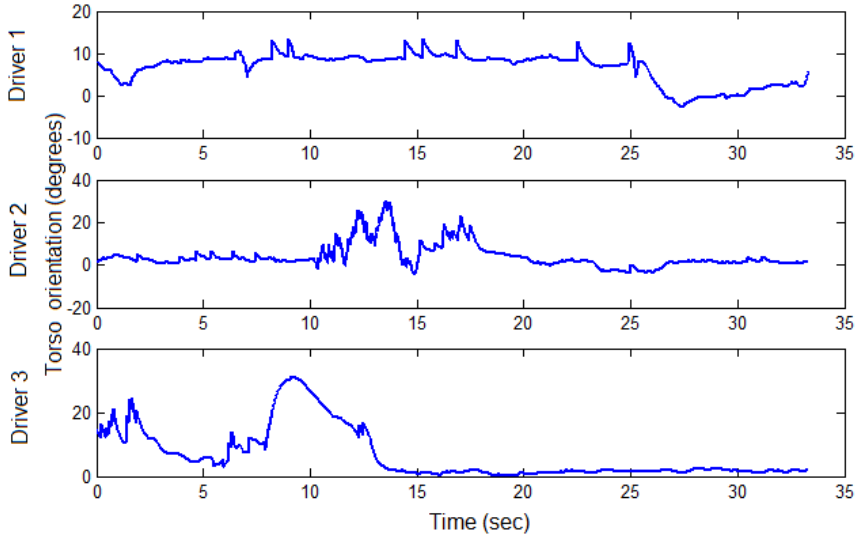


Figure 7. Comparison of the torso orientation during a merging maneuver at the same ramp merge junction between the three participants.

A comparative analysis of the time-series of the torsi orientation shows considerably different profiles of the torsi during the merging maneuver task. Comparable to Driver 3, Driver 2 also displayed increased torso rotation during the merging task; however, the profile of the torso rotation is different, as this is evident by the unsmooth time-series plot. Driver 1 torso orientation differs significantly from the other two. As shown in Figure *7*, Driver 1 does not seem to have an apparent torso rotation, but there is some rotation throughout the merging task (there is a small detected rotation at around 10 degrees for the first 25 seconds of the task).

Similarly to the torso orientation, a comparative analysis of other parts of the participants' body motion can be performed. Figure 8 shows the time-series of the X, Y, Z coordinates of the wrists, elbows, shoulders for Driver 3, during the entire duration of the driving task. Using the data shown in Figure 8 it is easy to obtain instances where there is significant body activity by identifying spikes in the respective graphs, and further analyze the underlying conditions for these instances.

By observing Figure 8 it is evident that there is a more frequent arm motion detected during the arterial segments compared to the freeway segments as we anticipated. For example in this dataset the driver started merging onto the freeway at 100 sec. and exited at 300 sec. which correspond to intense arm activity as indicated by a significant change to the coordinates of the wrists and the elbows. During the freeway segment (between 100 sec. and 300 sec.) there was no significant change of posture detected and the coordinates of the traced joints change only occasionally as it was also anticipated. This smooth driving pattern is significantly different compared to the one observed during the arterial segments which corresponds to 0 sec. - 100 sec., 300 sec. - 480 sec., and 650 sec. - 700 sec. During these segments the driver stopped at red traffic lights and followed a path that included many 90-degree turns as it is shown in Figure 3. All of these instances were naturally associated with arm activity, which corresponds to changes in the coordinates of the wrists slightly as it is shown in Figure 8. Finally, the segment from 480 sec. - 650 sec. corresponds to the northbound freeway segment, which was associated with occasional body motion according to

the plots in Figure 8.

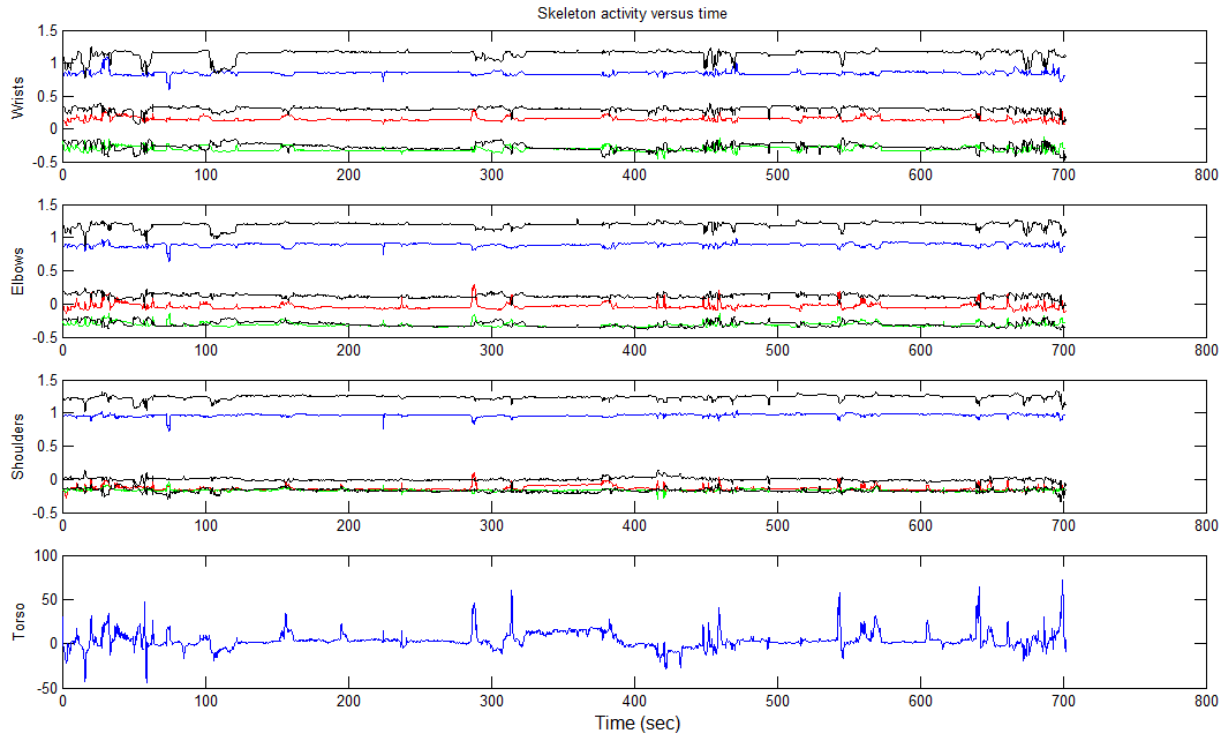


Figure 8. Skeleton activity versus time for Driver 3.

CONCLUSIONS AND RECOMMENDATIONS

In this paper a pilot study was conducted to capture the 3-D posture and activity of three drivers while performing various driving maneuvers. A low-cost infrared depth sensor was used to capture the 3-D posture. The participants were observed to perform merging and lane changing maneuvers on freeway and arterial segments in Gainesville, Florida. A comparative analysis of the participants' body movements while performing the maneuvers was conducted which revealed differences between the participants' body activity when performing the same maneuvers. More specifically, it was observed that there are considerable differences in terms of the torso activity for both lane changing and in merging maneuvers between different drivers. It is recommended to further examine the torso activity in conjunction with the activity of the shoulder, elbows and wrists, in order to obtain more complete representation of the entire body posture. It is further suggested to evaluate the effect of actual traffic conditions, such as proximity of the subject vehicle to the adjacent lead/lag vehicles, to the body posture while performing these maneuvers. The findings of this research can provide significant insights regarding which body movements may hide unsafe situations while performing a driving maneuver that requires the attention of the surrounding environment.

ACKNOWLEDGMENT

Funding for this study is provided by US DOT/RITA through STRIDE, the Alabama Department of Transportation (ALDOT) and the Florida Department of Transportation (FDOT). The support of the study sponsors is greatly appreciated.

REFERENCES

Barmpoutis, A. (2013), "Tensor Body: Real-time reconstruction of the human body and avatar synthesis from RGB-D," IEEE Transactions on Cybernetics, Volume 43, No. 5, pp. 1347–1356.

Braathen, B., Bartlett, M., Littewort-Ford, G., Movellan, J. (2001), "3-d head pose estimation from video by nonlinear stochastic particle filtering," in Proceedings of the 8th Joint Symposium on Neural Computation.

Cheng, S., Trivedi, M. (2010), "Vision-based infotainment user determination by hand recognition for driver assistance," IEEE Transactions on Intelligent Transportation Systems, Volume 11, No. 3, pp. 759–764.

Dingus, T. A., Klauer, S., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., Knipling, R. (2006), "The 100-car naturalistic driving study, phase ii results of the 100-car field experiment, report no. DOT HS 810 593," National Highway Traffic Safety Administration (NHTSA), Tech. Rep.

Doshi, A., Morris, B., Trivedi, M. (2011), "On-road prediction of driver's intent with multimodal sensory cues," Pervasive Computing, IEEE, Volume 10, No. 3, pp. 22–34.

Fatollahzadeh, K. (2006). "A laboratory vehicle mock-up research work on truck driver's seat position and posture", Doctoral thesis, Stockholm, Sweden.

Huang, K., Trivedi, M., Gandhi, T. (2003), "Driver's view and vehicle surround estimation using omnidirectional video stream," Proceedings in Intelligent Vehicles Symposium, IEEE, pp. 444–449.

Huang, K., Trivedi, M. (2004), "Robust real-time detection, tracking, and pose estimation of faces in video streams", Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004,Volume 3, pp. 965–968.

Kondyli, A., Sisiopiku, V., Barmpoutis, A. (2013), "A 3D experimental framework for exploring drivers' body activity using infrared depth sensors", IEEE International Conference on Connected Vehicles and Expo, Las Vegas, NV.

Kovacenic, S., Vucinic, J., Kirin, S., and Pejnovc, N. (2010) Impact of anthropometric measurements on ergonomic driver posture and safety, Period Biol, Volume 112 No 1, pp. 51-54.

McCall, J., Trivedi, M., Wipf, D., Rao, B. (2005), "Lane change intent analysis using robust operators and sparse bayesian learning," Computer Vision and Pattern Recognition - Workshops, CVPR Workshops, IEEE Computer Society, pp. 59–59.

Murphy-Chutorian, E., Trivedi, M. (2010), "Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness," IEEE Transactions on Intelligent Transportation Systems, Volume 11, No. 2, pp. 300–311.

NHTSA (2014), NHTSA Announcement available at http://www.safetypilot.us/images/documents/NHTSA_Announcement_2-3-14.pdf

Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A., Jarawan, E., Mathers, M. (2004), "World report on road traffic injury prevention," World Health Organization, Tech. Rep.

Reed, M. P., Manary, M. A., Flannagan, C. A. C., Schneider, L. W. (2002), "A Statistical Method for Predicting Automobile Driving Posture", Human Factors Volume 44 No. 4.

Shaout, A., Colella, D., Awad, S. (2011), "Advanced driver assistance systems - past, present and future," in Computer Engineering Conference (ICENCO), 2011 Seventh International, pp. 72–82.

Tijerina, L., Stoltzfus, D., Parmer, E. (2005), "Eye glance behavior of van and passenger car drivers during lane change decision phase," Transportation Research Record: Journal of the Transportation Research Board, Volume 1937, p. 3743.

Tran, C., Trivedi, M. (2009), "Driver assistance for "'keeping hands on the wheel and eyes on the road"',," IEEE International Conference on Vehicular Electronics and Safety (ICVES), pp. 97–101.

Tran, C., Trivedi, M. (2010), "Towards a vision-based system exploring 3d driver posture dynamics for driver assistance: Issues and possibilities," in Intelligent Vehicles Symposium (IV), 2010 IEEE, pp. 179–184.

Tran, C., Doshi, A., Trivedi, M. (2012), "Modeling and prediction of driver behavior by foot gesture analysis," Computer Vision and Image Understanding, Volume 116, pp. 435–445.

Trivedi, M., Gandhi, T., McCall, J. (2007), "Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety," Transactions on Intelligent Transportation Systems, IEEE, Volume 8, No. 1, pp. 108– 120.

Wu, J., Trivedi, M. (2008), "A two-stage head pose estimation framework and evaluation," Pattern Recognition, Volume 41, pp. 1138–1158.

Driving Behavior at Intersections Equipped with Red Light Cameras – A Monte Carlo Simulation Based on Driving Simulator Data

Cornu Joris, Brijs Kris, Daniels Stijn, Brijs Tom, Hermans Elke and Wets Geert

Transportation Research Institute – Hasselt University
Wetenschapspark 5, bus 6
BE-3590 Diepenbeek, Belgium

ABSTRACT

Background: Various impact evaluation studies of red light cameras (RLCs) showed an increase of rear-end collisions (up to 44%). **Objective:** To evaluate the driving and looking behavior of drivers at intersections equipped with RLCs in an urban area. **Method:** 63 participants passed an intersection without RLC, with RLC, and with RLC + warning sign (RLCWS) in a driving simulator. These data were used to estimate the risk of rear-end collisions by means of a Monte Carlo Simulation. **Results:** The relative risk of rear-end collisions for the intersection without RLC, with RLC, and with RLCWS was 1.97%, 12.65%, and 7.89% respectively. A higher percentage of participants observed the RLC when they stopped for the yellow light at the intersection with RLC, compared to the participants who did not stop (62% vs. 28%). **Conclusion:** Based on the Monte Carlo Simulation, risk of rear-end collisions increases when RLCs are installed. However, this risk decreases when a RLCWS is placed upstream. **Application:** Road administrations must consider the installation of RLCs very carefully. When RLCs are needed at certain intersections (e.g. to prevent red light running), road administrations should also deliberate about the potential of RLCWS in order to reduce the risk of rear-end collisions.

Keywords: Enforcement, Simulator study, Red light cameras, Eye-tracking

INTRODUCTION

Red light running is a worldwide road safety problem. In 2009, it was the main crash cause in 5.3% of the fatal accidents and 3.7% of the seriously injured accidents (SWOV, 2011). Red light running at signalized intersections has a significant impact on road safety since this leads to more serious accidents, being side collisions or collisions with vulnerable road users (Kloeden, Ponte, & McLean, 2001; R. A. Retting, Ulmer, & Williams, 1999; Shin & Washington, 2007; SWOV, 2011). Therefore, road authorities most of the time place red light cameras (RLCs) at signalized intersections to prevent red light running and improve road safety (De Pauw, Daniëls, Brijs, Hermans, & Wets, 2012; Martinez & Porter, 2006). However, rear-end collisions tend to occur more frequently at these intersections. This is often the result of a sudden braking maneuver of the leading vehicle, resulting in the fact that the following vehicle cannot stop in time (Shin & Washington, 2007).

Some suggested countermeasures to decrease red light running are: increasing the duration of the yellow sign (McGee et al., 2012; R. A. Retting, Ferguson, & Farmer, 2008; R. Retting & Greene, 1997), improving visibility of traffic lights, reducing clutter around intersections, installing RLC warning signs, reducing speed limits near intersections, and improving consecutive signal coordination (Martinez & Porter, 2006; Quiroga, Kraus, van Schalkwyk, & Bonneson, 2003; R. A. Retting, Williams, Preusser, & Weinstein, 1995). Furthermore, RLCs are an

effective tool to reduce red light running. Studies have shown that this reduction can be up to 44% (R. A. Retting, Ulmer, et al., 1999; R. A. Retting, Williams, Farmer, & Feldman, 1999). The effectiveness of RLCs is discussed in more detail below.

Urban areas are at greater risk for red light running crashes (De Pauw et al., 2012; R. A. Retting et al., 1995). Red light running even accounts for 22% of all urban crashes (R. A. Retting et al., 1995).

Effectiveness of red light cameras

In general, the effectiveness of RLCs appears to be studied less extensively than the effectiveness of speed cameras. The studies which have been carried out primarily focus on the effects of RLCs on red light running and collisions (i.e. rear-end and side) at intersections. Erke (2009) performed a meta-analysis about the effectiveness of RLCs. After installing a RLC, an average increase of 13% was found in the number of injury crashes. When making a distinction between the accident type, an increase of 43% and a decrease of 10% was observed in the number of rear-end crashes and side collisions respectively.

Shin and Washington (2007) investigated the effectiveness of RLCs on 14 locations in Arizona by means of the Empirical Bayes method. They found a decrease of 20-45% in the number of side collisions. On the other hand, the rear-end collisions significantly increased by 41%.

Another study (Høye, 2013) found a non-significant decrease of all injury crashes by 13%. The number of right angle side collisions decreased by 13%, but the rear-end collisions increased by 39%. Furthermore, the effectiveness of RLCs tended to be higher when warning signs for the RLCs were set up at main entrances to areas with RLC enforcement than when each intersection with a RLC was signposted separately.

De Pauw et al. (2012) found an increase of 5-9% of the total injury crashes after the installation of RLCs at intersections. However, the serious accidents (i.e. seriously injured or death) showed a significant decrease of 14-18%. The increase in the number of injury crashes can be mainly attributed to the increased number of rear-end collisions (44%). This number of rear-end collisions had a stronger rise at intersections in urban areas (+70%) than intersections outside the built-up area (+33%).

Persaud et al. (2005) conducted a before and after study with 132 locations equipped with RLCs. They found a significant decrease in the number of side collisions (16%) and increase for the rear-end collisions (24%). Thereby, they examined the spillover effects with a separate analysis for intersections where no RLC was installed. Spillover effects are those effects on the number of crashes at untreated intersections (i.e. no RLCs) that are close to an intersection with RLCs. There was a limited spillover effect (i.e. a slight decrease) for the number of right angle side collisions at intersections near intersections with RLCs. However, the number of rear-end collisions did not increase at the untreated intersections. Therefore, the authors concluded that the effect of RLCs is mainly restricted to the intersections equipped with RLCs (i.e. only a *local* effect).

Pulugurtha & Otturu (2014) also performed a before and after study to analyze the effectiveness of a RLC enforcement program. They indicated that RLCs lead to an increase in sideswipe and rear-end collisions at more than 50% of the signalized intersections.

A time series analyses showed that right angle collisions decreased by 46% at RLC intersections, but that there had also been an increase of 42% in rear-end crashes (Vanlaar, Robertson, & Marcoux, 2014). Results indicated that there were significantly fewer red light running violations after installing a RLC. Furthermore, RLCs had a protective effect on speeding behavior (also during green phases) at the intersections.

According to McCartt & Hu (2014) red light running declined significantly with 39% and 86% when traffic lights turned red for at least 0.5s and 1.5s respectively.

Porter et al. (2013) investigated the impact of a removal of the RLCs on red light running. The rate of red light running nearly tripled immediately after the removal of the RLCs and even quadrupled 1 year later. This means that there exists no time-halo effect concerning the use of RLCs.

There are significant variations with regard to the safety effects of RLCs between the different studies. Several factors that can play a role are: the visibility of the RLCs, the offense rate before placing the RLC, use of warning signs to announce the presence of a RLC, etc. (Erke, 2009). In conclusion, it can be said that many studies have

found a decrease in the number of side collisions after the RLC was installed. On the other hand, these studies also observed an increase in the number of rear-end collisions.

Dilemma zone

One of the main problems with signalized intersections is that drivers have to make a decision whether or not to stop at the yellow onset (Wilson, 2006; Yan, Radwan, Guo, & Richards, 2009; Zaal, 1994). This decision can be difficult, depending on the current speed and position of the vehicle. When the length of the yellow period is insufficient for the driver to stop comfortably, or to pass the stop line before the red phase has started, the driver is considered to be in the dilemma zone. The dilemma zone is a theoretical area of an intersection approach where a driver must take a decision (i.e. stop or go) when the traffic light is turned to yellow (McGee et al., 2012; Wilson, 2006; Yan et al., 2009). The actions of the driver (i.e. accelerate or brake) increase the potential for a dangerous situation (e.g. rear-end collision) to occur.

Especially when drivers approach a signalized intersection with a high speed, the dilemma zone problem results in that some drivers may stop abruptly while others may decide not to stop (or even accelerate). Such variation in driving behavior can lead to collisions (mainly rear-end collisions) on the intersection approach (Institute of Transportation Engineers (ITE), 2009; Yan et al., 2009)

The boundaries of the dilemma zone are approximately 2.0-5.5 seconds from the stop line (Bonneson, Middleton, Zimmerman, Charara, & Abbas, 2002; Federal Highway Administration, 2005; McGee et al., 2012).

Warning signs

RLC warning signs (RLCWSs) are in most cases set up to announce that the next intersection will be equipped with RLCs. Such warnings may have the potential to reduce the probabilities of collisions nearby intersections (Yan et al., 2009; Zaal, 1994). These warning signs can be placed at all RLC intersections or only at the main entrances to a specified area with multiple RLC intersections. If RLCWSs are only installed nearby these main entrances, spillover effects are more likely to occur because most drivers will not be aware of the exact locations of the RLCs (Høye, 2013).

Results indicated that the effects of RLCs are more favorable when not all RLC intersections are signposted separately (Høye, 2013). In general, drivers will have more respect for red lights when not all RLC intersections are signposted, which increases the favorable (i.e. prevention of red light running) and decreases the unfavorable (i.e. sharp braking maneuvers) effects of RLCs. These findings are supported by Zaal (1994). However, Zaal (1994) also indicated that generalized signposting can have a disadvantage. Generalized signposting may reduce the deterrent effect of site specific signposting, which possibly results in an increase of the number of accidents at potentially dangerous intersections.

Yan et al. (2009) investigated the effectiveness of a pavement marking with the word message “Signal Ahead”, which was placed upstream of signalized intersections. Results indicated that this marking had positive effects on the driving behavior at signalized intersections. The marking reduced the probabilities of both conservative-stop and risky-go decisions. Moreover, the marking contributed to a lower red light running rate and resulted in a lower deceleration rate for drivers who stopped at higher speed limit intersections.

Because rear-end collisions are the most common accident type at signalized intersections, Ni & Li (2013) studied the effectiveness of green signal countdown devices (GSCD). GSCD display the remaining seconds of the current signal status. These devices have been widely installed in Asia, but have both advantages and disadvantages. GSCD shortened the dilemma zone and reduced the number of rear-end collisions near the stop line during the yellow phase. On the other hand, these devices evoked risky car following behavior and resulted in higher rear-end collision probabilities during the flashing green phase (i.e. end of green phase). Therefore, Ni & Li (2013) recommended that GSCD should be installed cautiously.

OBJECTIVES

The primary objective of this study is to evaluate the driving and looking behavior of drivers at signalized

intersections equipped with RLCs in an urban area. For this purpose, 1 real world location was selected and rebuilt in the driving simulator at the Hasselt University's Transportation Research Institute. At this location, 2 conditions (i.e. RLC and RLCWS) and one control condition (i.e. no RLC) are examined. The driving parameters are used to estimate the relative risk of rear-end collisions under these 3 conditions. We address the following research questions:

1. Is there a difference in decision behavior (i.e. stop or go) between the 3 conditions?

2. Is there a difference in (relative) risk of rear-end collisions between the 3 conditions?

3. Is there a difference in looking behavior (i.e. regions of interest) between:
    a. the control and RLC condition?
    b. the RLC and RLCWS condition?

METHODOLOGY

Participants

Sixty-three volunteers (all gave informed consent) participated in the study. No outliers were identified based on the three interquartile distance criteria. Thus, the sample contained 63 participants (39 men), approximately equally divided over four age categories from 20 to 75 years old (mean age 46.2; SD age 18.1). All participants had at least two years of driving experience.

Driving simulator and eye tracker

The experiment was conducted on a medium-fidelity driving simulator (STISIM M400; Systems Technology Incorporated). It is a fixed-based (i.e. drivers do not get kinesthetic feedback) driving simulator with a force-feedback steering wheel, brake pedal, and accelerator. The simulation includes vehicle dynamics, visual/auditory (e.g. sound of traffic in the environment and of the participant's car) feedback and a performance measurement system. The visual virtual environment was presented on a large 180° field of view seamless curved screen, with rear view and side-view mirror images. Three projectors offer a resolution of 1024 x 768 pixels and a 60 Hz frame rate. Data were collected at frame rate.

The eye movements of the participants were recorded while driving through the scenario, making use of a camera-based eye tracking system (faceLAB 5 Seeing Machines). The recorded eye tracking data were analyzed with the EyeWorks software package.

Scenario

Road segment selection and description

The real world location was selected in consultation with the Flemish road authorities. At this intersection, the road authorities have recently (i.e. summer of 2013) installed a RLC. In this manner, it is possible to validate the driving simulator data with the real world data (i.e. data were recorded both before and after the installation of the RLC at this intersection). Moreover, this location is situated inside a built-up area, which is very interesting because most rear-end collisions appear to take place in urban areas (De Pauw et al., 2012; R. A. Retting et al., 1995). The road segment approaching the intersection has a speed limit of 50km/h (built-up area) with 2x1 lanes (cf. figure 1). The intersection is equipped with a 35m-long separate lane for the left turn traffic (cf. figure 2). Furthermore, an adjacent cycling path starts 50m before the intersection.

Road segment development

To rebuild the selected location in the driving simulator environment, a procedure called geo-specific database modeling (Yan, Abdel-Aty, Radwan, Wang, & Chilakapati, 2008) was adopted. This procedure consists of replicating a real world driving environment in a simulated virtual environment and is to be differentiated from

simulator research where often the driving scenarios are fictive. In order to reproduce the existing situation as realistic and detailed as possible, we made use of photographs, videos, detailed field measurements, AutoCAD drawings, and Google Street View. A picture of the real world environment and the simulated replica can be found in figures 1 and 2.

Scenario design

The overall scenario is a systematic combination of the real life replicated section with a set of 2-4km long filler pieces, differing from the analysis sections with respect to design, speed limit, and surrounding environment and meant to provide some variation throughout the scenario.

Figure 4 gives an overview of the intersection and the approaching segment, including the placing of the RLC and RLCWS. The analysis zone has a length of 500m, whereby the stop line at the intersection is set at the relative distance of 0m (cf. figure 4). The traffic lights are placed 5m beyond the stop line (i.e. down the road). Participants have always been confronted with a leading vehicle (at 65m) and a following vehicle (at 25m) when approaching the intersection. These vehicles did not influence the stop/go decision of the participants, since the distance headway was sufficiently large and the leading vehicle always drove through the green phase. The signal light turned from green to yellow when participants were 2.5s removed from the stop line (i.e. time headway of 2.5s). In this way, each participant had an equal time interval to react to the yellow onset. The yellow time interval remained for 3.8s. This is in line with the yellow time interval standards from the literature (Bonneson et al., 2002; Federal Highway Administration, 2005; Institute of Transportation Engineers (ITE), 2009; McGee et al., 2012).

All participants are exposed to 3 conditions (i.e. passed the intersection 3 times):

- Control condition: no RLC was installed
- RLC condition: a RLC was installed 15m before the stop line
- RLCWS condition: a RLC was installed 15m before the stop line and a RLCWS (cf. figure 3) was placed 50m before the stop line.

Procedure and design

Participants were asked for their voluntary cooperation and requested to fill out a form with some personal data (e.g. gender, driving experience, date of birth, etc.). After a general introduction, drivers acquainted themselves with the driving simulator by handling various traffic situations (e.g. highway, curves, traffic lights, urban and rural areas) during two practice trips of 4km each. Subsequently, the eye tracking equipment was calibrated. Then participants completed the experimental trip of 14.8km, resulting in a randomized within (3 conditions: control, RLC, RLCWS) subjects design. Subjects were asked to drive as they normally would do with their own car and apply the traffic laws as they would do (or would not do) in reality. A GPS voice instructed them during the trip.

Figure 1. Real world vs. simulator image at approaching road segment.

Figure 2. Real world vs. simulator image at intersection.

Figure 3. Red light camera warning sign (RLCWS).

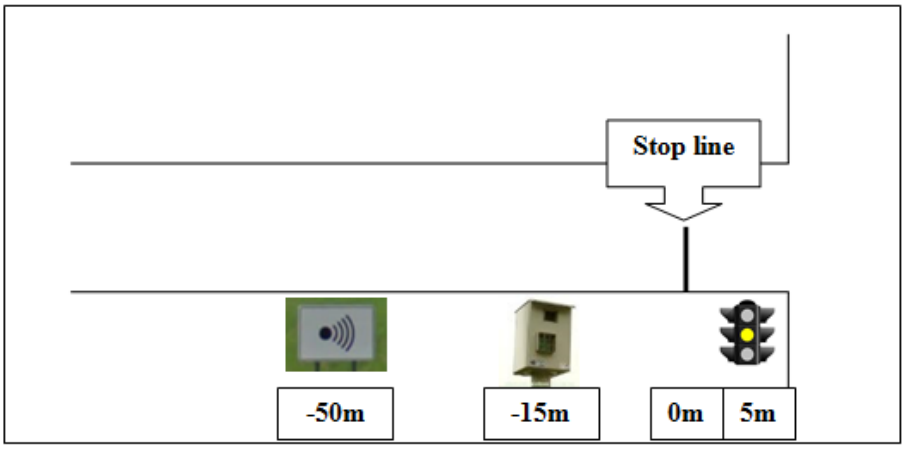


Figure 4. Scenario overview.

Data collection and analysis

Dependent measures

Driving performance measures for both longitudinal and lateral control were recorded by the simulator. For this study, measures for longitudinal control are particularly of interest. The speed [m/s] of the participants at the yellow onset is used in the Monte Carlo Simulation below. Furthermore, mean acceleration/deceleration (acc/dec) [m/s²] is also an important measure regarding the chance of rear-end collisions. Another dependent measure (i.e. distance headway) needed for the Monte Carlo Simulation was gathered at the real world location by means of observations.

Concerning the eye tracking data, the percentage of participants that fixated on the regions of interest (ROI) and the mean fixation duration are analyzed when the participants approached the intersections. Here we make a distinction between the participants who stopped for the yellow light (i.e. 'stop') and drove through (i.e. 'go').

Data analysis

The risk of rear-end collisions for each condition is estimated by means of a Monte Carlo Simulation using a normal distribution. These risk values are not expressed as absolute values, but as relative values (i.e. to compare the 3 conditions). The distance headway data observed at the real world location (i.e. field observations) are gathered by means of a (semi-automatic) conflict observation technique during 24 hours. For practical reasons, the data could only be collected for the control (n = 18) and RLC (n = 9) conditions. Therefore, the data from the RLC condition is also used for the RLCWS condition. Only situations where 2 vehicles (i.e. a following and a leading vehicle) were approaching the intersection while the traffic lights turned from green to yellow and the leading vehicle stopped were taken into account. To calculate the risk of rear-end collisions for each condition, the following parameters are used:

- Following vehicle
    - V_0: mean speed at the yellow onset (based on simulator data). Since the participants were the leading vehicle and we have no speed data for the following vehicle, the speed of the following vehicle will be equal to the speed of the leading vehicle.
    - a: maximum deceleration value (based on simulator data), assuming that the driver in the following vehicle can adjust his/her behavior based on the leading vehicle and even can make an emergency stop when necessary. Since the highest overall value is selected, the maximum deceleration rate will be constant for all conditions.
    - $t_{reaction}$: reaction time (based on the literature; (Caird, Chisholm, Edwards, & Creaser, 2007; Liu, Bonsall, & Young, 2003; Yan et al., 2008)), with respect to the decision of the leading vehicle to stop.
- Distance headway: distance between the end of the leading vehicle and the front of the following vehicle (observed at real world location).
- Leading vehicle
    - V_0: mean speed at the yellow onset (based on simulator data).
    - a: mean deceleration value (based on simulator data).

For the leading vehicle no reaction time was included in the Monte Carlo Simulation because the reaction time of the following vehicle was selected with respect to the stopping maneuver of the leading vehicle. The Monte Carlo Simulation was performed using Microsoft Excel with 100,000 iterations for each condition. The stopping distance is calculated for both the following and leading vehicle. A rear-end collision will occur when the sum of the stopping distance of the following vehicle and the distance headway is larger than the stopping distance of the leading vehicle.

For the looking behavior, several regions of interest (ROI) are selected: leading vehicle, traffic light, rear view mirror, speedometer (on screen, below the rear view mirror), RLC, and RLCWS. The mean fixation duration for these ROI are analyzed using the EyeWorks software package. Fixation durations of less than 0.05s are not taken into account. Subsequently, paired samples t-tests on a 5% confidence level were conducted using SPSS.

RESULTS

Decision behavior

First of all, we give an overview of the decision behavior (i.e. 'stop' or 'go') for the 3 conditions. In the control condition 7 (i.e. 11%) participants stopped for the yellow sign. For the RLC and RLCWS conditions the number of participants who did not drive through was 8 (i.e. 13%) and 19 (i.e. 30%), respectively. This means that most drivers drove through the yellow phase in the control condition (n = 56), followed by the conditions RLC (n = 55) and

RLCWS (n = 44).

Risk of rear-end collisions

Table 1 presents the parameter values (including SD) that were used for the Monte Carlo Simulation. The stopping distances for both the following and the leading vehicle were calculated. A rear-end collision occurred when the sum of the stopping distance of the following vehicle plus the distance headway was larger than the stopping distance of the leading vehicle. Given the fact that there were 100,000 iterations for each condition, the number of rear-end collisions that took place was 1,973; 12,646; and 7,984 for the control, RLC, and RLCWS condition respectively. This means that the relative risk for a rear-end collisions equals:

- 1.97% for the control condition;
- 12.65% for the RLC condition;
- 7.98% for the RLCWS condition.

Table 1: Parameter values (mean and SD) used for the Monte Carlo Simulation

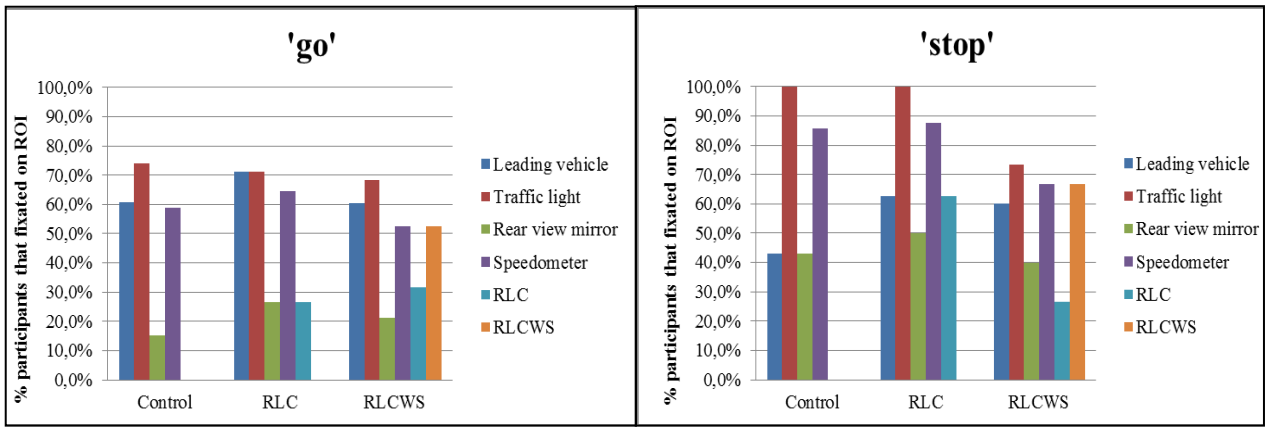
	Control condition	**RLC condition**	**RLCWS condition**
Following vehicle			
- V_0 [m/s]	12.69 (1.42)	12.29 (1.95)	11.03 (1.84)
- a [m/s²]	-7.14	-7.14	-7.14
- $t_{reaction}$ [s]	0.75 (0.25)	0.75 (0.25)	0.75 (0.25)
Distance headway [m]	19.81 (8.56; n = 18)	14.01 (5.51; n = 9)	14.01 (5.51; n = 9)
Leading vehicle			
- V_0 [m/s]	12.69 (1.42)	12.29 (1.95)	11.03 (1.84)
- a [m/s²]	-2.83 (1.42)	-4.28 (2.15)	-3.45 (2.36)

Looking behavior

Number of participants that fixated on ROI

Figure 5a shows that the number of participants that fixated on the ROI is approximately equal for the 3 conditions when participants drove through (i.e. 'go'). However, in the RLCWS condition participants fixated more on the RLC compared to the RLC condition. Remarkable is that 50% of the participants who did not stop at the intersection fixated on the RLCWS and that approximately 70% fixated on the traffic light (in each condition).

Figure 5b depicts the looking behavior of the participants who stopped. Here, 100% of the participants fixated on the traffic light in the control and RLC condition, compared to 72% in the RLCWS condition. Furthermore, more participants fixated on the RLC in the RLC condition than in the RLCWS condition. In the RLCWS condition almost 70% of the participants fixated on the warning sign.



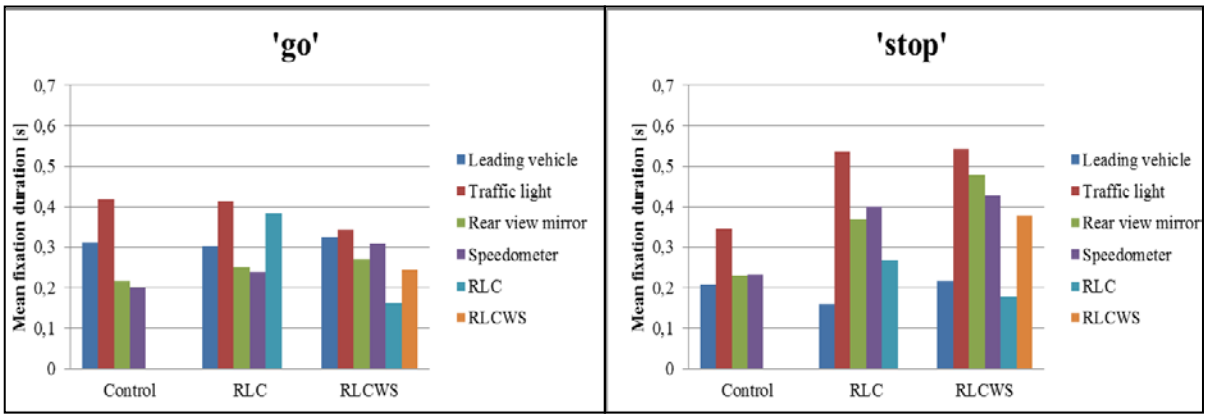
(a) **(b)**

Figure 5. Number of participants that fixated on ROI (a) 'go' and (b) 'stop'.

Mean fixation duration

The mean fixation duration for the participants who drove through and stopped are visualized in figures 6a and 6b, respectively. For the 'go' situation, no significant differences for the ROI between the conditions exist. The mean fixation duration for the RLC tends to differ between the RLC and RLCWS condition, but this difference is not significant ($t(13) = 1.47$; $p = 0.167$).

For the participants who stopped (cf. figure 6b) there seems to be a difference in mean fixation time for the traffic light, rear view mirror and speedometer between the control and RLC condition. Although, these differences are not statistically significant at a 5% confidence level. Between the conditions RLC and RLCWS also no statistically significant differences appear at a 5% confidence level. However, a significant difference in mean fixation time for the RLC exists between the RLC and RLCWS condition at a 10% confidence level ($t(4) = 2.46$; $p = 0.070$).



(a) **(b)**

Figure 6. Mean fixation duration for (a) 'go' and (b) 'stop'.

DISCUSSION

In this study, a real world intersection was selected and replicated in the driving simulator. The participants passed this intersection 3 times: control condition (i.e. no RLC), RLC condition, and RLCWS condition. The results

concerning the decision behavior (i.e. 'stop' or 'go'), the risk of rear-end collisions, and the looking behavior are discussed below.

Decision behavior

Results showed that 11%, 13%, and 30% of the participants decelerated and stopped at the intersection in the control, RLC, and RLCWS condition respectively. This means that slightly more participants stopped when a RLC was installed. When a warning sign (i.e. RLCWS) was placed at the roadside, it led to a further increase of the stopping maneuvers. These findings are in line with several studies (Høye, 2013; McCartt & Hu, 2014; Pulugurtha & Otturu, 2014; R. A. Retting, Williams, et al., 1999; Zaal, 1994) which found a significant decrease in the red light running behavior at intersections.

Risk of rear-end collisions

The relative risk of a rear-end collision is equal to 1.97%, 12.65%, and 7.89% in the control, RLC, and RLCWS condition respectively. This indicates that the presence of a RLC increases the risk of a rear-end collision. Several international studies support this increase (up to 44%) in rear-end collisions (De Pauw et al., 2012; Erke, 2009; Høye, 2013; Persaud et al., 2005; Pulugurtha & Otturu, 2014; Shin & Washington, 2007; Vanlaar et al., 2014). However, when a warning sign is placed in the approaching segment towards the intersection, this risk decreases but is still higher compared to situations where no RLC is present (i.e. control condition). Other studies (Høye, 2013; Ni & Li, 2013; Zaal, 1994) also found a lower risk of rear-end collisions when a warning sign was installed before (i.e. upstream) RLCs. In general, such RLCWS decreases the unfavorable (i.e. sharp braking maneuvers) effect of RLCs.

Concerning the parameter values used for the Monte Carlo Simulation, there are some interesting conclusions. Firstly, the mean driving speed at the yellow onset is the highest in the control condition (12.69m/s) and the lowest in the RLCWS condition (11.03m/s). Both values lie below the speed limit of 50km/h (i.e. 13.89m/s). The warning sign is probably responsible for this difference in speed, but no evidence is found in the literature. Secondly, the maximum deceleration rate from the simulator study is equal to -7.14m/s². Yan et al. (2008) observed a slightly higher value of -8.21m/s². Subsequently, mean deceleration values of -2.83m/s², -4.28m/s², and -3.45m/s² were found for the control, RLC, and RLCWS condition respectively. We can conclude that the deceleration value is the highest for the RLC condition, but decreases to a more 'normal' value in the RLCWS condition. A normal, comfortable braking deceleration value that is recommended is -3m/s² (Koppa, 2003; Liu et al., 2003; McGee et al., 2012; Yang, Han, & Cherry, 2013). Høye (2013) also found a smaller deceleration value when a warning sign was installed. Finally, the distance headway (observed at the real world location) for the control (19.81m) and RLC (14.01m) condition differs slightly. However, these values are based on a very small dataset and are both lower than the average distance headway (25-35m) found in the literature (Liu et al., 2003; Yan et al., 2008).

Looking behavior

To our knowledge, no (simulator) study investigated the looking behavior of drivers nearby intersections equipped with RLCs and RLCWSs to date. Concerning the looking behavior no statistically significant differences were found between the 3 conditions. However, some interesting conclusions regarding the looking behavior can be listed up:

- 'Only' 70% of the drivers who did not stop fixated on the traffic light. However, this does not mean that the other 30% of the drivers did not notice the traffic light (cfr. peripheral vision; (Dewar & Olson, 2007)).
- A higher percentage of the participants who stopped observed the RLC (62% vs. 28%) and the RLCWS (68% vs. 51%) compared to the participants who did not stop.
- In general, a higher percentage of the participants who stopped fixated on their rear view mirror and speedometer than the participants who did not stop.
- Mean fixation duration for both rear view mirror and speedometer was longer in the RLCWS condition compared to the RLC condition. On the other hand, mean fixation duration for the RLC was longer in the RLC condition than in the RLCWS condition. This can possibly be explained by the fact that participants who have already noticed the RLCWS, do not look at the RLC anymore.
- Mean fixation duration for both rear view mirror and speedometer was longer for the participants who stopped in comparison with the participants who did not stop.
- Participants who stopped had a longer mean fixation duration for the RLCWS compared to drivers who did not stop (0.38s vs. 0.23s).

LIMITATIONS AND FUTURE RESEARCH

The issue of external validity often is raised when discussing the results of research employing driving simulators. However, the geo-specific database modeling technique increases the reliability and validation of the experiment and the results (Yan et al., 2008). In addition, the simulator used in this study is equipped with a 180° field of view, which satisfies the prescribed minimum of 120° field of view for the correct estimation of longitudinal parameters (Kemeny & Panerai, 2003). Besides that, it would be possible to fill out the parameters for the following vehicle more accurately if the data observed at the real world location should be more elaborated.

Future research about the use of RLCs can be done concerning the spillover effects and the speeding behavior during green (i.e. speed on green). This study is mainly focused on the braking maneuvers at intersections. Furthermore, the impact of a varying distance between the RLC and the RLCWS could be investigated. The distance between the RLC and the RLCWS was held constant within this study (i.e. 35m). Finally, it should be interesting to carry out more research on the looking behavior at intersections equipped with RLCs and RLCWSs.

CONCLUSION AND POLICY RECOMMENDATIONS

Based on the Monte Carlo Simulation, the relative risk of rear-end collisions increases when a RLC is installed (12.65% compared to 1.97%). However, this risk decreases (7.89%) when a warning sign announcing the presence of a RLC is placed upstream. In conclusion, installing a RLCWS means that less drivers will look to the RLC, but more drivers will stop (safely) for the yellow light. Although RLCs lead to an increase in rear-end crashes, they can be useful to prevent red light running. This means that road administrations must consider the installation of a RLC very carefully. When RLCs are needed at certain intersections (e.g. to prevent red light running), road administrations should also deliberate about the potential of a RLCWS in order to reduce the risk of rear-end collisions.

REFERENCES

Bonneson, J., Middleton, D., Zimmerman, K., Charara, H., & Abbas, M. (2002). Intelligent Detection-Control System for Rural Signalized Intersections (No. FHWA/TX-03/4022-2). FHWA.

Caird, J. K., Chisholm, S. L., Edwards, C. J., & Creaser, J. I. (2007). The effect of yellow light onset time on older and younger drivers' perception response time (PRT) and intersection behavior. Transportation Research Part F: Traffic Psychology and Behaviour, 10(5), 383–396. doi:10.1016/j.trf.2007.03.002

De Pauw, E., Daniëls, S., Brijs, T., Hermans, E., & Wets, G. (2012). Effectevaluatie van snelheids- en roodlichtcamera's op gewestwegen in Vlaanderen (No. RA-2012-001). Diepenbeek, Belgium: Steunpunt Verkeersveiligheid.

Dewar, R., & Olson, P. (2007). Human factors in traffic safety. Tucson: Lawyers & Judges Publishing Company, Inc.

Erke, A. (2009). Red light for red-light cameras?: A meta-analysis of the effects of red-light cameras on crashes. Accident Analysis & Prevention, 41(5), 897–905. doi:10.1016/j.aap.2008.08.011

Federal Highway Administration. (2005). Traffic Control Systems Handbook. Washington, D.C.: FHWA.

Høye, A. (2013). Still red light for red light cameras? An update. Accident Analysis & Prevention, 55(0), 77–89. doi:10.1016/j.aap.2013.02.017

Institute of Transportation Engineers (ITE). (2009). Traffic Engineering Handbook (6th ed.). Washington, D.C.: ITE.

Kemeny, A., & Panerai, F. (2003). Evaluating perception in driving simulation experiments. Trends in Cognitive Sciences, 7(1), 31–37. doi:10.1016/S1364-6613(02)00011-6

Kloeden, C. N., Ponte, G., & McLean, A. J. (2001). Travelling speed and the risk of crash involvement on rural roads (No. CR 204). South Australia: Adelaide University - Road Accident Research Unit.

Koppa, R. J. (2003). "Human Factors" Chapter 3 of: Gartner, N.H., Messer, C.J. and Rathi, A.K. (Eds) Traffic Flow Theory. Washington, D.C.: Transportation Research Board Monograph, National Research Council.

Liu, R., Bonsall, P., & Young, W. (2003). Modelling safety related driving behaviour: impact of parameters values. Association for European Transport.

Martinez, K. L. H., & Porter, B. E. (2006). Characterizing red light runners following implementation of a photo enforcement program. Accident Analysis & Prevention, 38(5), 862–870. doi:10.1016/j.aap.2006.02.011

McCartt, A. T., & Hu, W. (2014). Effects of red light camera enforcement on red light violations in Arlington County, Virginia.

Journal of Safety Research, 48(0), 57–62. doi:10.1016/j.jsr.2013.12.001
McGee, H., Moriarty, K., Eccles, K., Liu, M., Gates, T., & Retting, R. (2012). Guidelines for Timing Yellow and All-Red Intervals at Signalized Intersections (No. 731). Washington, D.C.: Transportation Research Board.
Ni, Y., & Li, K. (2013). Estimating Rear-end Accident Probabilities at Signalized Intersections: A Comparison Study of Intersections With and Without Green Signal Countdown Devices (Accepted Manuscript). Traffic Injury Prevention. doi:10.1080/15389588.2013.845752
Persaud, B., Council, F., Lyon, C., Eccles, K., & Griffith, M. (2005). Multijurisdictional Safety Evaluation of Red Light Cameras. Transportation Research Record: Journal of the Transportation Research Board, 1922, 29–37. doi:10.3141/1922-05
Porter, B. E., Johnson, K. L., & Bland, J. F. (2013). Turning off the cameras: Red light running characteristics and rates after photo enforcement legislation expired. Accident Analysis & Prevention, 50(0), 1104–1111. doi:10.1016/j.aap.2012.08.017
Pulugurtha, S. S., & Otturu, R. (2014). Effectiveness of red light running camera enforcement program in reducing crashes: Evaluation using "before the installation", "after the installation", and "after the termination" data. Accident Analysis & Prevention, 64(0), 9–17. doi:10.1016/j.aap.2013.10.035
Quiroga, C., Kraus, E., van Schalkwyk, I., & Bonneson, J. (2003). Red light running: a policy review (No. CTS-02/150206-1). Texas: Texas Transportation Institute.
Retting, R. A., Ferguson, S. A., & Farmer, C. M. (2008). Reducing red light running through longer yellow signal timing and red light camera enforcement: Results of a field investigation. Accident Analysis & Prevention, 40(1), 327–333. doi:10.1016/j.aap.2007.06.011
Retting, R. A., Ulmer, R. G., & Williams, A. F. (1999). Prevalence and characteristics of red light running crashes in the United States. Accident Analysis & Prevention, 31(6), 687–694. doi:10.1016/S0001-4575(99)00029-9
Retting, R. A., Williams, A. F., Farmer, C. M., & Feldman, A. F. (1999). Evaluation of red light camera enforcement in Oxnard, California. Accident Analysis & Prevention, 31(3), 169–174. doi:10.1016/S0001-4575(98)00059-1
Retting, R. A., Williams, A. F., Preusser, D. F., & Weinstein, H. B. (1995). Classifying urban crashes for countermeasure development. Accident Analysis & Prevention, 27(3), 283–294. doi:10.1016/0001-4575(94)00068-W
Retting, R., & Greene, M. (1997). Influence of Traffic Signal Timing on Red-Light Running and Potential Vehicle Conflicts at Urban Intersections. Transportation Research Record: Journal of the Transportation Research Board, 1595, 1–7. doi:10.3141/1595-01
Shin, K., & Washington, S. (2007). The impact of red light cameras on safety in Arizona. Accident Analysis & Prevention, 39(6), 1212–1221. doi:10.1016/j.aap.2007.03.010
SWOV. (2011). Fact sheet - Effects of police enforcement of safety devices, of moped helmet use, and of red light running. Leidschendam: SWOV.
Vanlaar, W., Robertson, R., & Marcoux, K. (2014). An evaluation of Winnipeg's photo enforcement safety program: Results of time series analyses and an intersection camera experiment. Accident Analysis & Prevention, 62(0), 238–247. doi:10.1016/j.aap.2013.09.023
Wilson, A. (2006). Handboek verkeerslichtenregelingen. Ede: CROW.
Yan, X., Abdel-Aty, M., Radwan, E., Wang, X., & Chilakapati, P. (2008). Validating a driving simulator using surrogate safety measures. Accident Analysis & Prevention, 40(1), 274–288. doi:10.1016/j.aap.2007.06.007
Yan, X., Radwan, E., Guo, D., & Richards, S. (2009). Impact of "Signal Ahead" pavement marking on driver behavior at signalized intersections. Transportation Research Part F: Traffic Psychology and Behaviour, 12(1), 50–67. doi:10.1016/j.trf.2008.07.002
Yang, Q., Han, L. D., & Cherry, C. R. (2013). Some measures for sustaining red-light camera programs and their negative impacts. Transport Policy, 29(0), 192–198. doi:10.1016/j.tranpol.2013.06.006
Zaal, D. (1994). Traffic Law Enforcement: A review of the literature. Monash University.

Electric Vehicle Drivers Overcome Range Anxiety Prior to Their First Trip

Sarah Mansbridge, Mark Burgess and Margaret Harris

Psychology Department
Oxford Brookes University
Oxford, OX3 0BP, United Kingdom

ABSTRACT

Consumers' concerns over the drive range of an electric vehicle (EV) are widely acknowledged as the key reason for the non-acceptance of EVs. Yet to-date, the underpinnings of the phenomenon 'range anxiety' remain under researched and poorly understood. Pre-trial interview data were analysed from drivers in the United Kingdom BMW MINI E trial in order to explore the psychosocial factors drivers experience vis-à-vis range and to understand how drivers perceive and mentally construct range. Results revealed a four-stage process of cognition, whereby drivers' *preconceptions* of the EV range precipitated *anticipatory concerns* regarding running out of charge, leading to one of four *strategies* that directly guided their *intended behaviour,* in a manner that enabled them to avoid experiencing range anxiety. The results are discussed in relation to challenging drivers' pre-trial mind-sets, in order to change the way in which the range is perceived.

Keywords: Electric Vehicles, Range, Preconceptions, Intended Behaviour

INTRODUCTION

A key feature of the internal combustion engine (ICE) vehicle is the notion of mobility and the freedom and independence this encapsulates (Jensen, 1999). Limitations of the approximate 100 mile battery capacity of the modern day electric vehicle (EV) poses obvious challenges to both usability and mobility. Therefore, it is not surprising that concerns regarding the EV drive range are widely acknowledged as the key reason for the non-acceptance of EVs (Beggs, Cardell, & Hausman, 1981; Bunch, Bradley, Golob, Kitamura, & Occhiuzzo, 1993; Carley, Krause, Lane, & Graham, 2013; Chéron, & Zins, 1997; Dagsvik, Wennemo, Wetterwald, & Aaberge, 2002).

Since first being reported in the press in 1997, concerns over the EV range have escalated into the phenomenon now commonly known as 'range anxiety' (Neilsson, 2011). Range anxiety is defined as an individual's 'concern of not reaching their destination while travelling in an EV' (Neilsson, 2011). It must be noted that range anxiety not only relates to the limited battery range, but also to the difficulties of recharging (i.e., limited availability of public charging points and relatively lengthy charge times, that render an EV unusable for the period of recharging).

Despite perceptions of EV range being reported as the main barrier to the successful uptake of EVs, what is notably lacking is a deep understanding of range anxiety from a psychosocial perspective. Indeed to-date, the underpinnings of 'range anxiety' remain poorly understood. Previous research has largely ignored perceptions of the range and focused instead on the battery capacity required to make the EV range acceptable to the consumer. This research (taken from both non-EV and EV users) shows the 100+ mile EV range to be sufficient to satisfy the mobility requirements of a large market segment (Carroll, Walsh, Burgess, Harris, Mansbridge, King, & Bunce, 2013; Pearre,

Kempton, Guensler and Elango, 2011;Turrentine, Garas, Lentz, & Woodjack, 2011). However, despite the reported goodness of fit between mobility needs and resources, this research also demonstrates the existence of a 'range paradox' or gap between what consumers consider as being adequate and acceptable drive ranges for an EV (Carroll et al., 2013; Cocron, Bühler, Neumann, Franke, Krems, Schwalm, & Keinath, 2011; Eggers, & Eggers, 2011; Franke, & Krems, 2013a; Franke, Cocron, Buhler, Neumann, & Krems, 2012; Neumann, Cocron, Franke, & Krems, 2010;Tamor, Gearhart, & Soto, 2013; Vilimek, Keinath, & Schwalm, 2012). Thus, this research has largely found that drivers state they require the battery capacity to increase to approximately 200 – 300 miles before they would be willing to invest in an EV (Carroll et al., 2013; Daziano, & Chiew, 2013; Kurani, Turrentine, & Sperling, 1996).

Not only has previous research found 'range anxiety' to be detrimental to the general acceptance of EVs, it also shows that range has a lasting negative affect on drivers' interactions with the energy resources of the battery. Drivers' concerns regarding the range cause them to be overcautious in their judgements of whether the EV would successfully reach their destination (Carroll, & Walsh, 2010). Thus, a common finding from this literature is that drivers are failing to utilize the full amount of battery power available to them (Carroll et al., 2013; Cocron et al., 2011; Franke & Krems, 2013b; Franke et al., 2012). This suggests that direct experience of the EV range alone is not enough to overcome range as a psychological barrier in the mind of the EV user.

The above research principally concludes that an increased battery capacity (with associated increases in battery cost and size), growth of public charging points and faster charging times are required for EVs to appeal to the mass consumer market. However, this research (and the recommendations it proposes) masks, rather than addresses, the underlying psychological issues that result in range anxiety. Therefore, it is important to obtain a deeper understanding of the underpinnings of range anxiety from a psychosocial perspective, as this is undoubtedly delaying the adoption of EVs by the mainstream consumer market. Evidently the quest to determine how to make the EV drive range acceptable has presently ignored the perceived personal relevance of range from the consumers' perspective. In other words it is important to understand consumers' perceptions of the EV range, as this ultimately guides their emotional reactions and cognitive structures (mental representations), and directly affects their overt behaviour with the EV.

This is a particularly salient issue in relation to the EV drive range, where individuals have to psychologically adjust to having a limited range vehicle of approximately 100 miles (in comparison to the +400 mile range of the ICE vehicle), a charge time of 4 – 10 hours (in comparison to a refuelling time of 5 – 10 minutes) and a public charging infrastructure in the early stages of development.

Aim of study

The present study aims to explore the psychosocial factors drivers experience vis-à-vis range in order to understand how drivers perceive and mentally construct the EV range, in order to understand range as a psychological barrier.

METHOD

Background to present study

Participants in this study were enrolled in a field trial that was both part of the wider International BMW MINI E field trial (involving trials in the United States, Germany, United Kingdom, France, Japan and China) and was also part of the United Kingdom Technology Strategy Board Ultra-Low Carbon Vehicle Demonstrator Programme (involving low-emission vehicles from 18 manufacturers). These programmes supported research, development and assessment of EV technologies, with an aim of providing a sustained real-world test of everyday usage. Throughout the trials, analytic research data were collected from both the cars (through on-board data loggers) and the drivers (through questionnaires and interviews). The focus of analyses for this study is the drivers' pre-trial interview data.

The UK MINI E Trial took place in the South-East region of the country between December 2009 and March 2011. The study was divided into two separate phases lasting 6-months (phase one from December 2009 to June 2010 and phase two from September 2010 to March 2011). The MINI E trials incorporated drivers in both a private and fleet setting. Twenty electric MINIs were available for members of the public to lease for the 6-month trial period. A

further 20 MINI E vehicles were distributed among members of the public sector organizations that were part of the MINI E consortium. The public sector organizations that formed the MINI E consortium included the South East England Development Agency (SEEDA), Scottish and Southern Energy (SSE), Oxford City Council and Oxfordshire County Council and enabled the testing of the MINI E within a fleet environment.

Participants

Participants were 30 drivers of the BMW MINI E (20 private users; 10 corporate users) in the TSB ULCV Demonstrator Programme Trial. There were 21 men and 9 women of 21 – 62 years of age (Mean = 44.73). All but one of these participants had access to at least one ICE vehicle for the duration of their trial, and none had had any prior experience with an EV. Private drivers applied to participate in the trial either through online or newspaper applications. To be eligible to be included in the trial, applicants had to (1) pay a monthly lease fee of £330, (2) reside within a geographical triangle running from Andover, Oxford and West London, (3) have a private garage or parking space that would enable the installation of a 240V charging unit, and (4) indicate that they drove 300 miles per month. Successful applicants had their MINI E for a period of six months.

Corporate users were recruited through one of the aforementioned public sector organizations. Corporate users had sole access to an electric BMW MINI E for the duration of the trial and had to fulfill the same criteria as private users (with the exception of being exempt from having to pay the monthly leasing fee).

Drivers' average journey length

Prior to EV exposure drivers were asked what their average journey length was in a typical week. 16 of the 30 drivers (14 private users; 2 corporate users) reported their average journey length as not exceeding 20 miles (round trip). The remaining 14 drivers (6 private users; 8 corporate users) reported having longer average journey lengths of 50 miles or over (round trip). Note these drivers either had a routine daily round trip of at least 50 miles, or regularly had to undertake business trips greater than 50 miles (i.e., corporate users who were required to cover a relatively large geographical area as part of their work remit).

Characteristics of the MINI E

The MINI E was a two-seat conversion of a MINI Cooper with a range of around 100 miles, depending upon driving style and driving conditions (Vilimek et al., 2012). It was powered by a 204hp electric motor and a 35kWh Lithium-Ion battery. The battery took 4 hours to charge at 32 Amps and 10 hours to charge at 13 Amps.

Procedure and interview schedule

During the trial, participants were interviewed at (1) pre-trial, (2) immediately after they had first driven the car at a technical instruction event, and (3) 1 week after integrating the MINI E into their everyday life. Longer term adaptation was assessed through interviews at (4) 3 months into the trial, and (5) at the end of the trial. Here we focus on the pre-trial interview data. These interviews were one-to-one and took 45-90 minutes to complete. The interview schedule dealt with all aspects of the drivers' expectations and experiences. For the current study we were interested in drivers' reports of their expectations and experiences. For the current study we were interested in drivers' reports of their preconceptions of the drive range and their expectations of adapting to the drive range. Although participants may have given such information spontaneously (and this information was included in our analysis), interview questions were also designed specifically to address these issues (e.g. *How will you deal with the range?*). The interviewers prompted participants to give concrete examples in response to each question. Also, in accordance with Wengraf's (2001) suggestions for optimal narrative-based interviewing, the interviewer encouraged each participant to elaborate on their initial response with follow-up narrative-inducing probes. These probes were extemporized and were rooted in statements which the participant had made that the interviewer wanted to know more about (e.g. *You said you think you will have to preplan your journeys more with the MINI E. Can you please tell me more about that?*). This technique leads to additional narrative data and allows the interviewer to explore emerging areas of interest in 'real time' (Smith & Eatough, 2007) while also maintaining a sense of participant agency (Parker, 2005) and recognizing the participant as an 'experiential expert' who is capable of steering the direction of the interview through accounts of their personal experiences (Smith & Osborn, 2008).

Data analysis

The 30 participant transcripts were analyzed according to the Thematic Analysis (TA) guidelines provided by Braun and Clarke (2006). The initial stages involved reading and re-reading interview transcripts to become familiar with the breadth and depth of information pertaining to drivers' comments regarding range. The analysis was conducted in an inductive 'bottom-up' way with earl readings aiming to identify patterns and meanings in the data without drawing upon pre-existing frameworks for coding or theory development. This ideographic emphasis ensured that eventual themes were grounded in the participants' data and not driven by researchers' preconceptions.

RESULTS

Overview of results

Results revealed a four-stage process of cognition whereby drivers' *preconceptions* of the limited range precipitated *anticipatory concerns* regarding running out of charge, leading to one of four strategies that directly guided their *intended behavior* in a manner that enabled them to avoid experiencing range anxiety.

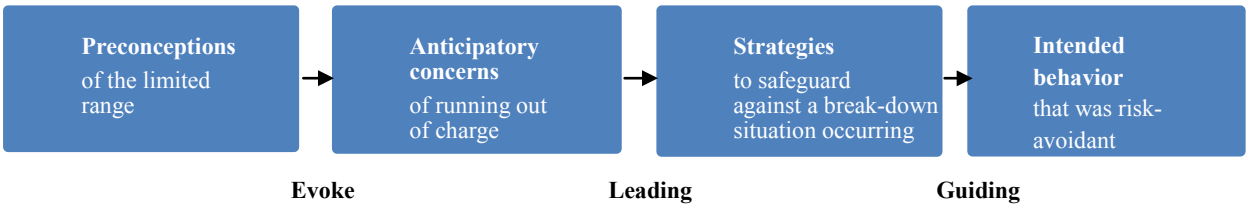


Figure1. Overview of four-stage process of cognition: drivers' preconceptions of the range evoke anticipatory concerns leading to strategies that guide their intended behavior.

Preconceptions

Prior to EV exposure, drivers saw EVs as short range vehicles, ideally suited for usage within a relatively small geographical area. Indeed, the majority of drivers (70%) were acceptant of the range and had already discounted undertaking journeys in the MINI E that exceeded its capabilities on a single charge. Only a small proportion of the drivers (30%) were open to exploring the range and had not entirely discarded travelling to destinations they felt were beyond the range. Therefore, none of the drivers anticipated the range to be a cause for concern on a daily basis, with the expectation that the range would be sufficient for their routine activities. However, prior to EV exposure, drivers' preconceptions of range estimates were principally based on hypothetical information and lacked real-world experience. Thus, drivers did not know with certainty how far they could travel on a single charge.

Anticipatory concerns

The limited range emerged as a critical issue for drivers in relation to exceptional, long, journeys. Indeed, 17 of the 30 drivers (56%) specifically referred to exceptional journeys believed to be beyond the limitations of the range, such as "popping to the coast", "driving up to Scotland at Christmastime" or "visiting family in Cambridge and Norfolk." Therefore, drivers did not feel the MINI E would satisfy all their potential travel demands, or be a viable transport option for covering a large geographical area.

Drivers' lack of certainty as to how far they would be able to go on a single charge was also found to be a cause for concern. Drivers perceived the range as more unpredictable to estimate with a battery than with a tank of fuel. They expected range to be affected by multiple factors (e.g., on-board heater and outside temperature). Drivers with longer average journey lengths (50 miles or above) were significantly more likely than drivers with shorter average journey lengths (20 miles or under) to express such concerns (Fisher's Exact, p=.046).

The present limited availability of public charging points, coupled with lengthy charge times were frequently mentioned in conjunction with the EV range as causing drivers' anticipatory concerns over the range (reported by 70% and 67% of drivers respectively). The present charging difficulties reinforced drivers' views of EVs as being acceptable for usage within the confines of the range only.

In sum, drivers' anticipatory concerns in relation to the EV range were found to originate from their ultimate fear of running out of charge and being stranded.

Strategies

Drivers constructed a variety of cognitive strategies that effectively enhanced their perceived control over the range and regulated their fear of a breakdown situation occurring. In the first instance, drivers set a 'comfort range' (i.e., distance they felt confident in travelling in the MINI E). 23 of the 30 drivers (77%) reported their 'comfort range' as being less than the distance they felt the MINI E was capable of achieving under optimal conditions. Thus, the vast majority of drivers imposed an energy reserve prior to their first trip in the MINI E.

The establishment of a 'comfort range' was associated with the construction of strategies to prevent energy resources being problematic within the limits of that 'comfort range'. Drivers' strategies for managing the battery capacity fell into 4 broad categories: *containing*, *maximizing*, *replenishing* and *conserving* energy resources. Table 1 provides an overview of drivers' strategies.

Table 1: Overview of drivers' strategies to manage energy resources within the limits of the range

Strategy 1: Containing the limited range	**Drivers responses**		**Total % of drivers**
	Short average journey length (<20 mile round trip) (Total drivers = 16)	**Longer average journey length (>50 mile round trip) (Total drivers = 14)**	
Preplanning of journeys	2 (13%)	11 (71%)	13/30 (43%)
Incorporation of contingency margin	0 (0%)	3 (21%)	3/30 (10%)
Not challenging energy resources	14 (87%)	3 (21%)	17/30 (57%)
Strategy 2: Maximizing State of Charge	**Drivers responses**		**Total % of drivers**
	Short average journey length (<20 mile round trip) (Total drivers = 16)	**Longer average journey length (>50 mile round trip) (Total drivers = 14)**	
Topping up to ensure max. State of Charge	7 (44%)	10 (71%)	17/30 (57%)
Strategy 3: Replenishing energy levels during trip	**Drivers responses**		**Total % of drivers**
	Acceptant of limited range (Total divers = 19)	**Open to exploring the range (Total drivers = 11)**	
Recharging in emergency situations only	7 (37%)	0 (0%)	7/30 (23%)
Recharging to extend distance travelled	1 (5%)	6 (55%)	7/30 (23%)
Strategy 4: Conserving battery power	**Drivers responses**		**Total % of drivers**
	Short average journey length (<20 mile round trip) (Total drivers = 16)	**Longer average journey length (>50 mile round trip) (Total drivers = 14)**	

Changing route selection	1 (6%)	1 (7%)	2/30 (7%)
Changing driving style (steadier speeds etc.)	2 (12%)	4 (29%)	6/30 (20%)

Containing the limited range: All drivers conceived of ways in which to effectively contain their use of the MINI E within their comfort range. 13 of the 30 drivers (43%) referred to making a direct assessment through pre-planning journeys to ensure energy resources would be sufficient to complete the trip successfully. 3 of these 13 drivers also explicitly mentioned incorporating a contingency margin (or safety buffer) around the perceived edge of the range to accommodate variations in the rate of battery depletion. 11 of these 13 drivers had longer than average journey lengths (50 miles or over).

The remaining 17 drivers (57%) did not report expecting to make such calculated assessments in order to manage energy resources. Instead, these drivers foresaw the MINI E as principally (or exclusively) being used as the household's short commuter vehicle, and therefore was not (or infrequently) anticipated to be used in situations that would be remotely challenging to the range. 14 of these 17 drivers had short average journey lengths (20 miles or under).

The above analysis indicates that none of the drivers anticipated making a journey in the MINI E without first of all having carried out some assessment to ensure the energy resources would be sufficient.

*Maximizing State of Charge:*17of the 30 drivers (57%) mentioned maximizing the State of Charge by charging on a routine basis (e.g., every night) or keeping energy resources topped up (regardless of the remaining State of Charge) in order to prevent the range from being an issue. These drivers anticipated starting their journeys with maximum battery power which increased their confidence in being able to cope with any unforeseen trips that may arise.

Replenishing energy resources during trip: 14 of the 30 drivers (47%) referred to the possibility of replenishing energy resources during a trip (primarily through the use of the 13A charger). 7 of the 14 drivers envisaged using such strategies in emergency situations only (e.g., if they had miscalculated the distance of a trip rather than increasing the distance it was possible to travel in the MINI E). All of these 7 drivers were found to be acceptant of the limited range (i.e., had dismissed the possibility of undertaking longer distance journeys in the vehicle).

In contrast, the remaining 7 drivers contemplated situations in which they could potentially recharge to extend the use of the MINI E beyond that of a single charge. 6 of these 7 drivers were found to be open to exploring the range (i.e., had not completely dismissed the possibility of undertaking longer journeys in the MINI E).

Conserving battery power: 8 of the 30 drivers (27%) referred to two strategies believed to effectively reduce the rate at which battery power was drained. In the first instance, 2 of the 8 drivers (25%) considered selecting different routes (e.g. to work) in order to conserve the State of Charge as much as possible In the second case, 6 of the 8 drivers (75%) felt they would adopt (or maintain) a conservative style of driving (e.g. steadier and smoother) in an attempt to prolong the battery life. 4 of the 6 drivers had a longer average journey length (50 miles or above).

Drivers principally anticipated avoiding using the MINI E if a particular journey exceeded the range limits, envisaging reverting to using their conventional vehicle(s).

Intended Behavior

Drivers' strategies were essentially defense mechanisms to the potential threat of running out of charge and being stranded. This manifested in drivers' intended behavior being conservative (or risk-avoidant). Indeed, drivers principally envisaged using the MINI E whenever possible, there and back, on a single charge. Thus, it was never drivers' intention to use the MINI E in isolation, but always in conjunction with the other vehicle(s) in the household. The integration of the MINI E into a multicar household meant that drivers did not anticipate any difficulties in adapting to the limited range. Indeed, drivers anticipated having to make very few (if any) behavioral changes, and consequently expected their daily routines to continue without any fundamental changes.

DISCUSSION

The main finding from this study is that drivers of the BMW MINI E had established *a priori* strategies for coping with range, enabling them to avoid experiencing range anxiety prior to EV exposure. Drivers with longer average journey lengths (50 miles or above) were more likely to anticipate having to preplan journeys to the point of destination, building in contingency margins to accommodate variations in the rate of battery depletion, and attempting to reduce the rate at which battery power depleted (i.e., by making adjustments to driving style). In contrast, drivers with short average journey lengths (20 miles or under) were less likely to construct such strategies to manage the energy resources of the battery. This is likely attributable to drivers with short average journey lengths anticipating their daily activities to be well within the limits of the battery capacity, therefore precluding the need for them to put any further strategies in place to safeguard against running out of battery power.

The above findings suggest that even at pre-trial drivers already had internal representations of how they anticipated dealing with the EV range. It also provides some evidence that drivers with longer average journey lengths had begun to develop a deeper (more complex) way of thinking in terms of their future interaction with the energy resources of the battery than had drivers with shorter average journey lengths. This may indicate that drivers principally anticipated having to make behavioral changes when travelling on a lower State of Charge only.

Drivers established an energy reserve, or safety buffer, even before they had got into their EV. Even at this point, the majority of drivers had established a 'comfort range' that was less than the optimum distance they believed the MINI E to be capable of achieving. This suggests that drivers were not confident in utilizing the full amount of battery power available to them. This is of importance as it lends support that instead of attempting to overcome range anxiety through additional battery capacity (with associated increases in battery cost and size), future research should explore ways of encouraging new consumers to use a greater proportion of the current battery capacity. One potential way of achieving this would be through driver training designed to get new consumers to push the limits of the range and to alter, or even challenge, their pre-trial mindset.

In summary, this study demonstrates the importance of research into explicating drivers' pre-trial beliefs and mental constructions the range.

CONCLUSIONS

Drivers of the BMW MINI E had established a priori strategies for coping with range, enabling them to avoid experiencing range anxiety prior to EV exposure.

REFERENCES

Beggs, S., Cardell, S., & Hausman, J. (1981). Assessing the potential demand for electric cars. *Journal of Economics, 17,* 1–19.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3,* 77–101.

Bunch, D.S., Bradley, M., Golob, T.F., Kitamura, R., & Occhiuzzo, G.P. (1993). Demand for clean-fuel vehicles in California: a discrete-choice stated preference pilot project. *Transportation Research Part A, 27,* 237–253.

Carley, S., Krause, R.M., Lane, B.W., Graham, J.D. (2013). Intent to purchase a plug-in electric vehicle: A survey of early impressions in large US cites. *Transportation Research Part D, 18,* 39–45.

Carroll, S., & Walsh, C. (2010). The Smart move trial: description and initial results. London: Cenex. Retrieved from http://www.cenex.co.uk/LinkClick.aspx?fileticket=yUKaRDJtWg%3D&tabid=60.

Chéron, E., & Zins, M. (1997). Electric vehicle purchasing intentions: The concern over battery charge duration. *Transportation Research A, 31*, 235-243.

Cocron, P., Bühler, F., Neumann, I., Franke, T., Krems, J. F., Schwalm, M., & Keinath, A. (2011). Methods of evaluating electric vehicles from a user's perspective – the MINI E field trial in Berlin. Germany. *IET Intelligent Transport Systems, 5(2),* 127–133.

Dagsvik, S., Wennemo, T., Wetterwald, D.G., Aaberge, R. (2002), Potential demand for alternative fuel vehicles, *Transportation Research Part B, 36,* 361–384.

Daziano, R.A., & Chiew, E. (2013). On the effect of the prior of Bayes estimators of the willingness to pay for electric-vehicle driving range. *Transportation research Part D, 21,* 7–13.

Eggers, F., & Eggers, F. (2011). Where have all the flowers gone? Forecasting green trends in the automobile industry with a

choice-based conjoint adoption model. *Technological Forecasting & Social Change, 78,* 51–62.
Franke, T., & Krems, J.F. (2013a). What drives range preferences in electric vehicle users? *Transport Policy, 30,* 56-62.
Franke, T., & Krems, J. F. (2013b). Understanding charging behaviour of electric vehicle users. *Transportation Research Part F, 21,* 75-89.
Franke, T., Cocron, P., Bühler, F., Neumann, I., & Krems, J.F. (2012). Adapting to the range of an electric vehicle – the relation of experience to subjectively available mobility resources. *Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems,* June 14-15, Valencia, Spain.
Jensen, M. (1999). Passion and heart in transport – a social analysis on transport behaviour. *Transport Policy, 6,* 19–33.
Kurani, K., Turrentine, T. & Sperling, D. (1996). Testing Electric Vehicle Demand in 'Hybrid Households' using a Reflexive Survey. *Transportation Research Part D, 1(2),* 131–150.
Neilsson, M. (2012). Electric vehicle: The Phenomenon of Range Anxiety. Retrieved from http://www.elvire.eu/IMG/pdf/The_phenomenon_of_range_anxiety_ELVIRE.pdf.
Neumann, I., Cocron, P., Franke, T. & Krems, J.F. (2010). Electric Vehicles as a solution for green driving in the future? A field study examining the user acceptance of electric vehicles. *Proceedings of the European Conference on Human Interface Design for Intelligent Transport Systems*, April 29-30, Berlin, Germany.
Parker, I. (2005). *Qualitative psychology : Introducing radical research.* Maidenhead : Open University Press.
Pearre, N., Kempton, W., Guensler, R.L. & Elango, V.V. (2011). Electric vehicles: how much range is required for a day's driving? *Transportation Research Part C, 19,* 1171-84.
Smith, J.A., & Eatough, V. (2007). Interpretative phenomenological analysis. In A. Coyle & E. Lyons (Eds.), *Analysing qualitative data in psychology: A practical and comparative guide* (pp. 35-50). London: Sage.
Smith, J.A., & Osborn, M. (2008). Interpretative phenomenological analysis. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods,* 2nd ed. London: Sage.
Tamor, M., A., Gearhart, C., & Soto, C. (2013). A statistical approach to estimating acceptance of electric vehicles and electrification of personal transport. *Transportation Research Part C, 26, 125 – 134.*
Turrentine, T., Garas, D., Lentz, A. & Woodjack, J. (2011). The UC Davis MINI E consumer study. Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-11-05.
Vilimek, R., Keinath, A., & Schwalm, M. (2012). The MINI E field study: Similarities and differences in international everyday driving. *Proceedings of the 4th International Conference on Applied Human Factors and Ergonomics*, July 21-25, San Francisco, CA.
Wengraf, T. (2001). *Qualitative research interviewing: Biographic narrative and semi-structured methods.* London: Sage.

Parking Behaviors for Inductive Charging

Stewart Birrell, Daniel Wilson, Gunwant Dhadyalla,
Chek Pin Yang, Andrew McGordon and Paul Jennings

WMG, International Digital Lab
University of Warwick
Coventry
CV4 7AL, United Kingdom

ABSTRACT

It is well established that the efficiency of transfer of electrical energy through induction is highly reliant on accurate alignment of the coils involved – in this case, between a primary coil in the parking bay and a secondary coil mounted on the vehicle. Whist inductive charging technology is market ready, understanding the issue of misalignment is an important human factors question and the focus of this paper. A retrospective assessment of parking behavior was conducted by recording parking orientation and alignment of 100 vehicles, parked in perpendicular bays, in three different car parks at the University of Warwick. Results showed on average vehicles tended to park 3.1 cm to the left of the bay, with an angle of practically zero. Parking was typically towards the rear of the bay; however the presence of a physical barrier led to vehicles being parked more centrally. The orientation of the parked vehicle had little impact on accuracy. Tolerances for misalignment with inductive charging systems are small in comparison to the distribution of parking accuracy observed in this study, at 15-20 cm verses 120 cm respectively. Conclusions were that only 5% of vehicles were parked sufficiently accurately to allow inductive power transfer to commence.

Keywords: Inductive Charging, Inductive Power Transfer, Parking Alignment, Electric Vehicles, Driver Behavior

INTRODUCTION

Inductive power transfer (IPT; also known as wireless power transfer (WPT) or inductive charging) has roots in Nikola Tesla's investigation of wireless transmission of power in the 1890's using some rather extreme methods. Modern IPT owes its resurgence to Prof's. André Kurs, Marim Soljačić, and the MIT team that demonstrated transmission of 60W across 2.5 m to illuminate a light bulb[1]. Many others worked on IPT even before this, notably the team at Auckland University in New Zealand[2]. The sheer convenience of inductive charging for electric vehicles (EV) coincided with standardization activities of conductive (or wired) charging of EV's so interest grew rapidly. In 2010, the Society of Automotive Engineers (SAE) has formed the SAE J2954 standard, which covers the static wireless charging on EV. In parallel with the SAE J2954, the IEEE Standard Organization has also established a workforce to define the standard specifically for in-motion (dynamic) charging.

The key area in this application is to transfer power over an air gap at an acceptable level of efficiency whilst meeting any legislative requirements. There has been significant work carried out by various companies on resonant inductive coupling. This is the near field wireless transmission of electrical energy between two coils that are highly resonant at the same frequency. The efficiency of the energy transfer is a function of frequency and alignment of the primary and secondary coils. Misalignment can occur across the air gap distance, lateral and longitudinal

[1] http://www.sciencecodex.com/mit_demonstrates_wireless_power_transfer
[2] http://www.engineering.auckland.ac.nz; http://www.qualcommhalo.com/

misalignment, and angular misalignment. The efficiency of energy transfer generally drops rapidly once the misalignment reaches 15 cm, as a result, the system control will cut off the power or the transfer will not start when the efficiency is below 80%. Recent development in coil design (i.e. multiple coils, field shaping etc.) has resulted in marginal improvements in the physical efficiency bandwidth. Figure 1 shows a typical IPT system. The power drawn from the 50Hz supply is first rectified through an AC/DC converter before transformed to a high frequency (kHz) resonant energy by the inverter. This energy is transferred across the air gap to the receiving end through the magnetic coils. At the receiving end, it is once again rectified to charge the battery in the vehicle. However whilst the technology is ready for adoption, the weak link in the chain is the driver and how accurately they can park, or drive over the coil to maximize energy transfer efficiencies.

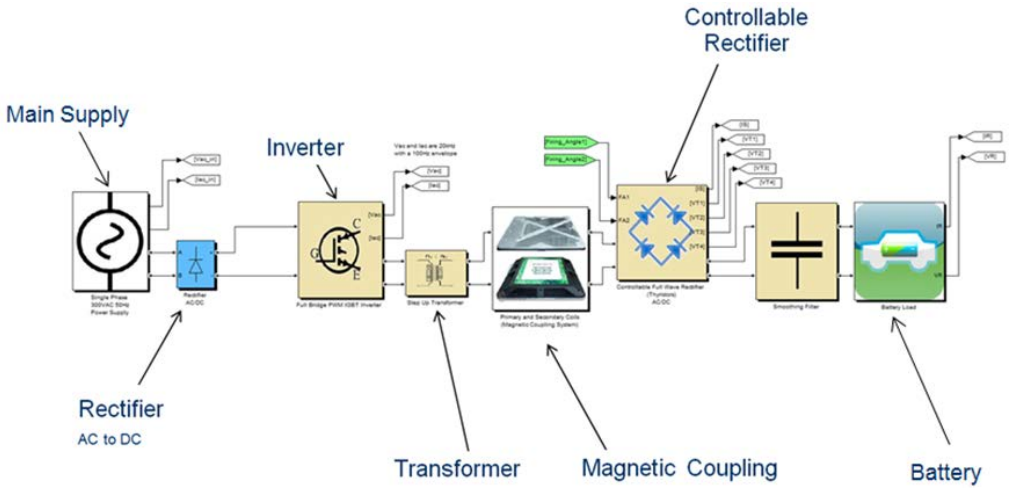


Figure 1: Single phase wireless IPT charger schematic diagram

A review of the literature reveals that limited research has been conducted looking into parking behavior, which has been primarily focused on parking orientation and preferences (Cullinane et al, 2004; Kobus et al 2012), gender differences (Wolf et al, 2010), desired clearance between vehicles (Gadgil and Green, 2005), parking related crashes and incidences (Green, 2006) or vehicle-to-vehicle gap in real-world parking (Thornton et al, 2014). The most relevant study for this current paper was conducted by Cullinane et al (2004) whose research for the University of Michigan Transportation Research Institute (UMTRI) evaluated parking accuracy of 102 vehicles in three different types of parking bays (angled, parallel and perpendicular). This was a retrospective analysis with participants not being aware that there parking would be assessed. Their research found no difference between the size of the vehicle and lateral parking accuracy, with drivers attempting to keep a constant amount of exit space on the driver's side of the vehicle in angled and perpendicular parking, and with parallel parking how close drivers parked to the curb. However vehicle size did have an effect on longitudinal clearance of perpendicular parking, with large vehicles more likely to overhang a ground barrier when parking a larger vehicle. The angle of parking tended to be quite small, generally less than 1 degree; however the range of angles was triple that for parallel parking verses angled and perpendicular parking (Cullinane et al, 2004).

As suggested above it is well established that the efficiency of transfer of electrical energy through induction is highly reliant on alignment of the coils involved – in this case, the alignment between a primary coil in the parking space, and a secondary coil mounted on the vehicle. If the integration of inductive charging technology into modern vehicles is to be accepted by drivers, engineers should seek to minimize the amount of disruption to current parking behavior. This raises a question – "Is current parking behavior compatible with inductive charging technology?"

The two main research questions to be answered by this paper are:

1. What correlations exist between the accuracy of parking and other factors?

2. What proportion of drivers currently park within the tolerances of inductive charging systems?

METHODOLOGY

Vehicles and Locations

100 vehicles were randomly selected for measurement from three different car parks on the main campus of the University of Warwick. Vehicles were already parked when measurement was taken, i.e. it was a retrospective analysis of static vehicle rather than a dynamic assessment. This obviously meant that drivers were not aware that their parking alignment was going to be assessed when they originally parked. Assessment took place over a two week period during July, with physical measurements being taken between 10:00 – 12:00 and 14:00 – 16:00.

Ethical approval was granted for the study by the University of Warwick's Biomedical and Scientific Research Ethics Committee (BESREC). It was agreed that no identifying features of the vehicle would be recorded (e.g. photographs of assessed vehicles, number plates, gender of driver etc.), and in the case of approach by a member of public the researchers would offer a verbal and written explanation of the study, and give the opportunity for the driver of a measured vehicle to have their data removed.

The three car parks selected for the study were all perpendicular bays, with at least white 'T''s to identify parking area at the head (or front) of the parking bay. Full car park characteristics are outlined below:

Table 1: Characteristics of the three cars parks from which vehicles were measured.

Car Park	Front Marker	Rear Marker	Rear Barrier	Side Barrier	Open / Covered	Max Stay Duration	N of Cars	Ave Length (cm)	Ave Width (cm)
8	T	None	High & Hard	Car or Hard	Covered	24 hours	41	476.1	228.9
8a	T	Dash	High & Soft	Car or Soft	Open	9 hours	47	483.4	234.5
10b	T	Dash	Low & Soft	Car or Soft	Open	1 hour	12	467.8	217.8

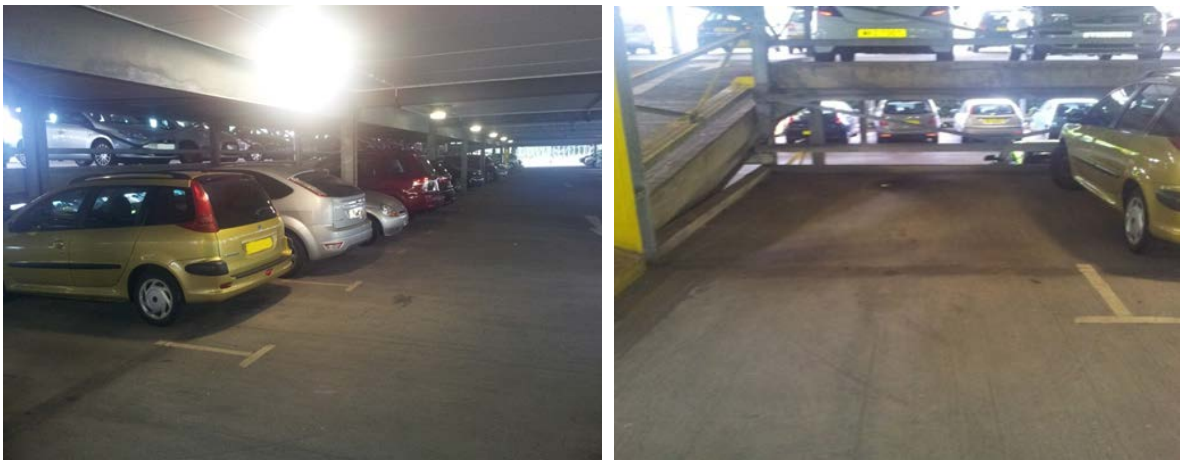


Figure 2: Car Park 8

Figure 3: Car Park 8a

Figure 4: Car Park 10b

Measured and Derived Variables

To measure the actual parking alignment of vehicles they were simply modelled as rectangles located in a rectangular bay. The principle method of data collection for this study was a retrospective analysis of parking accuracy, by determining parking bay and vehicle size by physical measurement and other relevant observations. From this numerous different parameters could be derived, including:

- **Length and width of parking bay:** These were measured by tape measure, from the back of the bay to the front (including line markings), and between the inside of both "T" or parallel marks.
- **The length, width and wheelbase of the vehicle:** These were measured by tape measure, between markers which were set down on the floor at the points at which the corners of the "vehicle" rectangle in the model would be located.
- **The distance from vehicle to side of bay:** Measurement to the markers at the respective corners on one side of the vehicle, using a tape measure.
- **The closest distance between the vehicle and the rear of the bay:** Measured using a tape measure.
- **The orientation of the vehicle**: Rear-parked or forward-parked.
- **The general category of vehicle:** From a pre-provided list of types.
- **Parking sensors:** If applicable.

- **Location of the center of the vehicle:** Derived with respect to a coordinate system with origin in the center of the bay.
- **Angle of the vehicle** to the bay.

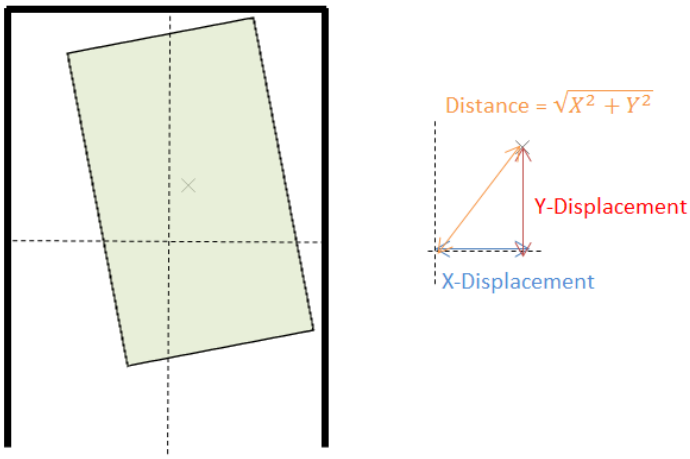


Figure 5: Example of the how the derived data was calculated; x represents the center of the blue vehicle 'rectangle' and dotted lines within the bold black lines the center of the bay 'rectangle. The displacement of the center of the vehicle to the center of the bay in both the x (lateral) and y (longitudinal) axis was calculated, as well as the distance from center of vehicle to center of bay.

Raw data were collected via physical measurement with a tape measure, collated on a data entry sheet, and then transferred to SPSS for analysis. Data for vehicles who reversed into the parking spaces were reversed only for analysis, but still considered separately, meaning that if the driver parked off-center laterally (i.e. leaving more space either to the drivers or passenger side then this would not be lost in the analysis).

RESULTS

Objective Measures

This section shows the results from the more objective measures of vehicle classification. Table 2 shows that the majority of vehicles evaluated were hatchback, with no parking sensors and forward facing into the parking bay. Table 3 shows that vehicles were tended to be parked towards the left-hand side of bay, by 3.1 cm, and 15.6 cm forward in the bay, regardless of orientation. The mean angle of parking was 0.02 degrees.

Table 2: Objective measures of vehicle classification.

Type of Vehicle	**N**	**Parking Sensors**	**N**	**Orientation**	**N**
Hatchback	64	Unknown	10	Reversed	27
Sedan	25	None	69	Forward	73
People-Carrier	5	Rear Only	17	Total	100
Coupe	5	Rear & Front	4		
Total	100	Total	100		

Table 3: Mean displacement and distance in the x (lateral) and y (longitudinal) axis, and distance away from the center of the vehicle to the center of the bay. Mean distance is simply the absolute value of the "X-Displacement", and the same in the Y case, i.e. if they were away from the center of the bay how far was it.

	Mean	**S.D.**
Centre X-Displacement (cm):	-3.21	14.64
Centre Y-Displacement (cm):	15.59	25.02
Centre Distance (cm):	29.30	15.13
Centre X-Distance (cm):	12.12	8.74
Centre Y-Distance (cm):	23.73	29.12
Angle to Bay (degrees):	0.018	2.27

Trends and Correlations

Figure 6 below show the distribution of X and Y coordinates of the center of vehicles in the sample, colored by size of vehicle and rear/forward parking. Size was determined by separating the data equally into three groups, based on the area of the vehicle base, "width x length". Figure 7 shows the angle of parking from different parking orientations and Figure 8 the distance to the rear of the parking bay for the different car parks.

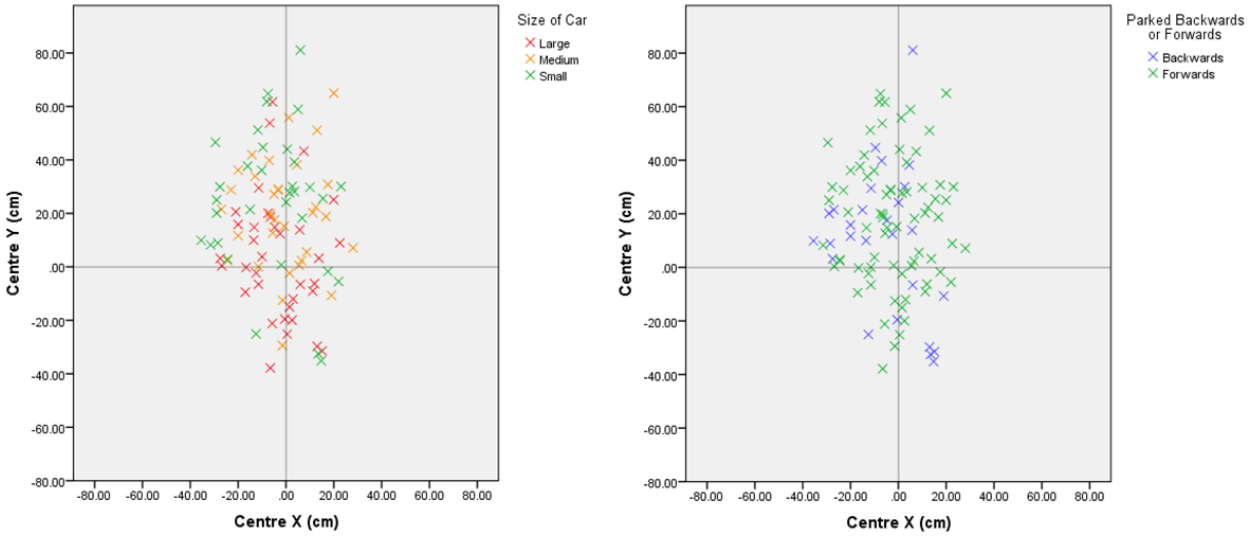


Figure 6: Distribution of vehicles to the center of the bay. Graph 1 shows alignment by vehicle size, graph 2 by parking orientation.

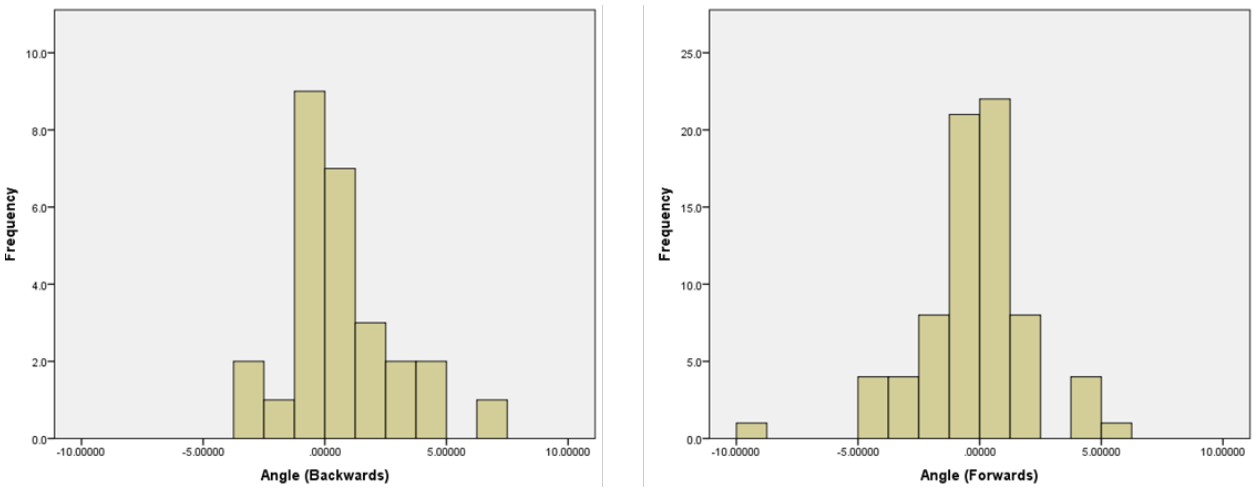


Figure 7: Distribution of the angle of parking for reversed (Backwards; graph 1) and forward (graph 2) parking alignment.

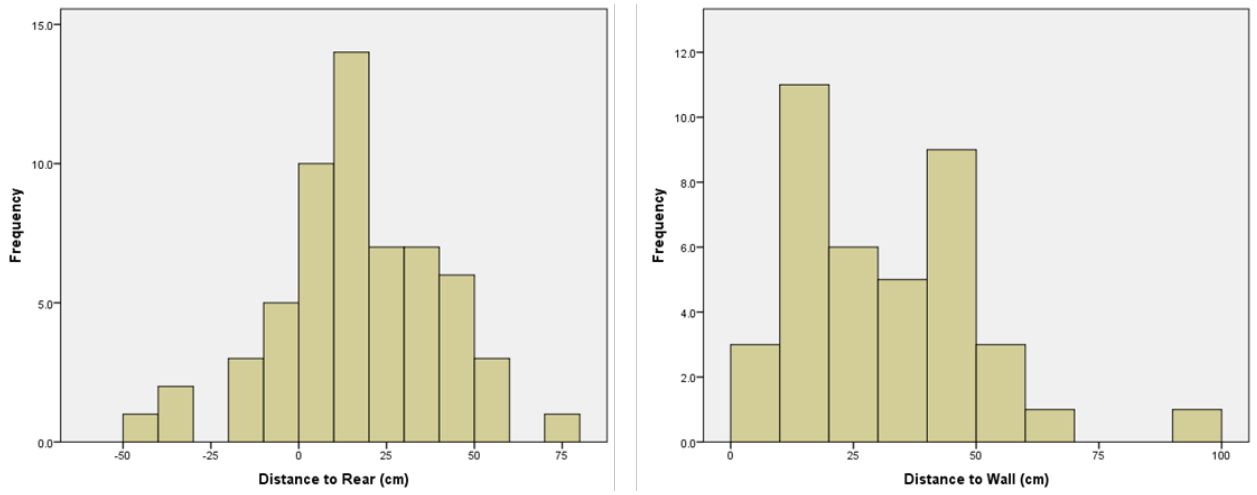


Figure 8: Distance from the most distal point of the vehicle to the rear of the parking bay. Data presented in graph 1 from Car Park 8a and 10b (low or soft rear barriers), graph 2 from Car Park 8 (high and hard).

DISCUSSION

Parking Alignment

The elliptical nature of the spread of the data in Figure 6, with the x-mean at approximately zero and y-mean at approximately 15 cm, suggest that vehicles were parked more accurately in the lateral (or x) direction compared to the longitudinal (y) direction. The spread in the data also reflects this, with the range of lateral accuracy at around 60 cm which was approximately half that of the longitudinal range. Table 3 shows that the mean distance from the center of the parking bay was 12.1 cm (either to the left or right), with a mean displacement of 3.2 cm to the left of the bay (i.e. more clearance on the driver's side than the passenger's side). Research conducted by Cullinane et al (2004) showed that in a comparable study conducted in Ann Arbor, Michigan, US, vehicles were parked with an average of 10.4 cm clearance on the driver's side (right in this case). This increased lateral offset may be a function of the average bay width in the US study being 47 cm wider than in the current study, or may simply be a cultural preference.

An assumption could have been made before the study was conducted that larger vehicles would be less well aligned in the parking bay in comparison to smaller vehicles. This is based on the fact that they would have less space for movement and correction, or reduced visibility and perception of the front and rear of the vehicle due to the length. This was not born out in the results as Figure 6 shows that the distribution of 'large' vehicles is far more evenly spread around the center of the bay. Small vehicles in particular tended to park more towards the rear of the bay compared to large vehicles. The spread of data in the lateral direction is similar for all sizes of vehicles. Possible reasons for the increased 'accuracy' of larger vehicles may be that the vehicle to bay ratio is smaller, meaning they have less space allowance to 'play with' when parking, not parking centrally within a bay would increase the risk of the vehicle either overhanging the front or rear of the bay, both of which may result in physical damage to the vehicle. It is a generalization (but not unfounded) that larger vehicles are likely to be more expensive than their smaller counterparts. More expensive, premium vehicles are more likely to be fitted with parking aids to assist specifically with the longitudinal aspect of parking; thus making them more accurate. Another explanation could again be the vehicle to bay ratio. Smaller vehicles may choose to position themselves to allow for ease on exiting the vehicle, or choose to give a vehicle parked at the side more clearance and hence reduced risk of damage to their vehicle, but resulting in reduced accuracy. Cullinane et al (2004) also found that vehicle size affected longitudinal accuracy, with drivers of a large vehicle more likely to over hang a ground (low) barrier.

An interesting occurrence was the adaption to parking behaviors with the presence of a 'hard' rear barrier, in this case a steel structure in car park 8 (Figure 2). Figure 8 shows that the distribution of distance to rear of the bay for no physical barrier (graph 1) is normal, with the most frequent distance being 10 to 20 cm from the rear barrier. Whilst 10-20 cm is till the most frequent distance for when a hard barrier is present (graph 2), the distribution could not be considered normal, but either chi-squared or bimodal. A second peak at parking distance of 40-50 cm may indicate the difficultly that people have judging the location of the front of their vehicle in comparison to the parking bay. It is at this point that feedback should be given to the driver to encourage more accurate longitudinal alignment.

The mean angle of parking was 0.02 degrees, even in absolute form this was only 1.6 degrees. Whilst this implies that in general vehicles were parked straight on in the bay, the range of the data reveals greater differences being 15.8 degrees (6.7 to -9.1 degrees). Interestingly results from Cullinane et al (2004) showed that whist the mean angle was very small (and similar to the present study) at 0.1 degrees, the range of vehicles they recorded was significantly smaller at just 1.4 degrees. This difference may be a function of the smaller number of vehicles evaluated parking in a bay (or perpendicular) with the Cullinane study (36 verses 100), or again that the average bay with in their study was nearly 50 cm wider than the current study.

Inductive Charging Compatibility

The second research question posed at the beginning of the paper surrounds understanding the proportion of drivers who currently park within the tolerances of inductive charging systems. For this, two hypothetical systems were established; System A has tolerances of 15 cm in both the lateral (x) and longitudinal (y) axis and System B 10 cm in the x and 20 cm in the y axis. The dimensions were selected to reflect two different approaches to coil design and are representative of inductive charging systems currently on the market.

Table 4: System tolerances of two notional inductive charging systems, and number of vehicles parked within these tolerances.

			N of vehicles parked within tolerance					
	System tolerance (cm)		**Centre of Bay**			**Mean Position**		
	X	Y	X	Y	*Both*	X	Y	*Both*
System A	15	15	40	20	*5*	38	26	*9*
System B	10	20	48	13	*4*	51	19	*11*

Table 4 shows that in general only 5% of vehicles assessed in this study parked within alignment tolerances for both systems when the primary coil is located in the center of the parking bay. If the coil was located in the mean position for the center of the vehicle (as per Table 3) then this would double the number of vehicles included, but still only account for approximately 10% of vehicles measured. This is highlighted in Figure 9 which illustrates that only a

very small fraction of vehicles would be aligned correctly for power transfers of over 80% efficiency. Also shown is just the sheer range of parking alignment that developers of inductive charging systems for EVs need to contend with.

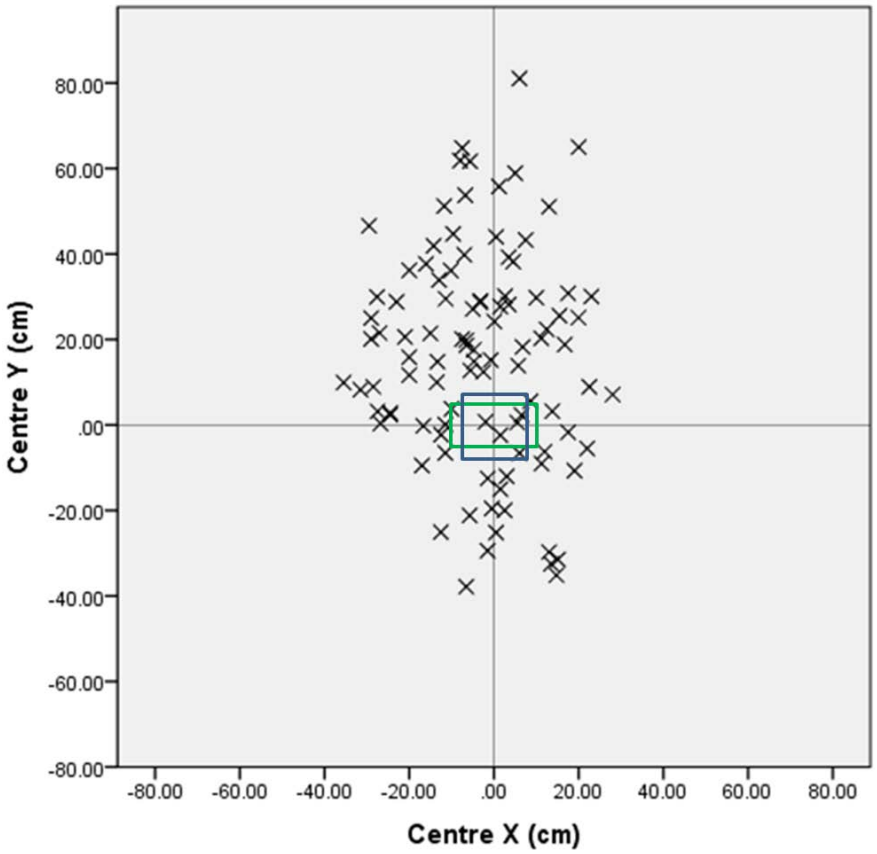


Figure 9: Distribution of vehicles to the center of the bay, with approximate dimensions of System A (blue) and B (green) tolerances overlaid.

How to facilitate users to park within tolerances of inductive charging systems is a difficult challenge. This cannot feasibly be solved by simply making the inductive coils larger, as this is not a cost effective solution also vehicle manufacturers will have packaging difficulties fitting a larger coil to the underside (and exposed aspect) of a vehicle. It may be addressed by technology, and specifically park assist systems. Automated parking is already a feature on many premium vehicles, which could be used to park the vehicle accurately over the primary coil in the bay. However, park assist typically is only available for 'difficult' parking maneuvers (i.e. parallel rather than bay parking), and it is still unclear if parking is between two vehicles or in the center of the bay, as this will have an obvious effect on alignment. A feasible solution would be to offer in-vehicle feedback to the driver during the parking maneuver. Results from this study show that whilst parking is more accurate in the lateral axis, between 40 and 50% of vehicles parked were still outside 15-20 cm alignment tolerances of the two systems (table 3). This means that providing only longitudinal feedback address only half of the issue.

Limitations and Future Research

This study adopted a retrospective measurement methodology with drivers being unaware that their parking accuracy was going to be assessed. This type of data collection offers the benefit of naturalistic, real-world data being collected. However, it is unknown to what effect users being aware that their parking accuracy is being assessed will alter actual alignment. It could be assumed that when users know they should be parking as accurately as possible, an improvement in both lateral and longitudinal control will be observed. In addition no context was

available as to why the driver chose to park as they did. It is entirely possible that drivers chose to parked 'misaligned' either to allow space to exit the vehicle, gain access to the boot, or even park off center to the right to allow the vehicle next to you more space to exit and hence reduce the risk of damage to your own vehicle. Future studies will need to need to address these issues, and also assess parking in a greater range of situations, such as parallel parking and 'free' parking which is not restricted by other vehicles and bay boundaries, a situation which would be typical on a driveway.

CONCLUSIONS

Parking accuracy to the center of a perpendicular bay was assessed retrospectively in this study. Results show that on average vehicles tended to be parked 3.1 cm to the left of the parking bay, with an angle of practically zero. Parking was typically further towards the rear of the bay; however the presence of a physical barrier led to vehicle being parked more centrally. The orientation of the parked vehicle had little impact on parking accuracy. However, results from this study showed that larger vehicles parked more accurately in comparison to smaller ones. Tolerances for misalignment with inductive charging systems are small in comparison to the distribution of parking accuracy observed in this study, at 15-20 cm verses 120 cm respectively. This study showed that only 5% of vehicles were parked sufficiently accurately to allow for inductive power transfer to commence.

REFERENCES

Cullinane, B., Smith, D., Green, P. (2004), *"Where, when and how well people park: A phone survey and field measurements"*, The University of Michigan Transportation Research Institute (UMTRI), US, Technical Report UMTRI-2004-18.

Gadgil, S., Green, P. (2005), *How much clearance drivers want while parking: Data to guide the design of parking assistance system"*, Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting, pp. 1935-1939.

Green, P. (2006), "*Parking Crashes and Parking Assistance System Design: Evidence from Crash Databases, the Literature, and Insurance Agent Interviews",* SAE Technical Paper 2006-01-1685,

Kobus, M., Gutierrez-i-Puigarnau, E., Rietveld, P., Van Ommeren, J. (2012), *"The on-street parking premium and car drivers' choice between street and garage parking"*, Tinbergen Institute, The Netherlands, TI2012-040/3.

Thornton, D., Redmill, K., Coifman, B. (2014), *"Automated parking surveys from a LIDAR equipped vehicle"*, Transportation Research Part C Volume 39 pp. 23-35.

Wolf, C., Ocklenburg, S., Oren, B., Becker, C., Hofstatter, A., Bos, C., Popken, M., Thorstensen, T., Gunturkun, O. (2010), *"Sex difference in parking are affected by biological and social factors"*, Psychological Research Volume 74 pp 429-435.

Section 16

Road and Rail:

Systems Thinking in Road and Rail Transport

Combining Human Factors Methods with Transport Data Recordings

Guy Walker and Ailsa Strathie

Institute for Infrastructure and Environment
Heriot-Watt University
Edinburgh, UK

ABSTRACT

Flight Data Monitoring (FDM) is the process by which data from on-board recorders, or so-called 'black boxes', is analysed after every journey to detect subtle trends which, if allowed to continue, could lead to an accident. An opportunity has been identified to advance the state of the art in FDM processes by coupling recorder data to established Human Factors methodologies so that issues arising from the strategically important human/machine-system interface can be better understood and diagnosed. The research has also identified a significantly underused source of recorder-data within the railway industry. Taking this data, the paper demonstrates how key areas of driver performance can be quantified using a simple behavioural cluster detection method coupled to sensitivity and response bias metrics. Faced with a class of operational accident that is increasingly human-centred, an underused source of data, and methods that can join it to established human performance concepts, the potential for detecting risks in advance of an accident are significant.

Keywords: Black box data recorders, leading indicators, process charts, signal detection theory

INTRODUCTION

The Spanish train crash in July 2013 reveals yet again that the interface of humans and machine systems is an area of strategic concern (Evans, 2011; CAA, 2011). Having closed off numerous other technical/engineering opportunities for accidents to occur what is left is a troubling class of accident which arises despite highly trained and motivated personnel, the presence of robust safety management practices, and fully functioning equipment. More worryingly, the methods and approaches that have helped us achieve the current high levels of railway safety seem to be less effective in the face of human/system problems such as these. Another feature of the press coverage around the Spanish rail accident was the presence of a black box data recorder. This paper describes a new way to deploy this data, new human-centered metrics for identifying trends and risks, and a predictive approach to accidents based on the concept of Human Factors leading indicators.

Regardless of measure, whether it takes into account exposure by distance or time, the risk to the travelling public and workforce of using and operating the railway is exceedingly low. In Europe the probability of a fatality is approximately 0.57 per billion miles (Evans, 2011), or two fatalities per 100 million person travel hours (EU, 2003). This figure arises despite the fact that exposure in time and distance have increased dramatically in some countries. In the UK, for example, between 1995 and 2012 the risk exposure by passenger distance rose by 25 billion kilometres or 58% (DfT, 2011). At the same time estimated mean fatal train accidents per billion train kilometres has fallen by approximately 9.1% annually (Evans, 2011). Risk exposure is accompanied by an increase in the overall intensity of operations. The UK railway system currently supports 1.3 billion passenger journeys (ORR, 2012) with 16% more trains timetabled in 2013 compared to 1995, most of which are running at higher passenger occupancy levels. This equates to 296.2 million train miles travelled in 2010/11 (ORR, 2012), all of which have at

one point or another been recorded on an On Train Data Recorder (OTDR) device.

The act of automatically recording data on system parameters over time is referred to as 'data logging' or 'data recording'. In the aviation industry the generic term data logging falls under the specific heading of Flight Data Recording, which itself comprises several individual procedures and devices. The most prominent of these is what is termed colloquially as the 'black box', which represents the combination of a Flight Data Recorder (FDR) and a Cockpit Voice Recorder (CVR). Other systems under the heading of Flight Data Monitoring include various Aircraft Condition Monitoring Systems (ACMS), such as engine health monitoring (e.g. the Rolls Royce EHM programme) and the wide range of parameters available from modern avionics (e.g. ARINC 573) via so-called 'Quick Access Recorders' (QARs). While the aviation sector has a long history with on-vehicle data recording devices for the purposes of safety and crash investigation, these are a much more recent innovation in the rail sector.

The Black Box Paradox

There are three key paradoxes inherent in this wider picture. Firstly, because so few major rail accidents occur there is no longer enough data to construct reliable forward looking estimates (e.g. Evans, 2011). When safety performance data reaches the level of that achieved in the rail sector it instead starts to become characterised by unpredictable periodicities, cycles or discrete events. This is becoming evident in EU rail safety data, with one large scale rail accident occurring on average every six years (EU, 2003). Put simply, safety data is 'levelling off' with a persistent class of human/machine-systems accident now elevated to the status of a key strategic risk (RSSB, 2009).

Secondly, "there is widespread concern within the industry that the background indicators – rather than the headline grabbing ones – have remained worryingly stable" (Wolmar, 2012). An example of this is UK data on Signals Passed At Danger (SPAD) incidents. In the period since the introduction of a countermeasure called the Train Protection and Warning System (TPWS), after which there were initial improvements, there has been comparatively little variation in the overall SPAD rate. For example, the rate for Quarter 4 2012 is the same as Quarter 3, and indeed the same (or very nearly the same) as on seven previous reporting periods since 2005 (e.g. ORR, 2012).

Thirdly, and most paradoxically, is that the opportunities to use black box data for their original purpose (i.e. post-accident analysis) are diminishing at the same time as the technical capabilities of data recorders are increasing. What this means is that enormous quantities of non-accident data is being collected day in and day out, but not currently used.

Data Monitoring

This paper is premised on best-practice techniques developed within the aviation domain, specifically a process called Flight Data Monitoring (FDM). This is a "a systematic method of accessing, analysing, and acting upon information obtained from digital flight data records of routine flight operations to improve safety. It is the pro-active and timely use of flight data to identify and address operational risks before they can lead to incidents and accidents." (CAA, 2003). FDM is mandatory for operators of aeroplanes of a certified take-off mass in excess of 27,000 kg. In effect, it is a way of using data collected from routine operations to detect trends which, if allowed to continue, might eventually lead to an accident. Changes are made to address issues, and the changes themselves are monitored for their possible effects on other parts of the system.

The traditional approach to FDM is focused on exceedence or event detection. Events are defined as: "deviations from flight manual limits, standard operating procedures and good airmanship" (CAP 739, p16). Computer software is used to automatically scan FDR data for instances of these deviations, and a set of core events that cover the main areas of interest are quite standard across operators. Event detection is commonly based on simple statistical techniques and automatic algorithms that detect different phases of flight and events therein. FDM is a highly successful way of making use of black-box data in a pro-active manner, but it too is challenged by the emerging class of human-systems problem that is the focus of this paper. Indeed, while having systems that automatically detect events, it is still incumbent on so-called FDM analysts to manually interpret the lower-order 'trace plots' that data recorders produce in order to derive meaning from them. As such, there is considerable value in being able to robustly transform these trace plots into higher order representations, to detect psychologically meaningful patterns therein, and to automatically derive human performance metrics that can help to assess risk.

Human Factors Leading Indicators

Leading indicators are measurable precursors to major events such as an accident. The indication of a precursor 'leads', or comes before, the actual event itself. Lagging indicators are the opposite. These are so called 'loss metrics' that can only become apparent after an event (Rogers, Evans & Wright, 2009). Leading indicators are said to be 'proactive' because they enable steps to be taken to avoid seriously adverse consequences. Lagging indicators are said to be 'reactive' in that a seriously adverse event needs to occur before it can be learnt from. For this reason, leading indicators are also sometimes referred to as 'positive performance indicators' and lagging indicators as 'negative performance indicators'. The concept of leading and lagging indicators originally derives from the field of economics and the need to understand 'business cycles' (i.e. growth, recession, investment, divestiture etc.) and to predict when one phase of a 'cyclical process' such as this will change to another (Mitchell & Burns, 1938). The terms have been appropriated more recently by the safety and risk field, particularly in view of developments in Safety Management Systems (SMS) since the 1990's. Leading indicators, in a Safety Management context, can be defined as "something that provides information that helps the user respond to changing circumstances and take actions to achieve desired outcomes or avoid unwanted outcomes" (Step Change, 2003, p. 3).

The basic 'research problem' can be stated thus: despite considerable improvements in safety performance in both rail and aviation sectors, a persistent class of accident/near accident continues to occur. This class of accident/near accident resides at the interface of people and machine-systems. What is required is a means to detect the presence/emergence of such problems before they manifest themselves as a serious operational incident. This paper describes how black box data can be coupled to existing Human Factors methods to provide leading indicators of trends residing at this interface. Specifically, it examines OTDR data on driver responses to an in-cab warning to reveal the types of errors that may be more likely to arise if the discovered trends are not addressed.

METHOD

Data File and Parameters

The OTDR data file is a continuous download for a single traction unit. The recording started at 05:34:57 on the 6th July 2012 and ceased at 21:36.32 on the 11th July 2012. This is a period of 136 hours, 1minute and 35 seconds during which the train made 107 journeys and travelled 1638 miles. The raw data takes the form of a Comma Separated (CSV) file containing a data matrix 191,021 time samples (rows) deep by 72 parameters (columns) wide: a total of 13,753,512 data points. The mean sampling rate is 2.56 seconds. The logger itself scans the parameters for changes at a rate of 20mS but, in the present system, to economise on memory requirements data are only logged when one of the 72 parameters changes. The OTDR device itself was a UK Railway Group Standards compliant Arrowvale unit which recorded 72 parameters, 25 of which are in addition to those mandated. In terms of data classification four of the parameters; time, distance and two speed signals derived from a driven and non-driven axle, are continuous ratio data. The remaining 68 are nominal/binomial (i.e. on or off).

Rolling Stock

The sample of OTDR data was obtained from a Class 153 'super sprinter', unit number 153 306. This is a single-unit diesel powered railcar built between 1987 and 1988. Class 153s are 23.2 meters in length and have an un-laden weight of 41.2 tons. They seat 72 passengers, comprise a riveted aluminium body shell affixed to a steel under-frame, and are equipped with four electrically powered single-leaf Bode doors. The prime-mover is an under-slung turbocharged 6 cylinder Cummins NT855 diesel engine producing 285bhp. A Cummins-Voith T211r hydraulic transmission drives both axles of the leading BT38 bogie via a Gmeinder final drive. The Unit's maximum operating speed is 75mph. It is fitted with electro-pneumatic clasp brakes, with cast iron brake pads acting directly on the tread of the wheel(s) via compressed air actuation. Air suspension is provided for additional passenger comfort and refinement.

Journeys and Routes

Data collection took place on the Great Eastern (Route 7) and West Anglia (Route 5) regions of the UK's rail network. The strategic 'backbone' of the Great Eastern region is the Great Eastern Main Line (GMEL) originating from London Liverpool Street and travelling North East to Norwich. There are numerous branch lines attached to

the GMEL providing services to commuter areas and freight hubs. Data for the present analysis was derived from three journeys between the towns of Ipswich and Felixstowe, a distance of 25 km, over which there are 14 AWS sites.

Automatic Warning System

The purpose of the Automatic Warning System (AWS) is described thus (McLeod, Walker & Moray, 2003):

"AWS serves two functions. The first function is to provide an audible alert to direct the driver's attention to an imminent event (such as a signal or a sign). The second function, linked to the first, is to provide an on-going visual reminder to the driver about the last warning. [AWS] *is there to help provide advance notice about the nature of the route ahead, and thus communicate to the driver the need to slow down or stop"* (p.4).

AWS alerts and reminders are triggered by an electro-magnetic device placed between the tracks approximately 200 yards prior to the signal, sign or other event to which it refers (although this distance can vary according to local circumstances). Sensors underneath the train detect the presence of magnetic fields and activate AWS accordingly. AWS has three system states:

State 1 – No Additional Action Required.

If the event to which an AWS activation is caused requires no action by the driver (such as a signal showing a green aspect) then a bell or simulated chime (at 1200Hz) will sound briefly. The visual display will also remain inactivated. The driver behaviour in this case is to proceed; there is no requirement to cancel AWS using the cancellation button nor is there any specific need to enact driving behaviours in addition to those that are current or planned.

State 2 – Attention is Required Towards Some Imminent Signal, Sign or Event.

If the event that activates the AWS requires (or potentially requires) the driver to slow down or stop then a steady alarm or horn sound (at 800Hz) will sound. The visual display will activate (turning from an all-black display to yellow and black, known as the 'sunflower').

State 3 – Acknowledge (and Continue to be Reminded of) the imminent or previous signal, sign or event.

The act of cancelling AWS (by pressing a button) stops the horn sound, and the sunflower continues to be displayed. Failure to cancel AWS within approximately 2 seconds leads to an immediate emergency brake application which cannot be cancelled (Railway Safety, 2001). This level of braking may cause some discomfort to passengers and the event will be logged on the OTMR equipment. Repeated failure to cancel the AWS within the time allowed is likely to lead to an investigation followed by some remedial training and coaching of the driver.

AWS is a legacy system that was originally conceived as a means to prevent Signals Passed At Danger (SPADs). Several major accidents resulted in the use of AWS, and the number of events it now refers to, being extended. AWS now provides warnings in six circumstances, for example, certain types of speed restriction, some level crossings, SPAD indicators and so on.

Unfortunately, the simple two state warning (bell/horn) and reminder (black/yellow) is unable to discriminate between these six different events. There are approximately 29,000 AWS sites around the UK railway network, which equates to a mean of 1.6 AWS indications received in the train cab every 1.6 route miles, or 2.7 activations (either a bell or horn sound) every minute when travelling at 100mph.

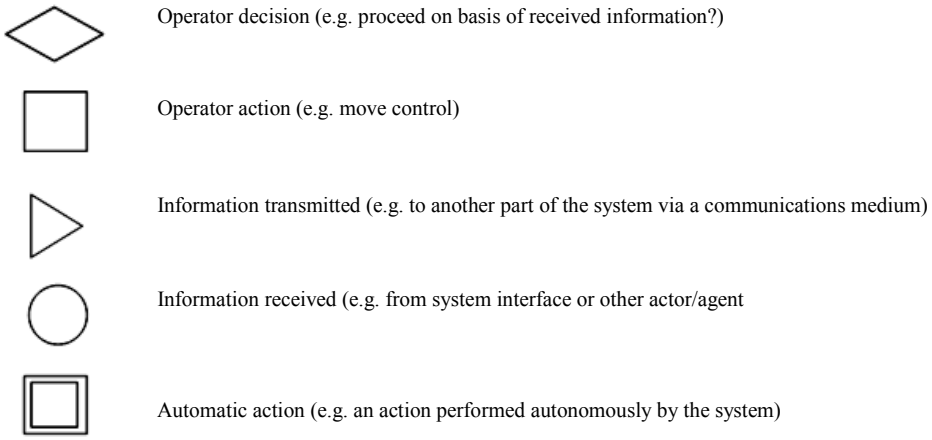
Many warning indications require no action from the driver, simply a press of the cancellation button. Many other warnings occur in situations when the correct behaviour at that time is to accelerate the train (it is moving at slow speed or departing from a station, for example). The task of the driver, therefore, is not a simple one of hearing the warning and pressing a button. It requires them to interpret the source of the warning and the context in which it is occurring, and then decide on the correct course of action. The confusion that this could cause for drivers has been cited in several accident inquiry reports (e.g. Cullen, 2001).

RESULTS AND DISCUSSION

Process charting techniques are used to represent complex real-world activity in an easy to read graphical format using standardised symbols and layout. The Process Chart methodology has a long application history, with early examples dating from the 1920's (e.g. Gilbreth & Gilbreth, 1921). It has been used extensively in military and high hazard domains as a way of understanding the interaction between people and systems, particularly in terms of identifying human error potential. The method has been used in both rail and aviation settings before. In this application, process charts offer a novel way of converting raw 'trace plots' derived from data recorders into an alternative representation, one that makes it easier to:

- Discern how larger journey phases break down into smaller component activities.
- The order and timing that component activities occur.
- Who is performing what activity.
- The presence of distinct activity clusters.

The 72 parameters extracted from the data recorder were classified into:



Once classified the process chart itself was constructed. This involved creating a timeline and columns for each 'agent' in the system. In the case of the railway example six such agents/columns have been used (Figure 1). As different recorder channels become active, the corresponding process chart symbol is inserted into the relevant column at the correct point on the timeline. The sequence of activities and their dependence on each other defines when these symbols are linked. Thus an activity/symbol that occurs after another activity/symbol becomes linked 'vertically'. Activities that are performed concurrently are linked 'horizontally'. Figure 1 shows the 'normative' structure of operations associated with receiving and cancelling an Automatic Warning System alert.

Having the ability to detect behavioural clusters grants the opportunity to assess whether such structures are typical or atypical. Indeed, whether they are one of a number of different behavioural responses within a wider repertoire, and whether one cluster of behaviour is implicated in risk outcomes more than another, and under what circumstances. From the collected data three different 'clusters' were detected. The first cluster is the normative 'perceive-decide-act' sequence. Here the infrastructure on the track triggers an in-cab warning horn. The driver perceives (hears) this, has enough time to classify it (0.89 seconds) and respond by pressing the cancellation button. The second cluster is the 'predictive cancel' sequence. In this case the infrastructure on track triggers the in-cab warning horn but the driver responds so quickly that it is not possible to have perceived, classified and responded to the warning. Instead, the driver has seen the track infrastructure and has anticipated the in-cab warning and timed

their response to coincide with it starting. The third cluster is the 'multiple predictive cancel' sequence. As in cluster two, the driver can see the track infrastructure ahead and is pressing the cancellation button numerous times before hearing the in-cab warning horn, and several times after the warning has sounded and been cancelled.

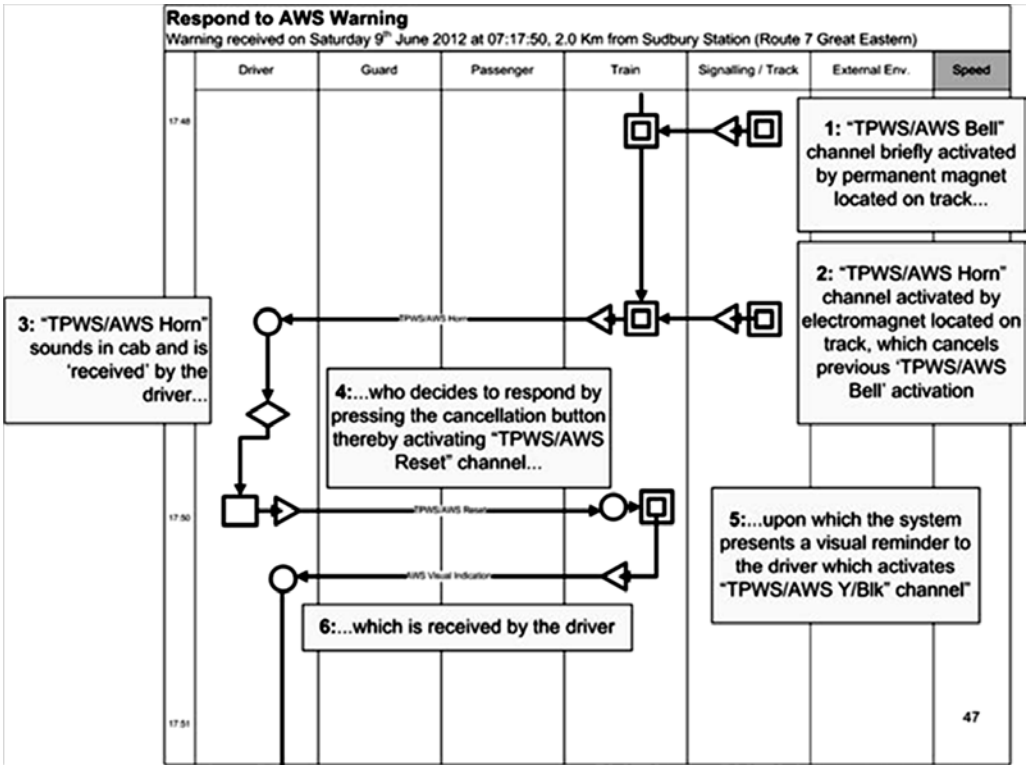


Figure 1. Annotated process chart showing the type and sequence of operations required to correctly respond to an AWS warning

Response Bias

Based on the analysis above it is clear, firstly, that cancelling an AWS warning is not merely a perceptual one of hearing and seeing the different system alerts and indications, it is also cognitive: drivers' not only have to discriminate a 'stimulus' from within a 'noisy' environment, but correctly classify it and respond. Secondly, there are different strategies that drivers employ to perform this apparently simple task. Signal detection Theory (SDT) helps to untangle these different aspects by separating out a person's sensitivity to stimuli (how easy it is to detect something) and their response bias (their preference for responding one way or another to the stimuli). SDT helps us to understand why a particular 'stimulus', which might be very loud, visible or unambiguous, is not always responded to in the ways we expect (or vice versa). Signal Detection Theory classifies human responses to stimuli in the environment in four ways. The responses that drivers made within each of these categories are shown in Table 1.

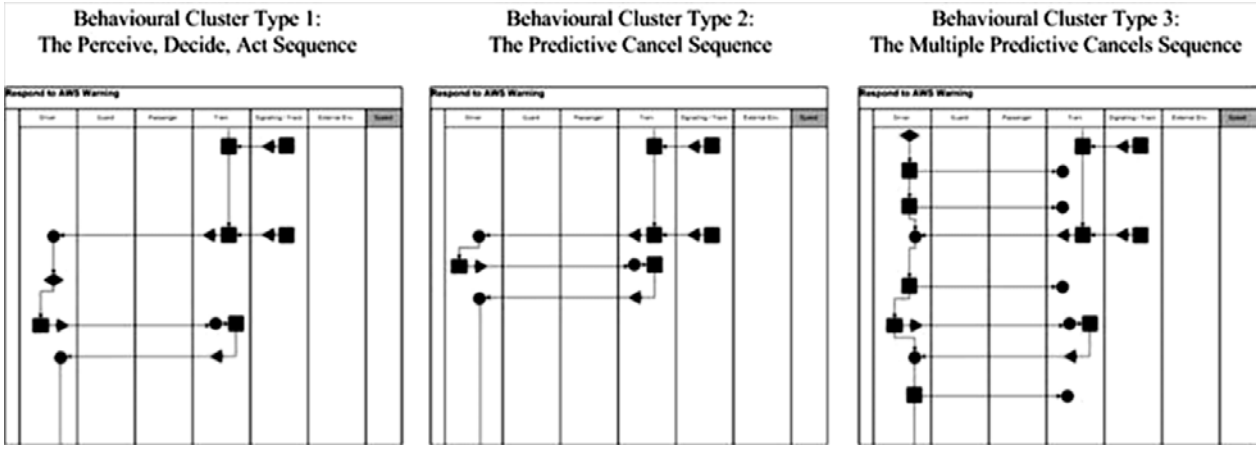


Figure 2. Three clusters of behaviour associated with cancelling an Automatic Warning System (AWS) alert were detected

Table 1. Driver responses to the Automatic Warning System (AWS) can be characterised according to the Signal Detection Theory (SDT) paradigm as follows

Status of AWS/TPWS System	Categorisation	Number	%
AWS Horn followed by reset	Hit	21	30%
AWS Horn followed by no response	Miss	0	0%
AWS Bell followed by reset or no activation followed by reset	False Alarm	22	31%
AWS Bell followed by no response	Correct Rejection	28	39%
	Totals	71	100%

The ability to accurately detect stimuli in the environment and correctly classify them is the desired outcome. Under the Signal Detection Theory (SDT) paradigm this requires a high number of hits *and* a low number of false alarms. For example, if the reset button was pressed in response to ANY warning indication this will ensure a 100% Hit rate but will also increase the rate of False Alarms. Accuracy in this case is low. If, on the other hand, the driver is trying to do the opposite, to avoid False Alarms and instead maximise Correct Rejections, they would not respond to ambiguous 'signals'. This would increase the number of Correct Rejections but it would also increase the number of 'Misses'. Accuracy in this case is also low. Signal Detection Theory enables us to separate sensitivity (d') from decision bias (C). Sensitivity is a measure of accuracy and tells us how easy it is to distinguish a particular environmental stimuli (e.g. an in cab warning). Decision bias tells us whether one response is more probable than another.

Human responses to certain stimuli vary because of incentives embodied in the environment. In train driving, for example, there is a strong incentive in normal operations to prevent unnecessary applications of the emergency brake because an AWS warning was missed. This is because an unwanted brake application cannot be cancelled and is highly inconvenient. On the other hand, the consequences of False Alarms are low because nothing happens when the cancellation button is pressed in error. Because of this it would be expected that train drivers might adopt a 'liberal' decision bias and be willing to accept False Alarms (redundant presses of the cancellation button) in order to achieve the maximum number of 'Hits' . Added to this is the discriminability of the 'stimulus' people are responding to, a stimulus that occurs in a 'noisy' real-world environment. By noise we refer to other stimuli, competing demands and distractions in the environment as well as the background noise inherent in human perceptual and cognitive processes. Decision bias and sensitivity interact with these 'noisy' transport environments to make some future responses more likely than others, in ways that are not always immediately apparent. For example, highly visible warnings that are apparently 'missed', or control actions that are at odds with the situation.

Sensitivity

Sensitivity to a stimulus is given by the metric d-prime, which was calculated as follows:

$$d' = z(FA) - z(H)$$

where z(H) is the number of Hits expressed as a z-value subtracted from the same Z-transformed False Alarm rate. The results obtained are shown in Table 2:

Table 2. Driver's sensitivity to the Automatic Warning System (AWS) alerts

Journey	Hits	Misses	Correct Rejections	False Alarms	d-prime
1	5	0	11	0	4.65
2	9	0	6	7	2.30
3	7	0	11	15	2.13
				Mean	3.03

The d-prime figure measures the strength of the stimulus, which in this case is an AWS warning. A value of 3.03 indicates that drivers are highly sensitive to it: in this situation it is unambiguous and easy to discriminate from the wider background of noise, distractions, other contextual factors etc. Expressed more formally, the responses drivers' are providing when an AWS warning is overlain on top of the 'contextual noise' is 3.03 standard deviations 'different' from the responses they give when the signal is absent (and only the 'contextual noise' is present). Sensitivity provides an important leading indicator concerning the discriminability of information needed for driver's to develop accurate situational awareness. The same 'stimuli' may yield different levels of sensitivity depending on external/contextual factors. A warning that was not expected, ambiguous, not fully understood or masked may lower sensitivity.

Decision Criterion

Decision bias/criterion is given by the metric c, which was calculated as follows:

$$c = -\frac{Z(H) + Z(FA)}{2}$$

The results obtained are shown in Table 3. Decision Bias is independent of sensitivity and relates not to the discriminability of the 'signal' but to the payoffs involved in making one response in favour of another. Thus, regardless of how easy it is to discriminate a stimulus a counter intuitive response may still be favoured. This is because the consequences of False Alarms, Misses and Correct Rejections vary with the context. Psychological research shows that decision bias is more unstable and situationally dependent than sensitivity and, therefore, a potentially valuable Leading Indicator.

Table 3. Results of decision bias (c)

Journey	Hits	Misses	Correct Rejections	False Alarms	Criterion
1	5	0	11	0	0
2	9	0	6	7	-1.21
3	7	0	11	15	-1.26
				Mean	**-1.24**

The mean decision bias value across the three sampled drivers was c = -1.24 which indicates a liberal bias. Driver's make more responses that indicate the AWS signal is present than it is absent. In other words, they are prioritising False Alarms over Correct Rejections which, in turn, provides a clue as to the sorts of error that may be more likely to occur in future (i.e. warnings that are cancelled incorrectly). Assuming that drivers' 'internal responses' to the AWS warning are normally distributed (as per Signal Detection Theory) it is possible to plot individual driver decision bias' into the chart below which provides an important diagnostic tool in defining risky psychological/decision making states.

According to Figure 3, Driver 1 shows no systematic bias in their responses to the AWS warning. They respond correctly to the AWS warning on every occasion and his/her False Alarm rate is zero. Drivers 2 and 3 are different. They are exhibiting a strong 'liberal response bias' meaning that they are much more inclined to exhibit 'false alarm' responses (and behavioural clusters 2 and 3). With the ability to detect these changes in decision bias comes the possibility to analyse a) the extent to which different biases interact with accident/incident rates (i.e. is a liberal bias of this magnitude associated with particular types of risk) and b) how the context influences human decision making (and therefore how that context can be modified to 'un-bias' human responses).

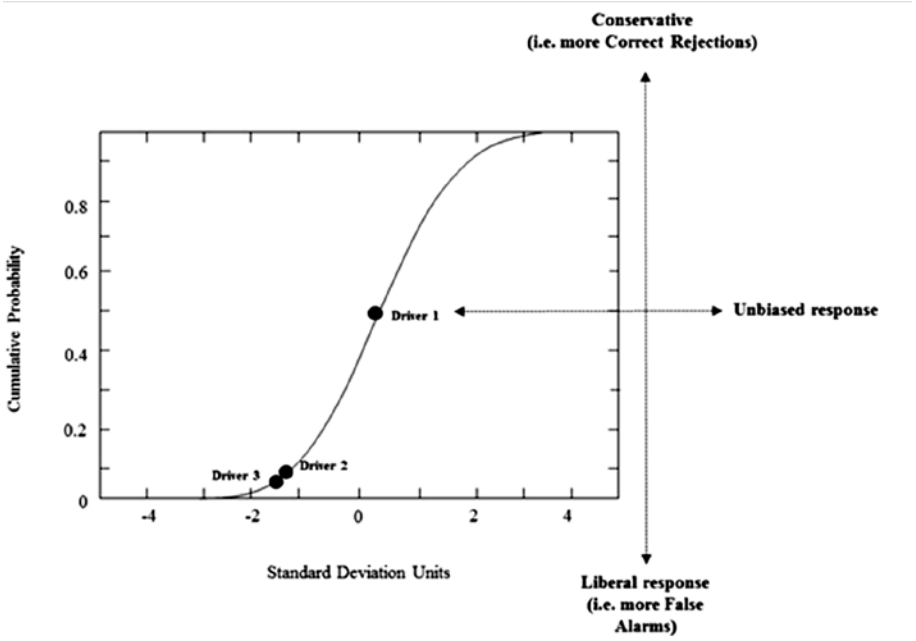


Figure 3 – Cumulative probability plot showing the extent and direction of response biases exhibited by the drivers.

CONCLUSIONS

This paper has described how the outputs from transport data recorders can be coupled to existing Human Factors methods to provide advanced indication of strategically important Human Factors risks. The 'black box paradox' is that the opportunities to use these devices for their original post-accident purpose are diminishing at the same time as their technical capability and data richness are increasing. In addition, the types of risks are changing, with progress in technical areas of reliability and performance exposing issues around human-machine system interaction. To continue to ensure safety in the face of increased risk exposure and operational intensity, better use of this data needs to be made. This paper demonstrates how black box data from the rail sector can be turned into useful 'information' in the form of Human Factors leading indicators of risks associated with the use of an in-cab warning device. Sensitivity provides a measure of how much useful information there is in the environment and the extent to which drivers can discriminate it from the background of contextual noise. Warnings, stimuli and so forth may, in an engineering sense, appear to be unambiguous, yet they may be considerably less so cognitively. Sensitivity provides a measure of this which can, in turn, be associated with changing risk. Decision bias reveals the

likelihood that one type of driver response will be favoured and how this interacts with risk. In a wider application it would be possible to examine decision bias in a systematic way looking at differences between drivers and between particular routes. This could provide insight into driving styles and indicate whether particular aspects of a route result in a shift in decision bias. For example, a specific AWS signal on a particular route may result in a high level of predictive pressing (high false alarms) relative to most others, identifying this as a more risky section of journey. Relationships such as these would need to be established based on large-scale future research but even on a smaller sample of data the method was able to detect potentially important differences between drivers, with some adopting a much more liberal response bias than others. The principle, however, is a much more important one. Human Factors methods like these can accept recorder data as an input, are amenable to the kind of software implementation that would be required in a full-scale application, and point the way towards Leading Indicators of strategically important Human Factors risks.

ACKNOWLEDGEMENTS

This work was funded by the UK Economic and Physical Sciences Research Council (EPSRC) under Grant Reference: EPSRC EP/I036222/1. The involvement and support of the UK Civil Aviation Authority, the Rail Safety and Standards Board (RSSB), the Association of Train Operating Companies (ATOC), Greater Anglia Trains and Aerobytes Ltd. is also gratefully acknowledged.

REFERENCES

CAA (2003) CAP 739*: Flight Data Monitoring. A Guide to Good Practice*. UK: Civil Aviation Authority.

CAA (2011) *Safety Plan 2011 – 2013*. UK: Civil Aviation Authority.

Campbell, N. (2007). *The Evolution of Flight Data Analysis*. In: ASASI 2007: Proceedings of Australian Society of Air Safety Investigators conference, Civic Square, Australia. Available online at: http://asasi.org/papers/2007/The_Evolution_of_Flight_Data_Analysis_Neil_Campbell.pdf

Clegg, C.W. (2000). Sociotechnical principles for system design. *Applied Ergonomics, 31*, 463–477.

Cullen, W.D. (2001). *The Ladbroke Grove Rail Inquiry: Parts 1 and 2 Report*. London, UK: HSE Books.

Department for Transport (DfT) (2011). *Department for transport statistics*. Retrieved 12 July 2012 from: http://www.dft.gov.uk/statistics.

European Transport Safety Council. (2003). *Transport safety performance in the EU: A statistical overview*. Brussels: European Transport Safety Council. Retrieved July 17, 2012, from: http://www.etsc.be/documents/statoverv.pdf.

Evans, A.W. (2011). Fatal train accidents on Europe's railways: 1980-2009. *Accident analysis and prevention, 43*(1), 391-401.

Ford, M. (2012, February). *FDM Technology and Data Recovery*. Presented at FDM Short Course, Cranfield University, UK.

Gilbreth, F. B. & Gilbreth, L. M. (1921). *Industrial efficiency*. New York: American Society of Mechanical Engineers.

McLeod, R. W., Walker, G. H., & Moray, N. (2005). Analysing and modelling train driver performance. *Applied Ergonomics, 36*, 671-680.

Mitchell, W.C., & Burns, A.F. (1938). *Statistical indicators of cyclical revivals*, NBER Bulletin 69, New York; reprinted as Chapter 6 of G.H. Moore (ed.) Business Cycle Indicators, Princeton, NJ, Princeton University Press, 1961.

Morcom, A.R. (1970). Flight Data Recording Systems: A brief survey of the past developments, current status and future trends in flight recording for accident investigation and operational purposes. *Aircraft Engineering and Aerospace Technology, 42* (7)12 – 16.

Office of Rail Regulation (ORR) (2012), *National rail trends yearbook, 2010 -2011*. London: ORR.

Qureshi, Z.H. (2007). *A review of accident modelling approaches for complex socio-technical systems*. In Proceedings of the twelfth Australian workshop on Safety critical systems and software and safety-related programmable systems, 86, 47-60. Adelaide, Australia: Australian Computer Society.

Rail Safety & Standards Board (RSSB) (2009). *The railway strategic safety plan 2009 – 2014*. London: Rail Safety and Standards Board.

Railway Safety (2001). *Automatic warning system (AWS) GE/RT8035*. London: Railway Safety

Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science, 27*, 183-213.

Rogers, A., Evans, R., and Wright, M. (2009). *Leading indicators for assessing reduction in risk of long latency diseases*. Greenstreet Berman Limited for Health and Safety Executive.

Step Change (2003*). Leading performance indicators: guidance for effective use*. Aberdeen: Step Change in Safety.

Uff, 2000

Warren, D.R., (1954). *A device for assisting investigation into aircraft accidents*. ARL-MECH-ENG-TECH-MEMO-142.

Wolmar (2012, May) *Rail Magazine, 697: safety is the hidden success story* http://www.christianwolmar.co.uk/2012/05/rail-697-safety-is-the-hidden-success-story

Woods, D.D. & Cook, R.I. (2002). Nine steps to move forward from error. *Cognition, Technology, and Work, 4*(2), 137- 44.

Key Players and Their Interactions Within the Young Driver Road Safety System: Perspectives and Partners

Bridie Scott-Parker[a], Paul Salmon[a], Natassia Goode[a] and Teresa Senserrick[b]

[a]University of the Sunshine Coast Accident Research
University of the Sunshine Coast
Sippy Downs, QLD 4558, AUS

[b]Transport and Road Safety
University of New South Wales
Sydney, NSW 2052, AUS

ABSTRACT

Young driver road safety has persisted as a global problem for over 50 years, despite a plethora of intervention. Interventions have emerged from a driver-centric approach, and as such predominantly focus upon "fixing-the-driver". An alternative approach with traction in other safety critical, high risk, domains is systems thinking, which recognises there are a multitude of influential actors (eg, government agencies, community groups) within the broader road safety system that influence young driver safety. Premised on a description of the young driver road safety system (YDRSS) that describes the actors across six levels, interviews were held with 42 actors important in young driver road safety. Synthesis of semi-structured interview transcripts revealed some horizontal integration (interaction amongst actors in the same level) and limited vertical integration (interaction between actors in lower and higher levels), suggesting the current YDRSS is ineffective. A systems approach integrating findings from the driver-centric approach whilst deepening understandings of the complex web of interrelationships amongst actors offers a unique solution to a complex problem. Thus the interviews have revealed avenues of intervention in the YDRSS that can lead to improvements in their road safety and in the road safety of all who share the road with them.

Keywords: Systems Thinking, Road Safety, Young Driver, Actor, Intervention

INTRODUCTION

Road crashes are the leading cause of death for persons aged 15-29 years (World Health Organization, 2013). The earliest period of independent driving (typically ages 16-18 years) is associated with the most risk. This higher level of risk has been attributed to driving inexperience (McCartt et al., 2009) and age- or development-related factors (Johnson & Jones, 2011). Accordingly, young driver road safety interventions have primarily focused upon "fixing the driver" through mechanisms like driver education and training, and graduated driver licensing (GDL), which is a system designed to allow the novice to commence driving in lower risk conditions, allowing "graduation" with increased experience and maturity. Whilst GDL has proven to be the most effective intervention to date in reducing the young driver road toll, this appears principally to be through exposure control, and a point of diminishing returns seems to have been reached (e.g., Williams, 2011). Rather than persisting in "fixing the driver", Scott-Parker et al. (in press) have argued that it is time to reform how young driver road safety is understood and managed by imple-

menting systems thinking which seeks to address the factors that influence and shape young driver behaviour, thus building upon and integrating knowledge and experience gained through previous research.

To comprehensively understand the young driver road safety system (YDRSS), Scott-Parker et al. (in press) operationalised a popular systems-orientated model of accident causation, Rasmussen's (1997) risk management framework (RMF) in the realm of young driver road safety. Figure 1 depicts Rasmussen's RMF's characterisation of the social, organisational and technical components influencing safety, applied within the YDRSS. It is important to note that for the *young driver road system* to maintain *safety*, vertical integration is vital: that is, actions and decisions at higher levels should influence the actions at lower levels, whilst information generated at lower levels should inform actions and decisions at higher levels (Rasmussen, 1997; Svedung and Rasmussen, 2002). The key drivers of vertical integration amongst the actors (eg, government agencies, community organisations, high schools) at the various levels are communication and feedback.

Government policy and budgeting Decisions and actions regarding young driver road safety including the development of behaviour-regulating laws and legislation; provision of funding for interventions; policy development.
Regulatory bodies and associations Conversion and informing of young driver legislation by regulatory bodies, research organisations and others with financial interest in young driver road safety.
Local area government, planning, budgeting, company management Integration of government policy including road rules into driving. Parents are the 'local government' by virtue of intra-familial dynamics.
Technical and operational management Other influential and authoritative bodies and organisations with a direct influence on young driver behaviour and decision-making.
Physical processes and actor activities The young driver themselves and the psychosocial influences upon their driving behaviour.
Equipment and surroundings The actual physical environment in which the young person drives.

Figure 1. Rasmussen's risk management framework in the context of the young driver road safety system (Rasmussen, 1997)

Scott-Parker et al. (in press) classified the contributing factors in young driver crashes identified in previous research according to Rasmussen's RMF. Perhaps unsurprisingly given the predominant fix-the-driver approach, the majority of factors were positioned at the *physical processes and actor activities* level (level 5). Classifying the young driver road safety countermeasures according to Rasmussen's RMF similarly revealed a predominance of interventions that focus directly on changing young driver behavior through training or education. Finally, to gain an understanding of who exactly is sharing the responsibility for young driver safety, an actor map using Rasmussen's RMF was created that identified key actors in young driver road safety in the Queensland context. Building on this, the authors conducted a study that attempted to identify (a) the roles of, and (b) the interactions between, the multiple actors identified in Queensland's YDRSS, and more broadly across Australia as appropriate. The aim of this paper is to report the findings of this study, which took the path less travelled by focusing upon how actors other than the young drivers themselves view the YDRSS.

METHOD

Participants

Forty-two participants represented 25 actors (see Figure 2) identified as playing a key role in the YDRSS. Frequently one individual spoke for an actor, therefore individual opinions/experiences are likely to have been represented.

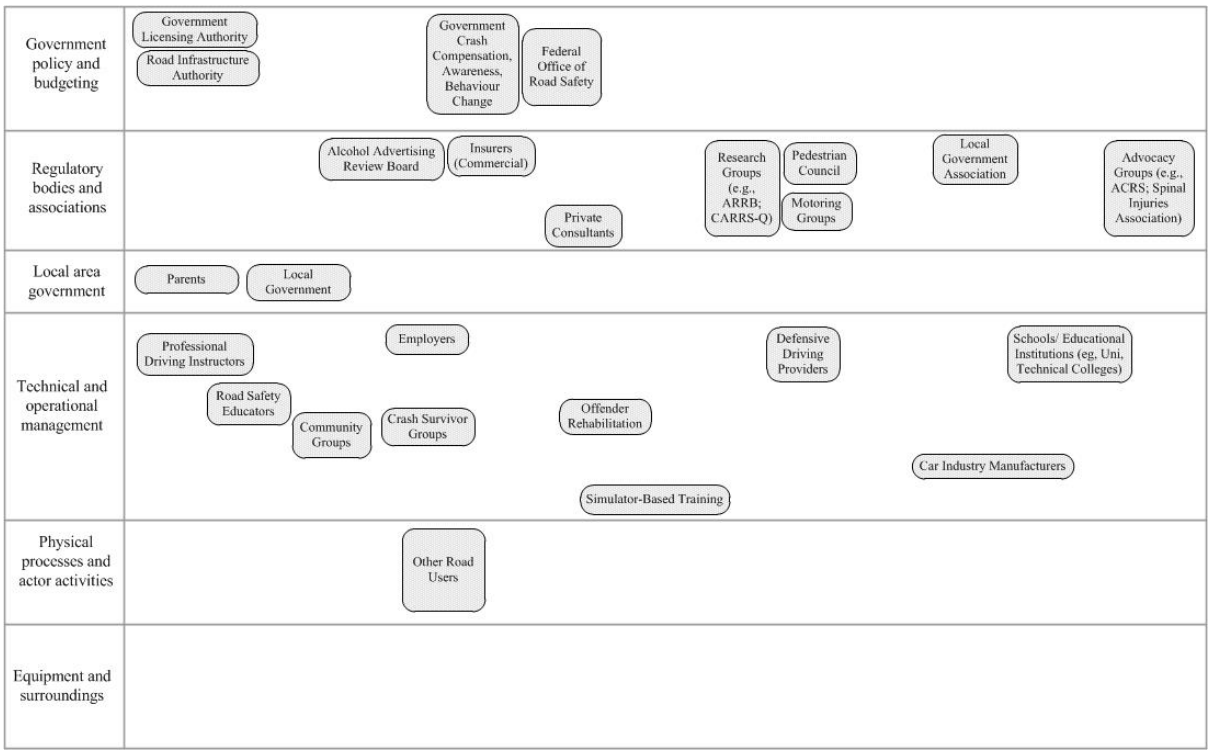


Figure 2. Actors who participated in the research project

Design, Procedure and Analyses

Emails were sent (and if required, online enquiries were submitted) to more than 140 individuals and organisations, in Queensland specifically and more broadly across Australia, inviting participation in the study. In the instance of government departments and larger non-government organisations, multiple individuals were contacted using the extensive young driver networks of the first and fourth authors. Telephone and semi-structured face-to-face interviews were conducted by the first author at mutually-convenient times. Questions included "*What role does your organisation play in the YDRSS?*", "*Who are the other organisations or stakeholders (actors) in the YDRSS?*", "*Which of these organisations/stakeholders do you interact with?*", and "*Which of these organisations/stakeholders don't you interact with*?" Finally, participants were asked if they could suggest what was needed to improve young driver road safety. The question "*What are the roles of these other organisations/stakeholders in the YDRSS?*" was abandoned early in the interview process as actors struggled to respond. Mp3 recordings of interviews facilitated verification of the accuracy and completeness of notes taken during the interviews. Perceived roles of, and interactions between, actors were synthesised and summarised according to the respective Rasmussen's RMF sociotechnical level. The study was approved by the University of the Sunshine Coast Human Ethics Committee (A/13/510).

RESULTS: ACTOR ROLES AND INTERACTIONS

Delineating interactions and non-interactions at the individual actor level, and more generally at a macro sociotechnical level, are vital to understanding and effectively managing the YDRSS. As such, the first step is to more fully understand who the actors are, and at what level in Rasmussen's they are positioned. Based on the participant responses, an Actor Map was created positioning the 79 identified actors (see Figure 3).

Further, Table 1 summarises the interactions and identified non-interactions among and within the various sociotechnical levels of Rasmussen's RMF reported by the participants. The interactions and the roles of the participating actors will be elaborated upon in the following paragraphs.

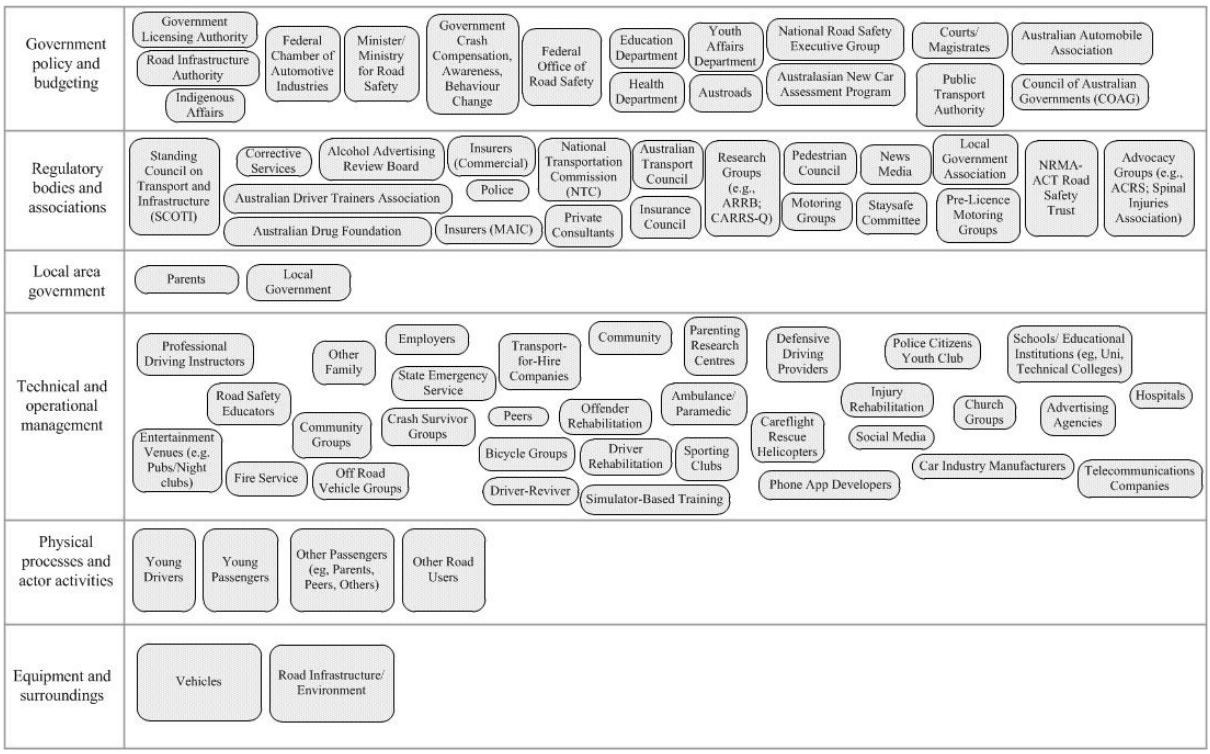


Figure 3. Actor Map of the young driver road safety system

Table 1: Summary of interactions reported by actors in the young driver road safety system

RMF Level	Other Identified Actor	RMF Level of Actor Reporting Interaction					
		1	2	3	4	5	6
1	Government licensing authority (State) *	✓	✓	✓	✓	X	
	Road infrastructure authority (State) *	✓	✓				
	Courts/ Magistrates *	X	✓		✓		
	Minister/ Ministry for Road Safety *	✓	✓				
	Council of Australian Governments *	✓					
	Austroads *		✓				
	National Road Safety Executive Group*	✓	✓				
	Department of infrastructure and Regional Development	✓	✓				
	Government Crash Compensation, Aware-ness, and Behaviour Change	✓	✓				
	Federal Office of Road Safety	✓	✓				
	Department of Youth Affairs *						
	Department of Health *	✓	✓				
	Department of Education *	✓	✓		✓		
	Indigenous Affairs *		✓				
	Australian Automobile Association *	✓	✓				
	Australasian New Car Assessment Program*						
	Federal Chamber of Automotive Industries *						
	Public Transport Authority *	✓					
2	Standing Council on Transport and Infra-structure *	✓					

	National Transportation Commission *						
	Alcohol Advertising Review Board		✓				
	Police *	✓	✓	X	✓	X	
	Corrective Services *		✓				
	Insurers (Commercial)	X	✓	X			
	Insurers (Government) *	X	✓		✓		
	Advocacy Groups		✓		✓		
	Research Groups	✓	✓		✓		
	Motoring Groups	✓	✓	X	✓		
	Staysafe Committee *		✓				
	News Media *	✓	✓		✓		
	Australian Transport Council *						
	Private Consultants	✓	✓				
	Australian Drug Foundation *	✓					
	Pre-licence Driving Motoring Groups (speedway) *	X					
	Insurance Council *						
	NRMA-ACT Road Safety Trust *		✓				
	Pedestrian Council						
	Australian Driver Trainers Association*				✓		
	Local Government Association	✓	✓				
3	Parents	✓	✓	✓	✓	X	
	Local Government	✓	✓		✓		
4	Other family					X	
	Peers *		✓		✓	X	
	Schools, Other Educational Institutions	✓	✓	X	✓	X	
	Professional Instructors	✓	✓	X	✓		
	Smart Phone App Developers *						
	Road Safety Educators	✓	✓	✓	✓	✓	
	Offender Rehabilitation		✓	✓	✓	✓	
	Employers *				✓		
	Car Industry Manufacturers				✓		
	Sporting Clubs *				✓		
	Community Groups	✓	✓		✓		
	Entertainment Venues *		✓				
	Social Media *		✓				
	Advertising Agencies *	✓	✓			X	
	Parenting Research Centre *	✓					
	Defensive Driving Providers	X	✓	X			
	Church Groups *	X					
	Hospitals *	✓	X				
	Ambulance *	✓		X	✓		
	Rehabilitation – Driver Behaviour	X	✓		X		
	Fire Service *	✓		X	✓		
	State Emergency Service *						
	Off-road Vehicle Groups *		✓				

	Transport for Hire Companies*	✓					
	Bicycle Groups	✓					
	Police Youth Citizens Club *		✓		✓		
	Careflight Helicopters (sponsored by Motoring Group)		✓				
	Crash Survivor Groups		X		✓	X	
	Driver-Reviver Stops *		✓				
	Telecommunications Companies *			X		X	
	Injury Rehabilitation *				X		
	General Community	✓	✓				
	Simulator-Based Training				✓	✓	
5	Young drivers *	✓	✓	✓	✓	✓	
	Young passengers *	✓	✓		✓		
	Other passengers *		✓				
	Other road users		✓			✓	
6	Vehicles		X		X		
	Road infrastructure *	✓					

X Actor(s) stated aware of other actor but do not interact with actor within young driver road safety system.

✓ Actor(s) stated interact with other actor within young driver road safety system.

☐ Indicates that interaction between actors was not mentioned.

* Actor did not participate in research (but was mentioned by other Actor participants).

Level 1: Government Policy and Budgeting

The government level includes agencies, organisations and stakeholders that are involved in the development of behaviour-regulating laws and legislation, provide funding for interventions, and develop young driver policy at a state and/or federal level. As can be seen in Figure 3, 17 level 1 actors were identified as playing a role in the YDRSS, the government departments and bodies operating at either a state or federal level. Four individuals represented the four participating actors in level 1. Half of these participants stated that they did not play a direct role in the YDRSS; rather, they believed that their role in the system was through indirect efforts undertaken to assist other actors with a direct role in YDRSS, and as such the majority of their interactions occurred with other level 1 actors. Indirect efforts reported by all participants included funding education programs that target young drivers (e.g., SDERA, http://www.sdera.org.au/, Keys 2 Drive, https://www.keys2drive.com.au/default.aspx); an advisory role in which they make evidence-based recommendations to government Ministers and departments; and educating the broader driving community through mechanisms such as the media regarding driving risks for all. Actors with a national scope advised that the commonwealth role in GDL focuses upon higher level policy focus. Similar to the experience of the United States (http://www.ghsa.org/html/stateinfo/laws/license_laws.html), in Australia, there is no national GDL program, rather each state and territory is responsible for the individual components at each level (Senserrick, 2009), and the enforcement of both general and GDL-related driving rules. Despite this a number of actors operating at a state/territory level (Australia is comprised of six states and two territories) advised that a national best-practice GDL program has been under consideration. As can be seen in Table 1, the majority of interactions involving actors from level 1 involve other actors in level 1, representing horizontal integration, followed by interactions with actors in level 2 and level 3.

Level 2: Regulatory Bodies and Associations

The regulatory bodies and associations level includes agencies that convert and inform young driver legislation, with agencies frequently having a vested financial interest in young driver road safety. As can be seen in Figure 3, 21 level 1 actors were identified as playing a role in the YDRSS, with 16 representatives of 6 level 2 actors sharing their perspectives regarding their role in the YDRSS. A number of themes emerged. Whilst a more active role was reported by actors in level 2 than by actors in level 1, many indirect roles were also reported, in both pre-crash and post-crash phases (see Haddon's matrix; Runyan, 2008). Some actors emphasised their capacity for an independent voice, i.e., that they could 'freely speak' for young drivers as they were not obligated by funding arrangements, in

contrast to other actors who relied upon government funding for day-to-day operations and/or specific research and other activities. The majority of actors were actively engaged more broadly in road safety research, often including young driver road safety research specifically, and research ranged from commissioned to higher degree research projects. At this level there was also an emphasis upon translating research into policy and practice through advocacy, education, media, publications, community engagement, business decisions such as reduced insurance premiums, and funding activities of community groups. Much focus remained at the state and local level, with one actor only identifying a role at an international level, in which there were collaborations with prominent research centres based in the United States (such as the Centers for Disease Control and the University of Michigan Transportation Research Institute). Contrasting to the majority of the interactions reported by the level 1 actors, level 2 actors reported greater vertical integration with other actors in the YDRSS, including upwards to level 1 actors and downwards to levels 3, 4 and 5 actors in particular (see Table 1).

Level 3: Local Area, Government, Planning and Budgeting

Level 3 includes agencies involved in integrating government policy, such as general road rules and GDL, into driving, and two actors were identified as playing a role in the YDRSS (see Figure 3). Perhaps unique to the safety critical domain of young driver road safety, intra-familial power and authority means that parents of young drivers are perceived as operating as a form of local government for their child. Two of the four participating level 3 actors were parents (one male) of Learner drivers who shared their perspectives regarding their position and role in the YDRSS. The role of parents was seen as primarily being operationalised in the motor vehicle as their Learner child drove. Roles included a combination of active direction and guidance in driving logistics (such as safe vehicle use, safe road use, and safe interaction with fellow drivers); education regarding driving risks and managing these (such as maintaining lane position on single lane country roads); and reinforcement of safe driving messages and behaviours (such as 'don't drink and drive'). Parents were generally isolated nodes within the YDRSS, with limited interaction with other actors in any other level, and thus limited vertical and horizontal integration in the YDRSS. In contrast, the other two participants representing local government actors interacted with other local, state and commonwealth departments and ministers in focusing upon building capacity at the local level to address road safety generally whilst developing statewide networks, representing upwards vertical integration. Engineering, education and enforcement were reported as key components in the YDRSS, and as such these components also guided funding expenditure, further exemplifying vertical integration within the YDRSS (see also Table 1).

Level 4: Technical and Operational Management

The technical and operational management level comprises actors who are involved in crash prevention and post-crash efforts, and as such can directly influence young driver road safety, including organisations such as crash survivor and community groups, and rehabilitation providers. As can be seen in Figure 3, 33 level 4 actors were identified by the participants. Ten participating representatives of 16 actors in level 4 reported a mix of prevention and post-crash roles in the YDRSS. Prevention roles principally involved instruction and education. A breadth of practical, on-road driver training and instruction providers reported that they directly engage with young drivers, including providing education regarding vehicle and body capabilities (e.g., defensive driving providers), community-based supervision (Learner mentor programs), and reinforcement of learning messages off-road (e.g., simulator-based learning). Education of young drivers occurred directly through road safety programs, injury prevention programs (often including presentations by individuals who have sustained a permanent traumatic injury from a road crash), and education as part of secondary health class curriculum. Actors such as vehicle manufacturers believed that they played a role in the YDRSS via such mechanisms as ensuring that young drivers operate as safe a vehicle as possible through vehicle programs such as the Australasian New Car Assessment Program (ANCAP). In addition, there was a strong sense of corporate responsibility, with employers mindful that young employees such as apprentices were also young drivers, and that older drivers indeed were likely to be parents of young drivers, thereby supporting organisational efforts to improve driver safety more broadly. Further, community responsibility was evidenced by actors supporting evidence-based young driver road safety education programs (by providing financial and social capital), in addition to sponsoring community-based road safety awards. These roles suggest some horizontal and vertical integration within the YDRSS by actors involved in crash prevention.

Post-crash roles in the YDRSS included young driver rehabilitation (for purposes such as reducing criminal sentences and to improve future driving behaviour) and restorative justice (which brings together incarcerated offenders and victims of road crashes as part of the healing process for grieving families). For a number of actors involved in the YDRSS during both pre- and post-crash stages, a key role was imparting the message to young drivers that driving

is a privilege and a responsibility, rather than a right to which they are entitled. Pre- and post-crash roles also included advocacy and advisory roles, lobbying actors in levels 1 and 2 for systemic change, particularly decision-makers in level 1 who can directly influence young driver road safety policy, demonstrating upwards vertical integration. Advocacy for safer vehicles exemplifies downward vertical integration. It is noteworthy also that there was some horizontal integration within level 4, such that actors like community groups and road safety educators approached schools as a way to access young drivers (see also Table 1).

Level 5: Physical Processes and Actor Activities

The physical processes and actor activities level includes the young drivers themselves and the way in which they behave on the road. As can be seen from Figure 3, it also includes other influences upon young driver behaviour, such as young and other passengers, and other road users. Three experienced drivers who share the road with young drivers discussed their position and role in the YDRSS. Key roles included teaching younger family members to drive (siblings, children); modelling safe driving behaviour to younger family members; and modelling safe driving behaviour to young drivers, actors stating that it is a responsibility for mature drivers to set good examples, particularly if they see "if he can get away with it (risky driving behaviour) I can too". Interestingly, a participant with substantial country driving exposure (with commensurate high driving speeds, poor roads and poor night visibility) advised that one of their main roles in the YDRSS was to safely get out and/or stay out of the close proximity of young drivers. Thus not only would the personal safety of the experienced driver be maintained as they would not be involved in a crash with a young driver (a key priority), but the safety of the young driver would not be influenced by interaction with the experienced driver. No three level 5 actors reported YDRSS vertical integration (see Table 1).

Level 6: Equipment and Surroundings

Level 6 represents the physical and technological infrastructure surrounding young drivers, including the road infrastructure and the vehicle they drive. This section of Rasmussen's RMF was represented by Car Industry Manufacturers (level 4) and Local Government Associations (level 2) (see Figure 3).

RESULTS: WHAT NEEDS TO HAPPEN TO IMPROVE YOUNG DRIVER ROAD SAFETY?

Participants were asked if they could suggest what was needed to improve young driver road safety, in light of the roles and (non)interactions they had earlier divulged. Interestingly, apart from a handful of actors in Level 1, there was much support for the research project, evidenced as strong consensus regarding the need to:

- Identify who are the actors;
- Delineate the roles played by the actors;
- Clarify the extent/non-extent of interactions amongst the actors;
- Eliminate unnecessary 're-invention of the wheel' through coordination of actors;
- Ensure evidence-based practices and policies are operationalised by actors, and that such information is shared in a manner that can be easily understood by actors;
- Improve the transparency of funding arrangements (and thus identifying potential control of research, policy and practices) for actors;
- Promulgate the need for the safety system of young drivers to incorporate both targeted and general road safety policy and practices by actors;
- Recognise that responsibility for young driver road safety should not be borne by young drivers and their parents alone, and that there are many other highly-influential actors;
- Engage with youth, which was reported as highly challenging for many actors;
- Reconsider state-based approaches when the national problem requires a national approach by actors; and
- Recognise that young driver road safety is not limited to the realm of road safety, and that domains such as parenting and health (e.g., drugs and alcohol) can be highly influential and as such, key actors.

As such, the current research findings regarding identifying actors in the YDRSS, their role, and interactions and non-interactions between actors and within the YDRSS more broadly is vital. Most participants across all levels of the Actor Map advised that there were likely other informal (and thus indirect) interactions with other actors of which they were unaware of, however they could not suggest at the time of the interview exactly what these may entail. There was recognition by most actors in all levels of the concept that "good partnership is important", how-

ever were less certain in how to initiate and maintain such partnerships/collaborative efforts, particularly when there was much unknown about who is out there, and what is currently being done in the young driver space.

Furthermore, most actors in the higher levels (1 and 2 especially) believed that they were "already interacting effectively with everyone" in the YDRSS. The pervasive issue of young driver road crashes is clear evidence against this argument, however. The lack of interaction may simply be because actors are do not consider the breadth of actors involved in the YDRSS, particularly beyond their own level, or what exactly their role should or even could be relative to these; as reflected by one actor who characterised their role as: "so embedded, [yet] so innocuous". Some actors in levels 2 and 4 expressed exasperation with their own "failed attempts" to improve young driver road safety. A "silo mentality" (level 4 actor reflecting upon competitive nature of research and intervention efforts) was reported in the realm of the YDRSS, and "no one thing is a silver bullet" was recognised. Further, a number of actors in levels 2 and 4 reported there was a need for a coordinator of the YDRSS, possibly at a national level.

A number of participants representing actors in level 1 and 2 advised that it was difficult to gain traction and maintain momentum for interventions in the YDRSS, with much interest in GDL reforms during the late 20th/ early 21st century giving way to recent declines of interest in other forms of intervention which could augment and/or improve GDL in an era of competing financial priorities. Actors in level 2 argued that "leaders in the debate are needed to maintain and make changes". Actors in levels 1 through 4 believed schools and other educational institutions were an ideal and oft-under-utilised path to access young people with various interventions to improve their road safety. Teachers (level 4) specifically stated it was difficult to access driver education/ training resources reliably for their students. "Information sharing is essential", with many actors in levels 4 reflecting that "this is not always happening" when they were very keen to operationalise best practice for pre-crash intervention but did not know what that actually entailed. This point is particularly relevant to commissioned and other research conducted by actors in level 2, which typically is targeted at a government or academic audience, rather than a lay audience. A "national centre of excellence" in young driver road safety, including best-practice resources and additional supporting information such as research grants, potential collaborations etc, was suggested as a repository that could be freely and readily accessed by, and relevant to, all actors at all levels of the YDRSS. Competitive funding appears to have fostered competitiveness within the YDRSS, with many passionate actors striving to make a difference on a small budget that may preclude them from integrating best practice, further alienating them from other actors within the YDRSS.

Parents were recognised by actors in levels 1 through 5 as playing an important part in their young drivers' road safety, further supported by the extant peer-reviewed literature. For example, participating parents (level 3) reported (a) that they are usually the most common driving supervisor during the Learner phase; (b) that parents and other road users (level 5) are a model of driving behaviours and attitudes during the Learner phase and beyond (a number of actors in levels 1 and 2 also recognising that parents are a model for a decade and a half before the child actually starts to drive themselves); and (c) that parents can and do reinforce good driving behaviour. However, based on the limited number and variety of interactions reported by parents – despite many actors across the Actor Map reporting they interact with parents – actors in the YDRSS do not seem to be engaging effectively with parents nor sustaining that engagement. In addition, actors in levels 1, 2 and 4 in particular recognised the need for community support (pre- and post-crash actors in level 4) and engagement within the YDRSS; however actors in level 2 who were actively engaged in translating research into practice also noted some past difficulties engaging with the community and were apprehensive about repeating these efforts.

There was some debate regarding what could and/or should be the role of professional driving instructors (level 4), and the extent and nature of actor-interaction with instructors. Representatives of the instructor actor advised that a potential obstacle to changing the role and any interactions with instructors, for example efforts to align instructors with other actors at higher, the same and lower levels through vertical and horizontal integration, would need to consider the competitive nature of the professional driving instruction industry. Specifically, undercutting of costs is commonly reported in an effort to sustain business in tough economic times, and there was also some concern regarding the consistency and completeness of instruction in a (perceived-to-be) largely unregulated industry. Instructors also believed their industry was an "old school network" and thus was not receptive to innovation and change.

Some actors in level 2 (particularly those involved in advocacy and policy and practice translation) and level 4 (particularly those interacting with young people in a number of risk-related domains) expressed dismay that young drivers were seen as an end-point, rather than as an integral actor in their own road safety. Young drivers need support, and teachers were viewed as sounding boards for a wealth of issues ranging from unsatisfactory or incomplete interactions with licensing authorities, to resisting negative peer pressure.

GENERAL DISCUSSION

A number of important observations regarding the identified actors, roles, interactions, and posited needs to improve young driver road safety can be made.

Firstly, in comparison to the breadth of actors within the six Rasmussen's RMF sociotechnical levels identified by Scott-Parker et al. (in press), considerably more actors – not only at a state or local level – were identified as playing a role in the YDRSS in Queensland. This was either through direct (such as licensing) and indirect (such as an exemplar of best practice) efforts. This means that previous efforts to map the YDRSS, based on young driver road safety expertise, appear to have underestimated the structure and size of the YDRSS. For an optimal systems approach to young driver road safety, *all* actors in the YDRSS should be identified. In addition, interviews with actors who have not contributed to the current research project (see Table 1) are likely to identify additional, heretofore-unrealised, actors important in the YDRSS. In addition, the need for alternatives for youth to "let off steam" apart from in a car on the road was also mentioned by a number of actors in level 4 in particular, consistent with development-related factors in adolescence (e.g., Bingham et al., 2008; Keating, 2007). Actors within these potential domains were not captured in the current research project, and therefore should be considered more fully in future research. Transparency regarding who is an actor, the role they play, and interactions (both direct and indirect) within the YDRSS can also assist the efforts of actors 'on the coal face' in level 4 who report struggling with limited financial and social capital. Furthermore, many actors reported that they were "already interacting effectively with everyone" in the YDRSS, however the findings of the research project suggests otherwise.

Secondly, whilst at a macro-level it appears that there is much interaction between actors within and across levels on the YDRSS, there is substantial variability in the extent and nature of any interactions. For example, within level 1, one actor only reported interacting with two other identified actors across two levels, whilst other actors in level 1 reported interacting directly and actively with more than one dozen actors across the six sociotechnical levels. In addition, actors generally varied widely in their reports regarding other actors in the young driver road safety with whom they *do not* interact. Rasmussen's RMF describes the importance of vertical integration for safety management: that is, actors at different levels of the system need a constant level of interaction with each other. Without this, systems lose control of safety. Delineating interactions and non-interactions at the individual actor level, and more generally at a macro sociotechnical level, are vital to understanding and effectively managing the YDRSS. Therefore to operationalise, and then maximise, the YDRSS, such diverse interaction experiences need to be recognised. Indirect interactions merit further consideration, particularly given the frequent reported use of intermediary actors 'closer' to the young driver by actors perceived as 'further' from the young driver. Whilst an optimal YDRSS remains to be determined, it is possible that direct interactions may be of more importance, especially in the case of actor efforts being diluted or waylaid through the involvement on an intermediary actor. Further, the unidirectional relationship suggested by the interactions summarised in Table 1 suggests that vertical and horizontal integration may not be occurring within and across the six sociotechnical levels of Rasmussen's RMF. It is noteworthy that it remains unknown whether bidirectional relationships amongst all actor pairings is desirable, possible or practical. Nonetheless, the diminishing progress in reducing youth road trauma suggests there are improvements to be achieved and the present results, particularly the unidirectional relationships and lack of vertical integration, offer potential avenues to explore in future intervention efforts. In addition, given that a recent survey of Learners in Queensland found that over 90% of them had used the services of a professional driving instructor (Scott-Parker et al., 2011) it appears that the road safety of the young driver may be enhanced through maximising the effectiveness of their interactions with young drivers and their parents.

Thirdly, as shown in Table 1, based on the feedback from the participants, there appears to be incomplete knowledge and knowledge sharing regarding the YDRSS within and across the six sociotechnical levels. Actors are aware of a number of other actors who play a role in the YDRSS, but for whatever reason they are not interacting with them. Reasons for interaction and non-interaction merit further investigation, particularly whether certain barriers to interaction need to be removed (or at a minimum, ameliorated) for the YDRSS to be effective. In addition, the YDRSS needs to maximise both mobility and safety, two oft-conflicting goals for transport policy-makers (level 1 actors) (Bates et al., 2010). A limitation of the current research is the small sample of participants representing approximately one-third of the identified actors, and that as such the views of the participating actors may not be representative of actors within the specified sub-categories (e.g., community groups, schools and educational institutions), or the six sociotechnical levels of Rasmussen's RMF. It is noteworthy however that there is no reason to suspect that those with more integrated views – i.e., in terms of the unidirectional and lack of horizontal and vertical interactions findings – have been systematically excluded. Wider sampling efforts are required, and as such this will help clarify

whether the knowledge-sharing issue pervades the YDRSS or, less likely, is specific to the current sample only.

Fourthly, the prevalence of empty boxes in Table 1 suggests that the majority of actors in all six sociotechnical levels of the YDRSS interviewed for this study were unaware of or did not come to mind as prominent 'who is sharing the responsibility' for young driver road safety in their own level, as well as in higher and/or lower levels. Whilst the pressures of recalling during an interview may have contributed to this deficit, it is possible that the actors simply were not aware. Systems thinking suggests that increased awareness of who is actually sharing the responsibility for young driver road safety is a vital first step for both horizontal and vertical integration and for the system to foster and maintain young driver road safety.

Fifthly, the literature abounds with tomes regarding the importance of parents in young driver road safety (e.g., Beck et al., 2003; Brookland and Begg, 2011; Goodwin et al., 2010; Hellinga et al., 2007; Lahatte and Le Pape, 2008; McCartt et al., 2007; Scott-Parker et al., 2012; Sherman et al., 2004; Simons-Morton and Ouimet, 2006; Taubman-Ben-Ari et al., 2005; Williams, 2006), further illustrated by the interactions and roles reported by the research participants. Yet the two parents interviewed as part of the project did not report interacting with other actors in the YDRSS. It is important to note however that two parents is not representative of the experiences of parents generally, and that future research will examine the perspectives and experiences of a larger parent sample. Notwithstanding this sampling limitation, the parents did not feedback upwards (through levels 1 and 2), sideways (through level 3), and through levels 4, 5 and 6 apart from their interactions with their young driver child. Thus it appears that these two parents are not embedded in a system characterised by vertical integration, which has ramifications for the efficient and effective operation of the YDRSS. Difficulty engaging with parents in the Queensland-context is consistent with recent parent-engagement difficulties experienced in the United States (e.g., Simons-Morton et al., 2006); however engagement with parents is possible and young driver road safety benefits have been demonstrated (e.g., Simons-Morton, 2007). Notwithstanding this, it is important to note that – as evidenced in the AcciMap and the summary table – "parents do not shoulder all the responsibility" (level 2 actor) for young driver road safety. In addition, whilst the current research project focus was upon the perspectives of partners in the YDRSS, the next step is to widely sample young drivers themselves, including the mandatorily-supervised Learner driver, and the independent Provisional (intermediate) driver, regarding their perceptions and experiences in the YDRSS.

CONCLUSIONS

Rather than persisting in "fixing the young driver", with only moderate success, it is time to reform how young driver road safety is understood and managed by implementing systems thinking as way to understand and address the factors that influence and shape young driver behaviour. Interviews with 42 participants representing 23 Actors who were identified as important players in the YDRSS further elucidated the role these actors believe they play. In addition, the interviews revealed a wealth of other actors in the YDRSS, some of whom the actors interact with directly or indirectly. Application of Rasmussen's RMF revealed some horizontal integration amongst actors, whilst in general there was a paucity of vertical integration, particularly feedback upward through the actors represented across the Rasmussen's RMF's six sociotechnical levels. Further, many suggestions were made by actors regarding research and operational needs consistent with systems thinking within the domain of young driver road safety, ranging from eliminating unnecessary 're-invention of the wheel' through coordination of actors to recognising that young driver road safety is not limited to the realm of road safety, and that other domains such as parenting and health (e.g., drugs and alcohol) can be highly influential. Research directions include exploring the YDRSS from the perspective of the young driver themselves, additional efforts to seek the perspectives of other actors identified as playing a role in the YDRSS, and wider sampling and investigation of the experiences and perspectives of parents.

REFERENCES

Bates, L., Watson, B., & King, M. (2010). Mobility and safety are conflicting goals for transport policy makers when making decisions about graduated driver licensing. *International Journal of Health Promotion and Education, 48*(2), 46-51.

Beck, K. H., Shattuck, T., Raleigh, R., & Hartos, J. (2003). Does graduated licensing empower parents to place greater restrictions on their newly licensed teens' driving? *Health Education and Behavior, 30*(6), 695-708.

Bingham, C. R., Shope, J. T., Zakrajsek, J., & Raghunathan, T. E. (2008). Problem driving behaviour and psychosocial maturation in young adulthood. *Accident Analysis and Prevention, 40*, 1758-1764.

Brookland, R., & Begg, D. (2011). Adolescent, and their parents, attitudes towards graduated driver licensing and subsequent risky driving and crashes in young adulthood. *Journal of Safety Research, 42*, 109-115.

Goodwin, A. H., Foss, R. D., Margolis, L., & Waller, M. (2010). *Parents, teens and the Learner stage of graduated driver licensing. Washington*. AAA Foundation for Traffic Safety.

Hellinga, L. A., McCartt, A. T., & Haire, E. R. (2007). Choice of teenagers' vehicles and views on vehicle safety: Survey of parents of novice teenage drivers. *Journal of Safety Research, 38,* 707-713.

Johnson, S. B., Jones, V. C. 2011. Adolescent development and risk of injury: using developmental science to improve interventions. Injury Prevention 17: 50-4.

Keating, D. P. (2007). Understanding adolescent development: Implications for driving safety. Journal of Safety Research, 38, 147-157.

Lahatte, A., & Le Pape, M. C. (2008). Is the way young people drive a reflection of the way their parents drive? An econometric study of the relation between parental risk and their children's risk. Risk Analysis, 28(3), 627-634.

McCartt, A. T., Hellinga, L. A., & Haire, E. R. (2007). Age of licensure and monitoring teenagers' driving: Survey of parents of novice teenage drivers. Journal of Safety Research, 38, 697-706.

McCartt, A. T., Mayhew, D. R., Braitman, K. A., Ferguson, S. A., & Simpson, H. M. (2009). Effects of age and experience on young driver crashes: Review of recent literature. Traffic Injury Prevention, 10, 209-219.

Rasmussen, J., 1997. Risk management in a dynamic society: A modelling problem. Safety Science 27(2/3), 183-213.

Runyan, C. W. (1998). Using the Haddon Matrix: Introducing the third dimension. Injury Prevention, 4, 302-307.

Scott-Parker, B., Bates, L., Watson, B., King, M. J., & Hyde, M. K. (2011). The impact of changes to the graduated driver licensing program in Queensland, Australia on the experiences of Learner drivers. Accident Analysis and Prevention, 43(4), 1301-1308.

Scott-Parker, B., Goode, N., & Salmon, P. (in press). The driver, the road, the rules.....and the rest? A systems-based approach to young driver road safety. Manuscript for a Systems Thinking special issue of Accident Analysis and Prevention.

Scott-Parker, B., Watson, B., King, M. J., & Hyde, M. K. (2012). "They're lunatics on the road": Exploring the normative influences of parents, friends, and police on young novice's risky driving decisions. Safety Science, 50, 1917-1928.

Senserrick, T. (2009). Australian graduated driver licensing systems. Journal of the Australasian College of Road Safety, February, 20-26.

Sherman, K., Lapidus, G., Gelven, E., & Banco, L. (2004). New teen drivers and their parents: What they know and what they expect. American Journal of Health Behavior, 28, 387-397.

Simons-Morton, B., & Ouimet, M. C. (2006). Parent involvement in novice teen driving: A review of the literature. Injury Prevention, 12(SI), i30-i37.

Simons-Morton, B. (2007). Parent involvement in novice teen driving: Rationale, evidence of effects, and potential for enhancing graduated driver licensing effectiveness. Journal of Safety Research, 38, 193-202.

Simons-Morton, B. G., Hartos, J. L., Leaf, W. A., & Preusser, D. F. (2006). The effects of the checkpoints program on parent-imposed driving limits and crash outcomes among Connecticut novice teen drivers at 6-months post licensure. Journal of Safety Research, 37, 9-15.

Svedung, I., Rasmussen, J., 2002. Graphic representation of accident scenarios: Mapping system structure and the causation of accidents. Safety Science 40, 374-417.

Taubman-Ben-Ari, O., Mikulincer, M., & Gillath, O. (2005). From parents to children – Similarity in parents and offspring driving styles. Transportation Research Part F, 8, 19-29.

Williams, A. F. (2006). Parents' views of teen driving risks, the role of parents, and how they plan to manage the risks. Journal of Safety Research, 37, 221-226.

Williams, A. F. (2011). Commentary: Graduated licensing – Moving forward or standing still? Traffic Injury Prevention, 12, 207-209.

World Health Organization. Road Traffic Injuries. Fact Sheet No 358. March 2013. Available at: http://www.who.int/mediacentre/factsheets/fs358/en/ . Accessed April 15, 2013.

Pathways to Failure? Using Work Domain Analysis to Predict Accidents in Complex Systems

Paul M. Salmon[a], Michael G. Lenne[b], Gemma Read[b], Guy H. Walker[c] and Neville A. Stanton[d]

[a]University of the Sunshine Coast Accident Research
Maroochydore, QLD, Australia

[b]Monash University Accident Research Centre
Monash University, Clayton, Victoria, Australia

[c]School of the Built Environment
Heriot-Watt University, Edinburgh, UK

[d]Civil, Maritime, Environmental Engineering and Science
Faculty of Engineering and the Environment
University of Southampton, Southampton, UK

ABSTRACT

Forecasting accidents before they occur is the final frontier for safety science. Although this has long been recognized, the discipline of human factors has yet to produce an appropriate methodology for achieving this. This paper presents some of the findings from an exploratory study in which the abstraction hierarchy method from the work domain analysis phase of cognitive work analysis was used to predict potential accidents. Using rail level crossings as a test case, the exploratory study revealed that the abstraction hierarchy method was able to predict a range of failure pathways that could potentially lead to a collision between a road user and a train at rail level crossings. In addition, certain features of the abstraction hierarchy method were found to make it highly consistent with contemporary systems level views on accident causation, including that it provides a systems level analysis of potential accident pathways, that is does not support a focus on broken human components (since the abstraction hierarchy model is actor independent), and that the primary focus is on the relationships between components rather than the components themselves. Further testing of the approach is recommended, including sensitivity and validity testing whereby the predictions made are compared to real world events.

Keywords: Rail Level Crossing Safety, Accident Prediction, Cognitive Work Analysis

INTRODUCTION

Forecasting accidents before they occur is the final frontier for safety science (Moray, 2008, Salmon et al, 2011, Stanton and Stammers, 2008). It is the breakthrough needed in order to make the next step in reducing risks in human-technology systems. Although there have been various attempts at developing accident prediction models (e.g. Deublein et al, 2013), most are statistical models that are unable to identify and describe how behavior across overall sociotechnical systems might combine to create failure scenarios. Despite decades of research into accident causation and prevention, we do not yet have a methodology that can be used to determine how interactions across

sociotechnical systems will lead to emergent behavior that creates accidents. Using rail level crossings as a test system, the aim of this paper is to take the first steps in filling this critical capability gap.

Why rail level crossings? Worldwide, the problem of collisions between road users and trains at rail level crossings remains resistant to current countermeasures. In Australia, between 2000 and 2009, there were 695 collisions between road vehicles and trains at rail level crossings, resulting in 97 fatalities (Independent Transport Safety Regulator, 2011). Despite various initiatives, in 2011 there were 49 collisions between trains and road vehicles at rail level crossing in Australia, leading to 33 fatalities (ATSB, 2012). In addition to the high levels of trauma and personal cost, the financial costs associated with rail level crossing collisions are significant. In Australia they have been estimated to cost around $24 million a year. The problem is not exclusive to Australia, with unacceptable incident, fatality, and injury rates worldwide. Across Europe, for example, there were 604 fatalities and serious injuries at rail level crossings in 2011(European Railway Agency, 2013). Clearly rail level crossings represent a significant safety issue worldwide and one that has so far proven intractable to safety interventions. This makes them the perfect test case for this accident prediction exercise.

The problem is exacerbated by the fact that the intensity of operations in rail is increasing. The railway industry in Australia is a nationally important infrastructure asset and a key enabler of economic growth. The number of trips made by rail in Australia is increasing, and in 2009 equated to over 180 million train kilometers travelled (National Transport Commission Australia, 2011). Risk levels have improved but maintaining these levels, and improving them further, has to occur in a context of shifting system boundaries resulting from both higher operational intensity and risk exposure. The current approach to preventing rail level crossing incidents is reactive and involves waiting for the next catastrophe to occur before implementing new safety initiatives such as advanced warning devices and boom barriers. Are we willing to accept that this is 'as good as it gets'? Clearly not, but in order to make continued progress the problem-space needs to be looked at afresh.

Part of the problem is that, despite having access to a range of valid and exhaustive systems analysis methods, we are still only able to exhaustively describe failure scenarios after they have happened (e.g. Salmon et al, 2013). This is particularly the case with rail level crossings, where pro-active risk assessment tools currently in use have been questioned (Salmon et al, 2013). The step change required in accident research is the ability to predict entire accident scenarios, not just individual human errors. This paper reports some of the initial findings from an exploratory study that aims to solve this problem. We present a study in which we explored the ability of the Work Domain Analysis (WDA) phase of the popular Cognitive Work Analysis framework (CWA; Vicente, 1999; Jenkins et al, 2008) to predict the causes of collisions between rail and road users at rail level crossings. Specifically, an abstraction hierarchy model of an 'active' rail level crossing (i.e. where road users are controlled by active warnings such as boom gates and flashing lights) was used to predict the range of failure scenarios that could occur where the end result is a collision between a road user and a train.

A Systems Approach to Accident Prediction

There is a range of well used and tested methodologies available that can be used to predict the kinds of human errors that might lead to accidents (see Stanton et al, 2013). Indeed, various error prediction methods have been shown to achieve acceptable levels of reliability and validity (e.g. Stanton et al, 2009). The problem with these methods is that they predict what is likely the last behavior in a long and complex network of interacting and emergent behaviors occurring across various parts of the system in question. Whilst it is of course useful to examine what errors a human operator might make in a given system, accident prevention efforts are far better served by looking at the interactions that occur before the human operator makes the error. In short, it is the entire accident scenario, including interacting factors and emergent behaviors that is important for understanding how to prevent accidents.

Accidents are a systems phenomenon; they are emergent properties of complex sociotechnical systems. Effectively predicting accident scenarios therefore requires a systems approach. The systems approach argues that safety, and indeed accidents, are emergent properties arising from non-linear interactions between multiple components distributed across complex sociotechnical systems (e.g. Leveson, 2004). It has a long legacy in safety science, from the foundational work of Heinrich (1931) through to the evolution of a number of more recent accident causation models and analysis methods (e.g. Leveson, 2004; Rasmussen, 1997; Reason, 1990). As it stands currently, the most up-to-date methods and approaches rely on a form of systems thinking closely linked to current debates in the science of complexity.

Complexity is formally described in a number of ways (e.g. Walker et al., 2009), central to which is the property of emergence (e.g. Waldrop, 1992). Emergence is behavior that arises from dynamic human-technical systems which rely on a continuous throughput of information, resources, actions and behaviors. In everyday terms, emergence is the 'more than the sum of its parts' effect of synergy and the ability to yield disproportionately favorable outcomes from small, targeted interventions. It is also the property that gives rise to system errors and other pathologies (e.g. Perrow, 1999; Reason, 1990). Decisions and actions at all levels of a complex system interact with one another to shape system performance: safety and accidents are thus shaped by the decisions of all actors, not just the front line workers in isolation, and accidents are caused by multiple contributing factors, not just one bad decision or action. In either case, emergence describes behavior that is not readily deducible from its low level properties and cannot always be detected via reductionist, non-systemic approaches.

Existing error prediction methodologies can be thought of as reductionist (although they are not entirely reductionist as they do focus on man-machine interactions). Reductionist approaches to safety science, those which rely on taking the system apart in order to understand the components, then reassembling the components back into the complete system (on the tacit assumption that the whole cannot be greater or less than the sum of its parts) do not allow us to detect the emergent properties associated with the types of strategic risk issues we wish to make progress on (see Walker et al., 2009). Systems approaches do. One means by which they can enable forecasting and prediction is to consider the causal texture of the systems environment, and the system's movement through that environment. Rasmussen (1997) describes a set of 'boundaries'. There is a boundary of economic failure: these are the financial constraints on a system that influence behavior towards greater cost efficiencies. Then there is a boundary of unacceptable workload: these are the pressures experienced by people and equipment in the system as they try to meet economic and financial objectives. The boundary of economic failure creates a pressure towards greater efficiency, which works in opposition to a similar pressure against excessive workload. Because transport systems involve human as well as technical elements, and because humans are able to adapt situations to suit their own needs and preferences, these pressures inevitably introduce variations in behavior which are not explicitly designed and can lead to increasingly emergent system behaviors, both good and bad (Qureshi, 2007; Clegg, 2000). Over time this adaptive behavior can cause the system to cross safety boundaries and accidents to happen (Qureshi, 2007; Rasmussen, 1997). The key, then, is to detect in advance a) where those boundaries are and b) where the system is travelling in relation to them.

The systems approach has become popular in part because of various systems analysis methodologies that can, to some extent at least, do this (e.g. Rasmussen, 1997). These methods, for example Accimap (Rasmussen, 1997), are becoming increasingly popular for accident investigation purposes in safety critical domains such as aviation, rail, maritime, and process control. This arises from recognition among practitioners and industry stakeholders that new approaches are needed to tackle issues that are proving resistant to previous approaches. The major limitation of current systems-based accident analysis methodologies is that, so far, they cannot be used in a pro-active manner: organizations are effectively waiting for loss events to occur before they can work on prevention strategies. This is not acceptable. The lack of data resulting from improved safety trends combined with greater operational intensity and risk exposure means that, if anything, loss events are more likely to be large-scale and unexpected, meaning that 'learning from disasters' is becoming increasingly dubious from an ethical perspective. The need for systems-based prediction approaches is discussed extensively in the literature (e.g. Moray, 2008, Salmon et al, 2011, Stanton and Stammers, 2008) but as yet a credible approach has yet to emerge.

Cognitive Work Analysis (CWA): A Framework for Predicting Vulnerability at Rail Level Crossings

CWA (Jenkins et al, 2008; Vicente, 1999) is a systems analysis and design framework that is used to identify the constraints imposed on activities and then to design new systems that better support the activities of interest. To date CWA has been used in a range of systems analysis and design activities in various safety critical domains. Although the framework is becoming popular in safety applications, applications examining accident causation are limited in number.

The first CWA phase, Work Domain Analysis (WDA), is used to provide an event and actor independent description of the system under analysis, in this case rail level crossing systems. The aim of the WDA phase is to describe the purposes of the system and the constraints imposed on the actions of any actor working within that system (Vicente, 1999). The system under analysis is described at the following five conceptual levels using the abstraction hierarchy method:

1. Functional purpose – The overall purposes of the system and the external constraints imposed on its operation;

2. Values and priority measures – The criteria that organizations use for measuring progress towards the functional purposes;

3. Generalized functions – The general functions of the system that are necessary for achieving the functional purposes;

4. Physical functions – The functional capabilities and limitations of the physical objects within the system that enable the generalized functions; and

5. Physical objects – The physical objects within the system that are used to undertake the generalized functions.

The output is a detailed description of the system under analysis in terms of the constraints influencing behavior and the physical objects (and their affordances) and functions that enable the system to achieve its functional purpose. Importantly, the abstraction hierarchy model that is used for the WDA phases uses means-ends relationships to link nodes across the five levels of abstraction. Every node in the abstraction hierarchy should be the end that is achieved by all of the linked nodes below it, and also the means that (either on its own or in combination with other nodes) can be used to achieve all of the linked nodes above it. This representation of the relationships between nodes can be thought of as one way of representing interactions between system components (e.g. object x and y are used together to achieve process z, and process z happens with process a to achieve function b)

This way of describing complex systems appears to be suited to identifying the emergent behaviors that might lead to accidents. As all of the functions required to achieve the functional purposes of the system are described, it is possible to identify which functions can potentially create adverse outcomes. This can be through either the function not being achieved or through emergent behavior arising from achievement of the functions. For example, in the rail level crossing context, if the function 'Maintain road and rail user separation' is not achieved, then the outcome is a collision between a road user and the train. It is then possible to examine the other functions (or combinations of other functions) that might have contributed to the failure to maintain separation. For example, should the function 'Alert (road user) to presence of train' not be achieved, then a failure to maintain separation between the road users and train is the likely end result. Moving down the abstraction hierarchy, it is then possible to explore the affordances and physical objects that can potentially play a role in the function not being achieved. This enables both the individual behaviors and combinations of behaviors that could lead to the failure to alert the road user to the presence of a train.

An additional feature of the abstraction hierarchy is that it is actor and event independent and so does not focus on specific human actors when identifying accident pathways (e.g. road users, pedestrians, train drivers). This is a significant departure from existing human error prediction and accident analysis methodologies, which tend to focus first on human actors as the source of error and accidents. The focus on interactions between components rather than the behavior of specific human components is a feature of the WDA approach that gives it a high degree of consistency with contemporary accident causation models (e.g. Rasmussen, 1997). Dekker (2011), for example, argues that the focus on broken human components is one of the major flaws in accident and safety-related research.

By examining failed functions and the related means-ends links, all potential pathways to failure can be explored. This is represented in Figure 1, which shows selected failure pathways leading to the function 'Alert (road user) to presence of train' not being achieved.

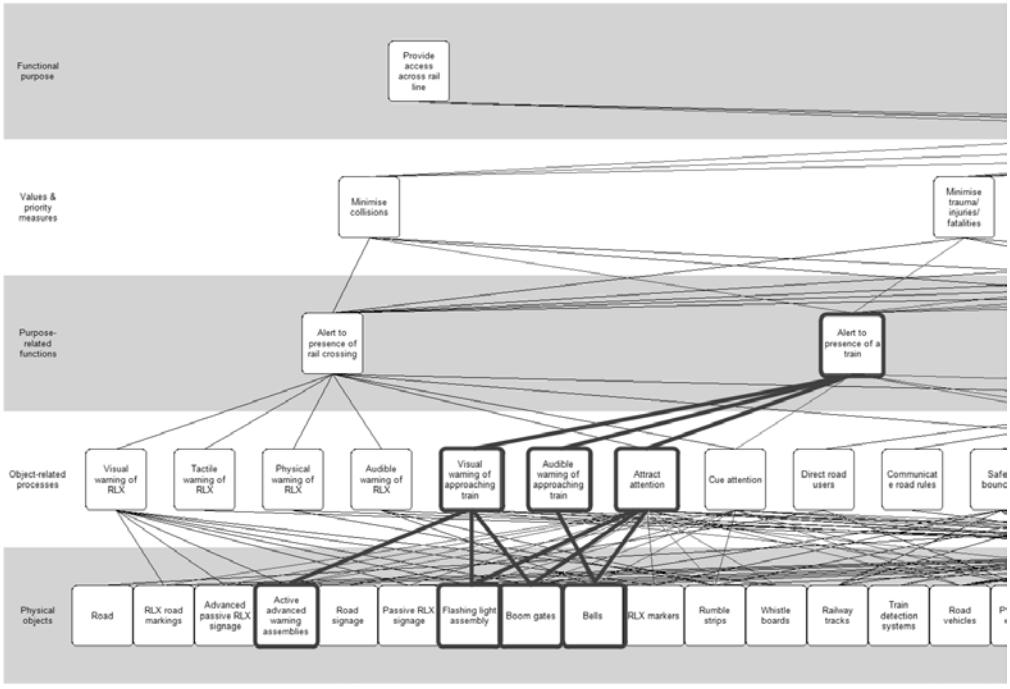


Figure 1. Work domain analysis extract showing accident pathways linked to failed 'alert to presence of train' function.

In Figure 1, any combination of the highlighted nodes at the physical object and object-related processes level failing or being absent at a particular rail level crossing environment could lead to the function 'alert (road user) to presence of train' not being achieved. For example, at a particular rail level crossing if there are no active advanced warning assemblies or boom gates, and the road user did not perceive the flashing lights or hear the bells, then the object-related processes 'visual warning of approaching train' and 'audible warning of approaching train' have failed, which in turn means the function 'alert to presence of a train' will not be achieved. This is precisely the scenario that occurred in the Kerang rail level crossing tragedy of 2007 in which the driver of a loaded semi-trailer truck continued toward a rail level crossing apparently unaware that a passenger train was also approaching. The resulting collision killed 11 train passengers and injured a further 15 people, including the truck driver (see Salmon et al, 2013).

METHOD

Work Domain Analysis Development

An abstraction hierarchy was developed for a generic 'active' rail level crossing. Active rail level crossings are those that have 'active' warning devices which provide a warning of approaching trains, such as boom gates, flashing lights, and bells. The purpose was to construct a WDA model of active rail level crossings generally, rather than for one specific rail level crossing system. The abstraction hierarchy was developed based on data collected from a range of activities, including on-road studies of driver behavior at rail level crossings (e.g. Young et al, 2014), interviews with road users regarding their behavior at rail level crossings, interviews with train drivers, surveys of road user behavior at rail level crossings (e.g. Beanland et al, 2013) and a workshop involving a range of rail level crossing subject matter experts (e.g. designers, rail safety practitioners, regulators, road safety practitioners, engineers). A summary of the abstraction hierarchy is presented in Figure 2.

Prediction of Failure Scenarios

One analyst with significant experience in accident analysis, error prediction, and applying CWA predicted the failure scenarios. This involved examining functions at the purpose-related functions level of the abstraction hierarchy to identify those that, if not achieved, could create a collision between a road user and train. The next step involved examining functions or combinations of functions that could potentially create the identified 'collision causing failed function'. Next, the means-ends links were used to trace all object-related processes and physical objects related to these functions. In addition, new links down from the purpose-related functions level to the

physical objects level were explored. Once relevant object-related processes and physical objects were identified, a simple failure mode taxonomy was applied to these objects and processes to determine all of the ways in which they might not work or be achieved. Each failure pathway was then rated in terms of probability of occurring (low, medium and high) and criticality (low, medium and high). This is based on the approach used by error prediction methods such as the Systematic Human Error Reduction and Prediction Approach (SHERPA, Embrey, 1986) and enables the most likely and critical scenarios to be focused on during accident prevention efforts. A representation of a simple failure pathway is presented in Figure 3.

RESULTS

Due to space constraints it is not possible to show the entire rail level crossing abstraction hierarchy. A summary is therefore presented in Figure 2. Figure 2 shows in full the functional purpose, values and priority measures, and generalized functions levels from the abstraction hierarchy, along with categories of the nodes found at the physical objects and physical functionality levels.

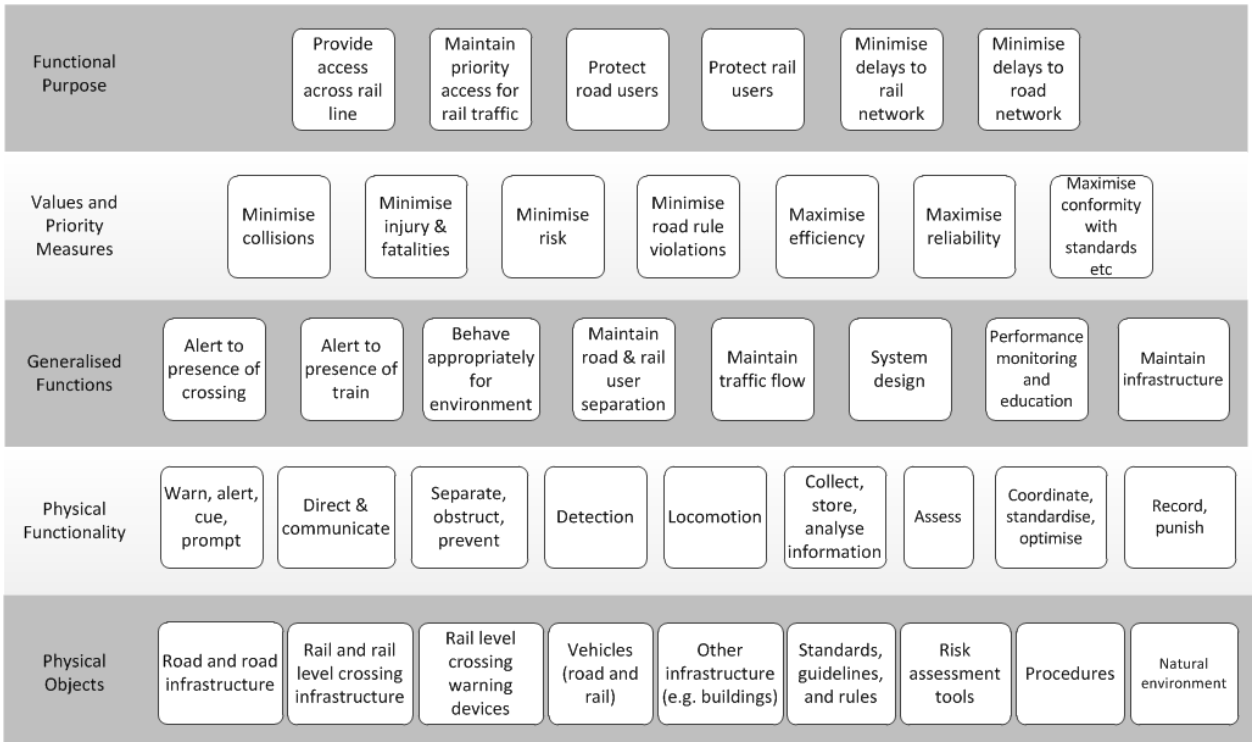


Figure 2. Summary of active rail level crossing abstraction hierarchy.

As shown in Figure 2, the primary function that is linked to a collision between a road user and a train is the function ‘maintain road and rail user separation’. Should this function not be achieved, then a collision between road user and train is the outcome. Of more interest are the functions and combinations of functions that can create the failure to maintain road and rail user separation. For example, if the function ‘alert to presence of a train’ is not achieved, then maintenance of separation may not be achieved. In turn, other functions can combine to create the failure to alert road users to the presence of a train. For example, a inadequate maintenance (maintain infrastructure function) combined with a lack of performance monitoring might lead to the warning devices at the physical object level not functioning properly, and in turn a failure to provide a warning at the next level up. This failure to provide a warning then creates the failed ‘alert to presence of train’ function.

This example failure pathway is presented in Figure 3.

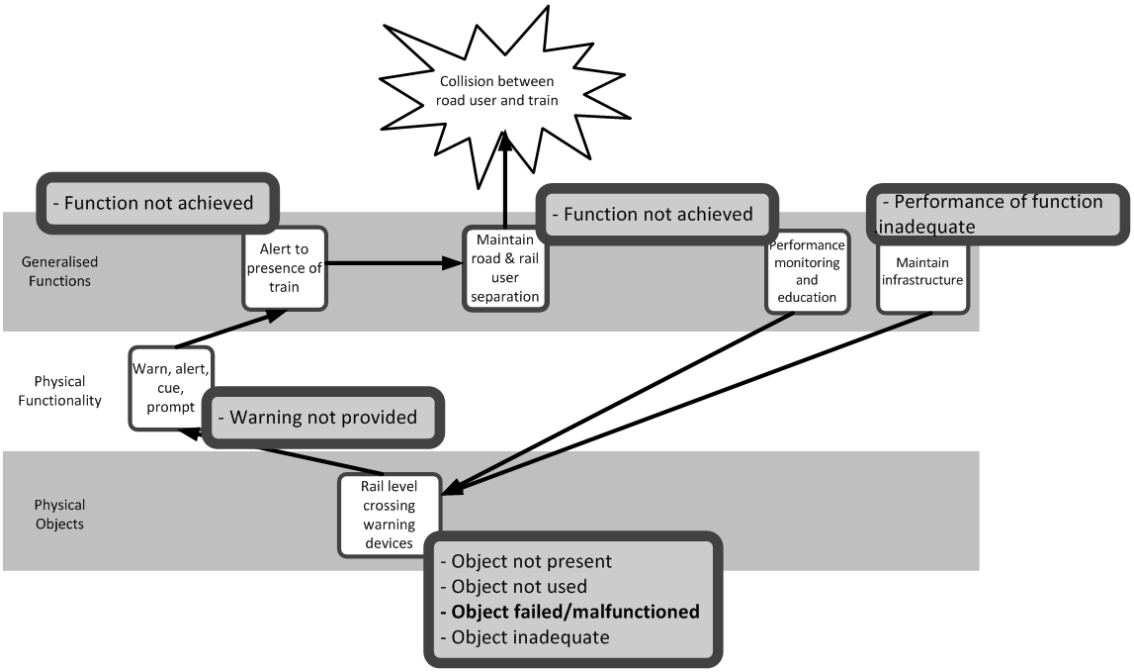


Figure 3. Example failure pathway causing a collision between a road user and train.

Extracts of the critical failure pathways that could potentially lead to a collision between a road user and train is presented in Table 1.

Table 1. Example failure pathways

Function not achieved	Related functions	Process not achieved	Object failure/absence	Description	Outcome	P	C
Maintain separation between road and rail user	Alert road user to presence of train			Road user does not perceive/comprehend warnings	Road user is not aware of approaching train and collides with train	L	H
Maintain separation between road and rail user	Alert road user to presence of train	Visual warning of approaching train	- Active advanced warning assembly (Object not present) - Flashing lights (Object failed) - Boom gates (Object not present)	Road user is not given warning of approaching train due to failed warning devices	Road user is not aware of approaching train and collides with train	L	H
Maintain separation between road and rail user	Alert road user to presence of train		- Flashing lights - Boom gates (absent)	Road user does not perceive flashing lights and rail level crossing does not have boom gates	Road user is not aware of approaching train and collides with train	M	H
Maintain separation between road and rail user	Alert road user to presence of train Maintain infrastructure Performance monitoring	Finance maintenance and upgrade Visual warning of approaching train	- Finances (Absent) - Flashing lights (Failed)	Due to financial constraints ail level crossing is not maintained to adequate standard leading to infrastructure failure Lack of performance monitoring means track operator is not aware of failed warning devices	Road user is not aware of approaching train and collides with train	M	H

A range of failure scenarios were identified. These will be presented in full during the conference presentation. The key findings from the exploratory study were that:

a. The abstraction hierarchy is able to identify a range of failure pathways that will lead to a collision between a road user and train; and

b. The failure pathways identified include instances where the 'maintain road and rail user separation' function is not achieved. The findings show that this function might not be achieved due to failed functions in isolation (e.g. alert to presence of train) or a combination of failed functions (e.g. maintain infrastructure, performance monitoring, alert to presence of a train). The failed functions in turn can be caused by individual physical objects (e.g. flashing lights failed) or combinations of physical objects.

DISCUSSION

This paper presents a summary of some of the findings derived from an exploratory study in which the first phase of CWA, WDA, was used to predict the range of failure pathways leading to a collision between a road user and a train at an active rail level crossing. The findings demonstrate that the approach was able to predict a range of failure scenarios, providing initial evidence that the WDA phase of CWA is capable of predicting accidents in a manner that is consistent with the systems level view on accident causation (e.g. Rasmussen, 1997).

The exploratory study revealed some notable features of this approach when used for accident prediction. First, provided the abstraction hierarchy is exhaustive, the majority of failure pathways present within a particular system can be identified before they happen. This is possible because all of the system components should be described at the physical objects level, and in turn, all of the affordances and functions should be described at the two levels above within the abstraction hierarchy. In addition, new failure pathways that might emerge following the introduction of new technologies or systems can also be identified by adding the new objects to the abstraction hierarchy and running the analysis again. Second, the abstraction hierarchy enables a systems level analysis of the potential pathways to failure. This is attractive because it ensures a level of comprehensiveness, but also because it is very much in line with contemporary models of accident causation which argue that accidents are a systems phenomenon and that the interactions between components are of interest rather than the components themselves. Accidents will be caused by a range of factors across the overall system; the abstraction hierarchy enables the overall system to be examined. Third, the fact that the abstraction hierarchy is actor independent ensures that there is no focus on the human operators as the broken component or major cause of the accidents identified. Rather, the way in which the system shapes human operator behavior is examined (i.e. in this case being unaware of an approaching train). The analysis is focused on objects, affordances and functions within the system and how they might interact to shape human operator behavior.

It is worth noting that the study described was highly exploratory. Further testing of the approach as an accident prediction tool is required. In particular, a validation study whereby the abstraction hierarchy is used to predict accidents in a particular system following which the predictions made are compared to actual accidents over a period of time in that system is recommended. In addition, using the other CWA phases (e.g. control task analysis, strategies analysis) to provide a deeper level of analysis is another pertinent line of inquiry.

In addition to predicting accidents, there is potential for the abstraction hierarchy to be used as an accident analysis template. This involves examining each node in the abstraction hierarchy during the accident analysis and investigation process in order to identify nodes and means and links that played a role in the accident. Specifically, the means ends links related to functions that were not achieved could be traced down the abstraction hierarchy in order to identify what object-related processes or physical objects played a part.

ACKNOWLEDGEMENTS

Paul Salmon's contribution to this research and paper is funded through the Australian National Health and Medical Research Council post-doctoral training fellowship scheme.

REFERENCES

ATSB. (2012). Australian Rail Safety Occurrence Data 1 July 2002 to 30 June 2012. Report number RR-2012-010. Australian Transport Safety Bureau, Canberra, ACT.

Beanland, V., Lenné, M. G., Salmon, P. M., Stanton, N. A. (2013). A Self-Report Study Of Factors Influencing Decision-Making At Rail Level Crossings: Comparing Car Drivers, Motorcyclists, Cyclists And Pedestrians. In Proceedings of the Australasian Road Safety Research, Policing and Education Conference 2013, Brisbane, Australia

Clegg, C. W. (2000). Sociotechnical principles for system design. *Applied Ergonomics*, 31, 463-477.

Dekker, S. (2011). Drift into failure: from hunting broken components to understanding complex systems. Ashgate, Aldershot, UK.

Deublein, M., Schubert, M., Adey, B. T., Köhler, J., & Faber, M. H. (2013). Prediction of road accidents: A Bayesian hierarchical approach. *Accident Analysis & Prevention*, 51, 274-291

Embrey, D.E., (1986). SHERPA: a systematic human error reduction and prediction approach. Paper presented at the International Meeting on Advances in Nuclear Power Systems, Knoxville, Tennessee.

European Railway Agency. (2013). Intermediate report on the development of railway safety in the European Union. European Railway Agency Safety Unit, 15 May 2013.

Heinrich., H. W. (1931). Industrial accident prevention: A scientific approach, McGraw-Hill, New York.

Independent Transport Safety Regulator. (2011). Level Crossing Accidents in Australia.ITSR, Sydney.

Jenkins, D. P., Stanton, N. A., Salmon, P. M., & Walker, G. H. (2008). Cognitive work analysis: coping with complexity. Ashgate, Aldershot, UK.

Leveson, N. G. (2004). A new accident model for engineering safer systems. *Safety Science*, 42:4, 237—270.

Moray, N. (2008). The Good, the Bad, and the Future: On the Archaeology of Ergonomics. *Human Factors*, 50(3), 411-417.

National Transport Commission Australia. (2011). Rail Safety National Law Draft Regulatory Impact Statement, September 2011

Perrow, C. (1999). Normal Accidents: Living With High-Risk Technologies. Princeton, New Jersey: Princeton University Press.

Qureshi, Z. H. (2007). A review of accident modelling approaches for complex critical sociotechnical systems. Conferences in Research and Practice in Information Technology Series; Vol. 336

Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science*, 27:2/3, 183-213.

Reason, J. (1990). Human Error. New York, NY, Cambridge University Press

Salmon, P. M., Read, G., Stanton, N. A, Lenné, M. G. (2013). The Crash at Kerang: Investigating systemic and psychological factors leading to unintentional non-compliance at rail level crossings. *Accident Analysis and Prevention*. 50, 1278-1288

Salmon, P. M., Stanton, N. A., Lenné, M. G., Jenkins, D. P., Rafferty, L., & Walker, G. H. (2011). Human factors methods and accident analysis: practical guidance and case study applications. Ashgate, Aldershot, UK.

Stanton, N. A. & Stammers, R. B. (2008) Bartlett and the future of ergonomics. *Ergonomics* 51 (1), 1 – 13.

Stanton, N. A., Salmon, P. M., Rafferty, L. & Walker, G. H. (2013). Human factors methods: A practical guide for engineering and design. 2nd Edition, Ashgate, Aldershot, UK.

Stanton, N. A., Salmon, P. M., Harris, D., Demagalski, J., Marshall, A., Young, M. S., Dekker, S. W. A. & Waldmann, T. (2009). Predicting pilot error: testing a new method and a multi-methods and analysts approach. *Applied Ergonomics*, 40:3, pp. 464-471.

Vicente, K.J. (1999). Cognitive work analysis: Towards safe, productive, and healthy computer-based work. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Waldrop, M. M. (1992). Complexity: the emerging science at the edge of order and chaos. New York: Simon & Schuster.

Walker, G. H, Stanton, N. A., Salmon, P. M. & Jenkins, D. P. (2009). Command and control: the sociotechnical perspective. Farnham, UK: Ashgate.

Young, K. L., Lenné, M. G., Beanland, V., Salmon, P. M., Stanton, N. A. (2014). Drivers' visual scanning and head check behavior on approach to urban rail level crossings. Applied Human Factors and Ergonomics conference, 19th-23rd July, Krakow, Poland.

How Much Route Guidance is Enough?

Guy Walker and Alastair Manson

Centre for Sustainable Road Freight
Heriot-Watt University
Edinburgh, UK

ABSTRACT

Four archetypal street patterns were modelled using traffic microsimulation, populated with vehicles possessing various levels of route guidance, and run multiple times to reveal outcomes in terms of journey length, duration, cost and carbon emissions. The findings show that the topology of urban street patterns interacts strongly with the amount of route guidance provided. In some network types the amount of route guidance provided led to consistent improvements in network and vehicle performance, whilst in other networks the same performance could be achieved with 100% route guidance as it could with 0%. A fundamental relationship is revealed between the universal network coefficient Beta (β) and the overall level of route guidance required to optimise performance across all variables. The implications for telematics strategies are profound: it seems every driver does not need to know everything in order to bring about optimal network performance. Indeed, there may be more self-optimization than we think.

Keywords: Street patterns, telematics, network metrics, carbon emissions

INTRODUCTION

Eighty percent of the UK population now lives in an urban area. This drives increasing urban freight movements which, in turn, interact with dense street networks and a strong planning incentive to maximise their capacity and reduce vehicle emissions (Hesse & Rodrigue, 2004). Telematics, in the form of route guidance, is a key enabler for this. There is good evidence for the positive benefits telematics can have (e.g. Asvin, 2008; Giannopoulos, 1996; Dutton, 2011 etc.) but as this technology continues along the s-curve towards full market saturation there are some fundamental questions that still need to be explored. Are some urban road network topologies more energy efficient when paired with route guidance technology than others? If so, to what extent might it influence a telematics strategy? Do all drivers have to have complete knowledge of traffic conditions? How realistic is this assumption anyway? Is it safe to assume that having invested in telematics that drivers will adhere to route guidance information in all cases? Research (e.g. Lyons, Avineri & Farag, 2008; Bonsall & Palmer, 1999; Bonsall, 1992; Chorus et al., 2006; Karl & Bechervaise, 2003 etc.) shows that between 30 and 50% of drivers do not: what happens then? Clearly, the success of telematics technology is heavily contingent on factors like these and this study provides an initial exploration.

Street Networks

Conventional transport network analysis methods use planar representations to reduce a complex transport network into a set of fundamental elements: nodes that represent junctions and links that represent roads (Lowe, 1975). A two dimensional set of systematically organised points and lines like this are referred to as a planar graph. These are the basis upon which various forms of spatial analysis normally proceed (e.g. Bowen, 2012; O'Kelly, 1998). Marshall (2005) identifies four 'archetypal' street patterns, the 'linear', 'tributary', 'radial' and 'grid' (as shown in Figure 1).

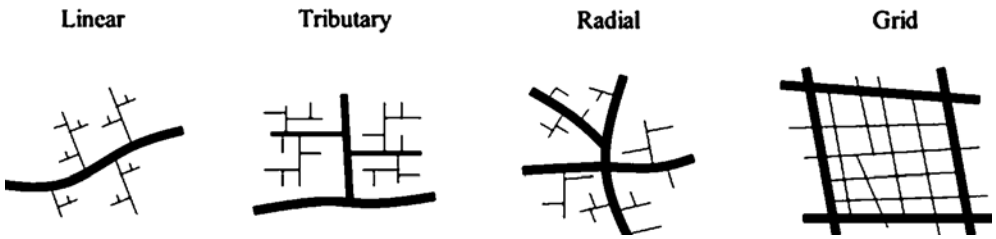


Figure 1. Urban street pattern archetypes

Network diagrams provide a visual representation that, in simple cases, makes it very easy to discern one archetype from another. In more complex real-world networks the visual complexity makes this task difficult to perform reliably and objectively. In these cases it is possible to turn to a number of formal metrics drawn from graph theory. These enable the connectivity of street patterns to be calculated. Three metrics, the 'Beta Index', the 'Gamma Index' and 'Network Depth', are used for this purpose. The Beta Index is a simple equation used to determine the relationship between the total number of links and the total number of nodes in a network and is calculated using the following equation:

$$\beta = e/v$$

Where e = number of links, v = number of nodes and $0.5 < \beta < (v-1)/2$. The Beta Index provides a measure of linkage intensity or "the number of linkages per node" (Lowe, 1975). Beta values generally lie between 0.5 and 3, with networks having values > 1 consisting of nodes within the network having more than one route between them, and the network being considered well connected. Beta values <1 indicate that the network is not as well connected and there being only one route between nodes.

The 'Gamma Index' helps to identify "the ratio between the actual and the maximum possible number of links" in the network (Lowe, 1975). This essentially determines whether or not every node is connected by a link and is derived from the following equation:

$$\gamma = e/(3(v-2))$$

Where e = number of links, v = number of nodes and $0 < \gamma < 1$. For example, if $\gamma = 0.5$ only 50% of the maximum possible number of links in a network are provided and not all of the nodes are fully connected.

The 'depth' of a network is a relatively simple concept. In essence, it accounts for the relative distance between the most minor route and the most major route. It establishes the idea of a hierarchy and the different interconnections that exist between levels. For example, minor roads providing access to houses are 'deep' whilst major routes are 'shallow'. Network depth is important if the goal is to impartially compare different street patterns. If one network was to have a depth of four and another a depth of two, the so called deeper network would have more connectivity. This would allow the road user to travel between points in a shorter period of time as there would be more route choice options for a road user to exploit. Different street patterns (with different Beta and Gamma coefficients) can be legitimately compared if the network depth is kept constant.

Abstracting street patterns to the level of planar networks puts them on the same level as communications and other network types, for which there is a rich literature in the fields of graph theory (Harary, 1994), sociometry (Leavitt, 1951; Monge & Contractor, 2003) and complexity (Watts & Strogatz, 1998). We know from this literature that network type is a strong contingency factor in how they perform (Leavitt, 1951; Pugh et al., 1968; Watts & Strogatz, 1998 etc.). This is in opposition to the tacit theories underlying telematics technologies, which tend to assume that more technology, more route guidance, and more driver knowledge of the wider traffic conditions, the better the

network will perform (Nijkamp, Pepping & Banister, 1997). We know from the wider literature on communication networks that this is not necessarily the case. Of course, roads are not communications networks in the way they have been previously studied, so in order to explore the hypothesis that street pattern is an important contingency factor in the benefits to be accrued from telematics technology, we apply instead an agent modelling approach called traffic microsimulation. Microsimulation enables us to create a set of virtual street layouts based on real towns and cities, populate them with virtual traffic having differing levels of route guidance, and observe the outcomes in terms of journey length, duration, cost, and most importantly, carbon emissions. This is an exploratory study which aims to examine the highly systemic nature of traffic and transport, revealing how strategies focused on individual drivers may (or may not) propagate through the system as a whole to give rise to effects that we want and expect.

METHOD

Design

The study uses traffic microsimulation to test the interaction between different levels of vehicle telematics and the outcomes achieved within different street patterns. There are four dependent variables: journey duration, journey length, journey cost and carbon output. These dependent variables are contingent upon two independent variables: driver knowledge of the traffic conditions in the network provided by telematics, and street pattern type. Driver knowledge had three levels. 100% telematics represented urban delivery vehicles in which every driver had complete knowledge of the traffic conditions on the network and was required to act upon the guidance given (a 'best-case' telematics implementation). 50% telematics represented other non-freight vehicles in the network equipped with telematics/route guidance/sat nav etc. but which, in accordance with the literature, only complied with/took notice of 50% of the information provided. 0% telematics represented vehicles with no route guidance, with the driver having only immediate local knowledge of traffic conditions based on what they can see. Street pattern had four levels: linear, radial, grid and tributary. The street patterns were based on real urban locations, with network depth held constant in order to control for non-systematic biases in connectivity. Network demand was based on real-life traffic count data from the relevant sites.

Real-World Problem

As determined in the previous section, there are four common forms of urban transport network layout: linear, radial, grid and tributary. It was necessary to develop microsimulation models which closely reflect the layouts of these urban networks and it was considered important to relate the layouts to real life towns and cities. This gives a more realistic approach and a more credible set of results. The Beta and Gamma coefficients were used to calculate the connectivity of the network archetypes shown in Figure 1 in order to select real-life road networks exhibiting the same properties. This approach allows networks to be categorised by a visual and a statistical approach. The depth of each model was set to three in order to control for the effects of network connectivity.

The Linear network required locating a small town in which there is one main road with the town located along its length. A settlement identified as meeting these criteria is Aviemore, a town located in the Scottish Highlands. The Radial network is one that has several roads intersecting or converging at a centre, analogous to the spokes of a bicycle wheel. These features can often be found in a larger town which has formed at a cross roads. These criteria are met by Dalkeith, a town located just to the south of Edinburgh. The Grid network, as its name suggests, is a network with straight roads intersecting other straight roads at right-angles to form a collection of squares or blocks. The characteristics of this network type are found in the center of Glasgow. Finally, the Tributary network is analogous to tributary rivers, with the smaller rivers feeding the bigger rivers. In road networks, it is the small roads that connect to the larger roads with 'Network Depth 1' only connecting to 'Depth 2' and 'Depth 2' only connecting to 'Depth 3' and so on. The result of this is that only the shallower roads are busy and the deeper roads are not used as shortcuts. An area meeting the description of a Tributary network is Livingston, a so called 'new town' in central Scotland. Planar graphs of these real-life street networks are shown in Figure 2 along with their respective network metrics.

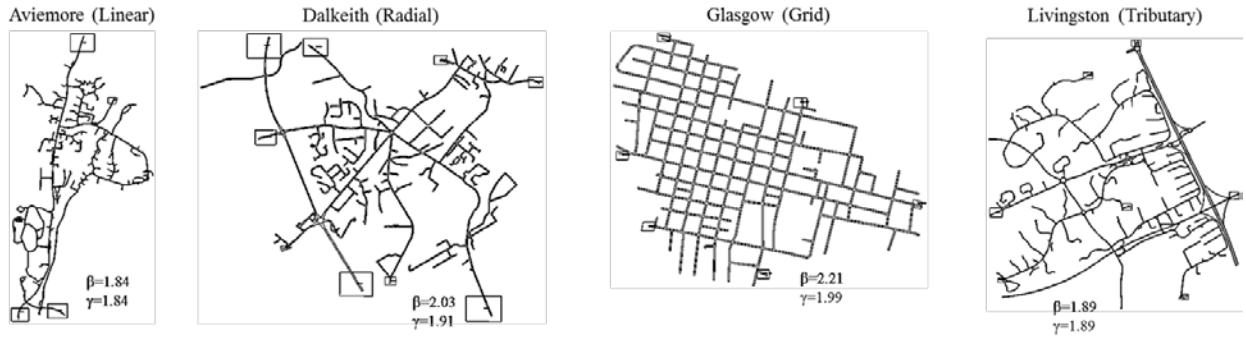


Figure 2. Traffic microsimulation models of Aviemore (tributary), Dalkeith (radial), Glasgow (grid) and Livingston (tributary).

Beta and Gamma values were calculated for each of the real-life towns to check the extent to which they conform to the linear, radial, grid and tributary archetypes. The real-life networks are considerably larger than the archetypes yet it can be noted how closely the network metrics align. The Beta values for the 'Linear' and 'Aviemore' networks, for example, are exactly equal as are the Beta values for the 'Tributary' and 'Livingston' networks. Even where differences do exist (and they are relatively modest) the rank-order of the network types is still preserved.

Development of Network Models

Planar graph representations of Aviemore, Dalkeith, Glasgow and Livingston were extracted from ArcGIS (a mapping and spatial analysis software) and imported into S-Paramics (the traffic microsimulation software) in order to create a basic model of each. An attempt was made to ensure that the modelled network was as true to life as possible but some simplifications were required to ensure uniformity. As such, junctions had simple priorities applied with the bigger roads having priority over the smaller roads. Traffic lights were avoided all together. There were also no buses or bus routes applied as not all networks had buses arriving at similar times and some networks had bus lanes whereas others did not. The so-called 'signpost distance' refers to the distance ahead of a hazard that the modelled road users become aware of it, and was standardised to 80 m to allow the traffic to use both lanes of the entrances and exits of roundabouts (anomalous behaviour would arise if not). Visibility was also varied. This relates to how far back from the stop-line vehicles begin assessing their gap distance. This ensured vehicles approaching roundabouts continued onto it in situations where nothing was coming and stopped in cases where something was approaching. The distance which meant vehicles followed this rule was found to be 20m. These rules allowed more life-like driver behaviour to be represented and were applied consistently across all the networks in order to isolate the effects of network and vehicle type.

Vehicle Types

The principle vehicle-based manipulation was the amount of route guidance provided to drivers, or as modelled, the amount of knowledge drivers had of wider network conditions. The three levels of familiarity were as follows:

- Vehicle Type 1 (Urban Delivery Vehicle) - 100% knowledge: drivers have full knowledge of the state of the current road network via telematics and they follow route guidance implicitly.
- Vehicle Type 2 (Private Car) - 50% knowledge: implies a vehicle fitted with telematics but one in which the driver chooses to ignore 50% of the network information provided.
- Vehicle Type 3 (Private Car) - 0% knowledge: implies a vehicle without telematics, hence the driver has no knowledge of the road network's state beyond what they can see ahead of them.

The three levels of familiarity were implemented in the microsimulation model by manipulating feedback. Feedback is information supplied to the road user, via telematics, about the current network conditions. Specifically, it lets them know of journey times on all routes so they can decide an optimum route to choose. Feedback is calculated using two aspects, the feedback interval and the feedback factor. The feedback interval refers to how often the information is updated and an interval of two minutes was used (this being a common value in similar models). The feedback factor is concerned with what percentage of delay information is taken from the previous feedback interval and was taken to be 0.5 or 50% as standard. A perturbation level of 5% was applied to

all three vehicle types. This helps to account for variability in travel costs, or a driver's perception of these travel costs. As perturbation increases, the road users' concentration tends to focus more on reducing journey cost. However, by applying a small percentage it means that road users will continue to focus on reducing journey length and journey duration as the key variants but will also look to reduce their cost simultaneously.

The same mix of traffic was applied to all the networks in order to capture the interaction between 'normal traffic' and urban freight delivery vehicles. 50% of the road users in all the network types were 'Vehicle Type 3' (0% telematics) with the remaining 50% being split in two; 25% of road users were 'Vehicle Type 2' (50% telematics) and 25% of road users being urban delivery vehicles, or 'Vehicle Type 1' (100% telematics). The normal traffic (i.e. Vehicle Types 2 and 3) were modelled as medium sized saloon cars, whereas the urban delivery vehicles (i.e. Vehicle Type 1) were modelled as car-based vans. This is a) because larger vehicle types would give unfavourable differential effects on networks that are less suitable for larger vehicles and b) the literature identifies vans as the 'dominant mode' in urban contexts (e.g. Cherrett, 2012).

Network Demand

The demand profile controls the number of vehicles in the network, the origin and destination of the vehicles, the percentages of vehicle type and the release rate of vehicles into the network. The number of vehicles released from the Origin to the Destination was varied from model to model based on actual traffic count data. The goal was to bring each network to its peak PM traffic flow and hold it there for the duration of the study. To do this, peak PM values from the nearest traffic counter site (Scottish Government, 2012) were used and multiplied by 24. This total value was then split evenly between each of the Origins and Destinations. The release rate of the vehicles over the 24 hour period was constant so that all the vehicles were not released at once and a steady flow was maintained. Each model was then subject to 30 'batch runs' between the network model hours of 1600 and 2000, which was the length of the PM peak flow period. The models had, of course, been established in this peak flow state for many hours previously hence any transient effects of the model being initially loaded with traffic were avoided. These 30 runs allowed a significant amount of data to be obtained on the outcome variables of journey duration, length, cost, carbon output and their contingency on network type.

RESULTS AND DISCUSSION

S-Paramics microsimulation outputs a set of raw data for each individual vehicle as it progresses through the network. The software calculates the position of each individual vehicle every half second and records its co-ordinates against time. This data was passed through an external programme known as 'AIRE' which calculates the resulting pollution output of each individual vehicle for a specific year, which was chosen to be 2012. An average value for Carbon Output was obtained for each vehicle type, along with data on journey duration, length and cost. A summary of the results is shown in Table 1, corrected for network size/distance, and collapsing some of the key variables into average speed, cost per km, carbon output in g/Km and total emissions.

Table 1 - Summary of corrected speed, cost and carbon output data

	Vehicles		Speed	Cost	Carbon		
	(Total N)	(Telematics)	(Km/h)	(pence/Km)	(g/Km)	(total Kg)	
Linear	3000	0%	42.28	28.05	60.39	390.00	
		50%	42.47	27.66	59.62	385.50	
		100%	42.65	27.31	59.15	384.32	
		Mean	42.47	27.67	59.72	**1159.82**	Total
Radial	4752	0%	42.30	26.34	60.23	654.40	
		50%	42.40	26.34	60.14	651.93	
		100%	42.48	26.26	60.02	651.95	
		Mean	42.39	26.32	60.13	**1958.28**	Total
Grid	6000	0%	31.84	75.79	98.88	773.99	
		50%	33.25	70.47	92.76	732.60	
		100%	34.74	65.32	86.67	690.02	
		Mean	33.28	70.53	92.77	**2196.61**	Total
Tributary	5995	0%	42.56	15.99	58.57	1285.97	
		50%	42.58	16.00	58.50	1281.77	
		100%	42.63	15.97	58.52	1285.29	
		Mean	42.59	15.99	58.53	**3853.03**	Total

Ignoring the effects of route guidance for a moment, it can be noted that the linear and radial networks are very similar, both being able to support average speeds in the region of 42Km/h, costs per Km of approximately 27 pence, with vehicles on the network each emitting approximately 60g/Km of carbon. The tributary is the same aside from cost, which is the lowest of all the road networks at 15.99 pence/Km. The slowest (33.28Km/h), most expensive (70.53 pence/Km) and carbon intensive (92.77 g/Km) network is the 'Grid' but it is within this network that the urban delivery vehicles (with 100% route guidance) performed the best. In this situation telemetry is raising average speeds by 2.9Km/h, reducing costs by a not insignificant 10.47 pence/Km and, most importantly, reducing carbon emissions by 12.21g/Km. The carbon value is of course determined by the physical size of the networks and the journey lengths therein, so in these examples the larger tributary network (i.e. Livingston) has the highest total emissions (3.8 tonnes of carbon per modelled PM peak), however, the network with the shortest average journey lengths (i.e. Glasgow / Grid) has the second worst carbon outputs (2.2 tonnes).

There are smaller differential effects present in the other network types that are also important. Although smaller at the level of individual vehicles, when multiplied by the number of vehicles in the networks (several thousand) and the number of times PM peak hour conditions occur (every weekday evening) these differences begin to magnify significantly. For example, in the tributary network, a per-vehicle difference in carbon emissions of only 0.7 grams as a result of route guidance still multiplies to an additional daily PM peak carbon output of approximately 4.2Kg, or approximately 1 tonne per year. Multiplied again by the number of settlements with tributary street patterns and these initially marginal differences start to accumulate rapidly. With this in mind, the following sections shift the focus from absolute values to relative values in order to discern the direction of these various effects, and what they might mean for an overall telematics strategy.

Route Guidance vs. Journey Duration

Urban delivery vehicles (with 100% telematics) operating within linear and tributary road networks are worse off than 'normal traffic' with 50% knowledge/acceptance of route guidance information, albeit not as bad as the population of vehicles with no route guidance. Indeed, the results show that the same outcomes on journey duration are achieved at both 100% telematics penetration and approximately 20%, the difference being that the latter is considerably more costly than the former. The trend line obtained for the 'Radial' network only slightly differs from the 'Linear' and 'Tributary' networks in respect of the duration increasing between 50% and 100%, and levelling off thereafter. This shows there is no further benefit of route guidance in this context. The relationship for the 'Linear', 'Radial' and 'Tributary' networks seem to show that there is an optimum level of route guidance. For the optimisation of journey durations a level of knowledge consistent with 100% telematics is not required. This principle does not hold for the 'Grid' network. Here the trend line shows increasing benefits as more telematics is provided, reaching a maximum benefit at 100% telematics penetration. The broader principle to be extracted here is that 50% route knowledge is optimum for networks characterised by one or two critical routes between the majority

of the origins and destinations. 100% telematics penetration reduces journey durations in networks where there are a large range of routes available to the road user. In this study, urban freight vehicles with 100% route guidance extract maximum journey time benefits in Grid networks.

Route Guidance vs. Journey Length

In the linear network, the journey length increases at a slow rate between 0% and 50% route guidance before it increases at a much greater rate between 50% and 100%. This relationship is very different from the Radial and Tributary networks, which both have a long journey length for road users with 0% but falling markedly between 0% and 50% before rising again very steeply between 50% and 100%. As with the results for journey duration, there is an optimum level of route guidance for Radial and Tributary networks (around 50%), with the same network performance achieved at 0% route guidance as achieved at 100%. Urban delivery vehicles with 100% route guidance would see little benefit in these settings. The Grid network differs again. The greater the amount of route guidance the greater the journey length as road users exploit the more numerous opportunities to divert. Taken together, the findings suggest that shorter journey durations are achieved with longer journey lengths. The broader principle to be extracted here is that urban freight vehicles in networks with more than one route between the origin and destination travel a greater distance but in a shorter time.

Route Guidance vs. Journey Cost

The Linear network has a more or less direct relationship between route guidance and journey cost (i.e. a straight line). It shows that as route guidance increases from 0% to 100%, the journey cost also increases. Interestingly, in this network type 0% route guidance is the optimum value for cost to be optimised. This is not the case for Radial and Tributary networks. Both of these undergo a reduction in cost between 0% and 50% with an increase then occurring between 50% and 100%. Once again, the optimum level of rote guidance in a Radial or Tributary network is 50%, with further increases not only having a negative effect on journey cost but 100% telematics penetration (as per the urban freight vehicles) yields the same outcome as 0%. The Grid network again performs differently to the other three network types. The relationship is linear (i.e. a straight line) between 0 and 50% route guidance, before tailing off slightly as 100% telematics penetration is reached. What this means is that 100% route guidance is required in Grid networks for meaningful journey cost savings to emerge, 0% for linear networks and 50% for Radial and Tributary networks.

Route Guidance vs. Carbon Emissions

The crux of the analysis is to see what effect these contingent values of journey length, duration and cost ultimately have on carbon emissions. The results show the Linear, Radial and Tributary networks all following a similar trend, with carbon emissions decreasing rapidly between 0% and 50% route guidance but with varying levels of diminishing further benefits as route guidance approaches 100%. The Linear and Radial networks level off beyond 50%, suggesting little (if any) further benefits of increasing the amount of route guidance. The results suggest the carbon emissions from the Tributary network worsen with increases beyond 50% route guidance. The 'Grid' network is once again quite distinct. It contains a directly proportional relationship between route guidance and carbon emissions, with the maximum value occurring at 0% and the minimum value at 100%. If reducing carbon emissions are the goal of urban freight vehicles then route guidance has the biggest role to play in Grid networks, and apart from Tributary networks, there are some benefits to be extracted before 100% route guidance is achieved.

Optimisation Values

Based on the data and relationships obtained from the previous sections, it is possible to create Table 2 to show the potential network performance trade-offs involved in minimising carbon emissions using route guidance technology. A mean level of route guidance (or acceptance thereof) is given, this representing a simple value by which the best compromise of journey duration, length, cost and carbon variables is achieved.

Table 2 – Optimum levels of telematics penetration for journey duration, length, cost and carbon emissions for linear, tributary, radial and grid street patterns

Network	Duration	Length	Traveller costs	CO_2	Mean
Linear	50%	0%	0%	100%	38%
Radial	50%	50%	50%	100%	63%
Grid	100%	0%	100%	100%	75%
Tributary	50%	50%	50%	50%	50%

Table 2 shows that, in theory, urban street networks can be designed to be sustainable with the smallest pollution outputs, designed to reduce the journey duration of road users travelling through them, or designed to reduce traveller costs. In practice, however, street networks have evolved and cannot be changed on the scale necessary to optimise these factors. This is where telematics comes in. Telematics interacts with network types to modify their inherent performance but the interaction is not a simple one. It is contingent on the level of telematics information provided to, and accepted by, drivers, combined with the topology of the network itself. Where some networks require varying amounts of route guidance to optimise different aspects of the network, some run at the optimum level for the majority of characteristics under one level of route guidance. This can be seen in Table 2 with the Tributary and Radial networks, which both run at their most efficient levels for all four characteristics with a 50% telematics rate. The table also shows that for a Grid network, 100% route guidance is the optimum value. Linear networks, however, do not reach their optimum level of efficiency for any one level of telematics. Table 2, therefore, becomes a useful tool when attempting to design a telematics strategy in order to modify the inherent characteristics of urban road networks. Stated simply, some levels of driver knowledge of network conditions (provided via telematics) are more optimal than others, and it depends on the street pattern urban road users are operating within.

Fundamental Relationships

The results shown and discussed above convey the idea that for each output characteristic (i.e. journey length, cost, duration, carbon emissions) different vehicles perform differently depending on the route guidance provided to drivers. What if a particular urban freight context does not conform to the archetypes presented? In this case it is possible to increase the generalizability of the results with recourse back to the connectivity coefficients discussed earlier. These can be applied to any transport network, of any size or type, in order to reveal its underlying level of connectivity.

Figure 2 – Relationship between the Beta Coefficient (β)and the level of telematics required to optimise journey length, duration, cost and carbon outputs.

A fundamental relationship emerges: as the amount of route guidance increases from 0% to 100% in a road network with Beta (β) ≥ 1.9, Journey Costs, Journey Duration and Carbon Output tend to improve (or at least do not worsen). In networks with β < 1.9 there is no added benefit of providing anything more than 50% route guidance. Figure 2 illustrates the relationship between β and telematics within the range of data observed. Figure 2 represents a simple

diagnostic for answering the question 'what are the benefits of increasing drivers' knowledge of the wider network conditions, via telematics, in a particular operating environment'.

CONCLUSIONS

Contrary to general belief, this study shows there is a point at which more route guidance does not lead to more efficient urban logistics. Indeed, in many cases the same outcomes can be achieved with 0% route guidance as can be achieved with 100%. Simply introducing an abundance of telematics into an urban freight situation, with planners and company policy enforcing 100% compliance with route guidance, however feasible that may be, will not always result in the outcomes expected. This paper shows that the topography of an urban street layout is an important contingency factor in how and when to deploy this technology.

This is an exploratory study but a number of potentially important implications arise from it. The first relates to previous research showing between 30 and 50% of drivers do not comply with telematics-based route guidance no matter how much of it is provided. As the results show, middle values like these represent an optimum on many outcome variables and street network types. This raises an interesting point. Does this established 30 to 50% driver acceptance of route guidance arise through repeated experience of traffic conditions on a network, the implication being that driver behaviour (and the network itself) are self-organising? If this is the case, then is the imposition of telematics based on a false premise? Does 'everyone' in the network need to know 'everything'? The results of this study would seem to suggest that, for some network types, they do not. If this is the case then the same outcomes could be achieved for considerably less cost.

The second implication relates to the different telematics strategies that urban logistics providers could adopt, and when. Is a costly strategy of providing complete knowledge of the network to drivers, and enforcing compliance with route guidance, optimum in all situations? No. Likewise, is a laissez fair approach to planning and route guidance, based purely on ad-hoc local knowledge brought to situations by individual drivers, optimum? Again, no. Optimisation is contingent upon the topology of the network being travelled upon. There are clearly some situations where it would benefit outcome variables such as carbon emissions to impose compliance with route guidance, and other situations where it would not be appropriate. The results of this paper are helpful in understanding what these relationships might be and what an 'adaptive telematics' strategy might look like. It would be a form of route guidance that would be cognitively compatible with drivers. One in which the timing and sequence of route guidance information would be oriented around different outcome variables at different times, but in all cases offering tangible journey based 'rewards' for the driver. These rewards would encourage telematics to be used in ways that exceed the current 30-50% acceptance rate where it is beneficial to do so, and as such, to accumulate some significant marginal gains.

ACKNOWLEDGEMENTS

This research was carried out under EPSRC Grant EP/K00915X/1 held jointly by Heriot-Watt University, Edinburgh and the University of Cambridge. This research also benefitted from a collaboration between Heriot-Watt University and SIAS Transport Planners, 37 Manor Place, Edinburgh, [UK]EH3 7EB, the developers of the traffic microsimulation software S-Paramics. Further details available at: www.sias.com.

REFERENCES

Asvin, G. (2008). Fleet telematics: real-time management and planning of commercial vehicle operations. Boston, MA: Springer.

Bonsall, P, (1992). The influence of route guidance on route choice in urban networks. Transportation 19, 1-23.

Bonsall, P.W. and Palmer, I. (1999). Behavioural response to roadside variable message signs: factors affecting compliance. In: Emmerick, R and Nijkampt, P. (eds). Behavioural and Network Impacts of Driver Information Systems. Farnham: Ashgate.

Bowen, J. T. (2012). A spatial analysis of FedEx and UPS: hubs, spokes, and network structure. Journal of Transport Geography, 24, 419-431.

Brindle, R. E. (1996). Urban Road Classification and Local Street Function. Australian Road Research Board.

Cherrett, T., Allen, J., McLeod, F., Maynard, S., Hickford, A. & Browne, M. (2012). Understanding urban freight activity – key issues for freight planning. Journal of Transport Geography, 24, 22-32.

Chorus et al. (2006). Use and effects of Advanced Traveller Information Services (ATIS): A review of the literature. Transport Reviews, 26(2), 127-149.

Dutton, G. (2011). Fleet management's magic box. World Trade, 24(2), 38-44.

Giannopoulos, G. A. (1996). Implications of European transport telematics an advanced logistics and distribution. Transport Logistics, 1(1), 31-49.

Harary, F. (1994). Graph Theory. Reading, MA: Addison-Wesley.

Hesse, M. & Rodrigue, J. P. (2004). The transport geography of logistics and freight distribution. Journal of Transport Geography, 12(3), 171-184.

Karl, C. A. and Bechervaise, N. E. (2003). The learning driver: issues for provision of traveller information services. 10th World Congress and Exhibition on Intelligent Transport Systems and Services, Madrid.

Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. Journal of Abnormal and Social Psychology, 46, 38-50.

Lowe, J. C. (1975). The geography of movement. Boston: Houghton Mifflin.

Lyons, G.,Avineri, E. and Farag, S. (2008). Assessing the demand for travel information: do we really want to know? Proc. European Transport Conference, Noordwijkerhout, Netherlands, October.

Marshall, S. (2005). Streets and Patterns. Oxon and New York: Spon Press.

Monge, P. R. & Contractor, N. S (2003). Theories of communication networks. New York: Oxford University Press.

Nijkamp, P., Pepping, G. & Banister, D. (1997). Telematics and transport behaviour. Berlin Springer-Verlag.

O'Kelly, M. E. (1998). A geographer's analysis of hub-and-spoke networks. Journal of Transport Geography, 6(3), 171-186.

Pugh, D. S., Hickson, D. J., Hinings, C. R. & Turner, C. (1968). Dimensions of organisation structure. Administrative Science Quarterly, 13(1), 65-105.

Ratan, P. (2013). Telematics: an emerging solution to streamline logistics. Cargo Talk, 13(5), 20-24.

Reggiani, A., Lampugnani, G., Nijkamp, P. & Pepping, G. (1995). Towards a typology of European inter-urban transport corridors for advanced transport telematics applications. Journal of Transport Geography, 3(1), 53-67.

Thomson, J. M. (1977). Great Cities and Their Traffic. London: Victor Gollancz Ltd.

Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. Nature, 393(4), 440-442.

Life in the Wild? Using Ergonomics Systems Methods and Field of Safe Travel Theory to Evaluate Intersection Designs

Miranda Cornelissen [a,b], Paul Salmon [c], Neville Stanton [d] and Roderick McClure [a]

[a] Monash Injury Research Institute
Monash University, Melbourne, Australia

[b] Griffith Aviation, School of Biomolecular and Physical Sciences
Griffith University, Brisbane, Australia

[c] University of the Sunshine Coast Accident Research, Faculty of Arts and Business
University of the Sunshine Coast, Maroochydore, Australia

[d] Civil, Maritime, Environmental Engineering and Science Unit,
Faculty of Engineering and the Environment
University of Southampton, Southampton, United Kingdom

ABSTRACT

A safe systems approach has been acknowledged as the underlying philosophy of contemporary road safety strategies. Despite this, systemic applications in road transport evaluation and design remain sparse. This paper explores the value of using Ergonomics design and evaluation methods such as Cognitive Work Analysis in conjunction with road transport theories such as the field of safe travel to provide easily interpretable analyses of road designs. The goal is that this would provide a facilitation platform for communication between Ergonomics analysts and road transport designers, aiming to facilitate systemic applications in road transport. The application of Cognitive Work Analysis and the field of safe travel theory in the evaluation of a new intersection design concept demonstrated that this proves a promising cross method collaboration. Cognitive Work Analysis provided the analytical detail of road user behavior possible as a result of the interaction between the intersection design, road users, vehicles and the environment. Subsequently, the field of safe travel theory provided a visual means to communicate these findings directly related to the intersection designs. The application furthermore provided additional insights into the constraints acting upon the field of safe travel and the paths that road users can possibly take within this field.

Keywords: Cognitive Work Analysis, Field of safe travel, Systems safety

INTRODUCTION

A safe systems approach has been acknowledged as the underlying philosophy of contemporary road safety strategies (Corben, Logan, Fanciulli, Farley, & Cameron, 2010; Johansson, 2009; Koornstra, Mathijssen, Mulder, Roszbach, & Wegman, 1992). Despite this, systemic applications considering more than just individual road user groups (e.g. drivers) or single countermeasures (e.g. road design) are sparse.

Systems-based methods from the discipline of Ergonomics have a key role to play in road transport design and evaluation. Outcomes of Ergonomics methods analyses are, however, often abstract representations. These include for example, hierarchies of goals (Hierarchical Task Analysis (Annett, 2004)) or complex networks and diagrams (e.g. Cognitive Work Analysis (CWA; Rasmussen, 1997; Vicente, 1999)) and Event Analysis of Systemic Teamwork (EAST; Stanton, Salmon, Walker, Baber, & Jenkins, 2005)). First explorations of representing such analysis in a real world context were promising (Cornelissen, Salmon, Stanton, & McClure, 2012), however, the absence of a visualization method translating abstract analyses to real world concepts made it difficult to conceptualize the outcomes and to communicate with, for example, design teams.

The field of safe travel theory (Gibson & Crooks, 1938) is a theory of driving relating the driver to the road transport environment and has underpinned application of Ergonomics methods such as CWA in road safety, e.g. designing in-vehicle warning systems (Jenkins, Stanton, Walker, & Young, 2007; Stoner, Wiese, & Lee, 2003; Young & Birrell, 2010). While current applications of the field of safe travel theory and CWA have focused mainly on interface design, the field of safe travel theory shows potential to translate Ergonomics methods analyses output to the road transport design context.

The aim of this paper is to examine the value of using Ergonomics systems analysis methods in conjunction with the field of safe travel theory to provide easily interpretable analyses of road environments and road design concepts. The ambition is that using a combination of these methods will provide a facilitation platform for communication between Ergonomics analysts and road transport designers.

Cognitive Work Analysis

In the field of Ergonomics, the systems approach to road safety is beginning to be actualized through the application of systemic modeling methods such as CWA (Cornelissen, Salmon, McClure, & Stanton, 2013; Cornelissen, Salmon, & Young, 2012). CWA is a popular framework used for designing and evaluating complex sociotechnical systems. It describes system constraints and potential behavior possible within those constraints (Vicente, 1999). The framework's five phases, as outlined by Vicente (1999), all model a different constraint set. First, Work Domain Analysis describes overall system constraints from physical objects to the functional purpose of the system. Second, Control Task Analysis describes situational constraints and decision making requirements. Third, Strategies Analysis models different ways in which activities can be carried out within a system's constraints. Fourth, Social Organization and Cooperation Analysis describes communication and coordination demands resulting from organizational constraints. Fifth, Worker Competencies Analysis describes skills, rules and knowledge required by actors within the system.

CWA has been used successfully to examine the interactions of road users and road system components (Cornelissen et al., 2013; Cornelissen et al., 2012). Past applications, however, represent evaluations of road user behavior induced by existing road transport systems (Cornelissen et al., 2013; Cornelissen et al., 2012; Stoner et al., 2003) and design of driver support systems (Birrell, Young, Jenkins, & Stanton, 2011; Hilliard & Jamieson, 2008; Lee, Hoffman, Stoner, Seppelt, & Brown, 2006; Seppelt & Lee, 2007). Applying systems based methods such as CWA during the road design process will provide the opportunity to proactively address design issues and produce road design concepts that alight with the systems approach and support all road users. Whilst having a safety benefit there is also a significant cost benefit since road design concepts can be modified early in the design lifecycle before they are implemented in the real world. Moreover, changing the role of Ergonomists from 'operational system problem solvers' to 'design concept problem solvers' is one of the major challenges facing our discipline. One of the key facilitators of this is the development of Ergonomics methods that can identify safety-related problems early on in the design process.

Field of safe travel

The field of safe travel theory (Gibson & Crooks, 1938) is one of the earlier road transport theories attempting to describe the relationship between driver and the road transport environment. Driving is regarded a type of locomotion through a terrain or field of space. The function of locomotion is to move an individual from one point in space to another, the destination. To reach the destination, location must be modified to avoid obstacles. The field of safe travel stretches out in front of a road user, exists within the boundaries of the road and consists of the possible paths a road user may take unimpeded, see Figure 1.

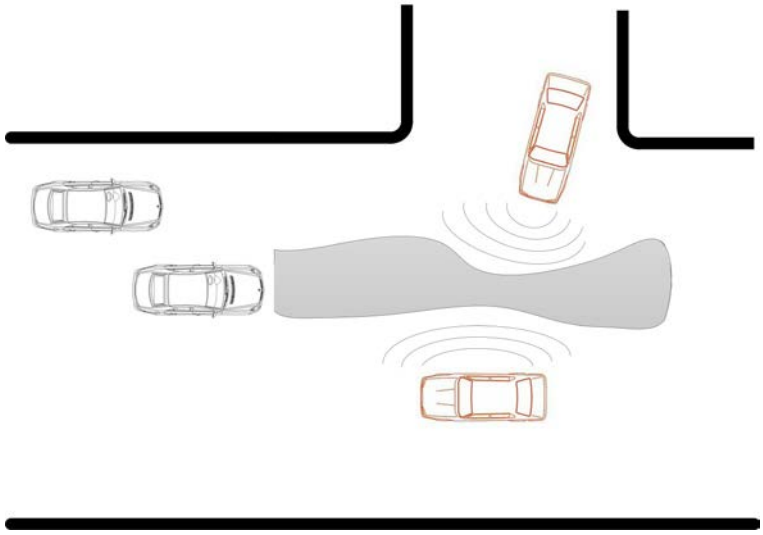


Figure 1. Field of safe travel (Adapted from Gibson & Crooks, 1938)

Cognitive Work Analysis & Field of safe travel

Gibson was a forefather of the ecological psychological movement of which CWA is a product. The field of safe travel has been cited to explain the relevance of constraint based approaches to describe complex systems such as road transport. Applications of field of safe travel and CWA to design in-vehicle systems have focused mainly on the first phase of CWA, Work Domain Analysis. These applications aimed to visualize obstacles, system constraints or the field of safe travel to drivers to help them appropriately modify and avoid obstacles (Jenkins et al., 2007; Stoner et al., 2003; Young & Birrell, 2010).

To support a systems approach, application of CWA beyond WDA is encouraged. This would explore a wider range of constraints influencing the task of negotiating a road transport system, the possible paths within the field of safe travel and the interaction between constraints and paths. CWA helps describe the field of safe travel of the road user based on the interaction of the road user with the road transport environment. The field of safe travel theory acts as a visual mediator between CWA outputs and road transport evaluations and design.

METHOD

Intersections

Two road designs were evaluated; an existing traditional Melbourne intersection and a newly proposed cut-through intersection design concept that was developed based on road transport safe-system principles, see Figure 2.

Traditional, arterial intersections in Melbourne are typically signalized, carry multiple lanes of traffic and speed limits on approach are between 60 and 80 km/hour. Often these intersections, if space permits, have a slip road to carry left turning traffic that can turn without using the signals. In Australia, road users travel on the left-hand side of the road. For right hand turns, road users that travel on the main road approach the intersection as close to the center of the road as possible and turn just right of the center of the intersection. Pedestrians and cyclists that turn right approach the intersection on the left hand side and use the pedestrian crossings to travel to the far right opposite corner of the intersection.

As part of a major road safety project, a new intersection design concept was produced aimed to improve intersection safety through infrastructure design (Corben, Candappa, Van Nes, Logan, & Peiris, 2010). The intersection design, named the cut-through intersection, has traffic islands fitted in the middle of the intersection. Right turning traffic is expected to use the cut-through lane created by the distribution of these traffic islands. Their turn is protected from oncoming traffic by the traffic islands and it changes the angle at which traffic meets to less than 90 degrees. To use the cut-through lane, road users have to be in the right hand lane before they approach the intersection. An additional traffic island separates right hand turning traffic wanting to use this lane from straight

through traffic upon approach. The intersection has a similar footprint to the traditional intersection and due to the circular shape of this intersection no slip lane is fitted. Left turning traffic merges with and diverges from straight through traffic in the intersection.

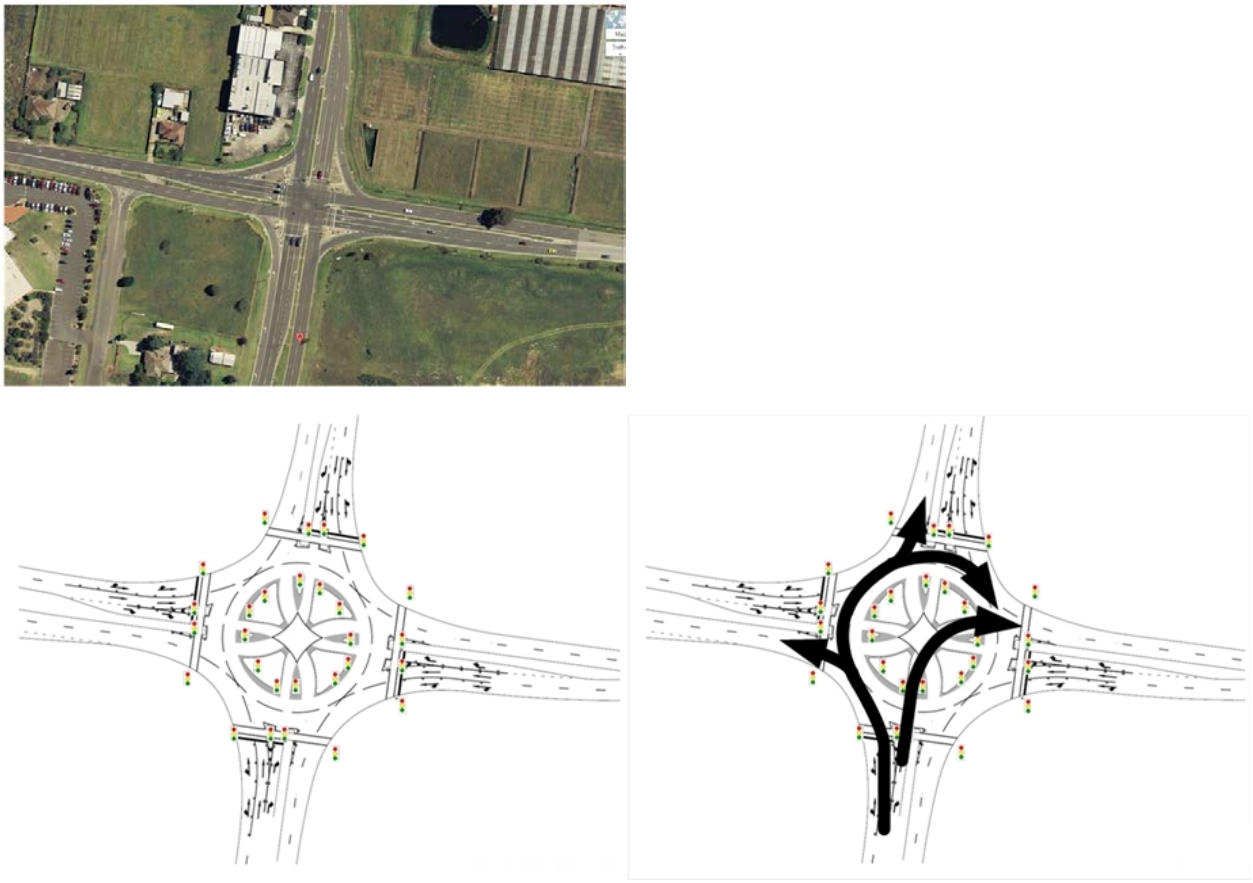


Figure 2. Traditional Melbourne intersection (above) and cut-through design (below).

Data collection

The analysis of the traditional intersection was based on data derived from an on-road study of different road users' behavior at intersections (c.f. Cornelissen, Salmon, McClure, & Stanton, 2013; Cornelissen, Salmon, & Young, 2012) as well as document analysis of publically available documentation on Victoria's road system (VicRoads, 2009, 2010, 2011). The on-road study data comprised verbatim transcripts of Verbal Protocol Analysis (VPA) and Critical Decision Method (CDM) obtained from drivers, motorcycle riders, cyclists and pedestrians as they negotiated intersections as part of a predetermined route through Melbourne.

The analysis of the cut-through design was based on the CWA analysis of traditional intersections. The analysis was complemented by information on those elements that were different as obtained from the design documents and discussions with the designers. For example, the removal of the slip lane, addition of traffic islands and change in traffic light design.

RESULTS

Cognitive Work Analysis

The Abstraction Hierarchy (AH) was used to assess the system constraints of both designs. System constraints are the constraints set on the system by its purpose, functions that need to be carried out and physical objects. The main difference between the system constraints of the traditional intersection and the cut-through intersection is the

physical objects of which the system is made up off. In particular, the slip lane is no longer available and the lane markings across the intersection are now replaced by traffic islands.

Situational constraints on road user interaction with the intersections were assessed using the Contextual Activity Template (CAT). The CAT was used to describe when certain functions can be executed (e.g. on approach, in the intersection or when exiting the intersection). This is where differences between the intersections became apparent. The distribution of the physical objects through the cut-through intersection, e.g. the traffic islands, results in changes to when decisions have to be made. For example, when turning right 'determine a path' across the intersection and 'take a lane' now have to occur on approach before road users encounter the additional traffic island and enter the cut through lane. Once in the intersection the traffic islands in the intersection prevent changes to these decisions.

Decision ladders were used to evaluate the decision making processes. Decision ladders were used to analyze the options road users have to execute the function, the information elements used to decide between the options and subsequently how the functions can be executed. For example, the options to turn right have changed due to the circular shape of the cut-through intersection and the creation of the cut-through lane. To turn right in traditional intersections, road users that travel on the road (drivers, motorcycle riders and cyclists) could do a hook turn, travel in a different direction and then do a u-turn or take the right hand turning lane. In the cut-through intersection, road users can also travel the long way around such as when travelling in roundabouts. Decision ladders further showed that while decisions to 'determine a path and lane' have to be made early on approach in the cut-through intersection, the information elements, e.g. traffic lane arrows, have not been provided at that time and place.

The strategies analysis evaluated the effect of intersection changes on road user behavior that was possible using the Strategies Analysis Diagram (SAD), see Figure 3 for an extract. The SAD was used to describe different ways in which functions could be executed within system constraints. The analysis shows that the range of behavior possible is different across the intersections. For example, the traffic islands in the cut-through intersection make it easier for drivers, motorcycle riders and cyclists to avoid conflict with other users while in the intersection as they are protected by the traffic islands as physical barriers. The traffic islands also influence behavior possible by pedestrians and cyclists (if they take the pedestrian route). The traffic islands, for example, allow protection and movement for these road users. This is the way they are designed to be used for pedestrian crossings on the outer ends of the intersection. The traffic islands in the middle of the intersection however afford the same protection and movement. Therefore pedestrians may use the traffic islands to travel diagonally across the intersection. These strategies are likely when traffic volumes are low or when traffic is stopped and pedestrians feel that this strategy satisfies both the efficiency and safety values. The SAD was used to evaluate complementary nature and redundancy of road user behaviors. For example, the CAT showed that decisions to 'determine a path and lane' had to be made earlier on approach in the cut-through intersection but that the information elements to make these decisions, e.g. traffic lane arrows, were not available. The SAD analysis demonstrated that road users can use alternative strategies to make the same decision, see Figure 3. For example, road users can assess other road users' indicators, traffic lights or recall what the directional signs displayed to anticipate whether there will be a right hand turning lane. These strategies are however more complex and provide less certainty to road users about the lay out of the infrastructure and their expected behavior. Understanding such trade offs is essential in understanding the impact of intersection designs on road users.

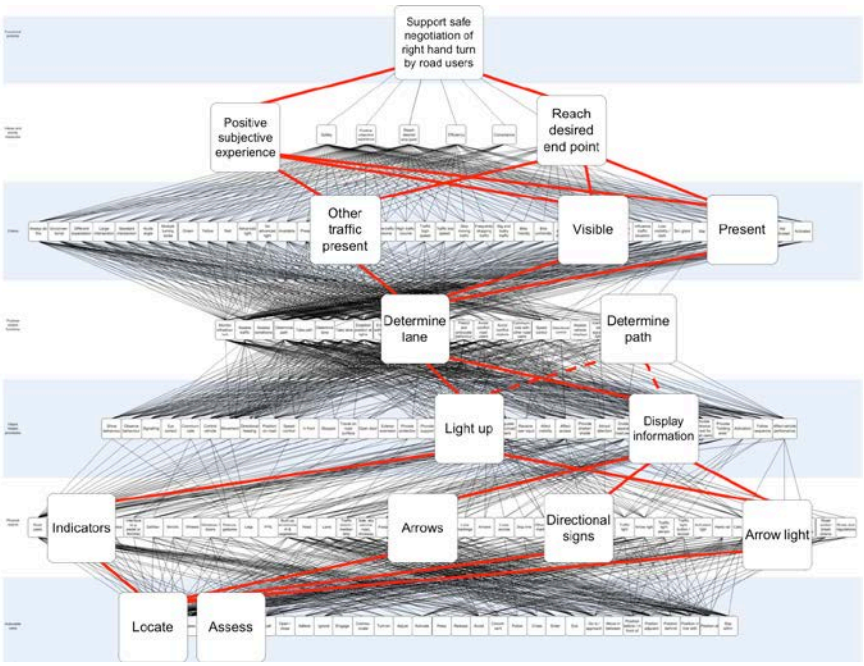


Figure 3. Extract Strategies Analysis Diagram

Field of safe travel

The synergy of the CWA outcomes and the field of safe travel theory will now be explored. Through the CWA analysis we have demonstrated the presence of non-linear emergence within the intersections examined; that is, how small changes in system constraints can have large consequences for the behavior possible (decision making processes and strategies). Therefore small changes in system constraints will have large consequences for the field of safe travel and possible paths within it. The SAD analysis provided a comprehensive insight into the behavior possible within the cut-through intersection. We therefore now understand paths possible within the field of safe travel. Further through the values and priority measures and criteria, the CWA analysis demonstrated the different paths that can be taken by different road users under different circumstances. This further specifies the factors of influence on the paths possible and paths that are likely to be chosen within the field of safe travel. These high level outcomes are depicted in Figure 4.

The field of safe travel can be used to visualize specific findings. For example, figure 5 communicates visually the CWA findings of emergent behavior; behavior that was not necessarily considered by the designers for evaluation. The designers intended that drivers, motorcycle riders and cyclists use the cut through lane to turn right. However, the CWA analysis demonstrated that road users have a range of other options, including using the intersection as a round about. The CWA analysis further demonstrated that cyclists and pedestrians can use the traffic islands in the middle of the intersection to cross diagonally. This behavior and the consequences of this behavior in interaction with other road users was not considered by the designers, and represents a form of emergent behavior that could potentially create conflicts between the different road user groups. Using the field of safe travel visualization quickly demonstrates how the paths these road users can possibly take is much larger than initially considered in the concept design phase. The field of safe travel then acts as a visual means to communicate the findings to the designers, making explicit design flaws

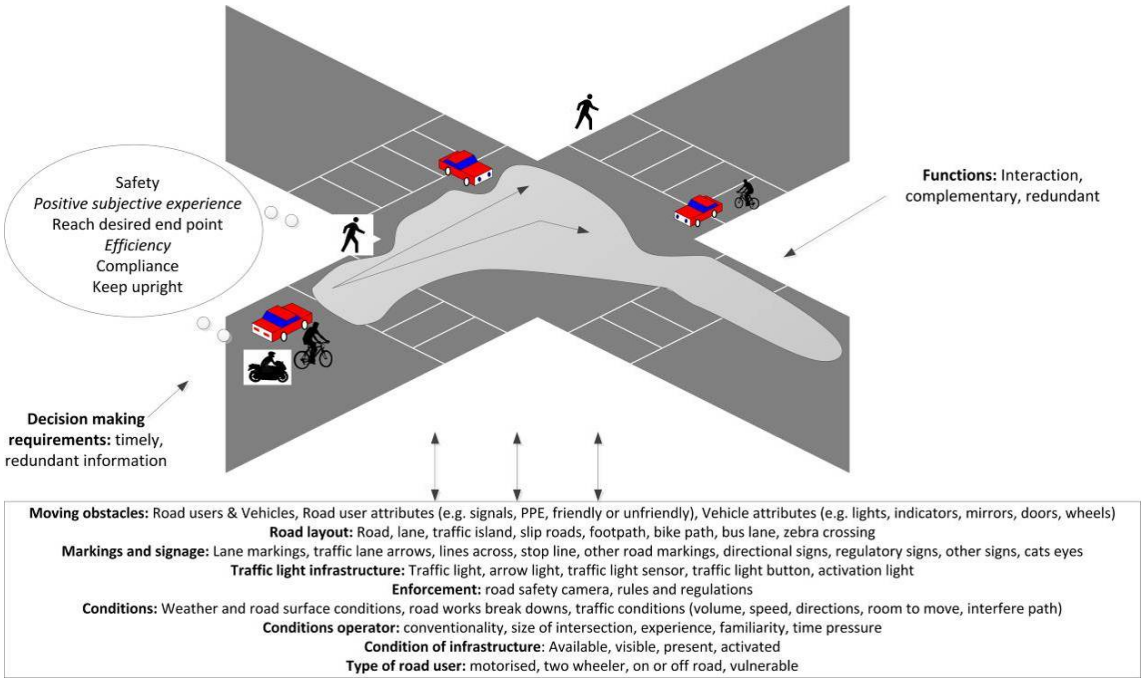


Figure 4. Insight factors shaping the field of safe travel.

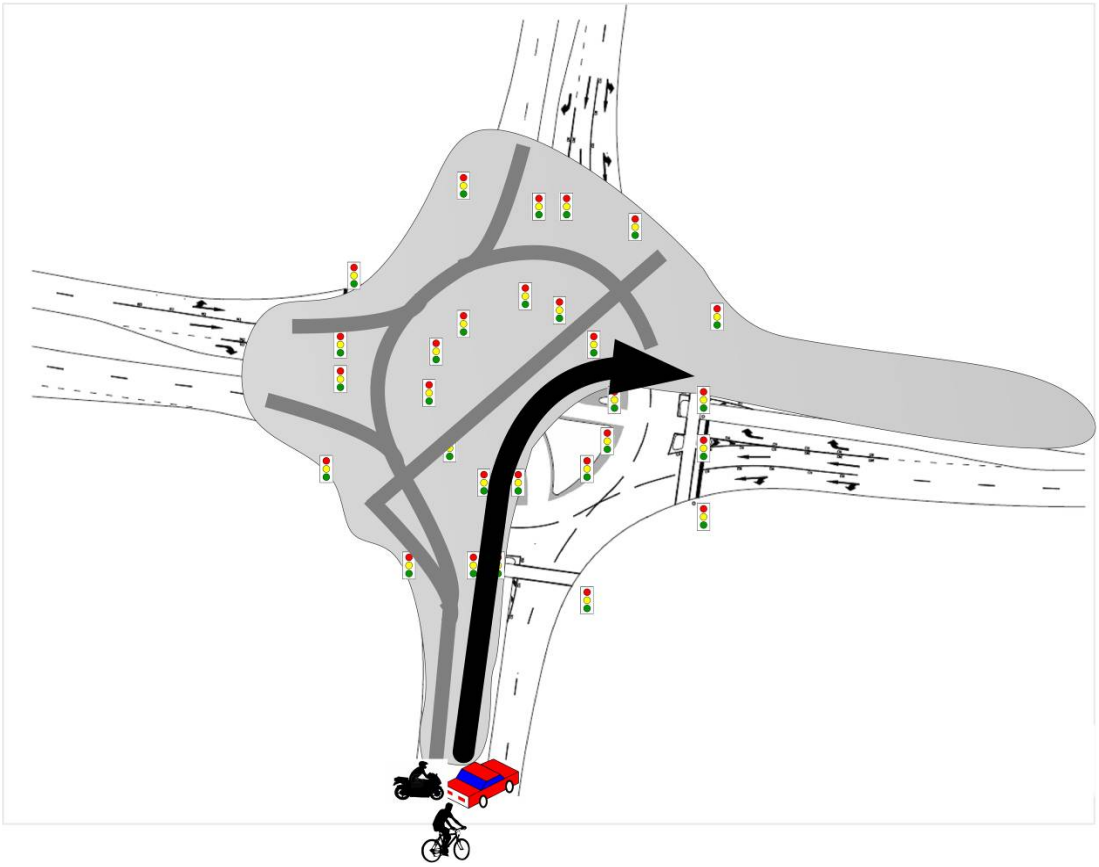


Figure 5. Field of safe travel Cut through design.

CONCLUSIONS

The aim of this paper was to explore the value of Ergonomics systems analysis methods in conjunction with the field of safe travel theory to provide easily interpretable analysis of road designs. Such synergy will facilitate communication between Ergonomics analysts and road transport designers. In particular, the combination of CWA and the field of safe travel theory were explored here.

The CWA analysis provided detailed insight into the system and situational constraints as well as decision making processes and possible road user behavior. The field of safe travel visually mediated the communication of these results by directly mapping it onto the intersection. The field of safe travel allows communication of the overall findings of constraints shaping the behavior as well as more specific findings, e.g. emergent behavior. The CWA analysis provided the analytical rigor to analyze the interaction between intersection design, road users, vehicles and the environment and the field of safe travel provides a visual platform to communicate these findings to a non-Ergonomics public, e.g. the designers.

This research represents a first step in better integrating Ergonomics methods in the early phases of the road design lifecycle. Following this successful exploration a full-scale application of the CWA analysis and the field of safe travel is encouraged. Future applications can also be used to further develop the field of safe travel theory.

REFERENCES

Annett, J. (2004). Hierarchical task analysis. In D. Diaper & N. A. Stanton (Eds.), *The handbook of task analysis for human-computer interaction* (pp. 67 - 82). Mahwah, NJ: Lawrence Erlbaum Associates.

Birrell, S.A., Young, M.S., Jenkins, D.P., & Stanton, N.A. (2011). Cognitive Work Analysis for safe and efficient driving. *Theoretical Issues in Ergonomics Science, 13*(4), 430-449.

Corben, B., Candappa, N., Van Nes, N., Logan, D., & Peiris, S. (2010). Intersection Safety Study: Meeting Victoria's Intersection Challenge. Task 5 Generation of Intersection Designs within the Safe System Context. Melbourne, Australia: Monash University Accident Research Centre.

Corben, B., Logan, D., Fanciulli, L., Farley, R., & Cameron, I. (2010). Strengthening road safety strategy development 'Towards Zero' 2008–2020 – Western Australia's experience scientific research on road safety management SWOV workshop 16 and 17 November 2009. *Safety Science, 48*(9), 1085-1097.

Cornelissen, M., Salmon, P.M., McClure, R., & Stanton, N.A. (2013). Using cognitive work analysis and the strategies analysis diagram to understand variability in road user behaviour at intersections. *Ergonomics, 56*(5), 764-780.

Cornelissen, M., Salmon, P.M., Stanton, N.A., & McClure, R. (2012). Human factors now showing in 3D. In N. A. Stanton (Ed.), *Advances in Human Aspects of Road and Rail Transportation* (pp. 653-662). Boca Raton, Fl.: CRC Press.

Cornelissen, Miranda, Salmon, Paul M., & Young, Kristie L. (2012). Same but different? Understanding road user behaviour at intersections using cognitive work analysis. *Theoretical Issues in Ergonomics Science, 14*(6), 592-615.

Gibson, James J., & Crooks, Laurence E. (1938). A Theoretical Field-Analysis of Automobile-Driving. *The American Journal of Psychology, 51*(3), 453-471.

Hilliard, Antony, & Jamieson, Greg A. (2008). Winning Solar Races with Interface Design. *Ergonomics in Design: The Quarterly of Human Factors Applications, 16*(2), 6-11.

Jenkins, D.P., Stanton, N.A., Walker, G.H., & Young, M.S. (2007). A new approach to designing lateral collision warning systems. *International Journal of Vehicle Design, 45*(3), 379-396.

Johansson, Roger. (2009). Vision Zero – Implementing a policy for traffic safety. *Safety Science, 47*(6), 826-831.

Koornstra, M.J. , Mathijssen, M.P.M. , Mulder, J.A.G. , Roszbach, R., & Wegman, F.C.M. (1992). *Naar een duurzaam veilig wegverkeer [towards sustainable safe road traffic]*. Leidschendam: SWOV.

Lee, J.D., Hoffman, J.D., Stoner, H.A., Seppelt, B.D., & Brown, M.D. (2006). Application of ecological interface design to driver support systems *Proceedings of the 16th World Congress on Ergonomics (IAE 2006)*. Maastricht, the Netherlands: Elsevier.

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety Science, 27*(2-3), 183-

213.
Seppelt, Bobbie D., & Lee, John D. (2007). Making adaptive cruise control (ACC) limits visible. *International journal of human-computer studies, 65*(3), 192-205.
Stanton, N.A., Salmon, P.M., Walker, G.H., Baber, C., & Jenkins, D.P. (2005). *Human factors methods: a practical guide for engineering*. Aldershot: Ashgate Publishing Limited.
Stoner, H.A., Wiese, E.E., & Lee, J.D. (2003). *Applying ecological interface design to the driving domain: the results of an abstraction hierarchy analysis.* Paper presented at the Human Factors and Ergnomics Society 47th Annual Meeting, Santa Monica, CA.
Vicente, K.J. (1999). *Cognitive work analysis: Toward safe, productive and healthy computer-based work*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
VicRoads. (2009). *Driving in Victoria: Rules and responsibilities*. Kew, Australia: VicRoads.
VicRoads. (2010). *Road to solo driving*. Kew, Australia: VicRoads.
VicRoads. (2011). Safety & Rules. Retrieved 17/02/2011, from http://www.vicroads.vic.gov.au/Home/SafetyAndRules/
Young, M.S., & Birrell, S.A. (2010). Ecological IVIS design: using EID to develop a novel in-vehicle information system. *Theoretical Issues in Ergonomics Science, 13*(2), 225-239.

Extracting Design Information from the Outputs of Systems Analysis: A Case Study in Public Transport Ticketing

Gemma J. M. Read[a], Paul M. Salmon[b], Michael G. Lenné[c] and Daniel P. Jenkins[d]

[a]Human Factors Group, Accident Research Centre
Monash Injury Research Institute, Monash University, Victoria, Australia

[b]University of the Sunshine Coast Accident Research
University of the Sunshine Coast, Maroochydore, Queensland, Australia

[c]Human Factors Group, Accident Research Centre
Monash Injury Research Institute, Monash University, Victoria, Australia

[d]DCA Design International
Warwick, UK

ABSTRACT

Analysis and design activities underpinned by the systems thinking are required for the design of modern complex sociotechnical systems. Cognitive work analysis is a commonly used analysis framework for understanding complex systems. Although the framework was developed with a view to then using analysis outputs to inform system design, questions about its direct application to design remain. This paper will describe an approach that was developed to support the translation of cognitive work analysis outputs into design-relevant information. The approach was applied to a case study analysis of a public transport ticketing system, with the insights documented during the analysis being used to prepare design materials for a participatory design workshop. The process of identifying design-relevant information from the analysis outputs and applying this to the preparation of design materials was found to be structured and efficient, providing the potential for traceability between analysis and design without constraining the creativity required for design innovation.

Keywords: System design; Cognitive work analysis; Systems analysis; Transport; Ticketing system design; Transport ticketing system

INTRODUCTION

The application of systems-based approaches to analysis and design is becoming more important in the modern age of increasingly technologically complex, distributed, high-risk domains (Dekker, 2011). Cognitive work analysis (CWA) is an analysis framework underpinned by systems thinking and open systems principles. While CWA has been used in many design applications (e.g. Jenkins, Salmon, Stanton, & Walker, 2010a; Naikar, Pearce, Drumm, & Sanderson, 2003), like all human factors analysis methods, the outputs of CWA provide information to support the design process rather than creating the design concept itself.

This paper will describe an approach to the translation of design-relevant information arising from the application of CWA to design. A case study focusing on the public transport ticketing domain will be used to demonstrate an aspect of this approach, the identification of insights and their application in the preparation of materials for participatory design workshops.

Cognitive Work Analysis

CWA is a framework of methods that supports the analysis of complex socio-technical systems with the aim of improving system design (Vicente, 1999). It is unique in its formative, constraint-based approach, in that it can be used to model the possibilities for behaviour, rather than describing actual behaviour (i.e. how work is done), or prescribing normative behaviour (i.e. how work should be done). The framework encompasses five phases of analysis that commence by modelling the environmental constraints within the domain (though work domain analysis), and then considers the tasks (through control task analysis), strategies (through strategies analysis), allocation of functions amongst human or technical actors (through social organisation and cooperation analysis) and the competencies required by actors operating within the domain (through worker competencies analysis) (Vicente, 1999).

CWA has been used to analyze many varied complex systems including nuclear power generation (e.g. Burns et al., 2008), military command and control (e.g. Jenkins, Stanton, Walker, Salmon, & Young, 2008), air traffic control (e.g. Ahlstrom, 2005), road transport (e.g. Cornelissen, Salmon, McClure, & Stanton, 2013) and rail transport (e.g. Stanton et al., 2013). It is an established analysis framework, with evidence showing that its application can improve system design (e.g. Naikar, et al., 2003). Yet despite the framework's increasing use, questions remain over its use as a design tool, and the details of how it is used in design applications (e.g. Jenkins, et al., 2010a; Lintern, 2005). Design guidance for CWA tends to focus on the development of displays and interfaces and there is a need to support the use of CWA outputs in design processes for all system elements, such as the design of teams, procedures, and training systems.

In response to the need for further support and guidance for using CWA in integrated systems design, the CWA design tool (CWA-DT) was developed. The tool aims to provide guidance for a design team wishing to use the outputs of CWA for design purposes. Acknowledging the close relationship between analysis and design, and the iterative nature of the process, the tool provides guidance for both analysis and design.

The CWA Design Tool

The CWA-DT provides guidance for various stages of a design process from planning the analysis process through to implementation. It consists of guidance as well as structured documentation to assist users to plan for and execute their design activities. Due to space constraints, this paper focuses the analysis phase and the design planning phase and omits discussion about the subsequent design activities.

Analysis stage

The CWA-DT provides some guidance to assist the analysis process. However, it does not intend to replace authoritative texts on CWA and refers readers to such guidance available for completing the analysis phases. The analysis phase is driven by an *Analysis Brief* (adapted from Liedtka & Ogilvie, 2010to suit CWA projects) developed prior to the analysis commencing. The main contribution of the CWA-DT in the analysis stage is the guidance and templates for recording insights derived by the analysts regarding the system under investigation. Insights in this context include both non-obvious inferences from the evidence provided in the analysis (Klein, 2011), as well as more obvious findings about the system that the analyst considers important for the design process.

The template for documentation of insights prompts the analyst to provide a description of the insight identified and how it was identified (i.e. which phase of CWA was being undertaken and the thought process surrounding the insight development). This enables an explicit linking between the analysis, the insight, and any proposed design subsequently stemming from the insight. When documented, insights are categorized into *Assumptions*, *Leverage point*s, *Metaphors*, *Scenario features* or *Design solutions*. A description of each is provided in Table 1.

Table 1: Types of insight documented in the CWA-DT

Insight type	Description
Assumption	An insight relating to the underlying assumptions upon which the system, or part of the system is based. Could include hypotheses and beliefs about how the system functions.
Leverage point	An aspect within a system which if changed in a small way, could produce big changes across the system (Meadows, 1999).
Metaphor	The comparison, interaction or substitution of two subjects on a symbolic level. To promote creativity through metaphor, the differences between the subjects should be significant but there should be some similarity (Morgan, 1980).
Scenario feature	A feature of a potential scenario such as a type of actor, attributes of an actor, a type of task, an environmental disturbance or influence, etc.
Design solution	A proposed design or feature of a design identified by the analyst/s.

Design planning stage

In addition to developing documentation to guide the design process, the design planning stage involves the development of materials for participatory design workshops. This enables the communication of insights usually only available to the analysts to those not directly involved and immersed in the analysis. This is important for participatory design processes where it is impracticable for stakeholders and users to be involved in the details of the analysis, but it is those details that make CWA such a unique and valuable analysis framework. The insights are not simply communicated to the participants, but are incorporated into participatory design activities involving, for example, scenarios, metaphorical design, and brainstorming.

ANALYSING THE PUBLIC TRANSPORT TICKETING SYSTEM

The public transport ticketing domain

Public transport systems are funded by public revenue (i.e. taxes) as well as contributions by the end user. The user contribution is collected through the sale of tickets, with various means employed to enforce the requirement to hold a valid ticket for travel. Internationally, the trend in ticketing systems has been a move towards durable smartcards that store value rather than disposable paper tickets in line with a general trend toward a cashless society.

An Australian state government has recently implemented a smartcard-based ticketing system. The smartcard system is used on multiple modes of public transport (e.g. train, bus, etc) with fares based on two geographically-based zones of travel. The durable plastic card is required to be purchased, and value (monetary amount) or a travel pass (a product that enables travel for a set period of time) to be purchased and loaded prior to travel. Purchasing and loading facilities are provided at many, but not all, locations at which passengers may begin their journey. Since it was introduced, the ticketing system has been subject to large-scale criticism from its users and the media regarding aspects such as the speed of processing, usability and convenience of ticket purchasing. As an example of a failure system design, it was considered that this example would provide an interesting case study.

Planning the analysis process

The analysis process was planned through the analysis brief developed at the beginning of the project. The fictional analysis brief noted that the current system was subject to criticism and described the need for the project as:

> *An election is held and a new government, not involved in any past decisions about the system, is elected to power. The new Minister for Transport has tasked this project team with reviewing the ticketing system and designing the next generation system for implementation in 5-10 years' time. The team is to identify and learn from the issues with the current system and design a new system that meets the goals of government and users.*

The analysis brief also described the deliverable of the project as *a proposed design solution for a ticketing system appropriate for a metropolitan public transport system in 5-10 years' time.* Other points from the analysis brief included that data should be collected regarding both regular and irregular users, and that the project is constrained to the design of a ticketing system, rather than alternative means of gaining funds for public transport operations.

Data sources

A work domain analysis was developed for the ticketing system drawing upon publicly available documentation and interviews with two subject-matter experts. The latter four phases of the CWA analysis were informed by interviews with six users of the ticketing system (three regular and three irregular users), using semi-structured interviews based on the critical decision method, as well as reference to publicly available documentation.

Work domain analysis

The work domain analysis phase of CWA describes the functional structure of the work domain (Vicente, 1999). The abstraction hierarchy (AH) is used in this phase of analysis. It describes the constraints of the work domain within which behaviour is possible. The AH encompasses five levels of abstraction, with means-ends links between nodes at adjacent levels. The representation identifies the physical resources available in the system, the processes afforded by those resources, the functions supported by the processes, the values and priorities that are measured and monitored within the system, and the overall purpose/s of the goal-directed work domain (Vicente, 1999).

During the development of the ticketing system AH, it became clear that there were discrepancies in how the work domain is viewed by the government stakeholders of the system and the users of the system. For example, while the primary purpose of the system from the government point of view was defined as *to collect revenue from public transport users*, for the users the primary purpose was *reach destination*. This suggests that while government may view the ticketing system as a sub-system of the wider public transport system, users experience ticketing in an integrated manner within the overall service being delivered. This discrepancy between different system stakeholders is dealt with in the work domain analysis of CWA by analyzing the object worlds, or perspectives, of different stakeholders (Naikar, 2013). This was achieved within the AH by delineating the parts of the work domain that belonged to only one stakeholder's object world with the remainder of the work domain being shared.

One of the insights derived during the development of the AH arose originally from reviewing the transcript of an interview with one of the subject matter experts. The interviewee had stated that '*(I) think they should have trusted the user a lot more than they did. People want to do the right thing and if people don't want to do the right thing they'll go out of their way to not do it*.' This statement resonated as a number of the users interviewed stressed that they wanted to comply with the rules around ticketing and make sure that they had a valid ticket for their journey. However, ensuring compliance with regulations is a key part of the government object view within the AH.

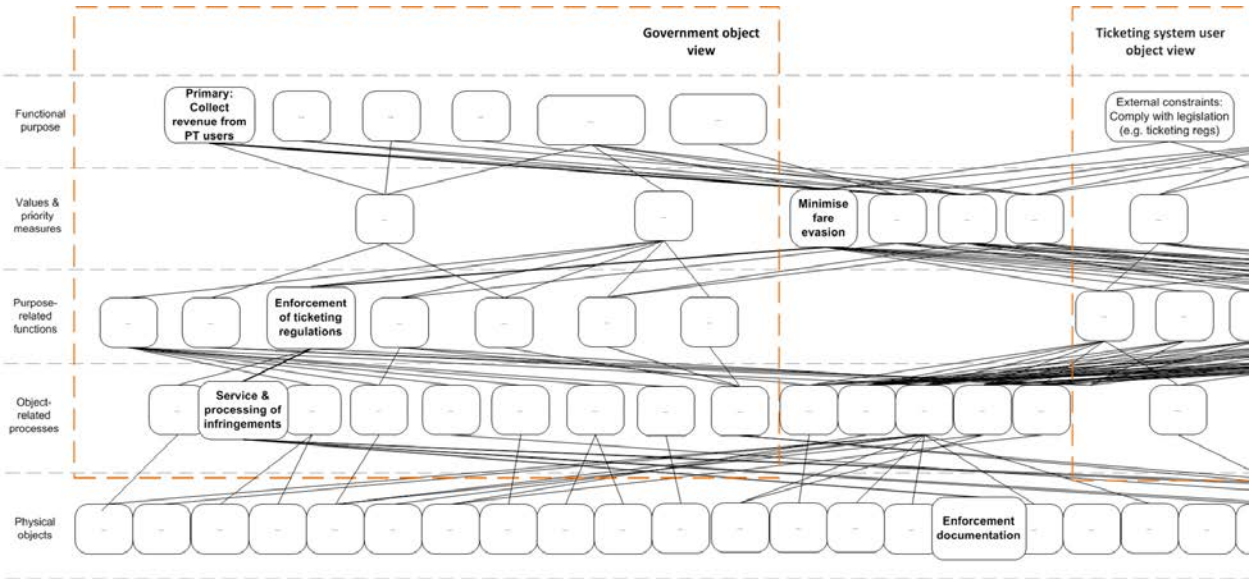


Figure 1. Extract of the AH showing stakeholder object worlds, focused on nodes associated with compliance and enforcement

Figure 1 provides a simplified extract of the AH; showing only those nodes relating to enforcement and compliance. It can be seen that the value of *Minimize fare evasion* was shared amongst both government and public transport users. Users also had a purpose to comply with legislation (e.g. ticketing regulations). Although users valued the need to avoid fare evasion, reviewing the nodes that are highlighted in bold, it can be seen that enforcement has a key role in the system. Reading from the bottom level up it can be seen that *Enforcement documentation* is a physical object that supports the process of the *Service and processing of infringements*. This process supports the function of *Enforcement of ticketing regulations*, which is related to *Minimizing fare evasion*, which in turn supports the government's primary purpose to *Collect revenue from public transport users*. This supported the initial insight that the system is designed on the assumption that users do not want to pay for transport and are not trusted to voluntarily comply with regulations. This assumption was documented in the insights template.

Control task analysis

The control task analysis phase of CWA identifies the activities and tasks that need to be carried out within the work domain. The tools used in this phase include the contextual activity template (CAT) which analyses activities in work domain terms, as well as the decision ladder which analyses activities in decision making terms (Naikar, Moylan, & Pearce, 2006).

Contextual activity template

The CAT developed for the ticketing system case study is shown in Figure 2. The template is a matrix with the situations encountered within the work domain represented by the columns and the functions occurring within the work domain shown in the rows. The circles with the attached bars in the matrix represent situations in which the functions typically occur. The dotted line boxes in the matrix indicate situations in which the functions could occur, although they may not typically occur there (Naikar, et al., 2006). Figure 2 identifies five situations that focus on the timeline of a public transport journey. The first situation is prior to travel, the next is when travel is beginning, the period during the journey is then analyzed, followed by the conclusion of the journey and then the period following travel. The functions represented in the first column of Figure 2 were transposed from the functions in the user object world of the AH. This was done because the focus of the analysis was on the user tasks and functions rather than those functions associated with customer service and administration.

The CAT shows that most functions occur either before or after the journey, with no functions typically occurring during the journey. This finding was documented as potential leverage point. This is because passengers generally have more time available during their journey for tasks such as validating their ticket, determining the cost of their fare, verifying the validity of their ticket, monitoring their account, registering their card, etc. Some of these activities could potentially be undertaken by passengers with portable electronic devices and internet access. However, there are no smartphone applications available to support this activity nor is wi-fi available on the public transport vehicles. Vitally, there is no equipment available on public transport vehicles to support users to undertake any of these functions.

A second insight that arose during the development of the CAT was that initial validation of the ticket cannot be undertaken prior to commencing the journey. However, airlines provide for online check-in prior to travel. This insight was documented as a potential metaphor for exploration in the design process.

Decision ladders

A decision ladder developed for the ticketing system case study is provided in Figure 3. Decision ladders outline the information processing activities (rectangular boxes) and resultant knowledge states (circles) that if followed from the bottom left to the bottom right represent the process of novice decision making (Vicente, 1999). As expertise develops, shortcuts can occur where for example, an alert is directly associated with a system state or the diagnosis of a system state leads to the selection and execution of a task without any intervening information processing steps.

For the case study, decision ladders were developed for the purpose-related functions in the AH (following e.g. Jenkins, Stanton, Salmon, Walker, & Rafferty, 2010b). However, this was done only for those purpose-related functions identified within the ticketing system user object world.

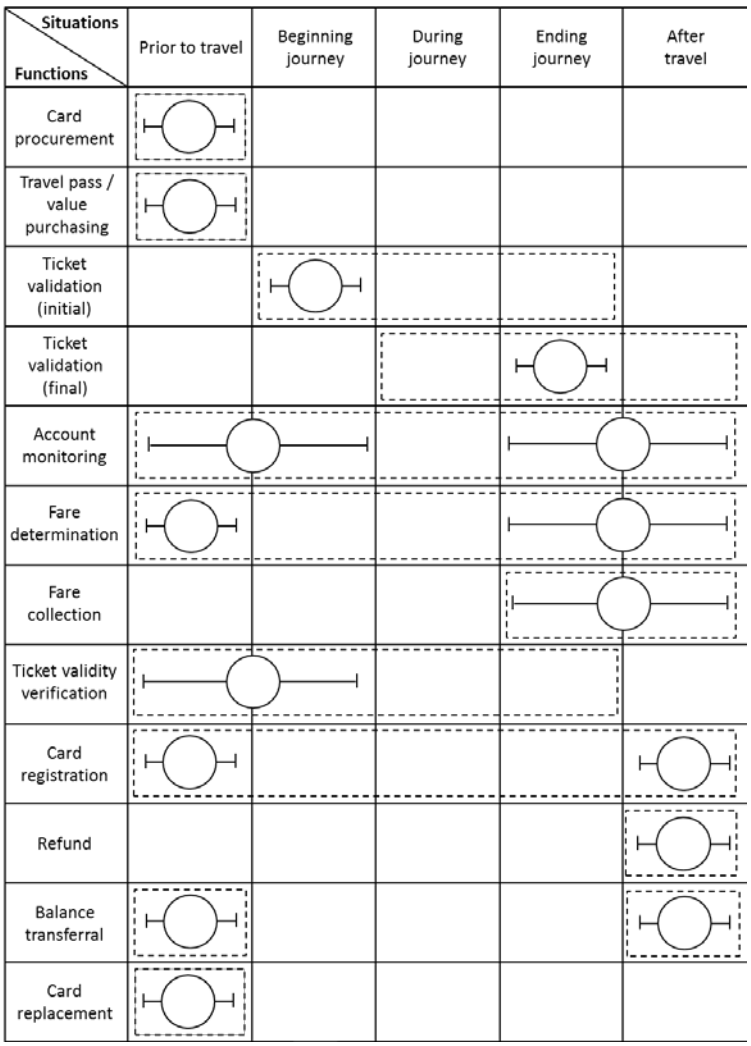


Figure 2. Contextual activity template

Figure 3 provides the decision ladder for the function of *card replacement*. The decision ladder shows the different ways that the system can alert the end user that a card requires replacement, the information then required by the end user to determine the system state (e.g. whether the card is defective, expired or lost / stolen), the options available to the end user, the possible goals the end user might have, etc. To assist in developing this decision ladder, publicly available user guidance and forms for card replacement were reviewed.

During the process of mapping this information on to the decision ladder it became clear that the guidance provided to the user was provided from an administrative perspective rather than a user perspective. For example, the same form is required to be filled out for replacement of damaged, defective, expired, and lost or stolen cards. However, the form has many caveats and exceptions depending on the reason for which the replacement is required, and depending upon whether the card was registered or unregistered. For example, the form asks for the unique identification number printed on the card. However, if the card is lost or stolen many users would likely not have kept a separate record of this information. During the process of trying to understand these documents for the purpose of populating the decision ladder, an insight was documented that the user guidance and forms are based around administrative processes rather than the problem being faced by the user. This insight was then framed as a design solution which was to amend the supporting materials to ensure they are user-focused. The documentation around this insight also noted that while this was potentially a low priority in terms of system change, it reflects the system administration (rather than user) perspective that has been taken in the design of the system.

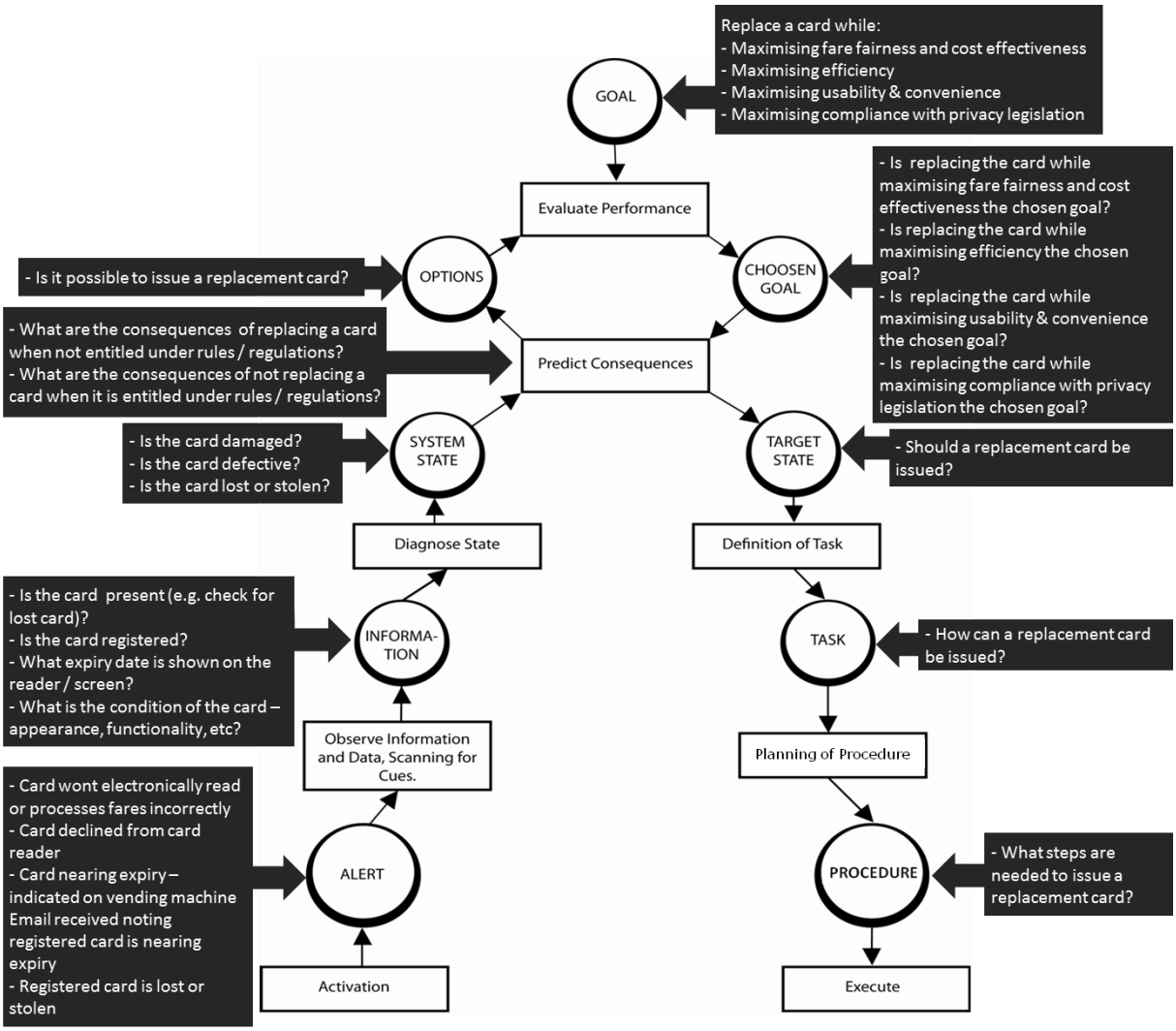


Figure 3. Decision ladder for the function *Replace card*

Strategies analysis

The strategies analysis phase identifies the range of strategies that can be employed to perform the activities and tasks in the system (Vicente, 1999). A common tool used for strategies analysis is the information flow chart. These flow charts identify the steps than can be used to move from a start state to an end state. Flow charts were developed for the key purpose-related functions identified in the user object view of the AH, however these are not presented in this paper due to space restrictions. Further, the flow charts tended to provide a very fine level of detail regarding task execution which was potentially more useful for a detailed design process following the selection of a system design concept.

In addition to information flow charts, a strategies analysis diagram (SAD, Cornelissen, Salmon, Jenkins, & Lenné, 2013) was developed for the ticketing system. This tool supports a more formative approach to identifying strategies. The SAD builds on the AH developed in the work domain analysis phase and involves the additional of two levels to the diagram: verbs and criteria. The verbs are used to specify how the physical objects can be used and the criteria specify the circumstances under which different strategies might be chosen. The SAD was completed for only parts of the AH that were relevant to the user domain. An extract of the SAD is provided in Figure 4. The extract focuses on the criteria identified for determining which strategy to use. For example, the purpose-related function of *Travel pass / value purchasing* may be undertaken when the user *Suspects value is low* to support the value of *Minimize fare evasion*. A total of 41 criteria were identified from the interviews with users and from document analysis. In documenting the criteria it became clear that they provided richness to the analysis that was not uncovered through the other phases of analysis. The criteria represented a number of circumstantial features that

were thought to provide important features for scenarios. For example, bad weather, a card having no or negative value, the presence or absence of ticket barriers, could all provide relevant scenario features.

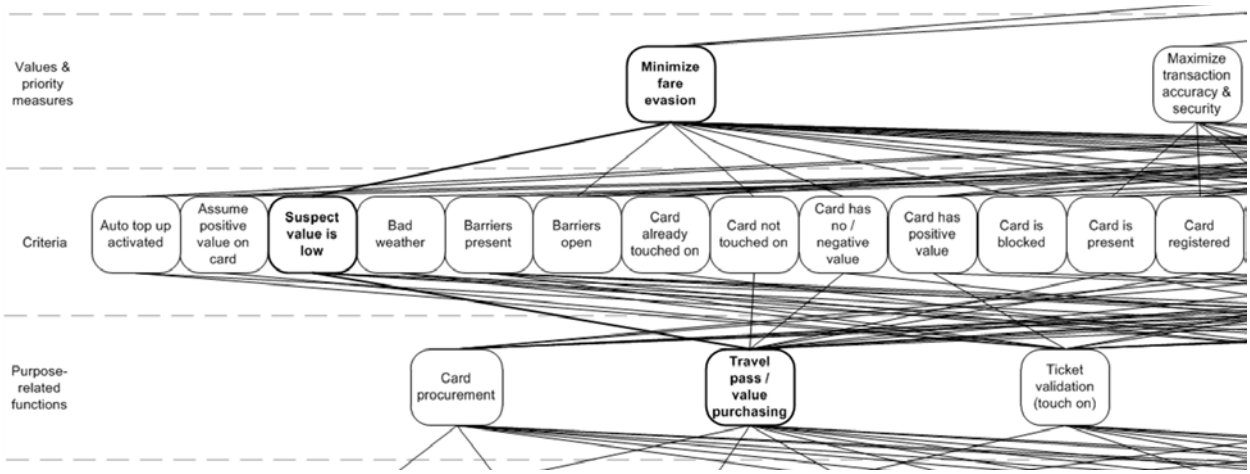


Figure 4. Extract of SAD showing some of the criteria identified within the analysis

Social organization and cooperation analysis

The social organization and cooperation analysis (SOCA) phase analyses how tasks and activities are distributed across agents within the system (Vicente, 1999). SOCA can be performed using any of the CWA tools already applied by overlaying the agents (human and technical) that could be involved in different aspects of the system's functioning. For the case study, SOCA was performed using the CAT. An extract of the SOCA-CAT is shown in Figure 5.

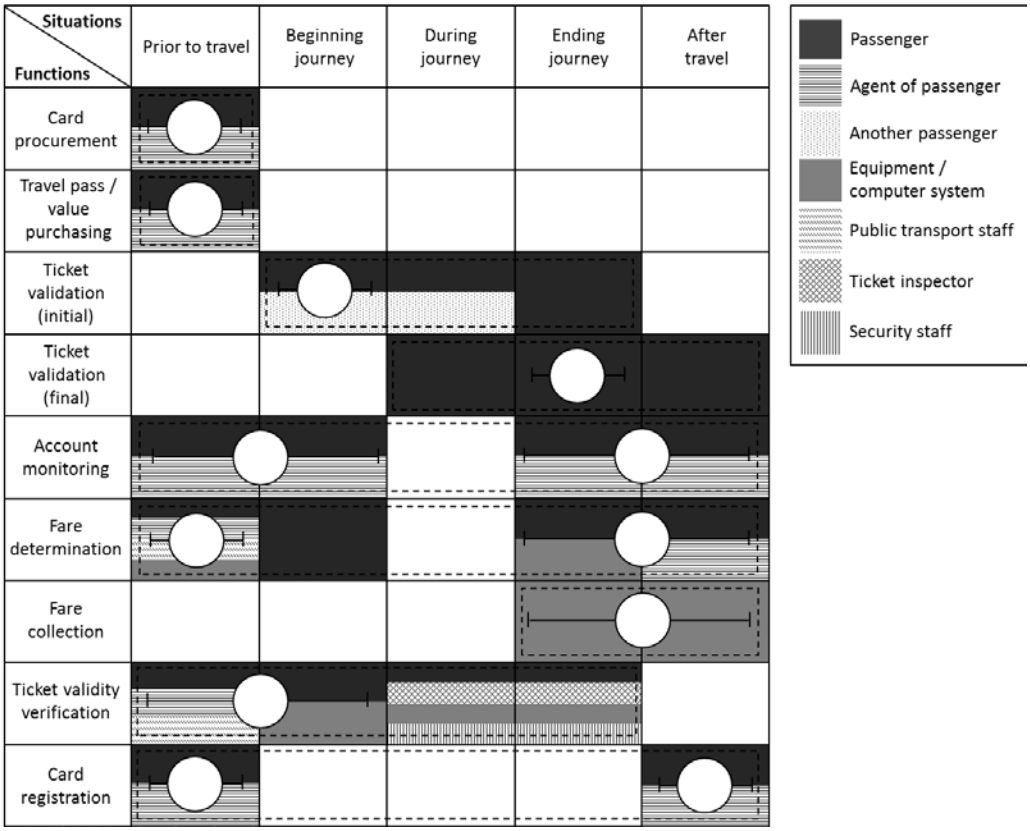


Figure 5. SOCA-CAT

Seven actors were identified within the SOCA phase: passengers, agents of passengers (e.g. parents or carers who perform functions for passengers but do not travel with them), other passengers travelling on public transport, equipment or the computer system (i.e. automation), public transport staff such as customer service personnel or drivers, ticket inspectors and security staff. The SOCA highlights that although the system is promoted as being more technologically advanced than previous paper ticket systems, most functions are still required to be performed by humans. Fare collection is the only automated function. Considering that the users interviewed stressed that their overall goal was to reach their destination, rather than to use a ticketing system, automating more functions around ticketing may be a beneficial design solution. An insight in the form of a design solution was recorded suggesting that functions such as ticket validation could be automated, for example, passengers could walk through a scanner which automatically detects the smartcard and registers it as beginning or completing a journey.

Worker competencies analysis

The final phase of CWA, worker competencies analysis, identifies the cognitive skills and processes required to perform tasks within the system (Vicente, 1999). The skill, rule and knowledge (SRK) taxonomy is generally applied in this phase. The taxonomy relates to levels of performance (Rasmussen, 1983): skill-based behaviour (SBB; sensory-motor performance which occurs without conscious control), rule-based behaviour (RBB; the application of stored rules in familiar situations and occurs on a more conscious level) and knowledge-based behaviour (KBB; explicit and conscious reasoning required for unfamiliar situations). The SRK inventory proposed by Kilgore and colleagues (2008) builds upon the decision ladder. The inventory was developed to support interface design activities, and has subsequently been utilized, and extended, by other authors (e.g. McIlroy & Stanton, 2011).

In this case study, the aim is to support integrated systems design, rather than the design of interfaces or information systems alone. In this case, the analysts determined that the SRK inventory was best completed in relation to the strategies identified within the strategies analysis phase. Strategies were drawn from the SAD by reviewing the relationships between the verbs and physical objects and transposing the strategies into a table identifying the SBB, RBB, and KBB required to execute each strategy. The inventory was also extended to identify the errors that could occur at each level of performance, the consequences of these errors and a judgment, based on the consequences, about the criticality of that behaviour. Where the behaviour is critical, consideration is given to whether the behaviour should be supported or restricted. An evaluation is then conducted regarding how the system currently supports and restricts this behaviour. This evaluation assists to uncover design problems and insights for re-design.

Figure 6 shows an extract of the SRK inventory developed for the case study in relation to the strategy *Query customer service desk / booth*. The focus is on the RBB which identified that when users experience problems associated with ticketing irregularities (for example, they believed they validated their card but subsequently it is not accepted by the equipment) they may apply use rules based on past experience to guide their interaction with customer service representatives to resolve the issue. Errors could be associated with misjudging the likelihood of being assisted, or the level of assistance provided, or with the likelihood that an excuse is accepted by different customer service staff members (e.g. expect consistency when it may not be present). The table identifies that this should be supported, but upon reviewing the outputs of the previous analysis phases no evidence was found that this is currently supported. It was found that the behaviour was restricted in two ways: with staff not always being available to query, and with the policies or guidance applied by staff in relation to dealing with ticketing inconsistencies not provided to users. The user guidance is written on the basis that the system operates as designed and does not, for example, provide examples of what a user should do if equipment is defective and will not accept an attempt to validate their smartcard at the beginning of a journey. Based on this insight, an assumption was documented that giving users too much information will enable them to abuse the regulations.

SRK behaviour	What errors could occur?	What are the potential consequences of error?	Is the behaviour critical?	Should the behaviour be supported or restricted?	How does system currently support?	How does system currently restrict?
SBB: - Ability to communicate with customer service staff, e.g. speak English, be able to hear response, etc.	...	...	...	...	...	...
RBB: - Use past experience (e.g. quicker if just ask customer service, have been let through barriers previously with certain excuse, etc.)	- Misjudge likelihood of being assisted by customer service - Excuse regarding non-valid ticket may not be accepted by all staff	- Delays in resolving the situation – hold up other passengers, miss service, frustration, etc. - Potential to be fined for fare evasion	Yes	Supported	n/a	- Staff not available at all stations / vehicles - Policies applied by staff regarding what is acceptable / not acceptable in terms of ticketing irregularities not communicated to passengers
KBB:	...	...	...	...	...	...

Figure 6. Extract of the SRK taxonomy table for the strategy *Query customer service desk / booth*

USING THE INSIGHTS IN DESIGN

The above discussion has described the background to seven insights derived from analyzing the system using CWA. These examples provide some understanding into the different types of insights documented and how they arose. It is worth noting that while some insights resulted from the final *content* of the analysis (such as reviewing the CAT and noticing that no functions are typically performed during the journey), other insights stemmed from the *process* of using data to develop the outputs (such as the insight that the user guidance documents were difficult to use to develop the decision ladder due to being based around the system rather than user needs). In total, 34 insights were documented while the analysis was completed; 26 were identified by the primary analyst and the remainder by a CWA expert who reviewed the analysis outputs.

Documented insights need to be transferred to inform the design process. The CWA-DT is predicated on participatory design processes and as such the insights were used to inform the development of activities within participatory design workshop. Table 2 summarizes the example insights and how they were used to develop materials for a design workshop. The activities included:

Assumption crushing: The process of presenting assumptions associated with the system, then challenging these assumptions and creating opposite statements (Imber, 2012). This enables the design process to transcend the traditional design solutions and promote innovation. The crushed assumptions can be used to brainstorm design ideas, e.g. through recording them on individual cards which are sorted and combined by workshop participants to create new ideas and design solutions. The materials developed were a list of assumptions uncovered during the analysis, and blank cards for recording the crushed assumptions.

Metaphor cards: This activity involved the presentation of cards to prompt metaphorical thinking, drawing from the potential metaphors identified in the analysis. Some metaphor cards presented symbols (such as the symbol of an airplane) while others presented brand logos of well-known companies (such as Virgin Blue airlines) to inspire thinking about the potential for application of initiatives from other domains as well as from various brand perspectives.

Pain point cards: Pain points (e.g. Clatworthy, 2011) are aspects of the system that cause frustration for users or stand in the way of users achieving their goals. For the design ideation phase, the leverage points and design solutions were re-phrased into pain points. This was done to avoid leading participants to certain solutions or constraining their thinking to certain types of solutions. The original insights are not lost, but are introduced later in the evaluation stage of the design process.

Scenarios: Scenarios are narratives of use situations relevant to the design process. Scenarios enable information discovered in the analysis to be communicated efficiently and effectively to design participants (Carroll, 2002).

They can promote understanding of user goals, experiences and challenges to prompt design solution ideas and to evaluate proposed designs by considering how the scenario might be different if that design was implemented. The analysis can provide many inputs to scenarios, including user goals, the presence of conflicting goals, different types of situations and circumstances, different strategies that can be employed and the different competencies of users.

Table 2: Summary of selected insights documented from the case study analysis

Insight	Analysis phase & tool	Insight type	Design process planning
The system is designed on the assumption that users do not want to pay for transport and are not trusted to voluntarily comply with regulations	Work domain analysis – AH	Assumption	Added to list of assumptions for assumption crushing exercise
Most functions occur either before or after the journey, with no functions typically occurring during the journey. During the journey users potentially have more time to undertake functions	Control task analysis – CAT	Leverage point	Pain point card noting that the lack of support for completing tasks during the journey
Airlines provide facilities for passengers to check in prior to arrival at the airport. Potentially public transport could provide an option for early validation of a ticket.	Control task analysis – CAT	Metaphor	Metaphor cards with symbol of an airplane and logo of Virgin Blue airlines.
User guidance and forms are based around administrative processes rather than the problem being faced by the user. A design solution could be to amend the supporting materials to ensure they are user-focused.	Control task analysis – Decision ladder	Design solution	Pain point card noting that guidance and forms are based around administrative processes
The criteria identified within the SAD can provide scenario features and circumstances (for example, bad weather, a card having low or negative value)	Strategies analysis – SAD	Scenario feature	Incorporated within scenarios
Functions currently completed by the user – e.g. validating the ticket – can be automated	Social organization and cooperation analysis – CAT	Design solution	Pain point card noting that most tasks (such as validating the ticket) much be completed by the user
The system is designed on the assumption that giving users too much information will enable them to abuse the rules.	Worker competencies analysis – SRK inventory	Assumption	Added to list of assumptions for assumption crushing exercise

CONCLUSIONS

The purpose of this paper was to demonstrate how design-relevant information from CWA can be applied to the preparation of materials for participatory design workshops. The aim of the CWA-DT is to promote the application of CWA findings in design, particularly in participatory design processes. The analysis terminology and representations can be difficult to communicate and share with users and system stakeholders which can affect their ability to meaningfully collaborate in the process. The challenge is to translate the analysis findings without compromising their utility and their validity through over-simplification.

The process of identifying insights while conducting the analysis was very efficient, taking little additional resources to complete in addition to the analysis process itself. The template provides structure for this process which promotes ease of use and efficiency. The documentation of analysis insights and how they were exposed can provide traceability between the analysis and the design activities, without constraining the creativity required for design innovation. It should however be noted that CWA-DT is not limited to the design activities described in this paper and that appropriate activities should be selected based on the needs and resources of the project.

Further development work on the CWA-DT will focus on the development of a set of questions or criteria that can be applied following completion of the analysis to trigger further insights that may have been missed during the analysis process itself. Further, additional insight categories may be added over time as the tool is applied to different domains. Additional design activities will also be added such as those that focus on positive or beneficial aspects of the system that could be enhanced. This would ensure that the tool builds on the positive aspects of current functioning and opportunities discovered through formative analysis.

ACKNOWLEDGEMENTS

Gemma Read's contribution to this article was funded through an Australian Postgraduate Award (Industry) provided by an Australian Research Council Linkage Grant (ARC, LP100200387) to Monash University in partnership with the University of Sunshine Coast and the University of Southampton, and the following partner organizations: the Victorian Rail Track Corporation, Transport Safety Victoria, Public Transport Victoria, Transport Accident Commission, Roads Corporation (VicRoads) and V/Line Passenger Pty Ltd. Dr Paul Salmon's contribution to this article was funded through an Australian National Health and Medical Research Council post doctoral training fellowship.

REFERENCES

Ahlstrom, U. (2005). Work domain analysis for air traffic controller weather displays. *Journal of Safety Research, 36*, 159-169.

Burns, C. M., Skraaning, G., Jamieson, G. A., Lau, N., Kwok, J., Welch, R., et al. (2008). Evaluation of Ecological Interface Design for Nuclear Process Control: Situation Awareness Effects. *Human Factors, 50*, 663–679.

Carroll, J. M. (2002). Scenario-Based Design. In W. Karwowski (Ed.), *International Encyclopedia of Ergonomics and Human Factors* (Second ed.). Boca Raton, FL: CRC Press.

Clatworthy, S. (2011). Service innovation through touch-points: Development of an innovation toolkit for the first stages of new service development. *International Journal of Design, 5*, 15-28.

Cornelissen, M., Salmon, P. M., Jenkins, D. P., & Lenné, M. G. (2013). A structured approach to the strategies analysis phase of cognitive work analysis. *Theoretical Issues in Ergonomics Science, 14*, 546-564.

Cornelissen, M., Salmon, P. M., McClure, R., & Stanton, N. A. (2013). Using cognitive work analysis and the strategies analysis diagram to understand variability in road user behaviour at intersections. *Ergonomics, 56*, 764-780.

Dekker, S. (2011). *Drift into failure: From hunting broken components to understanding complex systems*. Surrey: Ashgate.

Imber, A. (2012). Five ways to boost creativity. Melbourne.

Jenkins, D. P., Salmon, P. M., Stanton, N. A., & Walker, G. H. (2010a). A new approach for designing cognitive artefacts to support disaster management. *Ergonomics, 53*, 617-635.

Jenkins, D. P., Stanton, N. A., Salmon, P. M., Walker, G. H., & Rafferty, L. (2010b). Using the decision-ladder to add a formative element to naturalistic decision-making research. *International Journal of Human-Computer Interaction 26*, 132-146.

Jenkins, D. P., Stanton, N. A., Walker, G. H., Salmon, P. M., & Young, M. S. (2008). Applying cognitive work analysis to the design of rapidly reconfigurable interfaces in complex networks. *Theoretical Issues in Ergonomics Science, 9*, 273-295.

Kilgore, R., St-Cyr, O., & Jamieson, G. A. (2008). From Work Domains to Worker Competencies: A Five-Phase CWA. In A. M. Bisantz & C. M. Burns (Eds.), *Applications of Cognitive Work Analysis* (pp. 15-48): CRC Press.

Klein, G. (2011). A naturalistic study of insight. *Journal of Cognitive Engineering and Decision Making, 5*, 335–351.

Liedtka, J., & Ogilvie, T. (2010). *Designing for Growth: A Design Thinking Tool Kit for Managers*. New York: Columbia Business School Publishing.

Lintern, G. (2005). *Integration of cognitive requirements into system design*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting, Orlando.

McIlroy, R. C., & Stanton, N. A. (2011). Getting past first base: Going all the way with Cognitive Work Analysis. *Applied Ergonomics, 42*, 358-370.

Meadows, D. H. (1999). Leverage points: Places to intervene in a system. Hartland, VT: The Sustainability Institute.

Morgan, G. (1980). Paradigms, metaphors and puzzle solving in organization theory. *Administrative Science Quarterly, 25*, 605-622.

Naikar, N. (2013). *Work Domain Analysis: Concepts, Guidelines and Cases*. Boca Raton, FL: Taylor and Francis Group.

Naikar, N., Moylan, A., & Pearce, B. (2006). Analysing activity in complex systems with cognitive work analysis: concepts, guidelines and case study for control task analysis. *Theoretical Issues in Ergonomics Science, 7*, 371-394.

Naikar, N., Pearce, B., Drumm, D., & Sanderson, P. M. (2003). Designing teams for first-of-a-kind, complex systems using the initial phases of cognitive work analysis: Case study. *Human Factors, 45*, 202-217.

Rasmussen, J. (1983). Skills, rules and knowledge - signals, signs and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics, 13*, 257 - 266.

Stanton, N. A., McIlroy, R. C., Harvey, C., Blainey, S., Hickford, A., Preston, J. M., et al. (2013). Following the cognitive work analysis train of thought: exploring the constraints of modal shift to rail transport. *Ergonomics, 56*, 522-540.

Vicente, K. J. (1999). *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. Mahwah, NJ: Lawrence Erlbaum Associates.

Section 17

Road and Rail:

Electric Vehicles: New Challenges for Human Machine Interaction

Social Media Monitoring as a Tool for the Development of Disruptive Innovations

Andreas Keinath and Roman Vilimek

Concept Quality and Usability
BMW Group
Munich, D-80788, Germany

ABSTRACT

Understanding how electric vehicles (EVs) are driven in the real world has massively advanced with the implementation of the MINI E and BMW ActiveE field trials. The results of these studies served as a key learning project for the development of the BMW i3, the first purpose-designed EV produced by the BMW Group. Especially in the early phases of the fields trial in 2009, the design of user research and analysis methods posed a challenge as it was unclear how users would react to an EV. As with all disruptive innovations, the collection of immediate and unbiased feedback as well as long-term feedback was required to separate first-contact phenomena from permanent effects. Our research partners from academia therefore established a research schedule with repeated face-to-face interviews and diaries. In order to make sure that the questions asked also sufficiently reflect the customers' everyday life with the EV, social media monitoring was established as a means to participate in usual driving and charging experiences and to support the development of the methods tool set. This approach proved to be very fruitful as early tendencies in customer satisfaction and dissatisfaction were identified and later on systematically reviewed with quantitative methods. Several examples of relevant findings in the MINI E field trial are presented and social media monitoring is discussed as a tool for customer feedback in the development chain. Presented as a case study for the development of electric vehicles, the data basis for this discussion are 2242 Facebook and relevant blog comments of US MINI E users between April 2009 and October 2011.

Keywords: Electric Vehicle, Electromobility, MINI E, Field Study, Social Media

EARLY STEPS IN SERIES-PRODUCED ELECTRIC VEHICLES

With the introduction of several new electric vehicles (EVs) in the last years, the different underlying approaches in development become apparent. Most vehicles are based on existing product solutions by building conversion EVs. The obvious advantage is that well-established production processes can be applied to a large extent, adapting only for EV-specific components. Correspondingly, it allows for rather quick development progress. Purpose-designed EVs however, which are planned and developed for driving electric since product ideation, offer several advantages. Not only in terms of vehicle architecture with a secure and advantageous arrangement of the battery in order to give the vehicle a low center of gravity and optimum weight distribution, but also in terms of innovative use of materials, lightweight design and driving agility. Purpose design also enables to target two especially important aspects in the context of electromobility: a holistic approach on sustainability along the entire value added chain and tailoring the vehicle and associated services to fit the customers' needs.

Electromobility can truly be seen as a disruptive innovation. Not only the driving task in itself changes with a completely silent engine, limited range and a different acceleration behavior of the vehicle. The driving environment

also changes, with the need for charging stations and online information about charging locations. Designing for purpose in combination with user-centered design poses a particular challenge in that environment: Obviously, several years ago there was no broad prior customer experience any manufacturer could built on to develop electric vehicles. In order to explore new solutions for sustainable individual mobility, a BMW Group think tank was established in 2007 called *project i*. Thinking outside of the box, *project i* analyzed future developments and driving factors for mobility. It was borne out of the necessity to provide answers for changes in mobility needs associated with new opportunities like electromobility, which led to newly developed offers such as 'DriveNow' car sharing. One of the earliest steps of *project i* was the implementation of EV field trials in cooperation with academic, public and private partners providing a huge number of pilot customers for a substantial period of time with vehicles running purely on electric power. Starting in 2009, more than 600 MINI E – a conversion of the familiar MINI hatch specifically developed for field trials – were on the road in the USA, Europe and Asia to gather important feedback on customer experience and customer requirements. In 2011, a fleet of more than 1000 BMW ActiveE – which is based on a 1 Series Coupé and like the MINI E a conversion vehicle for pioneer customers – was launched in several countries for research on EV-related technology, charging solutions, service and sales processes. The BMW ActiveE can also be regarded as the first iteration loop on the way to the BMW i3. The BMW i3 is the BMW Group's first series-produced all-electric car. It is purpose-designed based on customer feedback of both field trials and more than 34 million test kilometers of experience achieved during these studies.

These MINI E and BMW ActiveE projects are unparalleled worldwide in their scope. A conventional market research approach would not have been possible in the early phases of EV development. Hardly any customers with EV experience were available, target customer groups unknown, the interdependencies between charging infrastructure needs and using an EV on an everyday basis were largely unclear and end-customer EV driving patterns have not even been touched by in-depth analysis on a valid large-scale basis. Therefore, the goal of the field trials was to gain knowledge on this new terrain in the context of disruptive innovations by widening the research focus with a large partnering network. Experts from universities and research institutions took care of the scientific monitoring of the field trial. Cooperation partners from infrastructure and energy provided the necessary means for regular charging, at home and in public. A close cooperation with governmental institutions involving all project partners helped to inform decision makers and relevant stakeholders. Several overview articles provide further details on the field trials (Vilimek & Keinath, 2014; Ramsbrock, Vilimek & Weber, 2013; Vilimek, Keinath & Schwalm, 2012).

To depict the role of social media monitoring during the field trials and the corresponding influence on the development of the BMW i3, the focus will reside on the MINI E field trials, especially on the early phases. The field trials started in 2009 in Germany, Berlin and in the United States, west and east coast. Customers applied online for participation in the field trial. They needed to fulfill certain selection criteria such as being willing to use the EV on regular basis and agree to pay a monthly leasing fee. Based on the application data, a sociodemographic and psychographic profile of potential customers seriously interested in becoming EV users was assembled. Customers selected from that pool, remarkably more than 15.000 applicants at the end of the international MINI E field trials, were chosen to represent that early adopter profile. The Berlin study can be seen as the methodological blueprint of the field trials in general. While all academic partners contributed their specific research background to the projects and in doing so made each single study partly unique, the methods tool set developed in cooperation between the Institute of Cognitive and Engineering Psychology at Chemnitz University of Technology and the Concept Quality and Usability department at BMW Group installed a common ground between all field trials. Bringing in their strong expertise in experimental designs, human-machine interaction and user research for in-vehicle systems, scientists from Chemnitz University of Technology proposed a very systematic research plan with repeated user interviews before, during and at the end of a usage period as well as additional instruments like charging and travel diaries. Whereas the basic set of questionnaire items remained constant over time, an initial review of blogs, social media posts and comments and newsgroup discussions conducted by the BMW Group's Concept Quality and Usability department shortly after the start of the field trial showed that many new topics arised directly from customer perspective that deserved deeper and systematic analysis. Weekly summaries on discussion topics were set up to provide input for research planning on a regular basis. This unfiltered feedback proved to be a very early and quick indicator on which characteristics of electric driving appeal to the customers and what they think must be improved in the future.

The largest customer group in the same country with 246 private customers and about 200 fleet vehicles resided in the United States. These customers were also most active in the web. Social media monitoring therefore concentrated on this part of the field trial. Scientific research activities in the US MINI E field trial were led by the Institute of Transportation Studies at the University of California at Davis. Their research with a strong qualitative component influenced by prior work of Axsen and Kurani (2012) on the effect of drivers' social networks in shaping

their understanding of plug-in hybrids was ideal to interpret many discussions among the customers during the phase of getting to know their new vehicle.

SOCIAL MEDIA EV CASE STUDY

Methods

While interviews, diaries or focus groups conducted during the MINI E field trial are classic active data collection methods, social media monitoring falls within the range of passive data collection (Sampson, 1996). Using direct unfiltered qualitative data from social media, newsgroups or online communities is an appealing idea since the early days of customer exchange via the web, it was critically reviewed and refined (cf. for instance Kozinets, 1999, 2002; Finch, 1999; Edwards, Housley, Williams, Sloan & Williams, 2013). Reading social media discussions can be a very helpful way to listen directly to the voice of the customers. But certain precautions must be taken to ensure valid use of this piece of information (Sampson, 1996; Klein & Spiegel, 2013). As the MINI E discussion groups were not restricted to MINI E users only, non-EV drivers could have been active contributors. Among the MINI E drivers, only a subset was part of the online community. Therefore, self-selection and a nonresponse bias must be kept in mind as confounding factors. The opinion of those active members of the community may not be misjudged as synonymous with the view of the target customer group. Finally, product-related discussions might not represent a balanced view of the customers' perception as they are often prompted to seek help when something does not work, while praises for positive product characteristics may only enter public debate in the case of unique outstanding features. Finding out about those extremely appealing features has of course a high value, but features that are simply regarded as well designed and appreciated in normal everyday use may go unnoticed – leaving a biased overall picture. Taking this into account, we used feedback from online communities and blogs besides methods development for two purposes.

1. Innovation. The aim was to get an impression about which ideas the customers exchange. When it comes to actively shaping features or functions during early phases of product development, even single opinions, impulse ideas or innovative wishes of customers can be very helpful from a practitioner's point of view (Füller, Bartl, Ernst & Mühlbacher, 2006). Although it is possible to involve online communities in terms of co-creation in innovation processes (cf. also Füller et al., 2006), we opted for a completely passive role to maintain neutrality. This was especially important for the to-be-performed interviews during the field trials.

2. Product optimization. Any discussion led with emphasis in public debates yields a very good indication that the respective topic seems to bother at least some customers. We concentrated on a small number of sources with high likelihood of being lead by real MINI E drivers. Several relevant topics discussed were translated into our systematic research instruments used for as many MINI E users as possible in order to get a representative vote from that early adopter sample. By doing so, we made sure that the voice of the customer expressed online is validated and finds its way into systematic scientific items in the questionnaires.

Starting with an open query via search engines and on Facebook it became clear that most MINI E drivers were active in an Facebook group called "MINI E". The group was established by a private MINI E driver. Soon after its establishment the group had a high number of members, staying at about 500 members during most of the time. This open group, i.e. requiring no further authentication to contribute than a Facebook login, was the most active community. Some MINI E drivers started blogs and a small number of other groups. All MINI E related sources were screened during the core field trial time between April 2009 and October 2011. In total, 2242 posts in 222 discussion threads were analyzed. Unlike in most social media analyses all data was processed manually. This approach was required because of the exploratory nature of the task without any prior experience or relevant categories for clustering and analysis. Furthermore, as Bartl and Ivanovic (2010) argue, language processing software is currently not capable to isolate all relevant statements or recognize, for instance, ironic comments. Finally, the risk of overlooking a salient single comment is extremely high in automatic analysis.

Results

All posts were read in detail and categorized manually. Twelve different categories emerged and were dichotomized according to the emotional valence of the post into "likes" and "dislikes" (see Figure 1). As discussed above, it would make absolutely no sense to treat the absolute number of positive or negative posts as results in itself. A

helpful approach starts with asking about the reason for a high number of posts and a qualitative analysis to the core of the discussion. Analyzed with the additional background of further analysis of MINI E US customer feedback (cf. Turrentine, Garas, Lentz & Woodjack, 2011) and MINI E worldwide results (cf. Vilimek, Keinath & Schwalm, 2013) several quite interesting highlights can be found in the likes/dislikes data. In the following, a small subset of these results is summarized to depict the role of social media monitoring.

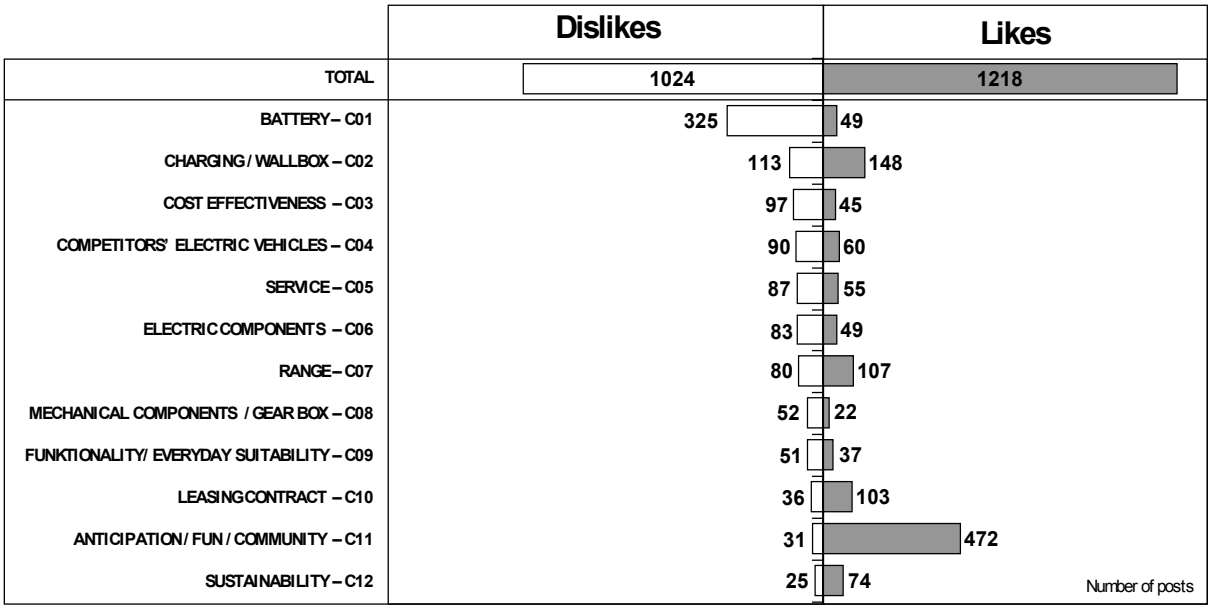


Figure 1. Categorization and number of analyzed posts.

The overall largest category is "C11: Anticipation, Fun, Community". Even before the actual handover, later MINI E drivers already formed a strong online community. Discussions mainly focused on expected vehicle features, assumed everyday suitability and service or infrastructure issues. After the handover customers shared experiences about how much fun it is to drive an electric vehicle with the sporty characteristics of a MINI E (150 kW engine, 220 Nm instant torque). The important role of an electric vehicle which is not only usable in everyday life but in fact fun to drive was later on confirmed in the UC Davis study (Turrentine et al., 2011). Also part of the fun and community aspects of electric driving were the self-organized driver meetings. Customers organized themselves not only online, but met on a regular basis bringing their EVs with them. It is obvious that a commercially available EV may not have the same potential to give the owner the feeling of being pioneers. However, it was significant to understand that community building is very important to EV customers. On the one hand, it reinforces the feeling of being part of an early adopter group, on the other hand there is the simple useful advantage of being able to share relevant information with other users that is not yet publicly available. For instance, when having difficulties to get high occupancy vehicle lane access stickers for EVs, MINI E drivers shared their experiences and suggestions on how to proceed.

Problems with the high-voltage system of the MINI E, "C01: Battery", includes the largest number of negative comments. The main complaint of the customers was that the function of the MINI E was massively compromised especially during very cold weather. In some cases, it took much longer than normal to charge the vehicle or was even not possible to charge if the battery was too cold. Cold temperatures also negatively affected the range while driving, partially because of a suboptimal working temperature for the battery, and partially because of the higher energy consumption of the heating system. The reason for this drawback can be found in the thermal management of the MINI E which relied on air cooling only. The air-cooled batteries performed somewhat better during very hot temperatures, but customers also complained about situations in which the regenerative braking was temporarily not available. Regenerative braking as implemented in the MINI E uses the electric motor as a generator when the foot is lifted off the accelerator pedal while driving, thereby feeding back deceleration energy into the battery. During very hot outside temperatures, it was possible that regenerative braking was temporarily deactivated by the vehicles' thermal management system in order to protect the battery from damage by charging an overheated battery. Some online customer feedback was very straightforward in terms of technical aspects: "*Note to BMW for the EV Citycar project - please water-cool the batteries! It's not so much the lost range, but the unpredictability...*" Turrentine et al.

(2011) analyzed that about one third of the MINI E drivers experienced difficulties due to hot or cold outside temperatures. The integration of liquid cooling/heating systems was technically not feasible at the time of the MINI E development. Prompted even more by customer experience reported in the field trials, a liquid thermal management system was introduced already in the BMW ActiveE in 2011, significantly reducing the reported problems and ensuring vehicle functionality even under extreme environmental conditions.

Category "C02: Charging / Wallbox" holds various interesting findings with a high number of positive but also negative comments. Particularly at the beginning of the field trial, technical difficulties complicated the installation of wallboxes at the customers' homes. Although difficulties like this are not unusual at the beginning of field trials and negative comments did not come as a surprise, the adverse situation provided valuable insights about customer expectations and potential for improvement. Most negative comments concerned the waiting time for the installation, uncertainty about installation costs or uncertainty about installation success as expressed in this statement: "*I just had my wall unit inspected and it failed. The inspector, really nice guy and very informative, told me he has failed three others and still has a few more to inspect (and fail)*". Without a wallbox, customers charged their vehicle on a standard wall socket with mach longer charging durations. From a customer's perspective these installation drawbacks were severe obstacles in intended EV use and they expected better service and assistance. For them, vehicle and wallbox are *one product* used in combination. From a car manufacturer's perspective, this situation very much reflects the magnitude of the shift from traditional sales processes with conventional cars. It means that it will not suffice to provide customers with a premium vehicle but that additional services beyond the car itself must be offered. Traditional structures in the automotive industry are not prepared for this. The BMW Group decided very early to close that gap as a part of the 360° ELECTRIC portfolio by offering wallbox solutions, installation checks and installation services and even green energy solutions, using information on future scenarios and customer requests from the field trials. Positive comments in category C02 were very beneficial to understand the meaning of home charging for EV customers. As could be expected and also reflected in later quantitative methods, the installation of a wallbox with shorter charging times greatly eased the charging process. One customer summarized it like this: "*I'm charged up - what normally takes 13 hours just took under 2 hours and I was at 98% from 62% in just over an hour. People - I've been with this 110 since June - I feel like I just exited the steam age.*" Much more as a surprise came the fact how much EV customers valued the independence from driving to the fuel station for gas. Opinions like this could be read frequently: "*I stop by my former gas station and wash my windows occasionally (but that's just because I like to rub it in) don't miss buying gas and would prefer to never do it again!*" Charging the EV can not only be performed at home but also in public. Public charging opportunities, however, were very limited during the MINI E field trials in the United States. Showing a strong sense of community and pioneer spirit, the MINI E drivers organized themselves: "[A MINI E driver] *has setup a pretty cool web site for people to sign up to be on a charger sharing list. This could be a very handy list for MINI E users if anyone is ever running low on battery a good distance from home. Or a good resource if you wanted to go on a longer trip.*"

As a final case study example the discussions about the already mentioned regenerative braking depict very well the advantages of social media monitoring in field trials. Regenerative braking allows an EV to use energy otherwise lost during deceleration and increases driving efficiency and range available. It also enables the customer to drive the electric vehicle very differently compared to a combustion engine vehicle: Regenerative braking slows the MINI E down up to -2.3 m/s, which, after a short while of practice, even allows to stop the car at a red traffic light without using the friction brake. The question was, however, and asked with emphasis from a development point of view in the light of introducing a substantial change in vehicle driving dynamics: Will customers accept the feature, as it helps to save energy? Even if so, how do they rate the resulting driving experience? Social media monitoring helped again to get unfiltered and very early customer opinions, especially as we did not want to ask explicitly for the customers' opinions in the very first approach to avoid any response biases. Shortly after the first contact with the new vehicle many customers commented online that although regenerative braking feels quite unusual it is very well accepted: "*I'm just back from test driving the MINI E!* […] *It will take a few drives to get use to the auto regenerative brakes. When your foot comes off the gas pedal the regenerative brakes kick on. It's a little jarring at first but once you learn to slowly take your foot off the gas you almost never need to actually use the brake pedal.*"; "*After only one day and 50 miles, I am pleasantly surprised at how much fun it is to drive* [… I] *rarely have to use the brakes thanks to the aggressive regeneration.*"; "*Most people have the same reaction. At first they are a bit put off by the strong regen* [=regenerative braking]*, but after a short while driving it, they love it.*" Customer interviews conducted by our research partners after at least three months of EV usage reveal an even stronger appreciation with growing driving experience. Following the interview plan, a random subset of customers was asked if they liked regenerative braking and the associated single-pedal driving. Feedback was overwhelmingly positive with agreement rates of 88% in China (n=49) to 92% in Japan (n=26) up to even 100% in the UK (n=33), Germany

(n=25) and the US (n=72). Between 74%-95% of customers over all countries involved stated that it has become a highly self-motivating almost game-like situation to stop the vehicle at the desired position with regenerative braking only. Given this rare unambiguity in customer feedback, regenerative braking and single-pedal driving were recognized as a positive EV-specific factor that was defined as a requirement for all following pure battery electric vehicles including the BMW i3. However, derived from the reactions on social media and validated with later customer interviews, it was also important to see the customers' surprise in first contact that necessitates adequate information about this newly developed feature.

SOCIAL MEDIA MONITORING AS AN EARLY FIELD STUDY TOOL

Social media monitoring proved to be very helpful to guide the development of qualitative and quantitative research instruments for later phases of the field trials. Using this feedback in early phases of a long-term study corresponds also to the highest activity rates of web users. It may not come as a surprise that almost half of the total number of comments was posted during the first half year of the trial (see Figure 2).

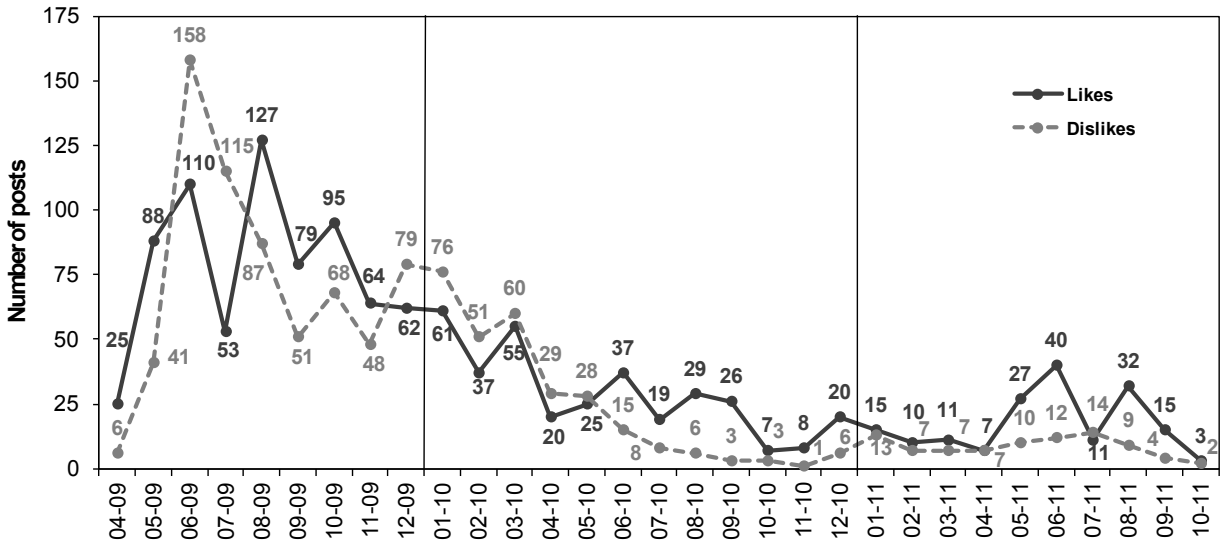


Figure 2. Number of posts during the analysis period 2009-2011.

Prior publications on that subject as listed in the first section of this document already pointed out the risks associated with confusing using social media comments with representative customer feedback or misunderstanding total numbers of likes and dislikes as a representative pattern of opinions. These precautions were also essential when looking at the overall MINI E social media feedback. At the same time it is obvious that a lot can be learned even from single comments. For instance, the report of several MINI E users experiencing difficulties in the installation process is a direct hint that actions need to be undertaken as this problem may occur to any customer. However, as this example also shows, adequate and cost-effective solutions always depend on deeper analysis: Are really all customers affected by a certain (no matter how obvious) problem or does it apply only to sub-groups? Which underlying causes led to the problem (missing service infrastructure, missing technological solutions, regulations restricting solution options, etc.) and are the explanations valid for all (sub-)groups? Can the same range of services be offered to all customers in all countries in order to solve the problem?

These questions point out that findings in social media can only be the first step in an wider chain of analysis methods. Used as a shaping tool for following systematic quantitative or qualitative research with controlled samples, social media monitoring in the context of disruptive innovation processes proved to be extraordinarily beneficial delivering the earliest possible indicators for further research in the field trials reported here.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the *project i* team at BMW Group for making this research possible. We would like to thank our international research partners for bringing in their expertise and for conducting the research with extraordinary commitment. Thanks to Juliane Schäfer for introducing social media monitoring to the project and to Andreas Klein for very helpful input to earlier versions of this document. Most of all we would like to thank all MINI E and BMW ActiveE customers for their outstanding support of the field trial research and their inspiring feedback.

REFERENCES

Axsen, J. & Kurani, K.S. (2012). Interpersonal influence within car buyers' social networks: applying five perspectives to plug-in hybrid vehicle drivers. *Environment and Planning A*, *44* (5), 1047-1065.

Bartl, M. & Ivanovic, I. (2010). Netnography - finding the right balance between automated and manual research. In P. Brauckmann (Ed.), *Web-Monitoring: Gewinnung und Analyse von Daten über das Kommunikationsverhalten im Internet* (pp. 157-174). Konstanz: UVK Verlags GmbH.

Edwards, A., Housley, W., Williams, M., Sloan, L. & Williams, M. (2013). Digital social research, social media and the sociological imagination: surrogacy, augmentation and re-orientation. *International Journal of Social Research Methodology*, *16* (3), 245-260.

Finch, B. J.(1999). Internet discussions as a source for consumer product customer involvement and quality information: an exploratory study. *Journal of Operations Management*, *17* (5), 535-556.

Füller, J., Bartl, M., Ernst, H. & Mühlbacher, H. (2006). Community based innovation: How to integrate members of virtual communities into new product development. *Electronic Commerce Research*, *6* (1), 55-73.

Klein, A. & Spiegel, G. (2013). Social media in the product development process of the automotive industry: a new approach. In M. Kurosu (Ed.), *Human-Computer Interaction, Part III, HCII 2013, LNCS 8006* (pp. 396-401). Berlin: Springer.

Kozinets, R.V. (1999). E-tribalized marketing? The strategic implications of virtual communities of consumption. *European Management Journal*, *17* (3), 252-264.

Kozinets, R.V. (2002). The field behind the screen: using netnography for marketing research in online communities. *Journal of Marketing Research*, *39* (1), 61-72.

Ramsbrock, J., Vilimek, R. & Weber, J. (2013). Exploring electric driving pleasure. The BMW EV pilot projects. In M. Kurosu (Ed.), *Human-Computer Interaction, Part II, HCII 2013, LNCS 8005* (pp. 621-630). Berlin: Springer.

Sampson, S.E. (1996). Ramifications of monitoring service quality through passively solicited customer feedback. *Decision Sciences*, *27* (4), 601-622.

Turrentine, T., Garas, D., Lentz, A. & Woodjack, J. (2011). *The UC Davis MINI E consumer study* (Research Report No. UCD-ITS-RR-11-05). Davis, CA: Institute of Transportation Studies, University of California.

Vilimek, R. & Keinath, A. (2014). User-centred design and evaluation as a prerequisite for the success of disruptive innovations: an electric vehicle case study. In M. Regan, T. Horberry & Stevens, A. *Driver acceptance of new technology: theory, measurement and optimisation* (pp. 169-186). Farnham: Ashgate.

Vilimek, R., Keinath, A. & Schwalm, M. (2012). The MINI E field study - similarities and differences in international everyday EV driving. In N.A. Stanton (Ed.), *Advances in human aspects of road and rail transport* (pp. 363-372). Boca Raton, FL: CRC Press.

Does Range Anxiety Exists? Experiences from the ELVIRE Project

Maria Nilsson [a,b]

[a]*Viktoria Swedish ICT*
Göteborg, SE-417 56, Sweden

[b]*Lindholmen Science Park*
Göteborg, SE-417 56, Sweden

ABSTRACT

This paper reports on the results from the EU-project ELVIRE regarding the phenomenon of range anxiety, the fear of not reaching the destination in an electrical vehicle (EV). Presented is the initial analysis of a questionnaire study with 42 Swedish EV drivers targeting their experience of range from a behavioural perspective (e.g., driving habits), as well as from a psychological perspective (e.g., experience of unease while driving). The following questions are addressed: (a) How is range anxiety portrayed by EV drivers? and (b) What is the role of experience of EVs for range anxiety? Regarding (a), findings show that the drivers' posses a behaviour indicating that they suffer from range anxiety, although, they explicitly state that they have a low level of range anxiety; highlighting the unconscious nature of range anxiety. Regarding (b), statistical analysis show that there is a significant difference in level of experienced range anxiety and understandings of vehicle information but no significant difference in driving style between drivers who has more than 6-months experience compared to those with less than 6-month experience; indicating that driving style are manifested early on and stays stable over time.

Keywords: Electric vehicle, User experience, Range Anxiety, driving behaviour,

INTRODUCTION

Electric vehicles (EVs) have the potential to truly transform the automotive industry. The environmental benefits of the success for the technology are widely emphasised. Range anxiety (the fear of getting stranded in an EV due to low battery) is typically mentioned as the main factor affecting the penetration of EVs on to the global market together with long charging times and a high purchase price (Hidrue et al., 2011). It is repeatedly maintained that range anxiety is the one obstacle that may hinder the success of EVs (Hidrue., et al. 2011; Wynn & Lafleaur, 2009; Wellings et al. 2011; Botsford & Szczepanek, 2009).

Range anxiety is, in many ways, an intuitive term, defined herein as the perceptions or the experience of drivers regarding the fear of not reaching destination while driving an EV. Both qualitative and quantitative studies present research findings on the topic and numerous blogs have debated its existence and importance. A report by Wellings et al., (2011), for instance, suggests that range anxiety is reduced as the drivers become accustomed (familiarized) to the EVs. Indeed, Taylor (2009) reports that range anxiety, in fact, decreases within just a few weeks of actual use of an EV. The positive effect of experience on range anxiety is also noted by Franke et al., (2012b). However, some other drivers have reported that their range anxiety increases as they experience that they cannot trust the feedback they receive from the EV (Wellings, et al., 2011). This is also consistent with findings reported by Morton (2011) showing that pre/post trials have revealed that drivers are over-cautions when planning journeys and that range anxiety can, indeed, increase. Pichelmann et al (2013) concludes, amongst other, that the average EV driver needs three months to be able to reach his or her "maximum" driving range. Indeed, appraoches that has the potential to decrease the amount of time needed to reach the "maximum driving range" has also been suggested by e.g,

Lundtröm (2012), Nilsson & Habibovic. A. (2013). Franke & Krems (2013) highlights the fact that other variables than experience affects range anxiety such as systems competence and personality traits. Interestingly, EV drivers are satisfied with the level of range and limited range incidents rarely occurred. It should be noted that previous studies typically only cover one single brand of EV. Considering the definitions of range anxiety (cf. Nilsson, 2011a), some definitions refer to the problem of recharging the EVs, that is, range anxiety may relate to the occurrence of the particular situation. Others mention the available actions needing to be performed when in a situation in which the destination can no longer be reached.

To further our knowledge on the effect of time and experience on range anxiety, a questionnaire study is presented in this paper. It compares the behavioural aspects (e.g. driving habits) and the psychological aspects (e.g., the experience of unease) of range anxiety of drivers with different levels of experience. The questionnaire study only targets experienced EV drivers in Sweden. This study thus excludes the opinions of prospective costumers, which does not have access to an EV. That is, the focus is on the *experience* rather than the *perception* of range anxiety. Also to be noted, the survey is not excluded to one particular EV brand.

About The ELVIRE project

The ELVIRE project (www.elevire.eu) had the goal to minimise the EV driver's "range anxiety" by developing an effective communication and service platform to help drivers to managing the charge of their EV and enable efficient use of energy. In details, ELVIRE approached range anxiety from a holistic viewpoint (cf. Nilsson 2011a). ELVIRE not only provided strategies to prevent a limited range incident to occur (i.e., re-route navigation based on energy demand), but also informed about the current status (i.e., range estimation visualisations). In addition, ELVIRE limited the risk for the occurrence of a limited range incident (i.e., provided services to find available charging infrastructure), provided solutions to solve the (emergency) limited range incident (i.e., provide service for charging spot reservations), as well as strategies to eliminate the consequences of a limited range situation (e.g., battery swop). Several studies regarding range anxiety have been performed within project: (1) defined the phenomenon by analysing the definitions of range anxiety (Nilsson, 2011a) and (2) characterised the phenomenon by analysing reported limited range incidents in interviews (Nilsson 2011b; Luettringhaus & Nilsson, 2012). As a complement to the previous investigations, a questionnaire study was performed and presented herein.

METHODOLOGY

Survey Design

The questionnaire investigated the drivers' experience of range from a behavioural perspective (e.g., driving habits), as well as from a psychological perspective (experience of unease while driving). A closed-ended questionnaire that included 21 statements to be judged on a 5-point Likert scale was used.

Descriptive statistics as well as statistical analyses (paired T-tests) were performed to identify differences and similarities among the data. The descriptive statistics was further explored by studying the comments made by the respondents in the free text field.

Participants

The questionnaire was answered by 42 Swedish EV drivers, with no restriction on EV brand. 8 women (19%) and 34 men (81%)) participated. 21 of the respondents had more than 6 months experience of driving an electric vehicle, while 16 of them had 1-6 months experience. 5 respondents had less than 2 weeks of experience. Most of the respondents had access to the electric vehicle via their work (20 respondents), 10 of them had participated in a research project, 8 had access via a car pool and one of them owned the vehicle himself. They had all had experience from different EV brands. Some respondents in the survey had experience from more than one brand. 18 had used a Volvo (43%), 9 had experience of Fiat (21%), 22 had experience of Mitchibishi (52%), 1 used a Peugeot (2%), 10 had experience of other brands not explicitly stated in the questionnaire (24%).

Material

The questionnaire is based on the previous research performed within the ELVIRE project on range anxiety, cf. Nilsson, 2011a and Nilsson 2011b. The questionnaire used a closed-ended format in which categorical, multi-response and rating (Likert) scales were used. The rating questions used a five-step rather than a seven-step scale due to people's tendency not to mark the edge numbers. For each question, it was possible for the respondents to add comments to explain their choice. The comments were later used to triangulate the findings to increase the reliability and validity of the results from the questionnaire. In the case of multiple-choice questions, the available options were defined based on the previous research. Also, the questionnaire was limited in size and published online to increase the response rate.

After an introductory text (outlining the purpose of study, the treatment of results etc.), the questionnaire had five main sections (A-F). The purpose of Section A was to collect the respondents' background/previous experience. The data collected were age, EV brand, experience, access to an EV, and their usage of information sources in the EV. The questions were of categorical and multiple choice type. Section B investigated weather the respondents had experienced any situation without battery power. The respondents were asked to respond to a set of statements and judge how well they fitted their experience on a 5-point Likert scale. The purpose of Section C was to investigate how respondents deal with various limited range situations. Similar to the previous section, the respondents were asked weather or not they have experienced the stated situation. Section D, focused on the psychological experience of driving an EV and the existence of range anxiety. As in previous sections, the participants used a 5-point Likert scale to indicate to what degree they agreed with the statements. Section E further investigated the problematic issues and the effect of range anxiety and limited range situations on drivers. Section F, in contrast, had the purpose of ensuring the correctness and validation of the answers provided in the previous sections. That is, the respondents were asked to judge their own understanding of and ability to answer the questions. That is, the respondents were asked to judge their own understanding of and ability to answer the questions.

Procedure

To be geographically independent, an online questionnaire was used to collect data. Lindholmen Science Park (Test Site Sweden) enabled the participation of respondents from several different Swedish national on-going projects (e.g., The Electrical Vehicle Initative, ELVIIS, Green Highway, E-mobility Malmö). The contact person at each project was provided with an introductory e-mail containing details of the study and a link to the electronic questionnaire to be sent out to the respondents. The respondents had 2 weeks from receiving the e-mail to answer the questionnaire.

RESULTS AND ANALYSIS

To investigate the effect of the experience on perceived range anxiety, the respondents were divided into two groups dependent on their level of experience: Group 1 = < 6 month experience; Group 2= > 6 months experience. That is, the effect of age, gender or EV brand on perceived range anxiety were not investigated; rather the population was treated as single group of participants represents the general EV community.

First of all, to investigate the factors influencing range anxiety the "level of agreement" has been calculated, and contrasted between the two groups. Table 3 list the different statements from the questionnaire and the level of agreement (i.e., respondents indicating 3 or higher on the likert scale). What can be seen in table 3 is that 13 of the 21 statements were indicated 3 or higher. This gives indication on what factors affect range anxiety. In particular, in Table 1 one can see that statements regarding driving style have a high agreement while statements of their psychological state have, generally a less high agreement level. The low agreement can be due to the fact that they did not experience range anxiety to a great extent.

Table 1. Statements and their level of agreement (percentage of respondents indicating 3 or higher on the Likert scale).

Statement	Level of agreement ALL	Experience <6 months	Experience > 6 months
B1. I have been in situations in which I got stranded along the way with no power in the engine	5%	5%	5%
B2. I have been in situations in which I have made an unplanned stopped to charge the EV due to low power	14%	19%	10%
B3. I have been in situations in which I was almost out of power	48%	48%	48%
B4. I have been in situations in which I have not taken the EV due to the uncertainty of the range of the vehicle	71%	52%	90%
B5. I have been in situations in which I needed a longer range	74%	62%	85%
C1. I am constant aware of the range	83%	90%	76%
C2. I usually drive to pre-defined destinations (A-B) rather than multiple destinations (A-B-C-D-A)	86%	86%	86%
C3. I carefully plan the travel before any trip	62%	57%	67%
C4. I always know the distance to the place I am going to	90%	90%	90%
C5. I adjust my driving according to the current battery level	83%	86%	80%
D1. I experienced the total driving range as (sufficient/not sufficient)	52%	57%	47%
D2. I have difficulties to understand how much range there is in the EV at any point	40%	47%	33%
D3. I find it difficult to understand how the different instruments effect each other	24%	43%	5%
D4. I have difficulties to understand how much energy it takes to get to the final destination	60%	71%	48%
D5. I experienced that information/values were changing without me knowing why	29%	47%	10%
E1. I feel worried for running out of energy while driving	36%	52%	19%
E2. I think it would be problematic to run out of energy as I would not reach the final destination on time	79%	85%	71%
E3. I think it would be problematic to run out of energy as I would need to fix alternative vehicle/transportation	86%	90%	80%
E4.i think it is problematic to run out of power as it takes so long before I can use the vehicle again	71%	81%	62%
E5. I would not feel comfortable in a situation in which I am without energy along the road	86%	95%	76%
E6. I do not know how I should act if I were to be out of energy along the road	28%	33%	24%

In the statistical analysis, the mean value from each question was compared between the two groups. As it can be seen in Table 2, statistical t-test reveals that three statements showed significant differences between the groups, with a probability of P=<0,05: D5. I experienced that the information changed without me knowing why [2,4 vs. 1,7] ; E1. I feel worried for being out of energy while driving [2,7 vs. 2,0]; E5. I would not feel comfortable in a situation without energy [4,6 vs. 3,7]. That is, significant differences in D5 show that participants that had less than 6 month experience thought that information changed to a higher degree without them knowing why, while those with more than 6 month experience knew why and how the information changed. This implies that there is a learning effect, dependent on level of experience. However, it should be noted that the mean value is on the lower end of the scale [2,4 vs. 1,7], indicating that they do not fully agree with the statement "I experienced that the

information changed without me knowing why". In addition, significant difference in E1 indicates that there is a difference in level of the range anxiety between the two groups. This imply that participants become more secure and confortable while driving EV. However, as similar to D5, the mean is on the lower range of the scale [2,7 vs. 2,0], indicating that the actual presence of range anxiety is low. Moreover, the significant differences in E5 shows that the participants experience that they would not feel confortable in a situation in which they are out of energy. The significant differences between the two groups show that the level of comfort changed along with the experience. Interestingly, there are no significant differences in statements regarding driving style or EV usage. This show that the driving style is, somewhat, stable and unaffected by experience.

Table 2. Overview of means between the two groups. The explicit statements corresponding to "B1" etc. are noted in Table 1.

B1	B2	B3	B4	B5	C1	C2	C3	C4	C5	D1	D2	D3	D4	D5	E1	E2	E3	E4	E5	E6
< 6 months (Mean/standard deviation)																				
1,2 / 0,9	1,8 / 1,3	2,8 / 1,5	3,2 / 1,8	3,2 / 1,7	4,3 / 1,0	3,9 / 1,2	3,1 / 1,4	3,8 / 0,9	3,7 / 1,2	3 / 1,5	2,8 / 1,4	2,3 / 1,3	3,2 / 1,1	2,4 / 1,3	2,7 / 1,3	4,1 / 1,4	4,4 / 1,2	3,9 / 1,4	4,6 / 1,0	2,1 / 1,0
> 6 months																				
1,3 / 0,9	1,2 / 0,8	3,0 / 1,6	3,6 / 1,2	4,0 / 1,2	3,9 / 1,3	3,5 / 1,0	2,8 / 1,3	3,7 / 1,1	3,6 / 1,4	2,9 / 1,5	2,2 / 1,3	1,4 / 0,6	2,6 / 1,4	1,7 / 1,0	2,0 / 0,9	3,6 / 1,6	3,9 / 1,4	3,4 / 1,7	3,7 / 1,4	1,7 / 0,9

When considering the standard deviation (cf. Table 2), some of the results highlight large differences within the groups. Therefore, to provide a better understanding of the data, Table 3 presents the results in more detail by providing descriptive statistics in the form of the percentage of users for each level of the Likert scale for each group.

Table 3. Overview of data as divided between the two groups. Noted is the percentage of drivers indicating each of the levels 1 to 5.

Level		B2	B3	B4	B5	C1	C2	C3	C4	C5	D1	D2	D3	D4	D5	E1	E2	E3	E4	E5	E6
Group 1: < 6 month experience (percentage)																					
1	**90**	**67**	24	33	24	0	5	14	0	5	19	19	**38**	5	**38**	**24**	10	10	10	5	**38**
2	5	14	**29**	0	14	10	10	**29**	10	10	**24**	**33**	19	24	14	**24**	5	0	10	0	29
3	0	5	14	14	14	10	19	14	24	**29**	19	14	24	24	19	19	14	5	10	5	24
4	0	5	14	14	10	19	29	19	**43**	**29**	14	19	14	**38**	29	**24**	10	14	24	10	10
5	5	10	19	**38**	**38**	**62**	**38**	24	24	**29**	**24**	14	5	10	0	10	**62**	**71**	**48**	**81**	0

Group 2: > 6 month experience (percentage)																					
1	**81**	**90**	24	5	0	0	5	24	5	10	19	**38**	**67**	24	**52**	29	14	10	24	10	**57**
2	14	0	**29**	0	14	24	10	10	5	10	**29**	29	29	**29**	38	**52**	14	10	14	14	19
3	0	5	5	**21**	19	14	29	**43**	**33**	29	10	19	5	24	5	10	14	14	5	14	24
4	0	5	14	10	14	14	**43**	14	29	14	14	5	0	10	0	10	10	19	14	19	0
5	5	0	**29**	14	**52**	**48**	14	10	29	**38**	24	10	0	14	5	0	**48**	**48**	**43**	**43**	0

CONCLUSIONS

Presented is the initial analysis of a questionnaire study investigating the effect of time and experience on range anxiety. The findings indicate that even though participants explicitly responded that they did not experience range anxiety, their responses to the driving style questions indicate that range anxiety, implicitly and unconsciously, exists and shapes their EV usage. Furthermore, there are indications that the respondents do not put themselves in a situation in which they may be out of power. The majority of the drivers had been in situations in which they chose not to take the EV due to its limited range. Likewise, a majority of the drivers had not been stranded or had had the need to make an unplanned stop because of limited energy. This is may explain why that the majority of the drivers did not experience range anxiety while driving, and further strengthen the fact that range anxiety is to some extent, unconscious. Indeed, statements regarding the explicit worry of range anxiety (e.g., "I feel concern of running out of energy while driving") received a mean below three while statements regarding their driving style (e.g., " I constant aware of range while driving; I typically driving to one destination) received a mean over three. This contradiction highlights the importance of investigation into the effect of range anxiety and the study of range anxiety as an unconscious phenomenon. The statistical analysis (T-test) shows that there was no significant difference in *driving style* (e.g., driving habits) between drivers who had more than 6-months experience and those who had less than 6-month experience. Indicating that experience does not affect driving style. However, one can assume, based on the results of the questionnaire, that there is a learning effect (experience effect) that affects (1) level of experienced range anxiety, (2) the level of comfort for situations with energy breakdown (i.e., getting stranded along the way due to limited energy), and (3) ability to understand changes in the EV's instrumentations. Moreover, even though the majority of the respondents agreed that they would not feel comfortable in a situation in which they were out of energy, yet they felt certain of how to act in such situations.

To conclude, the results from the presented questionnaire study show that range anxiety is a complex phenomenon that may implicitly and unconsciously shape drivers EV usage. Also the results indicate that driving style does not change with level of experience.

ACKNOWLEDGEMENTS

This work was performed by Lindholmen Science Park, Sweden, within the ELVIRE project, supported by EU FP7 PROJECT ID: ICT-2009.6.1

REFERENCES

Botsford, C. & Szczepanek, A. (2009) “Fast Charging vs. Slow Charging: Pros and cons for the New Age of Electric Vehicles”. EVS24 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium. Stavanger, Norway, May 13-16

Hidrue, M. K. Parsons, G., Kemton, W., Gardner, M. P (2011) Willingess to pay for electrical vehicles and their attributes” Resources and Energy Economics 33, pp. 686-705

Franke, T., Bühler, F., Cocron, P., Neumann, I., Krems, J. F. (2012a). “Enhancing sustainability of electric vehicles: A field study approach to understanding user acceptance and behaviour”. In: Sullman, M., Dorn, L. (Eds.), Advances in Traffic Psychology. Ashgate, Farnham, UK, pp. 295-306.

Franke, T., Cocron, P., Bühler, F., Neumann, I., Krems, J.F. (2012b). “Adapting to the range of an electric vehicle- the relation of experience to subjectively available mobility resources”. In: Valero Mora, P., Pace, J.F., Mendoza, L. (Eds.), Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems, Valencia, Spain, June 14-15 2012. Humanist Publications, Lyon, France, pp. 95-103.

Franke, T., Krems, J.F. (2013). “Interacting with limited mobility resources: Psychological range levels in electric vehicle use”. Transportation research part A: Policy and Practice 48, pp. 109-122.

Lundström, A., (2012) “COPE1 – Incorporating Coping Strategies into the Electric Vehicle Information System”. 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '12), October 17–19, 2012, Portsmouth, NH, USA

Luettringhaus, H., Nilsson, M. (2012) ”Elvire appraoches to mitigate EV driver´s range anxiety: technical paper by elvire-european research project on ICT for electric vehicles”. In Proceedings of 19th ITS World Congress, Vienna, Austria, 22-26 Oct, Paper Number: EU-00086

Morton, Schuitema and Anable (2011) “Electric vehicles: Will Consumers get Charged Up?” Technical document, Open University, Milton Keynes. In proceeding of: Universities' Transport Study Group Annual Meeting, At Milton Keynes. Available at: http://design.open.ac.uk/documents/Morton2011.pdf

Nilsson, M (2011a) “Electric vehicles: the phenomenon of range anxiety”. Public deliverable Task (1300) the ELVIRE project. Available at: http://www.elvire.eu/IMG/pdf/The_phenomenon_of_range_anxiety_ELVIRE.pdf [2014-01-25]

Nilsson, M. (2011b) “Electric vehicles: An interview study investigating the phenomenon of range anxiety". Public deliverable Task (5000) ELVIRE project. Available at: http://www.elvire.eu/IMG/pdf/An_interview_studyinvestigating_the_phenomenon_of_range_anxiety_ELVIRE-2.pdf [2014-01-25]

Nilsson, M., Habibovic. A. (2013) “Identifying EV drivers’ needs for information communication technology to ease the EV”. In adjunct proceedings of 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive UI´13), Eindhoven, the Netherlands, Oct 27-30.

Pichelmann, S., Franke, T., Krems, J., F (2013). “The Timeframe of Adaptation to Electric Vehicle Range”. In M. Kurosu (Ed.), Human-Computer Interaction. Applications and Services, LNCS 8005 (pp. 612-620). Berlin, Germany: Springer.

Taylor, D. (2009). The Differences and Similaritities between Plug-in Hybrid EVs and Battery EVs. 24th International Batteries, Hybrid and Fuel Cell Electric Vehicle Symposium (EVS24), Stavanger Morton

Wellings, T., Binnersley, J., Robertson, D., Khan, T (2011) “Human machine interfaces in low carbon vehicles: Market trends and user issues”. Dokument No. HMI 2.1. Low carbon vehicle technology project: workstream 13. Available at: http://www2.warwick.ac.uk/fac/sci/wmg/research/low_carbon/automotive/lcvtpevents/ws13_hmi_in_lcvs_market_analysis_and_user_issues_v2_1.pdf [2014-01-25]

Wynn, T & Leafeur, S. (2009) “A free market perspective on electrical vehicles”1. Technical Report, CascadePolicy Institute. Available at: http://cascadepolicy.org/pdf/env/FMPonElecVehicles_100109a.pdf [2014-01-25]

Attitude Towards Electric Vehicles in the Czech Republic

Petr Zámečník [a], Matúš Šucha[b] , Pavel Řezáč [a] and Kateřina Bohmová [a]

[a] Transport Research Centre
Czech Republic

[b]Palacký University Olomouc
Czech Republic

ABSTRACT

Present study deals with the attitudes of people in the Czech Republic towards electric vehicles. Research was carried out in the form of a questionnaire, which was filled by 245 persons aged 17-81 years from all over the Czech Republic. The results show that for most respondents is the acquisition of electric vehicle few conceivable. The most common reason is the cost. Other reasons are a small number of charging stations, small trunk and limited commuting travel. The results also show that between men and women is not a statistically significant difference in attitudes to electric vehicles in general. The exception is the attitude to the fact that electric vehicles are environmentally friendly and the area of the purchase price. Women, compared to men, perceive less the cost of electric vehicles as excessive. At the same time women (unlike men) consider electric vehicles as much more environmentally friendly.

Keywords: Attitude Towards Electric Vehicles, Electromobility, Green E-Motion, Alternative Fuel

INTRODUCTION

The European Union in the frame of the Europe 2020 Strategy undertook by 2020 to promote greener and more competitive economy which will be less resource demanding. Concrete expression of the priority of sustainable growth in the field of climate and energy is commitment "20-20-20", which aims by 2020 to reduce greenhouse gas emissions by 20% compared to 1990 level, to increase the share of renewable energy in total energy consumption in the EU on 20% and increase energy efficiency in Europe by 20%. Furthermore Transport White Paper aims to eliminate by 2050 conventionally powered vehicles in the cities and urban areas, along with other measures, plans to reduce emissions in transport by 60%. White Paper - Roadmap to a united European space aims to reduce greenhouse gas emissions by approximately 20% below 2008 level by 2030 and to reduce greenhouse gas emissions by at least 70% below 2008 level by 2050 (Žák et al., 2012).

Several surveys have been vehicleried out recently in EU to collect feedback from potential consumers on electric vehicles such as testing the familiarity of vehicle drivers with electric vehicles, investigating their interest to purchase an electric vehicle as well as inquiring about their priorities for improving the features of current electric vehicles (e.g. Ernst & Young, 2010; Bunzeck et al., 2011; EurotaxGlass, 2011; McMullan, 2012; Thiel et al., 2012). In the Czech Republic was held a survey on attitudes to electric vehicles for small, medium and large enterprises in the project "Elektromobil Prague". The results indicate that up to 70% of all respondents (Prague companies of different sizes) considering in next five years (till 2016) purchase of an electric vehicle, and up to 30% of them that count with that. More than half of the companies included in the survey had own fleet of over 10 vehicles and a quarter with more than 500 vehicles (Žák et al., 2012).

METHOD

We used quantitative processing of electronic form questionnaires. Statistical data processing was vehicleried out using SPSS.

RESEARCH SAMPLE

Our research is focused not on the company, but the individual residents of the Czech Republic. As part of the telephone survey was piloting the questionnaire, which showed that 9 out of 10 respondents refused to answer or continue to communicate when asked on household income. The questionnaire was subsequently distributed in electronic form so as to obtain a demographically and geographically balanced sample.

Total participants were 271 people. 26 participants were excluded from the analysis because of incompleteness (gender age etc.). Finaly we were working with 245 persons surveyed (134 women and 111 men). All persons surveyed were aged 17 - 81 years. The research sample was 128 persons employed 40 persons on parental leave, 29 entrepreneurs, 28 students, 13 pensioners and the unemployed 6. Three respondents did not specify their economic activity. The sample came from the entire country - from all 13 regions. Representation of individual regions, however, has been uneven. Most people (84) lived in the Southmoravian Region, the city of Prague was represented by 36 persons.

DATA GATHERING

The questionnaire was distributed via the link on social networks and presentation of links to websites dealing with the transport issues. Due to ensure acquiring subjects older than 50 years, a questionnaire was distributed among these people personally. Filling out an anonymous questionnaire took an average of 20 minutes.

DATA ANALYSIS

Data were analyzed using SPSS program. Before beginning the analysis, some questionnaires were excluded. Firstly were excluded incompletely filled questionnaires and secondly those, in which respondents indicate that they know nothing about electric vehicles (as their opinion in this case is rather hypothetical). In the analysis, we worked with gross scores, according to how respondents identified their testimony on the scale. The research sample was unevenly represented relative to the population except gender representation. First, we conducted a simple analysis of the frequency-response range of the testimony. Because of the uneven distribution of the examined sample, we also work with non-parametric test. To determine the difference in the variables between the groups, we used the nonparametric test Mann - Whitney U for independent samples, due to the failure to achieve the above normal distribution of the variables in the research files. For more detailed findings of the size differences between the variances, we also investigated the Cohen's d effect size - r (effect size). We additionally computed using the online calculator.

RESULTS

Driving license owned 90 % of the survey respondents. When asked what respondents know about electric vehicles, answers correspond mostly with three topics: environment friendly, run on electricity and that are expensive. There have been also answers that electric vehicles need long time to charge for a relatively short trip. Respondents were asked specific questions with the option to tick the answers to range "Definitely yes, rather yes, rather not, definitely not, I do not know." The conclusions of these issues could be summarized as follows. Except for two respondents are electric vehicles considered as expensive. Topic about the costs of electric vehicles is also related to the fact that only 12 % of respondents think that they would be able to buy an electric vehicle. 70 % of people surveyed think that they will not be able to buy an electric vehicle. The rest of those surveyed currently do not know if they could buy an electric vehicle.

Whether the sale of electric vehicles will in the next five years increase significantly, most respondents answered (40 %) probably not, but another 28 % think that more so, because this question is very unclear. On the question of whether electric vehicles are safe, most respondents replied that "Definitely yes and Rather yes" (61 %). Respondents in 13 % cases believe that electric vehicles can travel on great distances. 37 % of respondents believe that their friends consider them as a fools if they buy an electric vehicle. A total of 73 % of the respondents think that electric vehicles are environmentally friendly mode of transport. The research sample is considering to purchase some vehicle in next five years (not solely electric) in 55 %. When asked whether they would choose an electric vehicle when buying a vehicle, 96 (42%) out of a total of 226 respondents answered “Rather not”. 73 more persons (32%) said that they would definitely not buy an electric vehicle. Only 27 respondents would be willing to purchase an electric vehicle, 21 (9%) out of which chose “Rather yes” and merely 6 (2%) of them would definitely buy this type of a vehicle. The remaining 30 respondents did not know whether they would buy an electric vehicle or not. Reasons for not buying an electric vehicle are as follows: the most frequent was high purchase price of the vehicle (95 % of respondents), 91 % of respondents resent a small number of charging stations for electric vehicles, 85 % of respondents did not want electric vehicle because of the small vehicle go area (it is a “city car”), 77 % percent of respondents resent the total size of the vehicle, 76 % of respondents discouraged from buying limited battery life, 68 % of respondents matters little driving distance of the vehicle, 30% were discouraged from buying by low top speed.

A combination of several reasons was most often given by the respondents who said that they would not buy an electric vehicle. Multiple reasons for not buying an electric vehicle were stated by 60 respondents. A sole reason for not buying an electric vehicle was given by 57 respondents. The most common reason why respondents would not buy this type of a vehicle was its high purchase price. This was also the only reason stated by a total of 32 respondents (13 %). 44 respondents (18%) mentioned the purchase price along with another reason. The second most common reason for not buying an electric vehicle was its short driving distance mentioned in a total of 48 cases (20 %). According to the respondents, the third most common problem is the limited number of charging stations, as stated by 23 respondents (9 %). In sum, the reasons why the respondents thought about buying an electric vehicle are 87% of respondents low operating costs, 69% of respondents environmentally friendly, 74 % of respondents would be encouraged to buy an electric vehicle if they can recharge your batteries at home, with 72 % of respondents low noise, 66 % of respondents would purchase an electric when stimulated by state support. Some respondents would also appreciate enhanced performance of electric vehicles, simple charging facilities in the vicinity of their home, a possibility of having the vehicle repaired in a standard vehicle repair shop, availability of spare parts or the possibility to use an alternative unit in the case of a flat battery. Several respondents stated that they would buy an electric vehicle if it was largely subsidized.

An analysis of the Mann-Whitney U resulted in differences between the testimonies of men and women (see Table 1). It is evident that the majority of responses to the scale-response differences exist between men and women.

Table 1: Mann – Whitney U test (gender difference)

	Gender	N	Mean Rank	Sum of Ranks
Electric vehicles too expensive	Female	134	116,06	15552,50
	Male	111	131,37	14582,50
	Total	245		
Sales of electric vehicles will greatly increase	Female	134	122,14	16367,00
	Male	111	124,04	13768,00
	Total	245		
Electric vehicles are safe	Female	134	118,46	15873,00
	Male	111	128,49	14262,00
	Total	245		
You can travel great distances with electric vehicles	Female	134	122,83	16459,00
	Male	111	123,21	13676,00
	Total	245		
My friends would think I was crazy if I bought an electric vehicles	Female	134	118,54	15884,50
	Male	111	128,38	14250,50
	Total	245		
Electric vehicles are environmentally friendly	Female	134	142,81	19136,00
	Male	111	99,09	10999,00
	Total	245		

Test Statistics[a]

	Electric vehicles too expensive	Sales of electric vehicles will greatly increase	Electric vehicles are safe	You can travel great distances with electric vehicles	My friends would think I was crazy if I bought an electric vehicles	Electric vehicles are environmentally friendly
Mann-Whitney U	6507,500	7322,000	6828,000	7414,000	6839,500	4783,000
Wilcoxon W	15552,500	16367,000	15873,000	16459,000	15884,500	10999,000
Z	-1,999	-,439	-1,689	-,190	-1,690	-5,621
Asymp. Sig. (2-tailed)	,046	,661	,091	,849	,091	,000

a. Grouping Variable: Gender

The biggest difference is the attitude of men and women to the fact that electric vehicles are environmentally friendly. Women compared to men believe that electric vehicles are significantly more environmentally friendly. With the entire research group reached women in the analysis Mann-Whitney U rank average 142.81, compared to men who have achieved an average 99.09 respectively. These differences were demonstrated at the 1% level of significance. Statistically significant gender differences were also evident in the attitudes of people surveyed for the price of electric vehicles. Men against women, solves significantly more cost electric vehicles. Consider electric for very expensive. Men in analysis Man - Whitney U achieve average ranking 131.37. The average ranking of women reached 116.06.

DISCUSSION

Other results show that 83% of the examined group believes that the price of fuel will increase in future years. We nonetheless considered only 11% of respondents would purchase an electric vehicle. As the reason it can be seen the often quoted fact of the high purchasing price. Respondents probably consider the application of electric vehicles rather for cities where the annual mileage raid is relatively low and higher initial purchase price will be compensated by low operating costs. The most commonly reported information about electric vehicles were statements about their environmental friendliness, low operating costs and low commuting travel. The reason why respondents rather not think about buying an electric vehicle may also be that interested in protecting the environment is not such a significant factor that can outweigh the higher costs. These findings are consistent with the view Thiel et al. (2012).
The aspect of the high price appears as the primary interest of the respondents in this study. In the context of the question, what would have to happen that respondents thought about buying an electric vehicle, the most of answer has the form “that they had to reduce the purchase price of the vehicle”. It is interesting that there was not revealed more significant differences in the investigated areas between groups of men and women. At the same time was also not significantly higher propensity to buy electric vehicles with increasing household income. Economic status of the household does not play a role in thinking about buying an electric vehicle. Unlike in research of Thiel et al. (2012), where was presented strong agreement with the statement that electric vehicles are significantly safer, this was not mentioned by respondents in our study.

CONCLUSION

When thinking about future market of electric vehicles in the Czech Republic, one should see the company in Prague (capital) as a target group rather than individuals or household. Generally the electric vehicles are perceived as costly and that they provide less comfort. As the environmental issues do not play an important role in the Czech Republic, this view will be not changed in near future. As a surprise it can be seen more friendly attitudes towards electric vehicles by women than by men. As the buyers are usually men, women can consider wider context than just selling price as they see it in a bit more hypothetical way. Low willingness for buying is supported by the facts that there are no subsidies in the Czech Republic and that purchasing cost can be compensated by low operating cost only when using such a vehicle as company car (e.g. taxi). As this sample was representative only according to gender and people from big cities with higher education were overrepresented, there is need for further research to uncover more aspect and valid data from the Czech Republic. But above mentioned groups of citizens (high educated people in big cities) should be those who will buy electric vehicles due to their higher salaries and

environmental awareness. So their low interest portend grim times for the electric vehicles sales without further campaigns and state subsidies.

REFERENCES

Bunzeck, I.; Feenstra, C.F.J.; Paukovic, M., (2011), Evaluation of potential users of electric vehicles related to charging – A survey in eight EU countries.

Ernst & Young, (2010), Gauging interest for plug-in hybrid and electric vehicles in selected markets.

European Commission, (2011), Communication – Whitepaper: Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system. COM (2011) 144.

European Commission, (2012a). Communication - VEHICLES 2020, Action Plan for a competitive and sustainable automotive industry in Europe. COM(2012) 636.

European Commission, (2012b). Driving and parking patterns of European vehicle drivers –mobility survey.

European Environment Agency, (2012). Monitoring CO2 emissions from new passenger vehicles in the EU: summary of data for 2011.

EurotaxGlass, (2011), Electrification of the Automotive Industry – The European Consumer's View.

McMullan, G. (2012), Public Awareness of and Attitudes towards Electric Vehicles. Report of the Department for Regional Development (DRD), Belfast.

Thiel, C., Alemanno, A., Svehiclecella, G., Zubaryeva, A., Pasaoglu, G. (2012), Attitude of European vehicle drivers towards electric vehicles: a survey. Reference Report of the Joint Research Centre of the European Commission.

Žák, Z., Zadina, V., Votruba, Z. (2012), Manifest elektromobility v České republice. Národní spolekt pro elektromobilitu a podporu moderních technologií. Praha. Dostupné z http://www.narodni-spolek.cz/manifest-elektromobility

New Features on Noise, Vibration and Harshness Posed by the Next Generation of Vehicles

Nuria Campillo-Davo[a], Miguel Sanchez-Lozano[a], Ramon Peral-Orts[a], Bert Pluymers[b] and Wim Desmet[b]

[a]Department of Mechanical Engineering and Energy
Miguel Hernandez University
03202, Elche, Alicante, SPAIN

[b]Noise and Vibration Research Group
Department of Mechanical Engineering
KU Leuven, BE-3001, Heverlee, BELGIUM

ABSTRACT

Traffic noise is generally considered one of the most important sources of environmental pollution in urban and interurban areas. Each individual vehicle in the traffic flow contributes to the total emission by means of three different sub-sources: noise emitted by engine, exhaust and transmission; aerodynamic noise; and noise radiated during tire/road interaction. Those sub-sources play a different role depending on the characteristics of the traffic, i.e. for low speed traffic – less than 30km/h – mechanical noise coming from engine and drive chain is the prominent, whilst for high speed – above 30km/h – the main contribution is related to tire/road noise. The new generation of hybrid and electric vehicles (HEVs) will contribute to a reduction of noise pollution through reduced use or inexistence of the internal combustion engine. That feature, which is positive for environmental noise control, carries a negative connotation since low noise emissions of HEVs represent an actual danger to pedestrians. Also drivers of HEVs should adapt their driving style and behavior to the new features related to noise, vibration and harshness (NVH) inside the cabin. In order to solve that problem, warning sounds are being designed by HEVs manufacturers, but, it is not actually known how those new sounds will integrate into urban soundscapes. These and many other questions arise related to the increased presence of electric vehicles in city car fleets. The Research Group in Applied Mechanical Engineering of Miguel Hernandez University in Spain is currently coordinating the COST TU1105 Action "NVH analysis techniques for design and optimization of hybrid and electric vehicles", together with the Noise and Vibration Research Group of KU Leuven in Belgium. The aim of the Action is to engage NVH experts from vehicle industry and renowned research groups in the accumulation, development and dissemination of novel analysis techniques to be developed for the next generation of vehicles and their drivers. In this article, new features posed by HEVs are discussed and latest developments on the topic are presented.

Keywords: Electric and Hybrid Vehicles, Noise, Vibration, Driver and Pedestrian Safety

INTRODUCTION

Hybrid Vehicles will play a major role for long time, well beyond 2030. They are a major enabler to reach the CO2 targets, to reduce gas emission in general, to enable good air quality in urban areas and to spare the non-renewable energy resources. Since Hybrid Electric Vehicles (HEVs) do not have the range limitations of Full Electric Vehicles (FEVs, further referred to as Electric Vehicles – EVs) and also not the drawback of emissions like the pure Internal Combustion Engine (ICE) vehicles, they are more in line with the consumer's needs and the driving patterns of

today and of tomorrow. HEVs combine the advantage of two different propulsion systems, the possibility to drive with zero emission and to drive on long distances. EVs will dominate as pure city- and short distance solutions. Hybrids will be the major solution for sustainable mobility, for individual mobility, for goods transport and for public transport, suitable to enter cities as well.

Worldwide demand for HEVs will advance rapidly from 1.6 million units in 2010 to 4.3 in 2015 to nearly double that number by 2020 (The Freedonia Group, 2006). Currently, HEVs have become already an established solution in the US market. The Asia Pacific region, however, remains the hotbed of activity where a major portion of demand stems from Japan. HEVs and EVs are seeing widespread adoption in Europe too, where environmental consciousness has strongly increased over the last decade and strict adherence to EU environmental directives and Copenhagen Protocol requirements has become a requisite. In this context there is an increasing awareness around the issue of the environmental noise. The World Health Organisation is reporting that traffic noise causes every year the loss of at least one million healthy life-years in Europe and responsible for over 20.000 deaths each year via resulting heart problems, reduced sleep quality, etc. The road industry alone cannot entirely solve this problem: road users, public authorities, automotive constructors have their role to play as well. Therefore real and tangible effects in noise reduction could only be achieved by a common effort and a shared responsibility of all these actors.
On the other hand the very low noise levels due to an approaching electric or hybrid car can be a danger to pedestrians and cyclists. Of particular concern are small children and the elderly. Hybrids may generate a sufficient noise level to warn pedestrians under some driving conditions, but then suddenly reduce their level under other conditions, particularly as they slowdown in circumstances where they are more likely to be mixing with pedestrian traffic. Several manufacturers have developed external noise generators for such vehicles but there remain two outstanding issues. First, how a significant warning level can be generated in front of the vehicle, without generating significant environmental noise pollution by radiating high noise levels from the sides and rear of the vehicle. The second issue is concerned with the best character of the warning noise to use. This should be perceived to be coming from a vehicle but not as alarming that it will cause a startle reaction and thus increase the danger.

Concerning the vehicle interior aspects, a relevant current trend is the active control of the overall sound quality in a vehicle, rather than just the sound level. This may become important with the use of hybrid vehicles and vehicles with variable cylinder management, since the changes in the quality of the sound inside these vehicles, as well as the power source changing character can be disconcerting to the driver. The active control of sound quality generally involves a control system that drives the microphone signals inside the car towards a target, or command, signal, rather than just minimizing it. This has been termed “noise equalization”, “sound synthesis”, “active design” and “sound profiling” and can include the use of psychoacoustic models. Recent trends also include the use of active control systems to provide a smoothly changing sound profile with engine speed, but with an emphasis on sporty sound during acceleration, to make the vehicle “fun to drive”. Further development along these lines is also possible by providing an acoustic environment inside the vehicle that encourages the owner to drive in a more fuel-efficient way, for example. There has been some resistance to this trend towards of active control of sound quality in some parts of the automotive industry, where such electronic sound control is seen as ‘cheating’ compared to mechanical re-design. However, as more virtual systems are introduced in vehicles with active braking, stability and steering, and with a younger generation of customers, more used to audio manipulation, these objections are likely to die away.

Given the above NVH issues and the continuous development of next generation HEV and EV powertrains, whereas the currently released vehicles are still more or less prototyped systems build on existing vehicle platforms, the main objective of the TU1105 COST Action “NVH analysis techniques for design and optimization of hybrid and electric vehicles” (TU1105 COST Action website) is to acquire, unify and coordinate necessary information about vehicle dynamics, driveability and noise, vibration & harshness (NVH) analysis technologies in view of the challenges posed by the novel power train concepts (with HEV and EV including all types: series, split and parallel hybrid vehicles and range-extended). The main mean to achieve this objective is the generation of a network in order to form a critical mass of expertise. Although in the past a lot of research effort has been spent on the development of NVH analysis tools, these existing tools are mainly applied to conventional IC vehicles and the limited knowledge on electric and hybrid vehicles is scattered all over Europe. In order to identify shortcomings in current state-of-the-use and to pinpoint areas in which new technologies need to be developed, a review of the existing technologies has to be made. In such a way knowledge fragmentation can be reduced and the efficiency in developing novel methods dedicated to EV and HEV will be increased. TU1105 will provide the necessary coordination in order to set up the development of these novel methods.

This paper reports on the progress of the Action going into its third project year, summarizing the results obtained so far and the work planned for the second period of two years.

THE TU1105 COST ACTION

The TU1105 COST Action "NVH analysis techniques for design and optimization of hybrid and electric vehicles" is hosted in the intergovernmental framework for European Cooperation in Science and Technology (COST). The network started in April 2012 with 13 representatives from 8 countries, and nowadays it is composed by 36 entities from 15 countries (including 2 Non-COST Countries: New Zealand and Brazil) coming from the academia and also the industry. The consortium promotes the transfer of knowledge by means of meetings, training schools, short term scientific missions and dissemination of results at conferences and publications.

The main objective of the Action will be achieved by accomplishing a set of secondary objectives. In this respect, it will be necessary to analyse the current situation and the state-of-art, and to promote the networking among the partners, in order to bring together fragmented NVH knowledge and to adapt NVH technologies from ICE to EV and HEV when possible. It is also essential to collect customer's perception by means of questionnaires and to assess industrial requirements now and over the time, as these groups will be some of the end-users of the Action. Also, it will be required to stimulate users, transport companies and authorities responsible for the development of infrastructures about new existing interactions between EV/EV and the environment, and to bootstrap on existing and stimulate and seed new European initiatives in the field. Finally, the network will disseminate NVH research results towards likely end-users.

The scientific program of the TU1105 COST Action is structured into five Work Groups, which activities are focused on achieving the objectives stated above: WG1: state-of-the-art concerning NVH techniques on ICE vehicles and NVH characteristics of novel HEVs and EVs vehicles; WG2: development of experimental NVH analysis techniques; WG3: development of numerical NVH analysis techniques; WG4: development of sound quality metrics and sound engineering of HEVs and EVs, including safety aspects; WG5: transfer of the knowledge to likely end-users.

In the next chapter are explained the main areas of research and activity within the network that nowadays are being developed. These correspond to the activities within WG2, WG3 and WG4. The activities within WG1 have been already finalized and nowadays WG2 is the responsible of interpreting the data about customers' expectations. Finally, activities related to WG5 will be performed at the end of the Action lifetime.

AREAS OF WORK WITHIN THE TU1105 ACTION

Experimental NVH Analysis Techniques

The main objectives of NVH experimental techniques are to determine the characteristics of the noise and vibration sources, as well as the transmission paths to driver and passengers. In fact, the absence of ICE engines and other mechanical noise issues, change the relative importance of the other noise sources of the EV, so, that is one of the reasons why the identification of noise sources in an EV is so important to assess the interior sound comfort. In effect, although electric vehicles are significantly quieter, their interior noise is marked by high-frequency noise components which can be subjectively perceived as annoying. At the same time, the lack of external noise supposes a real danger for pedestrian, been necessary to use warning sounds that improve saving conditions without affecting interior sound comfort and exterior noise pollution.

Previous characterization of EV's justifies the importance of experimental techniques to analyze and identify NVH sources of Electric Vehicles. Some of the traditional tools are well-know devices to study ICE vehicles and to solve vibro-acoustics issues, it is the case of:

- Order tracking
- Modal analysis
- Noise mapping techniques such as sound intensity
- Near-field acoustic holographic and beamforming

But some of them will not be useful for experimental NVH analysis of the EV's due to the fact that they target

towards the mechanical engine of the vehicle, so they are being currently adapted to the electric vehicle conditions (Finger et al., 2012).

Interior NVH Techniques

A useful technique to study NVH of an EV is the Transfer Path Analysis (TPA) (Sarrazin et al., 2014). TPA models the connection between a noise source and one or more receivers (see Figure 1). The TPA method identifies the amount of excitation by measuring excitation force by introducing a force sensor (matrix and stiffness methods are estimative alternatives).

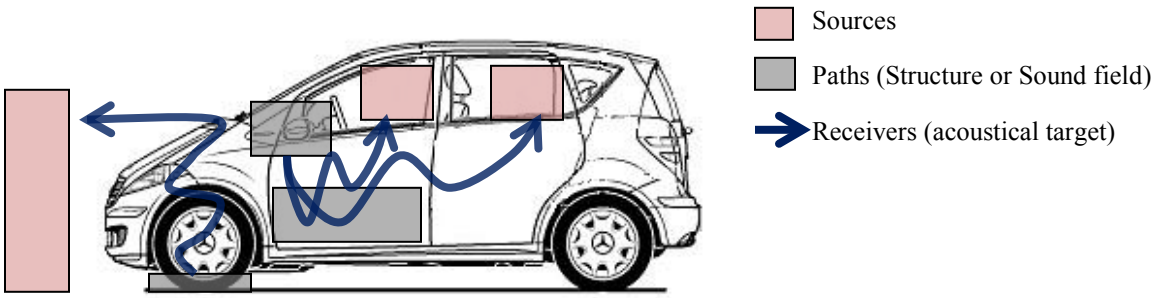


Figure 1. Transfer Path Analysis identifies the effect of NVH source to the potential receivers

Based on experimental dates, a variety of signal processing methods including statistical analysis, spectral analysis, time-frequency analysis and wavelet transform have been also used to separate close occurrences and analyze EV's sources noise and vibration.

Source identification techniques like sound intensity, acoustic beam-forming and acoustic holography are used for a number of reasons outside of troubleshooting including calculating sound power, deriving acoustic impedance, estimating sound radiation and calculating surface velocities for acoustic modelling support. But the primary purpose of noise source identification is to locate where the majority of the noise comes from, being of particular interest in a complex acoustic environment. The location of the source with these techniques is also useful in small electromechanical devices that operate with a cover, where the source is difficult to separate. Batteries with integrated pumps and cooling fans and small gear trains are examples of products where localizing the noise source would be very difficult without these techniques.

Structural Techniques are also used to understand the structure enough to make a design modification. The main methods used to describe the structure motions are operating deflection shapes (ODS) and experimental modal analysis (EMA). ODS can be performed as time, frequency spectra, or be RPM based. EMA is performed with hammer or shaker input. The goal of ODS is to determine the real-world forced deflection during operation through measuring transmissibility between a reference transducer and a group of roving response transducers. On the other hand, the objective of modal analysis is the construction of a mathematical model of the inherent dynamic properties and behavior of the structure through measuring the frequency response function between a force transducer at a driving point and group of roving response transducers (Park and Lee, 2014).

Exterior NVH Techniques

Electric vehicles at low speed are extremely quiet, because of the lack of tire/road noise. Pedestrians use auditory as well as visual cues as warning signals that a vehicle is approaching. Current regulatory requirements aim at limiting the noise emitted by a vehicle in its loudest mode of operation (see pass-by test in ISO 362), and there is still no provision for ensuring that quiet vehicles can be heard by pedestrians.

Some research groups are currently developing sound generator systems, and they are using experimental techniques to evaluate the behavior of the proposed systems and how they are going to affect to the vehicle perception and to the environment.

Other sources are studied using traditional experimental techniques such a pass-by test, drum tests or coast-by test, but alternative methodologies could be able to identify moving noise sources (Peral et al., 2010), the horizontal

directivity of EV's or the power level tire/road noise (Campillo-Davo et al., 2013).

Numerical NVH Analysis Techniques

Working Group 3 focuses on the development of numerical NVH analysis techniques. Numerical methods are very powerful in accurately modeling and predicting vibration and sound radiation from complex structures. The most widely used numerical methods are the Finite Element Method (FEM) and the Boundary Element Method (BEM). Both of them can be used for structural and acoustic analysis but due to certain advantages and disadvantages FEM is mainly used for structural problems whereas BEM is used for acoustical problems. These methods, together with others as multi-body or lumped simplified approaches, have been used extensively in the detailed design of vehicles to achieve optimum designs. The catalogue of commercial software for performing FEM, BEM or multi-body analyses has become very wide.

But electric and hybrid vehicles are characterized by different noise and vibration behavior. In addition, these vehicles can present simplified or unusual driveline configurations, and different vehicle architectures can be found. And finally, some new materials are being used in EV, with different behaviors from the traditional ones used in conventional vehicles equipped with internal combustion engines.

As a consequence, some of the existing numerical methods for NVH analysis may be invalid for those unusual operating conditions, and techniques may need to be adapted. In most cases, the existing methods should be valid, but the results to obtain (frequencies, intensities, etc.) will be in a range of values very different from the usual for ICE vehicles.

According to the Memorandum of Understanding of the Action, the deliverables obtained by this Working Group should be:

- Definition and development of numerical NVH analysis techniques to determine the modes, the characteristic frequencies and the relevant non linear effects of the driveline.
- Definition and development of numerical techniques to simulate the dynamic behavior of new materials used in electric vehicles.
- Integration of simulation and experimental measurements.

To accomplish those targets, a comparison of existing numerical methods will be made for some EV and component examples, and their utility will be verified compared to ICE. The study could cover finite elements models, boundary elements models, multibody and lumped mass models. The different research groups involved will share their methodologies for modeling, and experimental measurements realized by partners acting in WG2 will be used for validation of models.

The comparison between the numerical simulation methodologies will be based on several common benchmarks, with different levels of complexity. For example: one detailed transmission component model, one driveline subsystem model, one entire vehicle model, etc.

Depending on the complexity, models could include the following items:

- Electric machine models
- Gear noise and vibrations models
- Tire/road noise models
- Aerodynamic noise models
- Various damping material model for noise / vibration attenuation
- Interior noise propagation models
- Exterior noise propagation models

As a result of WG3 activities, a comparison of the utility and scope of various numerical methods will be obtained, and particulars of their application to the NVH analysis of electric and hybrid vehicles will be discussed. Some conclusions could also be obtained related to general results expected for EV noise and vibration, and sensitivity to different parameters and conditions.

Sound Quality Metrics for User's Perception

The noise signature of a traditional ICE vehicle presents quite differences from the noise footprint of an EV. As commented previously, beyond the reduction of noise provided by the electric engine, some other systems become more relevant in acoustical terms, as now those sources are not masked by the noise of the combustion engine. Such is the case of the whining noise coming from the gearbox, or the tire/road noise and wind noise. Also new sources appear, as possible high frequency tonal electric motor noise and the noise generated by complex transmission systems. Finally, warning sounds to increase pedestrians' safety are being developed to be implemented in EVs.

The features exposed above will condition the response of drivers and pedestrians facing an EV, both in terms of interacting with the vehicle and also the level of accomplishment reached related to the customers' expectations. The development and validation of sound quality metrics will allow to indicate which experimental and numerical techniques can give the best understanding of the NVH issues of EVs.

Regarding to the sound quality in the interior of the cabin, some studies on sound quality metrics applied to H-EVs sounds are found in the literature (Szatkowski, 2013). Those metrics can be improved by means of listening experiments using measured and modified interior noise recordings, in both conditions: stationary constant speed and acceleration. The comparison of both metrics results shall permit to propose a series of guidelines to ensure a high sound quality inside the cabin of an EV.

Concerning the exterior vehicle noise, and specifically in relation to warning sounds, previous work can also be found in the literature, for instance the eVADER project (eVADER Project) where it was analyzed the timbre features on detectability of EVs. However, further work must be developed such as a comparison between existing sounds and those recommended for instance by European regulations.

CONCLUSIONS

The TU1105 COST Action is running until April 2016 and the expected potential impacts will be based on achieving a uniform point of view of the new features on NVH posed by the H-EVs. In this sense, the completion of the Action will imply to obtain a clear understanding of NVH features of H-EVs and of involved subsystems, which will allow to propose the standardization of test rigs and experimental procedures to quantify external/internal H-EVs noise. Besides, the development of dedicated analysis numerical and experimental tools for H-EV NVH analysis will provide a collection of best practices for the research on this field. Finally, the development of common NVH metrics, in the form of guidelines, to characterize noise annoyance acceptance, will enable to analyze the customers' expectations taking into consideration the ongoing changes on the NVH features that may occur thanks to further research and innovation in the field of EVs.

The end-users of the Action are all those people concerned in the field of NHV of EVs, as experienced and early-stage researchers, European authorities for transport regulations, independent consultants, experienced representatives from industry and relevant associations. But also vehicle customers, pedestrians, cyclists, and indirectly the whole European society will benefit of the Action achievements.

REFERENCES

Campillo-Davo, N. Peral-Orts, R. Velasco-Sanchez, E. Campello-Vicente, H. (2013). "An experimental procedure to obtain sound power level of tyre/road noise under Coast-By conditions". Applied Acoustics 05/2013; 74(5):718–727.

eVADER Project. http://www.evader-project.eu/

Finger, K. Brandes, M. Schmidt, E. (2012). "TPA and NVH Prognosis – Application to Mercedes Benz Car Development of New Hybrid Methods Coupling Digital Simulation with Prototype Testing Results". SAE International Journal of Passenger

Cars - Mechanical Systems. 5(3):1091-1100, 2012.
Park, J. Lee, H. (2014). “Study on the Vehicle Cabin Noise Employing the Interfacial Friction in Double Layered Frames Used in Electric Vehicle Traction Motors”. SAE International Journal of Passenger Cars – Electronic and Electrical Systems. 7(1):2014.
Peral-Orts, R. Velasco-Sanchez, E. Campillo-Davo, N. Campello-Vicente, H. (2010). “Noise source detection system based on two independent microphone arrays”. INTERNOISE 2010. Lisbon, June 2010.
Sarrazin, M.M. et al. (2014). “Electro-Vibro-Acoustic Analysis of Electric Powertrain Systems”. FISITA 2014 World Automotive Congress.
Szatkowski, A. (2013). “Passenger Vehicle Hybrid Hydraulic Powertrain Sound Quality Investigation". SAE Technical Paper 2013-01-2004, 2013.
The Freedonia Group. (2006). “World Hybrid-Electric Vehicle”, Oct. 2006.
TU1105 COST Action. http://www.tu1105.ulg.ac.be/index.html

Pushing the Limits? Regenerative Braking in Takeover Situations

Claudia Witzlack, Peter Cocron and Josef F. Krems

Chemnitz University of Technology
Department of Psychology
Cognitive and Engineering Psychology
Chemnitz, Germany

ABSTRACT

The aim of the study was to investigate if individual experience in driving an electric vehicle (EV) influences the interaction with regenerative braking. Therefore, a field experiment was developed to explore driver behavior in takeover situations. Here regenerative braking reaches its limits, so drivers must use the conventional brake to come to a stop in a safety manner. Furthermore, it was tested, if trust towards regenerative braking changes as experience in driving an electric vehicle increases. Especially the changes in trust when regenerative braking meets its limits were investigated. Results showed that drivers with little EV-experience come closer to a stop in front of stop line and overshot this stop line more often compared to inexperienced drivers. Trust towards regenerative braking increased slightly after the subjects without driving experience gained experience in driving an electric vehicle und decreased after takeover situations. Altogether subjects with little EV-experience showed a higher trust in regenerative braking than subjects without EV driving experience. Finally, the impact of trust on driver behavior is discussed as a possible explanation for our findings.

Keywords: Regenerative braking, trust, takeover situation, human-machine interaction

INTRODUCTION

In times of shrinking fossil fuels one of the biggest challenges of our age is the extraction of renewable energy sources to control the emission of carbonic acid gas. One of the leading approaches to reduce carbonic dioxide emission in traffic is the electrification of power trains. Electric vehicles (EV) are typically equipped with regenerative braking systems (RB), which are focused on the present study. Those braking systems recapture the kinetic energy which is released while the EV decelerates and feed the battery of the car with electric energy.

During the last years extensive field studies that investigated human-machine interaction and adaptation processes in electric mobility (Ramsbrock, Vilimek, & Weber, 2013). Concerning RB, studies revealed that subjects appreciate RB to be a very comfortable feature of EV's (Krems, Weinmann, Weber, Westermann, & Albayrak, 2013; Turrentine et al., 2011). Nevertheless subjects reported decelerating with RB to be the biggest adjustment compared to other EV features (Krems et al., 2013). Objective data reported by Cocron, Bühler, Franke, Neumann, Dielmann and Krems (2013) revealed that drivers can quickly adapt to RB system which is triggered via accelerator. The authors found out that it takes about several hours up to one day until subjects mostly use RB. The portion of conventional braking maneuvers decreased mostly at the beginning of the EV usage followed by reaching a kind of constant level in dependence of driving duration. The process of learning to use an accelerator triggered RB system seems to be an important factor to transfer the system's usage to different traffic situations and to integrate it into every-day driving behavior (Cocron et al., 2013; Turrentine et al., 2011).

The degree to which people adopt new technologies like RB and are willing to use or rely upon them depends on psychological factors like acceptance and trust (Jian, Bisantz, Drury & Llinas, 1998; Sheridan, 1988). Muir and

Moray also supported this assumption and argued that the tendency of using a system depends on trust. If there is trust in a system, the probability to use it increases and vice versa. Those findings were mostly situated in the domain of automated or partly-automated systems but here we tried to transfer them to RB. Furthermore, Lees and Lee (2007) stated that trust is an important factor to understand any human-machine interaction. If new technology is brought to the human-machine interaction, the behavior of the human operator will change. For example, the lack of control over a certain event or action in an automated system is usually compensated by trust (Cahour & Forzy, 2009). In conclusion we assumed trust to be a decisive variable in human-machine interaction that is used by the human operator to decide if he delegates certain tasks to a system or not (Cocron et al., 2013).

The question that arises here is how do people set their level of trust? Can people trust too much in a system and what happens then? To describe this problem Parasuraman and Wickens (2008) introduced the constructs of overreliance or overtrust, which means an inappropriate level of trust towards a technological system. There are several studies which investigated the consequences of overreliance. Sparaco (1995) found that aviators of an Airbus relied too much on autopilot which was followed by an air crash. Another study in the field of shipping (Lee & Sanquist, 2000) showed that an inappropriate level of trust in a navigation system lead to a misrouting of the ship. In that sense a disproportion of trust can cause very serious consequences. Parasuraman and Riley (2007) described the disproportion of trust in human-machine interaction as disuse (level of trust is too low so people do not use the system) or misuse (inappropriate level of trust so people use the system unintentionally in a false manner). To sum up, trust is the variable which decides whether people misuse or disuse a certain system. Furthermore, there is a common ground because the pilots or captains missed the point to take back control from the system. Here we define takeover situations in RB systems, when the driver is forced to take back the control from the system, which means that he has to utilize the friction brake to decelerate.

A System which is often investigated in relation to takeover situations and trust is the Adaptive Cruise Control System (ACC). Rajanoah, Anceaux and Vienne (2006) investigated the ACC in takeover situations and also measured the level of trust. They found out that subjects who showed a high level of trust used the system more often (Jian et al., 1998; Lewandowsky, Mundy & Tan, 2000; Muir & Moray, 1996; Sheridan, 1988) than subjects with lower levels of trust. Those subjects who relied too much on the ACC used the conventional brake later in the deceleration process than subjects with lower levels of trust. As mentioned above, trust regulates the manner people use certain technologies and also the actions of a system can regulate the level of trust. Either the system works properly and therefore confirms the trust of the human operator or -in case of malfunction- trust decreases (Kantowitz, Hanowski & Kantowitz, 1997; Lee & See, 2004).

To sum up, the aim of the study was to investigate the behavior of drivers while using RB and experiencing takeover situations. Here we define takeover situations as a possible traffic incident, where RB reaches its limits so the driver is forced to use the conventional brake to come to a stop in a safety manner. Furthermore we wanted to compare different levels of trust in RB. We developed a deceleration task in which drivers had to come to a stop by only using the deceleration of RB. We tested inexperienced and drivers who were familiar with the test EV and assumed that drivers with EV experience come to a stop more closely in front of a stop line. Furthermore we expected that they would overshoot the stop line more often due to their higher level of trust. In addition we suppose that reaching the limits of RB should cause a bigger decrease in trust.

METHODS

Participants

In total, 29 subjects participated in this study, which differed in their EV experience. The group of subjects ($N = 14$) which was familiar with the EV drove the test vehicle (MINI E) for 24 hours before (in further paragraphs they are called *experienced* drivers). In that period drivers covered a mean distance of 165 kilometers (SD = 107.5 kilometers). The group of inexperienced drivers (N = 15) had no experience in driving an EV at all (in further paragraphs they are called *inexperienced* drivers). Both groups where matched using certain socio-demographic variables like age, gender or driven kilometers in the past year. The whole group of subjects consisted of 8 females and 21 males and had a mean age of 37.52 years (SD = 10.09 years).

Test Vehicle

The test vehicle was a MINI E, which is a two-seated EV with a bttery pack in the back of the vehicle. The vehicle had an engine power of 150 kW, a torque of 220 Nm, a top speed of 152 km/h and a maximum range of 250 km (under controlled conditions). The RB system was triggered by the accelerator pedal. As soon as the driver released the gas pedal, there was a noticeable deceleration. The RB system had a deceleration rate of -2.25 m/s² (Eberl, Sharma, Stroph, Schumann & Pruckner, 2012).

Furthermore, there was a notebook integrated inside the vehicle, which was placed on the right side of the car dashboard to show a stop signal.

Test Area

We used a rectangular sealed area (30 m x 50 m) to conduct the experiment (see figure 1). Three stop lines were marked in each direction. At the rightmost position of each linear slope there were yellow and red pylons in different positions, which indicated the investigator to give the stop signal.

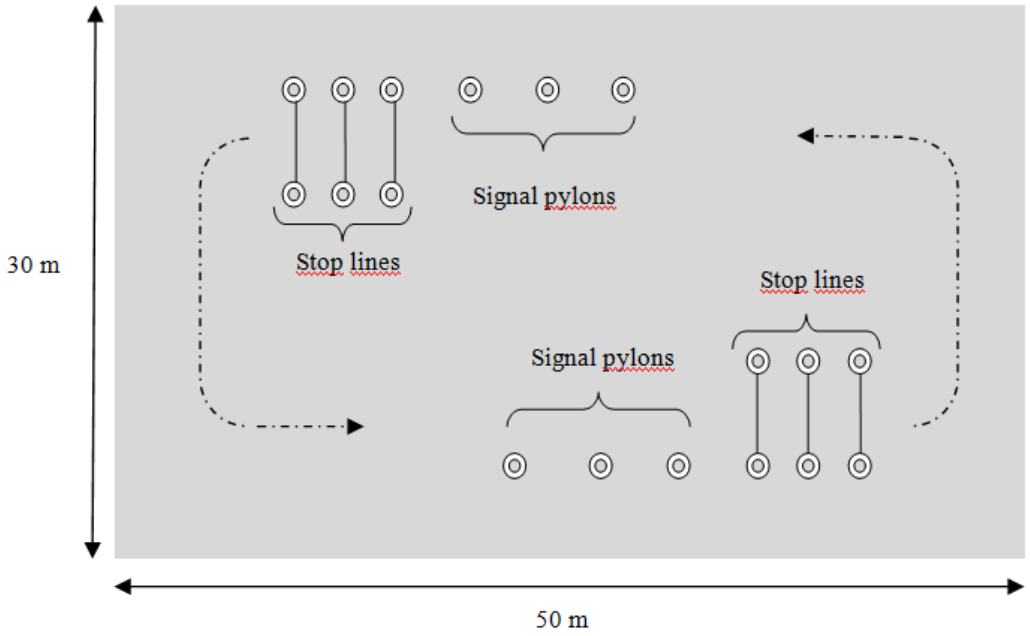


Figure 1. Test area in view from above (not true to scale)

Procedure

The whole experiment consisted of two parts. The aim of the first part was the investigation of learning processes while using the RB. Details are reported in Witzlack, Cocron and Krems (2013).

The second part of the experiment is reported here. Subjects were instructed to drive through the circuit and accelerate at the linear slopes up to 30 km/h respectively 40 km/h. Furthermore they were told that there will be a signal via the display of the notebook which tells them the number of the stop line at which they must come to a stop. Here subjects had to decide by themselves whether to use the conventional brake or RB. In total subjects drove nine rounds and had to stop three times per speed level.

The stop line at which they had to come to a stop, the number of round and the side of the course were randomly chosen. The investigator gave the signal to stop when the vehicle passed a specific pylon. The aim of this experimental setting was to provoke a takeover situation where the deceleration of RB reached its limits. When the signal indicated to come to a stop it was not possible to stop in front of the stop line without using the conventional brake. Therefore participants had to take over control and brake conventionally. This situation reflects an incident in traffic, when an unexpected event (e.g. the car in front of the own vehicle decelerates very quickly without a visible

reason) occurs and one has to stop very quickly.

When the subject came to a stop the experimenter measured the distance between stop line and the front bumper of the car. After passing the three braking maneuvers per speed level, subjects had to rate their trust based on items by Cocron et al. (2013).

RESULTS

To investigate if there is any different behavior between experienced and inexperienced drivers in takeover maneuvers, we compared the distances in which they came to stop in front of a stop line. We assumed that experienced drivers came more closely to a stop in front of the stop line than inexperienced drivers. Therefore we compared the different speed levels (30 or 40 km/h) and each takeover maneuver (see figure 2 and figure 3). In this part of the statistical analysis we only focused on those drivers who came to a stop in front of the stop line, so we had to compare groups of widely different size. Therefore we used several t-Tests, but corrected the accumulation of the alpha error using the method of Bonferroni-Holm (Holm, 1979).

At 30 km/h, results pointed in the assumed direction but revealed no significant differences. In the third takeover situation inexperienced drivers came marginally closer to a stop in front of the stop line than experienced drivers ($Mean_{inexperienced\ drivers}$ = -91.64 cm vs. $Mean_{experienced\ drivers}$ = -92.80 cm). Regarding the speed level of 40 km/h (see figure 3), results support the assumption mentioned above. Nevertheless, there are no significant differences between the distances of inexperienced and experienced drivers (see table 1; none of the calculated p-values falls below the corrected alpha-value of .017).

Table 1: Statistical analysis of mean distance in which inexperienced and experienced drivers stopped in front of a stop line

	30 km/h speed level			40 km/h speed level		
	1st takeover situation	2nd takeover situation	3rd takeover situation	1st takeover situation	2nd takeover situation	3rd takeover situation
t	-1.585	-1.512	.550	-1.869	-.554	-1.841
df	18	22	22	24	20	24
p	.065	.042	.037	.478	.293	.039

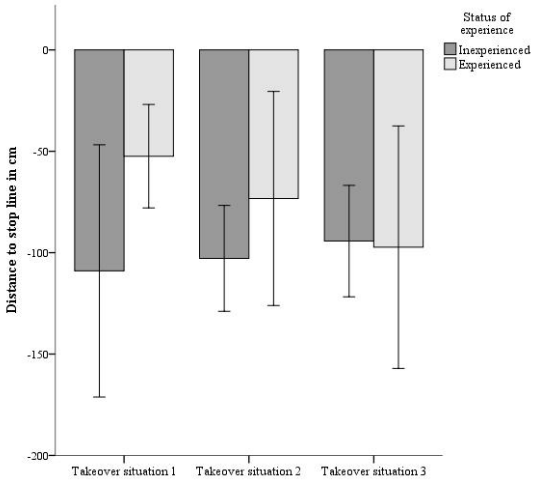


Figure 2.Comparison of distances in front of the stop line concerning speed level of 30 km/h (mean, confidence interval)

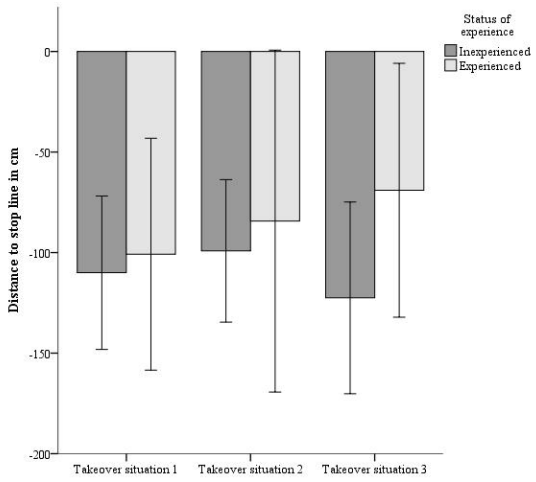


Figure 3.Comparison of distances in front of the stop line concerning speed level of 40 km/h (mean, confidence interval)

Another way to investigate differences in the braking behavior of inexperienced and experienced drivers concerning takeover situation is to analyze the amount of subjects who overshot the stop line. Here, we supposed that the experienced drivers would overshoot the stop line more often than inexperienced drivers. The results supported this assumption, because in each takeover maneuver irrespective of the speed level there were more experienced drivers who overshot the stop line (see table 1). Still, we couldn't find any significant differences in the frequency of inexperienced and experienced drivers overshooting the stop line (see table 2).

Table 2: Comparison of the amount of drivers who overshot the stop line and results of significance analysis

	30 km/h speed level			40 km/h speed level		
	1st takeover situation	2nd takeover situation	3rd takeover situation	1st takeover situation	2nd takeover situation	3rd takeover situation
Inexperienced drivers ($N = 15$)	4	1	1	0	1	0
Experienced drivers ($N = 15$)	4	3	3	2	6	2
p (Exact Fisher Test)*	1.000	.311	.311	.206	.029	.206

*footnote: We report the exact fisher test, because there often were often values below 5 in the cells.

Additionally, we investigated trust and to which extend trust was adjusted in takeover situations. For a better understanding we also report the level of trust in pre-test measurement and for all learning points of measurement

(see figure 3). We assumed that experienced drivers show a higher trust in RB at every point of measurement than inexperienced. Results support this assumption (see figure 3). Here the statistical analysis showed a significant influence concerning the experience $F(1,22) = 6.202$, $p = .011$ and a very big effect size $\eta_p^2 = .220$. There were neither significant influence in relation to the different experimental settings (learning vs. takeover situation) $F(3.392, 74.615) = 2.027$, $p = .055$ nor for the interaction of experience and experimental setting $F(3.392) = .998$, $p = .203$.

Furthermore, we supposed a decrease in trust when the RB reached its limits. Notably inexperienced drivers should show a higher decrease in trust compared to experienced drivers when they passed the takeover situations (see in figure 3). Here we compared the points of measurement after finishing the experimental part of learning with 40 km/h and the first part of takeover situations with 30 km/h. We found a significant influence of experience $F(1, 24) = 1.694$, $p = .043$ and a big effect size $\eta_p^2 = .125$. There were no significant results concerning the point of measurement $F(1, 24) = 1.694$, $p = .103$ or the interaction of experience and point of measurement $F(1) = .152$, $p = .350$.

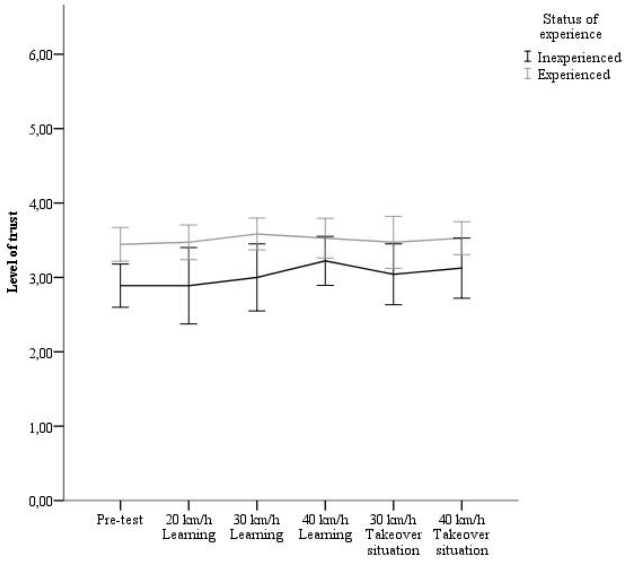


Figure 3. Comparison of trust in regenerative braking several points of measurement (mean, confidence interval)

DISCUSSION

The aim of the study was to investigate the braking behavior of inexperienced and experienced drivers in takeover situations. In our experiment inexperienced and experienced drivers went through a learning procedure where especially the inexperienced drivers had the chance to gain some practical experience in using RB. In the takeover situations subjects had to stop in front of a stop line deciding whether and when to step on the conventional brake. We assumed that inexperience drivers used the conventional brake earlier in the deceleration process and came to a stop earlier than experienced drivers. As a consequence, inexperienced drivers were expected to overshoot the stop line less often than experienced drivers. Results showed tendencies in direction of our hypothesis, but there were no significant differences. That means inexperienced drivers tended to come to a stop further in front of the stop line and tended to overshoot the stop line less often than experience drivers. One could argue that the level of trust the drivers developed towards RB might be a possible reason for our findings. (Lee & Sanquist, 2000; Sparaco, 1995). Inexperienced drivers had a lower level of trust in the RB system than experienced drivers. Especially in situations where RB reaches its limits, trust seems to decrease so inexperienced drivers might switch more often and earlier to systems they already know and perhaps trust in like the conventional brake. The other way to understand this explanation means that experienced drivers, who already know and eventually trust the RB system, might use the system more often and until it reaches its limits, so they also might overshoot the stop line more often. Another explanation for our findings might be the small sample seize and our definition of experience. Perhaps, our definition was not as selective as it should have been. That means we eventually compared groups of drivers who differed that much in their level of EV-experience or 24h EV-experience might not be enough to produce distinct

results in our study.

One can associate these results with studies of Parasuraman and Riley (2007). They described the mismatch in human-machine interaction as misuse or disuse, which is determined by the level of trust. Do people trust too much in a system, they might misuse it (or in this case maybe overshoot the stop line). A lower level of trust should lead more often to disusing a new system rather using an already known system (or in this case not using the conventional brake instead of RB).

Trust seems to be an influencing factor on how people use technologies (Lee & See, 2004). That is why we assessed the level of trust towards RB in both groups of drivers. As we mentioned above trust exceeds when people gain experience in using a certain system. Therefore we assumed that experienced drivers would show a higher level of trust irrespectively of the experimental setting. Furthermore we supposed that passing the takeover situation would decrease the level of trust, especially in the group of inexperienced drivers. The results supported our hypothesis and showed the biggest impact on the differences in trust and the decrease of trust for both experimental groups to be the experience. More experience with a certain technological system can lead to a higher level of trust and also to overreliance, which can cause serious problems, like accidents (Lee & Sanquist, 2000; Sparaco, 1995). Based on studies of Muir and Moray (1996), we supposed a relation between trust and the usage of a system like RB and found supporting results. In total, our results could be associated to studies, which postulate that the level of trust determines on the hand the acceptance of a certain system and on the other hand the extent to which people rely on a system (Jian, Bisantz, Drury & Llinas, 1998; Sheridan, 1988) and finally use a system. Considering the decrease of trust after experiencing the first takeover situations shows that the level of trust is adaptable towards actions of a system. If the systems does not work as excepted or like in this study it reaches its limits, people decrease their level of trust which supports results from Kantowitz, Hanowski and Kantowitz (1997) or Lee and See (2004). Following these thoughts the relation between trust and the usage of a certain system shows a kind of two-way influencing character. Trust determines the extent to which people use a system, but the other way round the usage of the system can adjust the level of trust. Nevertheless, to find more solid results that support a relation between trust and the usage of RB objective data like data loggers should be taken into consideration.

CONCLUSIONS

Regenerative braking as tested in the current study seems to be a trustworthy system, because all of our subjects showed levels of trust around the mean or higher values regarding the used scale for measurement. The minority of subjects overshot the stop line, so one could reason that the used takeover maneuvers were manageable, but still worked to provoke erratic behavior. However there is a real traffic meaning of our results. Takeover situations can happen for example when driver use the RB system more often than conventional braking and experience a situation similar to those we provoked in our study. Consequences might be heavy traffic accidents. Another possible situation might be that drivers use their conventional car which has no RB system, but maybe they rely upon this assumption (Cocron et al., 2013). We also found that experience seems to influence the level of trust towards RB and also a two-way dependence between trust and the usage of the system. Nevertheless both groups do not show clearly distinguishing results neither regarding their braking performance nor their levels of trust, which may have caused only occasional emerging significant results. With regard to future studies, it might be a solution to adjust the definition of experience in terms of e.g. more driven kilometers.

ACKNOWLEDGEMENTS

This study was funded by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety. Any views expressed herein are those of the authors and do not necessarily reflect those of the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety or of other partners involved in the project.We are grateful for the support of our consortium partners, Vattenfall Europe AG (Dr C.F. Eckhardt and F. Schuth) and the BMW Group (G. Schmidt, Dr M. Hajesch, Dr A. Keinath and Dr M. Schwalm) which made our research possible.

REFERENCES

Cahour, B. & Forzy, J.-F. (2009). Does projection into use improve trust and exploration? An example with a cruise control system. *Safety Science, 47*(9), 1260-1270.

Caroll, S., & Walsh, C. (2010). *The smart move trial – Description and initial results*. Research Report. Retrieved from: http://www.cenex.co.uk/wp-content/uploads/2013/06/2010-03-23-Smart-move-trial-report-v3-Compatibility-mode-11.pdf

Cocron, P., Bühler, F., Franke, T., Neumann, I., Dielmann, B., & Krems, J. F. (2013). Energy recapture through deceleration – regenerative braking in electric vehicles from a user perspective. *Ergonomics 56*(8), 1203-1215.

Cocron, P., Franke, T., Neumann, I., Wege, C., Bühler, F., & Krems, J.F. (2010). Ist das Fahren mit einem Elektrofahrzeug so besonders? Anpassung des Verhaltens beim Fahren mit Elektrofahrzeugen [What is so special about driving an electric vehicle? Behavioral adaptation while driving an electric vehicle]. In C. Frings, A. Mecklinger, D. Wentura & H. Zimmer (Hrsg.), *Beitrage zur 52 . Tagung experimentell arbeitender Psychologen* (S. 28). Lengerich: Pabst Science Publishers.

Eberl, T., Sharma, R,, Stroph, R., Schumann, J. & Pruckner, A. (2012). Evaluation of interaction concepts for the longitudinal dynamics of electric vehicles. In N.A. Stanton (Ed.): *Advances in Human Aspects of Road and Rail Transportation* (pp. 263-173). Boca Raton, Florida: CRC Press.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65–70.

Jian, J.-Y., Bisantz, A.M., Drury, C.G., & Llinas, J. (1998). *Foundations for an empirically determined scale of trust in automated systems*. Research Report, AFRL-HE-WP-TR-2000-0102. Buffalo: United States Airforce Research Laboratory.

Kantowitz, B. H., Hanowski, R. J., & Kantowitz, S. C. (1997). Driver acceptance of unreliable traffic information in familiar and unfamiliar settings. *Human Factors, 39,* 164–176.

Krems, J. F., Weinmann, O., Weber, J., Westermann, D., & Albayrak, S. (Hrsg.) (2013). *Elektromobilität in Metropolregionen: Die Feldstudie MINI E Berlin powered by Vattenfall. Fortschritt-Berichte VDI Reihe 12 Nr. 766 [Electric mobility in model regions: The field study MINI E Berlin powered by Vattenfall]*. Düsseldorf: VDI Verlag.

Lee, J. D., & Sanquist, T. F. (2000). Augmenting the operator function model with cognitive operations: Assessing the cognitive demands of technological innovation in ship navigation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 30,* 273–285.

Lee, J.D., & See, K.A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46,* 50-80.

Lees, M.N., & Lee, J.D. (2007). The influence of distraction and driving context on driver response to imperfect collision warning systems. *Ergonomics, 50,* 1264-1286.

Lewandowsky, S., Mundy, M., & Tan, G.P.A. (2000). The dynamics of trust: Comparing humans to automation. *Journal of Experimental Psychology, 6*(2), 104-123.

Muir, B. M., & Moray, N. (1996). Trust in automation: 2. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics, 39,* 429–460.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39,* 230–253.

Parasuraman, R., & Wickens, C.D. (2008). Humans: Still vital after all these years of automation. *Human Factors, 50,* 511-520.

Rajaonah, B., Anceaux, F., & Vienne, F. (2006). Trust and the use of adaptive cruise control: a study of a cut-in situation. *Cognition, Work & Technology, 8,* 146-155.

Ramsbrock, Vilimek, & Weber, 2013

Sheridan, T.B. (1988). Trustworthiness of command and control systems. *IFAC Man-Machine Systems,* 427-431.

Sparaco, P. (1995). Airbus seeks to keep pilot, new technology in harmony. *Aviation Week and Space Technology, 62*–63.

Turrentine, T., Garas, D., Lentz, A., & Woodjack, J. (2011). *The UC Davis MINI E Consumer Study*. Research Report, UCD-ITS-RR-11-05. Retrieved from: http://pubs.its.ucdavis.edu/publication_detail.php?id=1470

Witzlack, C., Cocron, P., & Krems, J. F. (2013). Verzögern mit Rekuperation - Lernprozesse aus der Nutzerperspektive [Deceleration using regenerative braking – learning processes from the user point of view]. In VDI (Hrsg.). *Der Fahrer im 21. Jahrhundert. VDI-Berichte 2205* (p. 107-118). Düsseldorf: VDI Verlag.

Examining User-Range Interaction in Battery Electric Vehicles – a Field Study Approach

Thomas Franke[a], Madlen Günther[a], Maria Trantow[a], Josef F. Krems[a], Roman Vilimek[b] and Andreas Keinath[b]

[a] Department of Psychology
Technische Universität Chemnitz
Chemnitz, D-09107, Germany

[b] Concept Quality and Usability
BMW Group
Munich, D-80788, Germany

ABSTRACT

In order to provide users with an optimal range-related user experience, it is essential to develop a comprehensive understanding of everyday user-range interaction. For this to happen, it is important to sample a high amount of situations where battery-electric vehicle (BEV) users must actively interact with the range resources of their BEV. The present contribution presents the methodology (i.e., a toolbox to study user-range interaction) and preliminary results of a field study that was designed to reach this objective: The BMW ActiveE long-distance commuter field trial in the area of Leipzig, Germany. Within this study, a sample of 75 customers was given the opportunity to drive a BEV for three months. Applicants were recruited via an online screening questionnaire. In order to be selected, participants had to drive at least 90 km per day. A comprehensive data set was generated based on qualitative interviews, questionnaires, diary methods, and data loggers. Relevant constructs were identified and translated into items/scales or specific score definitions. Initial results show that even users who have high daily mobility needs can cope with the limited range of a BEV. However, results also point to the potential of strategies aimed at supporting users in adapting to BEV range to reach an optimal user experience even under conditions of high range demand.

Keywords: battery electric vehicles, field study, user experience, behavior, range, methodology

INTRODUCTION

Battery electric vehicles (BEVs) are a promising form of sustainable mobility because of their potential to reduce CO_2 emissions and air pollution (Holdway, Williams, Inderwildi, & King, 2010), as well as their potential to mitigate risks associated with peak oil (Hirsch, Bezdek, & Wendling, 2005). However, limited range can be a challenge for BEV users. Although substantial improvements in electric vehicle battery performance are likely within the coming years, smaller but suitable battery sizes will always constitute a more sustainable battery layout because battery size is strongly related to the ecological footprint of an electric vehicle (Hawkins, Gausen, & Strømman, 2012; McManus, 2012). Cost-effectiveness of a battery also plays a major role in buying decisions. Even with declining battery costs, it is likely that a large share of customers will not automatically opt for the largest possible battery, but will instead compromise between their daily driving needs and a suitable battery size. As a consequence, the following question will remain relevant in research and development: How can the maximum usable range for the customer be achieved given a certain nominal battery capacity?

Research has shown that it is challenging for users to interact with limited battery resources in an optimal way. For example, it has been shown that users tend to avoid situations that would lead to range stress (i.e., range anxiety) by reserving substantial range safety buffers (i.e., around 20% of real available range; Franke, Neumann, Bühler, Cocron, & Krems, 2012c; Franke & Krems, 2013a; Caroll, 2010). Their limited range comfort zone (i.e., comfortable range; Franke & Krems, 2013a) results in the usage pattern in which a certain share of the battery capacity is lost as a psychological safety buffer.

Obviously, it is difficult to study critical range situations systematically in everyday situations. Not only because users seek strategies to avoid these situations as described above, but also because most potential customers would simply not buy a vehicle that only marginally fits their mobility needs in the first place. Isolated laboratory experimental methods are also not suitable as user-range interaction takes substantial time to stabilize (Franke, Cocron, Bühler, Neumann, & Krems, 2012b; Pichelmann, Franke, & Krems, 2013; Burgess et al., 2013). Therefore, a field trial was established that recruited long-distance commuters interested in electric mobility.

The objective of the present contribution is to explicate key elements of the methodology developed to study user-range interaction (i.e., the interaction with limited resources in human-machine systems; Franke, 2014). The aim is to provide researchers with a toolbox for future studies in this area in terms of constructs and operationalizations. To this end, the range-related aspects of the methodology developed within the project "BMW ActiveE Leipzig – long-distance commuters" are described and discussed in the following sections along with a short description of preliminary results.

FIELD STUDY RESEARCH DESIGN

The project "BMW ActiveE Leipzig – long-distance commuters" was a field trial in the region around Leipzig, Germany. This field trial was set up by a consortium consisting of the BMW Group, the Stadtwerke Leipzig (SWL, Leipzig municipal utilities), and Technische Universität Chemnitz. It was funded by the German Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety (BMUB) as part of the e-mobility showcase region Bavaria-Saxony ("Electric Mobility Connects"). Major objectives of the study were to examine: (1) acceptance of having to charging once or even more than once a day (e.g., at work and at home), (2) interaction with range and range stress (i.e., range anxiety) in everyday BEV use, (3) adaptation and learning effects in driving a BEV (i.e., related to range management), (4) eco-driving and additional strategies to cope with range, (5) need for support and potential for optimizing user-system interaction, (6) user acceptance and behavior regarding different layouts of the regenerative braking systems (research topic is separate from user-range interaction, therefore it is not discussed in this contribution), and (7) phase effects (i.e., grid interferences) and charging curves (technical research topic of the SWL, therefore this area is also not documented).

The field trial was designed as a longitudinal study that involved three main points of data collection (see Table 1): Before users received their BEV (T0) including a short additional data collection event after users had gained their first experience with the BEV after approximately one week of BEV use (T0+1), after six weeks of BEV use (T1) and at the end of the trial after twelve weeks (T2). At each point of data collection, users completed a 2- to 3-hour face-to-face interview including questionnaires. Additionally, participants filled out diaries and data loggers automatically recorded several parameters. The general study design was based on structural elements that had already proven fruitful in earlier BEV field trials (Cocron et al., 2011; Franke, Bühler, Neumann, Cocron, & Krems, 2012a). To yield a high sample size with the given number of available cars (15), this study design was repeated five times (i.e., usage phases 1 to 5) resulting in a total number of 75 customers who got the chance to drive the BEV for three months.

The field trial used a person-based main user data collection approach, meaning that only data from the (prospective) primary (i.e., main) user of the BEV was collected and analyzed. The main user of the BEV was defined within the screening process based on an estimated usage share (i.e., primary drivers' share of BEV vehicle miles travelled relative to the total miles driven by all drivers of the BEV) of greater than 70 percent. In terms of the logger data, personalized car keys were used to filter out data segments that were not generated by the main (i.e., primary) user of the BEV.

Several behavior-related variables (e.g., parameters of charging behavior, trip distances and other mobility parameters) were assessed with multiple methods (e.g., diaries and data loggers) for certain periods. This redundancy allowed for a comprehensive data triangulation (e.g., identification and quantification of weaknesses/biases of different methods), as well as data fusion (i.e., reduction of specific method biases through integrated multi-method assessment). As a result, for each variable the most accurate method(s) could be determined

and/or several data sources could be combined to yield a more accurate estimate. Moreover, this aspect of the research design made it possible to derive data quality indicators (e.g., regarding participants' compliance or precision of subjective diaries).

Table 1: Sequence of data collection events within the user study.

Timeline (weeks)	Data collection event	Acronym
–12	*screening finalized*	
–3	first telephone interview	T0–3
–2–0	baseline trip diary (1 week) & baseline GPS logging (combustion vehicles)	T0
0	face-to-face appointment: questionnaire, interview, test drive *vehicle handover*	T0
1	questionnaire & telephone interview	T0+1
5–6	travel diary (1 week)	T1
6	face-to-face appointment: questionnaire, interview, test drive	T1
10–12	range and recharging diary (10 days)	T2
12	face-to-face appointment: questionnaire, interview, test drive *vehicle return*	T2
0–12	logger data collection (GPS & car data)	-
0–12	car diary (event-based diary)	-

RECTRUITMENT AND SCREENING OF PARTICIPANTS

To recruit interested customers from the general public, information on the project was distributed via radio, local television, newspaper, online media, partner websites, public events, and informational stickers on the BEVs. People could apply by accessing a public website via one of several short links (e.g., www.bmw.de/testfahrer or www.tu-chemnitz.de/activee) and completing an online questionnaire on the linked website. More than 650 people applied for the three-month lease of the BMW ActiveE. Applicants were included in the final selection if they: (1) were willing to pay the monthly leasing rate of 450 € (reduced to 370 € as soon as the BMW i3 was available for sale) and the costs for electricity, (2) had at least one charging opportunity or the possibility of installing a charging station at home and/or work (dependent on the prospective daily distance driven with the BEV), (3) were willing to take part in data collection, (4) had a mobility profile that would be expected to result in a frequent active interaction with range (i.e., at least 90 km daily driving distance with the BEV), and (5) lived or worked in the area around Leipzig (max. 120 km away). Given that the average daily distance travelled in Germany is around 39 km (infas & DLR, 2010), 42 km for rural districts, a mobility profile resulting in a 90-km driving distance per day is relatively high (i.e., users with such a mobility profile can be readily labeled as daily long-distance commuters).

After screening for the inclusion criteria (see above), participants were primarily selected and prioritized based on their expected frequency of having to interact with range (i.e., combined analysis of daily driving distances and available charging opportunities). As restrictions for inclusion in the sample were similar to those for leasing a BEV (e.g., users paid an EV-adequate monthly leasing-rate, needed charging opportunity), we expect the sample to represent a population of early adopters (i.e., early customers) of BEVs in Germany.

BEV AND CHARGING INFRASTRUCTURE

The BEV used in this study was the BMW ActiveE, an electric conversion vehicle based on a BMW 1 Series Coupé with a maximum available driving range between 130 and 160 km in real terms, depending on driving style (Ramsbrock, Vilimek, & Weber, 2013). It took 4-5 hours to fully charge the battery using a 32 A charging station and 8-10 hours using a normal socket. Customers could access, amongst other data, information on battery charge level, range, energy consumption, and the charging process: The estimated remaining and full charge range were displayed based on energy consumption over the last 30 km (as stated in the user manual), the state-of-charge was displayed (0-100%), the current and average energy consumption were displayed (including an energy consumption

history display with 1-minute averaging intervals for the current trip). Furthermore, during charging processes the estimated time until the vehicle was fully charged was displayed. In addition, users had access to advanced online vehicle functions including the "My BMW remote" smartphone app (see also Franke et al., 2014b) that could be used to remotely control certain functions (e.g., starting/stopping a charging process, preconditioning of the battery and passenger compartment) and monitor certain parameters (e.g., remaining range, remaining time until fully charged, charging status). Participants could install the app on their own smartphone or tablet (Android or iOS). For this field trial, the standard series version of the BMW ActiveE was slightly modified. The series version comes with an ECO PRO mode that can be selected to automatically adjust the drive configuration and comfort functions to achieve a higher range. For this field trial, the ECO PRO mode was deactivated, partly to maintain greater control over the available range for the users, but also to allow for testing of different layouts of the regenerative braking system.

Participants used different charging infrastructure configurations depending on their mobility profile. There were two different settings, determined by the commuting trip distance, i.e. the one-way trip distance between the place of residence (home) and the workplace (work): Setting 1 for participants using only *one* charging point (at work or at home) and setting 2 for those using *two* charging points (at work and at home). In the first case (setting 1), the distance should not be more than 60 km, in the second case (setting 2), it should not be more than 120 km. In terms of charging stations (work/home), there were different conditions possible: a 32 A charging station (or 16 A, depending on the capacity of the grid connection) or a normal socket. Sometimes, the 32 A charging station was already installed at the workplace. Moreover, there were participants using public charging stations as a charging point. It was considered important not to include the combination of two sockets (at work and at home) because charging duration could have been too long. Apart from that all participants received a "charging card" from the SWL that enabled them to use all public charging stations in the city of Leipzig (around 30).

PROTOCOL METHODS AND DATA LOGGERS

Travel Diary

The travel diary was a person-based, self-report diary. Before T0 (i.e., baseline mobility with combustion vehicle) and before T1 (i.e., mobility with BEV) the (prospective) main user of the BEV was asked to record all trips made with every means of transportation for 1-week (5 workdays, 2 weekend days – meaning non-working days as clarified in the T0-3 telephone interview with users who did not have a typical Mo-Fr workweek). The following trip definition was applied: Each discrete instance of travel outside of the home was defined as a trip (i.e., had to be recorded). If the trip was interrupted (e.g., for an errand) the section after this interruption was defined as a new trip (i.e., also outbound and inbound trips were counted as separate trips). Also trips by foot >100 m had to be recorded. For each trip, participants answered the following items at T0: (1) date and time at the beginning and at the end of the trip, (1) trip distance in km, (2) purpose of the trip (categories: more work-related vs. more leisure-related trip), (3) main mode of transportation used (defined as the mode of transportation with which the user traveled the most distance; categories included: car or other motorized private transportation, public transportation including taxi and plane, non-motorized transportation, other), and (4) if the trip was logged by the GPS logger (i.e., if the trip was taken in the vehicle that had the GPS logger installed). At T1, the additional category "BEV" was introduced as a mode of transportion and additional items assessed (5) remaining range, and (6) charging status at the beginning of the trip, as well as (7) reason for non-use of the BEV (if applicable; categories included: general limited range of the BEV, currently insufficient remaining range, non range-related reasons).

A score that was only derivable from this method was, for example, the number of trips driven with the BEV relative to the total number of trips (or relative to the total number of trips with motorized private transport). Moreover, this method allowed for characterization of trip parameters (e.g., frequency, distance, mode of transport) that could not be taken with the BEV due to its limited range. Finally, by administering the diary in a similar layout before T0 and T1, comparisons of mobility profiles (i.e., BEV-related changes in general mobility patterns) became possible.

Range and Recharging Diary

The range and recharging diary was a car-based, self-report diary. Before T2, users were asked to record all trips and recharging events with their BEV over a period of at least 10 days of BEV usage within a diary (i.e., similar to a driver's logbook). The trip definition was: "Each instance of travel (i.e., distance driven) with the BEV is a trip." and "As soon as you step out of your car the trips ends and subsequently a new trip starts." For each trip,

participants logged the following variables at the beginning and at the end of the trip: (1) time, (2) odometer reading, (3) information on remaining range and (4) full charge range, (5) charge level, (6) the availability- and (7) the use-of a charging opportunity at the end of the trip, as well as (8) the user (main user of the BEV vs. secondary user). Data from secondary users was recorded to obtain a continuous (i.e., gap-free) record that allowed for more accurate plausibility checks and was more tolerant of single missing values (i.e., certain missing values could be reconstructed based on prior and subsequent data entries).

A score that was only derivable from this method was, for example, the *percentage of charging opportunities used* score (i.e., the amount of initiated recharging events divided by the amount of subjectively available recharging opportunities at the end of the trip). Together with information on trip length and available range, this also yielded a score for the *abundance of charging opportunities,* which is an important parameter when examining user-battery interaction (Franke & Krems, 2013b). Another score derived from this diary was the *time to charge* score. For each charging event, the available time in minutes until the next trip started was computed and divided by the amount of energy that had been discharged (i.e., 100 – the current state of charge). Results showed, for example, that there was a moderate negative relationship between this variable and the usage frequency of range-related remote access functions provided by the "My BMW remote" smartphone app (i.e., typically less time to charge = more intensive usage of remote access to range/charging status; Franke et al., 2014b).

Event-based Car Diary

The event-based car diary was used to record any events where the car behaved unexpected or car behavior was incomprehensible to the users. For each event, users were asked to record: (1) date, (2) time, (3) odometer reading, (4) how the car behaved (i.e., what was unexpected/abnormal), and (5) speculation as to what caused this car behavior.

Data Loggers

Different data loggers were used to record relevant parameters. First, a GPS logger recorded standard GPS variables (e.g., time, position, speed, height), trip distance (with map-matching algorithm correction), and information on the user (i.e., users were asked to press a button switch that changed a marker signal in the data logs from "main user" to "secondary user" and vice versa). The user-information was only collected for the GPS-logs before T0 (i.e., of the users' main combustion vehicle) because in this period no user-information was available from other logger data sources. For the BEV usage period (T0 to T2) it was deemed too burdensome and potentially error-prone to require users to operate the switch for the complete twelve weeks. By collecting GPS-information both before T0 and between T0 and T2 (i.e., for BEV usage), further analyses regarding BEV-related changes in mobility patterns were enabled (e.g., changes in activity space; Kurani, Turrentine, & Sperling, 1994).

The trip definition for the GPS logger and the other car-based data loggers was a function of turning the vehicle on/off (e.g., GPS logger received an on/off-signal from the cigarette lighter), and therefore, slightly differed from that of the diaries (see above). Hence, this resulted, in a tendency for more trips to be recorded via the loggers than through the diaries. However, such an on/off trip definition was not applicable for the diaries in most cases (e.g., given travel diary items) and would have resulted in too much burden on users, and therefore, potentially lower data quality.

Additionally, the BMW Group recorded several signals with different data loggers. Exemplary variables derived from these signals that were used to study user-range interaction were: date, time, odometer reading, trip distance, parameters of recharging processes, range, speed, and temperature. Moreover, the IDs of the personalized car keys were recorded along with trip, time, and odometer information to allow for filtering out all secondary user data segments from all the logger data. To identify user-related problems in using the car keys as instructed, users were asked in the interviews at T0+1, T1, and T2 to quantify and list events where they could not (or simply did not) use the car keys as instructed.

PSYCHOLOGICAL CONSTRUCTS AND OPERATIONALIZATION

The following sections will focus on the scales and items (i.e., scores) assessed within the questionnaires. As the aim of the development of the study methodology was to comprehensively examine user-range interaction, several steps were taken to ensure that the relevant psychological constructs were identified and operationalized appropriately.

First, in the *identification phase,* multiple strategies were employed to enable an anticipation of the everyday

experience and behavior of users who have high range demands, thereby enhancing identification of constructs which could play a role in everyday user-range interaction: (1) literature research, (2) deduction of research questions and hypotheses from the adaptive control of range resources (ACOR) model (Franke et al., 2012c; Franke & Krems, 2013a; Franke & Krems, 2013b, Franke, 2014), as well as from results of the previous field trials with the MINI E and BMW ActiveE in Berlin (Franke, Cocron, Bühler, & Neumann, 2013; Vilimek & Keinath, 2014), (3) weekly meetings of the method development team, (4) consortium workshops, (5) extended trips with the BEV (i.e., by members of the method development team) to learn about range-related vehicle behavior (e.g., range dynamics) from a drivers' subjective perspective, and finally (6) a pre-study with more than 70 users who drove the BEV in a critical range situation (≈ 100 km trip with ≈ 120 km range in the beginning given challenging driving conditions) that provided comprehensive subjective data. The result of this process was a research topic list that contained relevant psychological constructs and topics that were then prioritized to create a final consolidated list of research topics.

Second, in the *operationalizations phase,* those constructs were assigned to appropriate methods and, where appropriate, questionnaire items and scales were developed (for a list of final scales see Figure 1). In this process, a first step was to search for existing (established) items and scales that could serve as a basis for item construction or that could be (directly) transferred/translated for usage in the present study. This included, for example, a literature search for properly developed short scales assessing psychological constructs related to task motivation, stress resistance and self efficacy, as well as the screening of questionnaires that had already been applied (and had proven to be useful) in the previous BEV field trials with the MINI E and the BMW ActiveE in Berlin (see above). In sum, most items needed to be newly constructed, yet at least structural item elements (e.g., certain response scale characteristics and item formats) could be transferred for most of the items. In terms of the response scale, most items were formulated so that a standard 6-point Likert scale that had proven to be useful in previous field trials (completely disagree, largely disagree, slightly disagree, slightly agree, largely agree, completely agree; coded as 1 to 6) could be used. This scale also had the advantage that it allowed for dichotomization of responses (i.e., % of users who rate the item with scale values of 4-6 = % of users who tend to agree with the item text). Additionally, for some items a more fine-grained scale with numerically indicated scale gradation (0, 1,…, 9, 10) was used where only the end points of the scale were labeled (e.g., for a rating of confidence in one's retrospective estimation of average consumption: "not at all confident" vs. "absolutely confident"). Finally, frequency ratings were assessed with a 6-point scale (never, almost never, occasionally, often, almost always, always; coded as 1 to 6).

As the general psychological characteristics behind the specific BEV-related user-range interaction styles were of high interest (e.g., general interaction styles in managing limited resources), several scales and items were included that allowed for examination of: (1) the role of interaction style transfer from other human-machine systems and (2) the role of general personality characteristics (for the latter aspect see Figure 1). Regarding the role of cross-system transfer of certain key attributes/styles (e.g., user-battery interaction, range utilization, see Figure 1), additional items were constructed that were framed upon the non-BEV technical system (e.g., combustion vehicle and mobile phones), but formulated as similarly as possible to the BEV-related items. For example, the UBIS-1 and UBIS-8 scale (Franke & Krems, 2013b) were also assessed within the context of mobile phone charging. For instance, the BEV-related item of UBIS-8 "… I typically charge when I am below a specific buffer range that I always want to have in the battery." was translated to "… I typically charge when I am below a specific reserve capacity that I always want to have in the battery." Moreover, in terms of range-interaction in combustion vehicles, several items were administered to examine whether people who utilized the range of the BEV to a high extent (e.g., who frequently experienced situations with low battery, assessed at T1 and T2) had already been among the people (at T0) who frequently encountered low-fuel situations in their combustion vehicle. This was, for example, assessed by asking users at T0 to report the frequency of encountering situations with: (1) less than 5 liters vs. (2) less than 10 liters in the fuel tank and asking at T1 for the frequency of encountering situations with the BEV with: (1) less than 20 km remaining range vs. (2) less than 40 km remaining range. Of course those values cannot be interpreted and compared in an absolute sense (i.e., comparison of means). Yet, testing for a correlation of the combustion vehicle items (T0) with the BEV items (T1) can give some indication regarding possible transfer effects from combustion mobility to electric mobility. However, to correctly analyze this relationship, it is also important to include mobility-related variables in the analysis. For example, some BEV drivers may simply not have the freedom to decide whether to charge the car before reaching 20 or 40 km remaining range because of their unique configuration of charging opportunities and commuting distance.

Acceptance – purchase intentions | satisfaction & usefulness[1] | overall satisfaction | suitability for everyday use | recommendation[2]

Mobility needs fit – coverage of mobility needs by BEV | BEV-related barriers to optimal mobility needs fit | BEV-related flexibility loss | availability of alternative (combustion) car

Range evaluation – range preferences | range satisfaction

Range experience – range stress & range concerns (frequency & intensity) | annoyance of range interaction

Range levels – comfortable range (absolute/relative safety buffer, distance with good feeling, comfortable range scenario task[3]) | performant range | competent range

Range competence – subjective range competence (prediction & control) | mental model of range dynamics | general technical background knowledge

Range adaptation – subjective adaptation duration | adaptation intensity | adaptation strategy | task motivation | subjective task difficulty | need for assistance | active practice with BEV

Range utilization – frequency of low-battery situations | approach/avoidance of critical range situations | maximum distance per day & per charge | minimum displayed range/state-of-charge

Eco-driving – interest & motivation | behavior (intensity & frequency) | range related adaptation of driving style | general driving style | competence/knowledge | general technical background knowledge

Display interaction – interaction frequency with range & consumption displays | trust in range prediction | usage of remote access to range-related information & perceived support

User-battery interaction – User-battery interaction (UBI) motivation | User-battery interaction style (UBIS-1 & UBIS-8[4])

Charging – appraisal of charging duration & reliability | charging locations | charging behavior (frequency & omission of opportunities)

Mobility – daily distance estimates (average & weekly maximum) | confidence in daily distance estimates

Baseline mobility – frequency & distance of long distance trips | yearly mileage | daily distances | affinity to plan trips | regularity/predictability of daily mobility patterns

Personality – impulsivity (planning affinity)[5] | control beliefs in dealing with technology (KUT)[6] | self efficacy[7] | need for cognition[8] | affinity to technology | interaction with time buffers

[1] Van Der Laan et al., 1997 | [2] Bühler et al., 2014 | [3] Franke et al., 2014 | [4] Franke & Krems, 2013b | [5] Meule et al., 2011 | [6] Beier, 1999 | [7] Schwarzer & Jerusalem, 1999 | [8] Bless et al., 1994

Figure 1. Scales and constructs assessed in the questionnaires.

Range Evaluation and Range Experience

Within this and the following sections some exemplary range-related scales will be discussed. One central research question of the field study was how users with high range demand experience, and consequently, evaluate BEV range. Furthermore, it was of high interest under which conditions (e.g., mobility and personality profile) range stress occurs with high frequency and intensity. In terms of range evaluation, for example, two open-ended, stated preference items were adopted from previous research (Franke & Krems, 2013c) to assess range preferences (e.g., "Which EV range would you consider to be quite short, but just acceptable?"). In terms of range stress and range concerns, items applied previously (Franke et al., 2012c; Franke & Krems, 2013a) were used and further developed (e.g., new item for range concern frequency “How often is a range situation unpleasant for you?”).

Psychological range levels

At least within the context of relatively average mobility profiles in the German mobile population, user-range interaction is typically not characterized by the experience, but rather by the avoidance, of range stress (Franke et al., 2012c). This leads to the pattern in which there is a discrepancy between the maximum available range and the maximum range that the driver is comfortable using. As a consequence, the efficiency of user-range interaction (and therefore also the user experience) can be understood as a function of three psychological range levels (Franke & Krems, 2013a). In the present study a methodology was applied to assess range levels that is very similar to methodology utilized in previous research (Franke & Krems, 2013a). Yet, there were some further developments towards simpler scales to assess comfortable range (Franke, Günther, Trantow, Rauh, & Krems, 2014a).

Mobility and Mobility needs fit

Characteristics of the users' baseline mobility profile (i.e., resource demand profile) constitute a key variable set when studying user-range interaction. This is because mobility demand is strongly associated with the opportunities, challenges and the necessity of actively interacting with range. For example, drivers who have lower mobility needs will need to more actively search for critical range situations to push the limits of their range comfort zone while people with higher mobility needs will more often come into potentially critical range situations incidentally, in other words they will have a higher need to adapt their range comfort zone (i.e., a higher adaptational pressure).

However, not all parameters of the mobility profile are easily assessed with objective data like travel diaries (particularly not with a diary of only seven days). So we also targeted low-frequency events (e.g., long distance trips) and the regularity/predictability of daily mobility patterns (see Franke et al., 2014b) in the questionnaire.

Also, in terms of the mobility needs fit, there is more to evaluate than simply the percentage of mobility needs that can be covered with the BEV. For example, the need for planning and adjusting daily schedules (i.e., a loss in flexibility) is something that can go hand-in-hand with dealing with a more limited resource (i.e., resources that require substantial time to refill/recover). Hence, items like, "Planning car usage (trip planning, charging times) with the BEV is demanding" and "With the BEV, I am as flexible as with a combustion vehicle." were also included to study the subjective mobility needs fit of the BEV.

PRELIMINARY RESULTS

The following section shall give a preliminary impression of the pattern of results within the field study. For this purpose, only the data from the first two usage phases with 29 users who completed the whole study (one dropout before T1) could be analyzed. These had an average age of $M = 41.1$ years ($SD = 8.1$), 5 were female.

The aforementioned users estimated at T0 that their average daily car driving distance was around $M = 120$ km ($SD = 40$ km). Based on the travel diary at T1 it was found that people indeed drove, on average around $M = 109$ km a day with the BEV on days with BEV mobility including weekend days without commuting ($SD = 41$ km, $N = 28$ because one user had too few data points, additionally one user had an extremely low average daily distance of 2.5 km, without this user $M = 112$ km). Hence, users had, on average, even higher daily distances than what was required by the screening criterion. The longest driving distance with the BEV on one day that participants reported over the whole usage phase (the item was administered at T1 and T2; thus, the maximum of the two values of each user was taken) was on average $M = 176$ km ($Min = 90$, $Max = 265$).

In terms of general satisfaction (i.e., acceptance), the tendency to recommend the BEV to a friend with a similar mobility profile (item: "I would recommend electric vehicles like the BMW ActiveE to a friend with similar mobility demands", adapted from Bühler, Cocron, Neumann, Franke, and Krems, 2014) was very high throughout the study. While 90% of the 29 users agreed with the statement at T0 (dichotomization of 6-point Likert scale, see above), this number was the same at T1 (90%) and T2 (90%). Also, in terms of the average rating of this item (scale values from 1-6, see above) there was nearly no difference between the data collection time points ($M_{T0} = 4.86$, $M_{T1} = 4.79$, $M_{T2} = 4.72$).

In terms of learning processes, participants reported that they increasingly had the feeling that they were able to predict the range of the BEV under different conditions (item: "I can precisely estimate the range of the BEV under different conditions"). While 76% of the users agreed at T0+1 ($M = 4.10$), this value increased to 100% at T2 ($M = 5.14$). A similar result was revealed with the item, "I can precisely estimate the influence of different factors on range'' (T0+1: 69%, $M = 3.39$; T2: 100%, $M = 4.86$).

However, the increase in subjective ability to estimate (i.e., predict) the range under different conditions did not come automatically (i.e., without effort). Users who endorsed stronger agreement at T2 with the item, "I have tried to understand the factors that influence range" (i.e., a higher active exploration tendency), also reported a substantially higher increase in their subjective ability to predict range from T0+1 to T2. For this analysis, the difference between users' ratings of the two aforementioned items at T0+1 and T2 was computed and a mean score of the two new difference variables was computed (high values = high increase in subjective range prediction ability). The correlation between active exploration tendency and subjective range prediction competence was significant ($r = .41$, $p = .026$). This suggests that assistance strategies which engage the user in active exploration of range dynamics might have the potential to enhance range-related learning processes (i.e., adaptation processes), thereby, resulting in a better user experience.

Parallel to the development in subjective range competences, active preoccupation with the range (item "While driving the BEV I am often preoccupied with the range") continuously decreased with experience (T0: 90%, M = 4.55; T0+1: 83%, M = 4.34; T1: 79%, M = 4.28; T2: 76%, M = 4.34). Finally, we also examined the frequency of range-related stressful encounters (i.e., the frequency of range stress) with items that are similar to those used in previous research (Franke et al., 2012c; Franke & Krems, 2013a). Specifically, we administered the item, "How often do you feel stressed because of the range?" (exactly the same item as used in Franke et al., 2012c and Franke and Krems, 2013a) at T1 and T2. First, the correlation between range stress frequency assessed at T1 and T2 was strong. As normal distribution was not indicated for the frequency rating at T1 and T2 (p-values for Kolmogorow–Smirnov test < .05) Spearman's rank correlation coefficient (r_s = .88, p < .001) and Kendals tau were computed (τ = .78, p < .001) and, for completeness, also Pearson's product-moment correlation coefficient (r = .54, p = .003). This finding indicates an acceptable test-retest reliability and that range experience had stabilized to a larger extent within the period between T1 and T2. This stabilization was also evidenced by the average values at T1 (M = 2.40, Mdn = 0.50) relative to T2 (M = 2.34, Mdn = 0.50). However, as suggested by the difference between the Mean and the Median values, there was a high degree of variance in users' range stress frequency and there were also several users who never experienced range stress. In particular, at T1 45% of the users stated that they never experienced range stress in a typical month and 28% indicated more than one situation per month (same results at T2). These results are somewhat higher than what has been found previously in the MINI E Berlin field study, a study conducted in an urban region with a typical daily BEV driving distance of approximately 38 km; Franke & Krems, 2013b). Here the results at T2 were an average of approximately one stressful range situation per month (Franke et al., 2012c; Franke & Krems, 2013a). However, in this study only 34% of the users never experienced a stressful range situation and, similar to the present study, 24% experienced more than on situation per month (Franke et al., 2012c). Therefore, the story seems to be slightly more complicated than just "higher daily range demands lead to more range stress".

DISCUSSION

The objective of the present contribution was to describe a methodology that was designed to better understand how users interact with the limited range resources in BEVs. An extensive package of methods was developed and applied within a field study setting. Based upon experience in utilizing these methods within the field trial, it can be concluded that the developed body of methods has proven to be fruitful and the developed questionnaire items and scales produce meaningful data and results. Moreover, preliminary results indicate that even users who have high daily mobility needs can cope with the limited range of a BEV. However, the results regarding the role that active range exploration plays in the increase in subjective range competence also suggests that strategies aimed at supporting users' adaptation to BEV range have the potential to help users reach an optimal experience even under conditions of high range demand.

One final question may be whether we have learned something that is only specific to the examination of user-range interaction in BEVs based on the present research or if elements of this research methodology have broader applicability to other areas of human factors research. We think that transferability is ultimately dependent upon the similarity of the resource situation. In BEVs, the interaction with range resources essentially has the following structure: Sustaining mobility (i.e., reaching the trip destination, avoiding breakdown of resources) is the primary goal of users, as is the case with many other facets of driving behavior (e.g., Fuller, 2005). Resources are continuously consumed while the system is in use. The resource storage capacity is limited and resources can only be refilled through the investment of other resources (e.g., time resources). Within the resource consumption phase, there are only periodically convenient opportunities to replenish the resources. This description of the resource situation structure fits with several battery-powered mobile devices to some degree. And indeed many of the psychological constructs and questionnaire items used in the present study can be translated, for example, also to the domain of user-battery interaction in smartphones. Consequently, we hope that the presented research methodology can inspire further studies that examine other human-machine systems where users have to deal with limited resources, particularly because the sustainable interaction with limited resources is a vital topic of our time.

ACKNOWLEDGEMENTS

This research was funded by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (16SBS014B). Any views expressed herein are those of the authors and do not necessarily reflect those of the funding bodies or partners involved in the project. We are grateful for the support of our consortium partners, BMW Group and Stadtwerke Leipzig (SWL) who made our research possible.

REFERENCES

Beier, G. (1999). Kontrollüberzeugungen im Umgang mit Technik. *Report Psychologie*, *9*, 684–693.

Bless, H., Wänke, M., Bohner, G., Fellhauer, R.F., & Schwarz, N. (1994). Need for cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben. *Zeitschrift für Sozialpsychologie, 25*, 147–154.

Bühler, F., Cocron, P., Neumann, I., Franke, T. & Krems, J. F. (2014). *Is EV experience related to EV acceptance? Results from a German field study*. Manuscript submitted for publication.

Burgess, M., Harris, M., Walsh, C., Carroll, S., Mansbridge, S., & King, N. (2013, November). *A longitudinal assessment of the viability of electric vehicles for daily use.* Paper presented at the 4th Hybrid and Electric Vehicles Conference 2013 (HEVC 2013), London, UK.

Caroll, S. (2010). *The smart move trial: Description and initial results*. Retrieved from http://www.cenex.co.uk/LinkClick.aspx?fileticket=yUKAcRDJtWg%3D&tabid=60

Cocron, P., Bühler, F., Neumann, I., Franke, T., Krems, J. F., Schwalm, M., & Keinath, A. (2011). Methods of evaluating electric vehicles from a user's perspective – the MINI E field trial in Berlin. *Intelligent Transport Systems, IET, 5*(2), 127-133. doi:10.1049/iet-its.2010.0126

Franke, T. (2014). *Sustainable mobility with limited resources: Experience and behavior in dealing with electric vehicle range* (Doctoral dissertation, Chemnitz University of Technology). Retrieved from http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa-133509

Franke, T., Bühler, F., Cocron, P., Neumann, I., & Krems, J. F. (2012a). Enhancing sustainability of electric vehicles: A field study approach to understanding user acceptance and behavior. In M. Sullman & L. Dorn (Eds.), *Advances in traffic psychology* (pp. 295-306). Farnham, UK: Ashgate.

Franke, T., Cocron, P., Bühler, F., Neumann, I., & Krems, J. F. (2012b). Adapting to the range of an electric vehicle: The relation of experience to subjectively available mobility resources. In P. Valero Mora, J. F. Pace, & L. Mendoza (Eds.), *Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems, Valencia, Spain, June 14-15 2012* (pp. 95-103). Lyon, France: Humanist Publications.

Franke, T., Cocron, P., Bühler, F., & Neumann, I. (2013). Die Nutzerperspektive auf Elektromobilität: Ergebnisse der Feldstudie. In J. F. Krems, O. Weinmann, J. Weber, D. Westermann, & S. Albayrak (Eds.), *Elektromobilität in Metropolregionen: Die Feldstudie MINI E Berlin powered by Vattenfall. Fortschritt-Berichte VDI Reihe 12 Nr. 766* (pp. 48-79). Düsseldorf, Germany: VDI Verlag.

Franke, T., Günther, M., Trantow, M., Rauh, N., & Krems, J. F. (2014a). The range comfort zone of electric vehicle users – concept and assessment. In *Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems, Vienna, Austria, June 5-6 2014*. Lyon, France: Humanist Publications.

Franke, T., Neumann, I., Bühler, F., Cocron, P., & Krems, J. F. (2012c). Experiencing range in an electric vehicle: Understanding psychological barriers. *Applied Psychology, 61*(3), 368-391. doi:10.1111/j.1464-0597.2011.00474.x

Franke, T., & Krems, J. F. (2013a). Interacting with limited mobility resources: Psychological range levels in electric vehicle use. *Transportation Research Part A: Policy and Practice, 48*, 109-122. doi:10.1016/j.tra.2012.10.010

Franke, T., & Krems, J. F. (2013b). Understanding charging behaviour of electric vehicle users. *Transportation Research Part F: Traffic Psychology and Behaviour, 21*, 75-89. doi:10.1016/j.trf.2013.09.002

Franke, T., & Krems, J. F. (2013c). What drives range preferences in electric vehicle users? *Transport Policy, 30*, 56-62. doi:10.1016/j.tranpol.2013.07.005

Franke, T., Trantow, M., Günther, M., Krems, J. F., Vilimek, R., & Keinath, A. (2014b, June). User interaction with remote access to range-related information in BEVs. In V. Blervaque (Chair), *Proceedings of the 10th ITS European Congress*. Helsinki, Finland.

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis and Prevention, 37*(3), 461-472. doi:10.1016/j.aap.2004.11.003

Hawkins, T., Gausen, O., & Strømman, A. (2012). Environmental impacts of hybrid and electric vehicles – a review. *The International Journal of Life Cycle Assessment, 17*(8), 997–1014. doi:10.1007/s11367-012-0440-9

Hirsch, R. L., Bezdek, R., & Wendling, R. (2005). Peaking of world oil production and its mitigation. *AIChE Journal, 51*(1), 2–8. doi:10.1002/aic.10747

Holdway, A. R., Williams, A. R., Inderwildi, O. R., & King, D. A. (2010). Indirect emissions from electric vehicles: Emissions from electricity generation. *Energy & Environmental Science, 3*, 1825-1832. doi:10.1039/C0EE00031K.

infas, & DLR (2010). *Mobilität in Deutschland: Ergebnisbericht*. Retrieved from http://mobilitaet-in-deutschland.de/02_MiD2008/publikationen.htm

Kurani, K. S., Turrentine, T. S., & Sperling, D. (1994). Demand for electric vehicles in hybrid households: an exploratory analysis. *Transport Policy 1*(4), 244–256. doi:10.1016/0967-070X(94)90005-1

McManus, M. C. (2012). Environmental consequences of the use of batteries in low carbon systems: The impact of battery production. *Applied Energy, 93*, 288-295. doi:10.1016/j.apenergy.2011.12.062

Meule, A., Vögele, C., & Kübler, A. (2011). Psychometrische Evaluation der deutschen Barratt Impulsiveness Scale – Kurzversion (BIS-15). *Diagnostica, 57*(3), 126-133. doi:10.1026/0012-1924/a000042

Pichelmann, S., Franke, T., & Krems, J. F. (2013). The timeframe of adaptation to electric vehicle range. In M. Kurosu (Ed.), *Human-computer interaction. Applications and Services, LNCS 8005* (pp. 612-620). Berlin, Germany: Springer.

Ramsbrock, J., Vilimek, R., & Weber, J. (2013). Exploring electric driving pleasure – the BMW EV pilot projects. In M. Kurosu (Ed.), *Human-Computer Interaction. Applications and Services, LNCS 8005* (pp. 621-630). Berlin, Germany: Springer.

Schwarzer, R,. & Jerusalem, M. (Eds.) (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*.

Berlin: Freie Universität Berlin. Retrieved from http://userpage.fu-berlin.de/~health/self/selfeff_public.htm

Van der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies, 5*(1), 1-10. doi:10.1016/S0968-090X(96)00025-3

Vilimek, R., & Keinath, A. (2014). User-centred design and evaluation as a prerequisite for the success of disruptive innovations: An electric vehicle case study. In M. Regan, T. Horberry & A. Stevens (Eds.), *Driver acceptance of new technology: Theory, measurement and optimisation* (pp. 169-186). Farnham, UK: Ashgate.

Section 18

Road and Rail:

System Design and Evaluation II

Advanced In-Vehicle Applications to Mitigate Driver Aggression

Alessandro Soro[a], Sabine Wollstädter[b] and Andry Rakotonirainy[a]

[a]Queensland University of Technology
Brisbane - Australia

[b]Leuphana University of Lüneburg
Lüneburg - Germany

ABSTRACT

Aggressive behavior at the steering wheel has been indicated as a contributing factor in a majority of crashes and anger has been compared to alcohol impairment in terms of probability to cause a crash. It has been shown that being in a state of anger or excitement while driving can decrease the drivers' performances. . This paper reports the evaluation of 6 novel design alternatives of In-Vehicle Information Systems (IVIS) aimed at mitigating driver aggression. Each application presented was designed to tackle the following contributing factors to driver aggression: competitiveness, anonymity, territoriality, stress as well as social and emotional isolation. The 6 applications were simulated using computer vision algorithm to automatically overlay the real traffic conditions with 'Head-Up Display' visualizations. Two applications emerged over the others from participant's evaluation: shared music combined the known calming effect of music with the sense of sympathy and intimacy caused by hearing other drivers' music. The Shared Snapshot application provided an immediate gratification and was evaluated as a potential prevention of roadside quarrels. The paper presents Theoretical foundation, participant's evaluations, implications and limitations of the study.

Keywords: Automotive, Aggressive, Driving, Interaction, Design, Evaluation, IVIS.

INTRODUCTION

Aggressive behavior at the steering wheel has been indicated as a contributing factor in a majority of crashes (AAAFoundation.org, 2009) and anger has been compared to alcohol impairment in terms of probability to cause a crash (Cook, Knight, & Olson, 2005) . It has been shown that being in a state of anger or excitement (as opposite to a neutral mood) while driving can decrease the drivers' performances (e.g. (Cai, Lin, & Mourant, 2007)). The development of in-vehicle applications aimed at preventing aggressive driving is on its infancy, and generally involves the adoption of advanced in-Vehicle information Systems (IVIS) to provide the driver with socially relevant information, calming messages, entertainment. More details are provided further on in the paper; the design approach of such new IVIS is changing too, moving towards a more user centered approach (Cardoso & Clarkson, 2012; Gellatly & Hansen, 2010; Lindgren, Chen, Amdahl, & Chaikiat, 2007), and hence emphasizing the need to evaluate early prototypes, test alternatives, and observe actual performances (Marcus, 2004). However, evaluating early prototypes of automotive user interfaces involving real users in real contexts, which is a core practice of evaluating interaction design, is difficult to apply. Hence, most (if not all) of the above examples have been evaluated in laboratory studies, often supported by a driving simulator system. Very little is found in the literature on how to deploy and evaluate such concepts in a realistic (if not real) driving situation. Using a naturalistic driving setting to introducing a new, experimental technology raises the topic of risk management in a context in which the life of the user is at stake.

This paper reports on the evaluation of 6 novel design alternatives aimed at mitigating driver aggression and the methodological approach that was applied in the evaluation. Each application presented was designed to tackle one or more of the main contributing factors of driver aggression: competitiveness, anonymity, territoriality, stress as well as social and emotional isolation. Definitions and the rationale for each design alternatives are presented in the relevant sections. The methodology for evaluation was aimed at combining some of the advantages of a naturalistic driving study with the safety and predictability of a driving simulator environment: participants were asked to drive an instrumented research vehicle through several suburbs/traffic conditions. Real frustrating events from their test drive were captured on video and later re-played in a driving simulator, where early prototypes of the applications described here could be deployed and tested.

Predicting and Mitigating Driver Aggression

In a recent survey 50% of drivers admitted to have verbally abused another driver; remarkably 82% of them felt such act to be justified (AAMI, 2011). Driver aggression can take many forms; while larger majority of cases consist of mild aggression (such as verbal abuse, obscene gestures and tailgating), up to 18% of motorists reported severe aggressions, e.g. having been chased, run off the road, or assaulted (Soole, Lennon, Watson, & Bingham, 2011). Yet, it is difficult to estimate how aggressive driving contributes to crashes. This is partly due to a lack of consistency in the use of terms such as driver aggression, hostile or angry driving and road rage (Dula & Geller, 2003). The repertoire of behaviors that have been labeled as aggressive driving is just as wide, and spans from honking and flashing headlights to using the vehicle as a weapon to attack and possibly kill another road user (Soole et al., 2011). Several factors have been found to elicit episodes of driver aggression: personal factors such as age, sex, competitiveness and sensation seeking; as well as situational factors such as isolation, anonymity, previous states of anger or stress, road congestion and time pressure (O'Brien, Tay, & Watson, 2004; Soole et al., 2011). Furthermore, the physical and psychological distance between drivers makes it difficult to manage emotional arousal (Deffenbacher, Lynch, Oetting, & Swaim, 2002). It has been observed, that, the obstruction caused by the vehicle design blocks any attempt to express irritation, solicit and apology or even the attention of the supposedly offending person after a frustrating incident or episode (Parkinson, 2001). On the contrary, cars and their immediate surrounding have been demarcated as the driver's own territory (Fraine, Smith, Zinkiewicz, Chapman, & Sheehan, 2007), potentially defended by occupants or owners (Altman & Chemers, 1980).

Anger, when articulated, is an important component of social interaction; it has the function of soliciting cooperation or apology from a supposedly offending person, or to direct accordingly the blame of onlookers. In face to face interaction, anger typically cools down as soon as an acknowledgment is received (Parkinson, 2008). In the car, attempts of expressing anger and receiving feedback or apology are limited: social cues such as voice tone or face expressions are unavailable unless overly exaggerated, and cannot promptly reach other drivers. Similarly, the offending driver will only receive over-amplified expressions of anger. Such disproportion and the lack or delay of feedback exacerbates issues of anger (Parkinson, 2001), as a result, anger at the steering wheel is experienced more often and more intensely than in face to face interactions. Lupton has further observed how people tend to de-humanize other drivers, often referring to them as 'machines' (Deborah Lupton, 1999). Cars provide to the driver and passengers a protective and personal armor, which Leckie and Hopkins referred to as 'semiprivate metal containers' (Leckie & Hopkins, 2002). Marketing strategies have often insisted on such concept, presenting cars as an extension of the self or a member of the family (D. Lupton, 2002; Deborah Lupton, 1999).

The sense of social and emotional isolation that drivers experience has also been connected to aggressive and anti-social driving. Social and emotional skills seem to be compromised or obstructed when driving: for example a majority of drivers regard themselves as less risky and more skillful of the average fellow driver (Svenson, 1981). Despite this, drivers struggle to re-establish a channel of social communication, and invent means of exchanging non-verbal cues, using those tools that they have at hand: headlights, hazard lamps, blinkers, and of course, hand gestures, in what Renge has dubbed 'roadway interpersonal communication' (Renge, 2000).

There is evidence that such interpersonal communication can be mediated and augmented by means of proper technology capable of visualizing for example other drivers' emotions, goals and attention could discourage aggressive behaviors. Previous research has demonstrated that perceiving the eye gaze of other drivers, represented with digital avatars could promote a pro-social behavior (Rakotonirainy, Feller, & Haworth, 2009). Studies exist that explore the feasibility of automatically sensing and sharing the emotional state of the driver with other road users (Schroeter, Soro, & Rakotonirainy, 2013), or the possibility of visualizing common interests and similarities based on social media profiles (Mitrevska, Castronovo, Mahr, & Müller, 2012). Furthermore, several solutions have been

proposed to prevent or mitigate driver aggression by means of specially designed In-Vehicle Information Systems (IVIS). The scope of intervention includes for example inducing a re-appraisal of potentially frustrating events (e.g. (Harris & Nass, 2011)), adapting the interface to the mood of the driver (e.g. (Jonsson, Nass, Harris, & Takayama, 2005; Nass et al., 2005)), displaying additional information about other drivers' intentions or behaviors (e.g. (Rakotonirainy et al., 2009)), exploiting gamification to break anonymity (e.g. (Schroeter, Rakotonirainy, & Foth, 2012)), exploiting the calming effect of music (e.g. (Wollstädter, Pfister, & Höger, 2013)). Some applications (see e.g. (Diewald, Möller, Roalter, Stockinger, & Kranz, 2013)) aim at simultaneously achieving the goals of (i) encouraging the adoption of a new technology and (ii) rewarding the achievement of a safety related goal, such as having a greener driving style, given that the distinctive features of a green driver (e.g. anticipating and planning ahead, smooth accelerations and braking, sensible choice of speed) can as well describe a safe driver (Young, Birrell, & Stanton, 2011). Such game-like applications have the advantage of being particularly attractive to the young driver population who are known to have higher crash rates (see e.g. ((Fitz-Walter, Tjondronegoro, & Wyeth, 2012)). Yet, the design and assessment of in-vehicle applications to mitigate or counteract driver aggression is in its seminal phase, and the possibilities of exploiting driver-to-driver interaction principles is a relatively new and unexplored territory.

Research Aim

Our goal is to test whether drivers' behavior intention can be positively affected by in-car systems and to identify potential design alternatives for future in-car communication systems. For this we follow a qualitative approach: after completing a naturalistic test drive and several questionnaires, participants evaluate certain design alternatives of future in-vehicle communication systems (countermeasures). The design alternatives, described below, derive ideas from different research disciplines (e.g. social media, gaming and education).

METHOD

Participants

Eleven Participants participated in the study. However, the first two sessions were aimed to fine-tune the software and the method itself and were not included in the analysis. The demographics of the remaining nine participants are: 27,3% females, age from 23 to 58 with a mean of 34 years. All participants held a valid driving license and 36% drove to work in rush-hour traffic. Participants were recruited through mailing lists.

Material: Rationale and Description

The design ideas presented below focus on the attempt to address the situational factors that could influence aggressive behavior, with specific emphasis on *competitiveness, anonymity, territoriality, stress, social and emotional isolation*. The applications have been evaluated following the methodology presented further on. The descriptions presented hereafter were actually read out to the participant before the driving simulator session. However, the underlying rationale for each application, described further down, was not presented to the participants.

Eco-Driving Reward: This application, illustrated in figure 1, comprises the idea that drivers would earn 'green leaves' when they drive fuel-efficient. Thus, a green and eco-friendly driving state would be achieved by saving fuel (e.g. by slowing down and avoiding sharp accelerations/braking). More leaves would indicate a more eco-friendly driver. In this case the car would be fitted with an appropriate device capable of gathering the relevant data from the on board GPS and car electronics and of scoring the performances of the driver: for example calculating the saving with respect to average fuel consumption for the specific type of road. The leaves could be visualized using a special projection on the windscreen (Head Up Display).

Driving-Behavior Badges: In this example application, illustrated in figure 2, drivers would earn a 'color badge' based on their behavior on the road: a green badge would indicate a good driver, a red badge would point out a bad one (e.g. one that ignore speed limits, doesn't keep safe distances, etc.), an amber badge would indicate an 'average' driver (e.g. one that has received both good and bad votes). Such badges would be assigned on the basis of a 'social' score system. Each driver would have the opportunity to 'vote' other drivers' performances, either good or bad, on the road. Each individual score would be an average of the votes received. Every car could have a button on the

steering wheel that could be hit to assign a vote to another driver nearby. All the votes would then published and shared among all drivers. The badges could be visualized using a special projection on the windscreen (Head Up Display).

Shared Emotions: The following example allows the drivers to share their emotional states; it is illustrated in figure 3. This application would be one way to break the isolation between cars driving in close proximity. Drivers will be able to share their emotional state with other drivers nearby. The emoticons would be similar to the ones, which drivers may already use in digital communication (e.g. Chatting via Skype or Facebook, text messages, email). The visualization could be through a Head Up Display, i.e. overlaying graphics to the real street view.

Shared Music: Drivers perceive their car as a private and solid shield. However, the reality is that drivers, as users of the same road, are actually using a public space. This could be made explicit by adding the opportunity to share the music they are listening to with other drivers nearby. Every driver would have the option to tune in the music offered by others or not, while at the same time sharing his/her own music. This would be similar to sharing music with friends e.g. via Spotify or other music sharing social networks. The in-car entertainment system would be able to alert the driver when other drivers nearby want to share their music. This application is depicted in figure 4.

Shared Chat: In this scenario drivers would have the opportunity to talk to other drivers' nearby/in close proximity, as illustrated in figure 5. The conversation could be one to one or could involve many drivers. Drivers would have the option of keeping the conversation private (e.g., talking being in a private space) or to leave the conversation open (e.g., chatting in the corridor). Other drivers could then enter a conversation, just as it happens in face to face encounters. If a car is nearby a in-car application would indicate whether other drivers are having a conversation or not. The application would also indicate whether a conversation is open or private respectively whether other drivers are open for a conversation.

Shared Snapshots: This application would give drivers the opportunity to share (frustrating) scenes or pictures from (frustrating) driving events with their friends (e.g. on Facebook). It would be the choice of the drivers to share their annoyance virtually with their social network instead of using the horn to express their irritation or anger. The car could be fitted with a camera and a device capable of taking a picture and posting it together with a predefined dislike statement (e.g. a button on the wheel next to the honk to operate the system taking the picture and sending it seamlessly). The application is represented in figure 6.

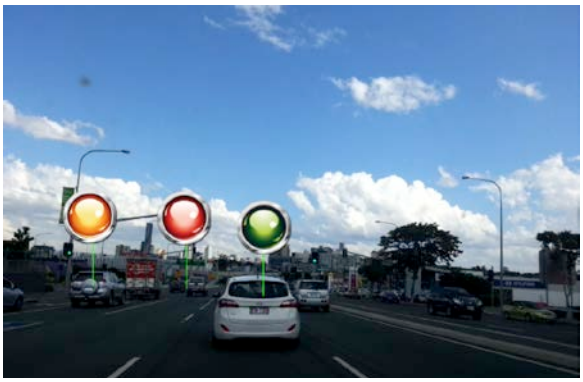
Rationale

Socially inspired gamification techniques, represented in Applications 1 (Eco-Driving Reward) and 2 (Driving-Behavior Badges) leverage people's natural desires for achievement, managing to direct the driver's **competitiveness** towards a safe goal (gaining eco-driving rewards or driver-behavior badges). Hence the hypothesis that a carefully designed gamified IVIS could provide an incentive towards a safe behavior, and, at the same time, provide visual feedback on other drivers' score (i.e. usual behavior).

Applications 3 (Shared Emotions), 4 (Shared Music) and 5 (Shared Chat) have been designed to address the sense of **isolation**, making visible and audible the presence, goals and emotions of other humans on the cars nearby. While they are similar to each other in how they try to create a shared information space to overlay the physical space, the 3 applications articulate different key concepts. Shared Emotions aims at fostering a sense of empathy towards the other road users, and at the same time allows the driver to express his or her own feelings, whether positive or negative, thus fighting the sense of **emotional isolation**. Shared Music tries to reduce **stress** building on the calming effect of music (Wollstädter et al., 2013, Wiesenthal, Hennessy, & Totten, 2003, van der Zwaag & Fairclough, 2011). Shared Chat aims at providing a channel for casual conversation, such as the generally polite and peaceful negotiation of the shared path that is common between pedestrians. Both Shared-Chat and Shared-Music expose the private space of the car to the influence of other road users, thus breaking **anonymity** and **territoriality**.

1. Eco-Driving Reward

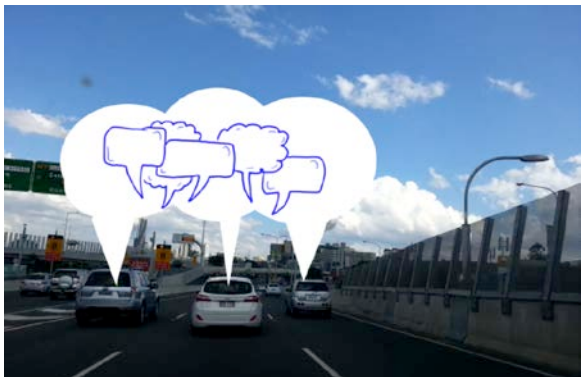
2. Driving-Behaviour Badge



3. Shared Emotions

4. Shared Music

5. Shared Chat



6. Shared Snapshots

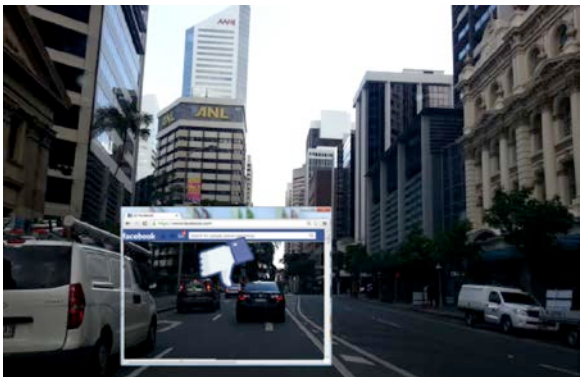


Figure 1: In-vehicle application to mitigate driver aggression

Finally, Application 6 (Shared Snapshots) is designed to provide the drivers with a mean of expressing their anger or irritation that is less prone to escalating or result in an open conflict. The action of exposing the offending behavior to friends and personal contacts could provide an immediate gratification, and replacing the need of aggressively seeking further apology from the offender. On the other hand, evidence exists that the presence of other people in the car is a deterrent to risky driving (Fleiter, Lennon, & Watson, 2010); this application explores the possibility to exploit the virtual presence of friends to remove the sense of **isolation**.

Procedure, Apparatus and Measures

The study consisted of two parts: First, participants drove a naturalistic drive in a CARRS-Q research vehicle which was equipped with video recording devices. The naturalistic test drive was to identify frustrating driving incidents,

which were experienced in a real driving situation for each participant and which were used as video footage for the second part of the study. The predefined test drive comprised specific segments of different predictable difficulty (e.g. standing in the parking area; exiting the parking ramp; city driving; residential areas; merging into traffic; driving in the CBD). For the entire duration of the test drive, which took approximately 45 minutes, a researcher sat on the back seat and did annotate frustrating events (e.g. merging conflicts, high traffic with weaving car in front etc.). Participants were asked to refrain from using radio or mobile phone, and to strictly obey traffic rules. Prior to the test drive informed consent was completed and each participant agreed to be video recorded and completed a questionnaire to assess the demographic.

The second part took place in a *SCANeR*^{TM}*studio* driving simulator, which was located in the CARRS-Q laboratory, Brisbane, Australia. The driving simulator was equipped with steering wheel, foot pedals and headphones for replaying traffic noise and with one projector which presented the traffic situation onto one wall screens with a 60° field of view. This second part of the study aimed to discuss the above-described six in-vehicle applications and lasted approximately 40 minutes. The 6 applications were simulated using computer vision algorithms to automatically render the HUD visualizations on top of the real traffic conditions and were presented in a random sequence for each participant. Figure 2 shows a snapshot from an evaluation session: the participant (left) seats at the steering wheel of the driver simulator. Although no steering is actually possible, he gets an immersive sense of the traffic conditions thanks to the large display and headset. The operator (right) controls the simulation choosing which one of the six applications to display, deciding the appropriate badge to assign the 'offending' car, and selecting the appropriate background noise and sounds. The researcher (not shown) takes a seat at the left of the participant and moderates the evaluation, requesting detailed comments when necessary.

Figure 2: Recreating stressful conditions and evaluating the applications in the driving simulator

As said above, the second part followed directly after participants finished their test drive. A short interview was conducted after the driving session in order to agree with each participant on a specific event, which occurred during the test drive and had been a cause of irritation. Often the video footage was rapidly reviewed to help the participant recall the event, or to re-evoke the sense of irritation experienced. Each participant was then asked to rate the base level of *anger, irritation* and *calmness* experienced during the selected event (see Table 1) on a 7-point scale from 1 'not at all' to 7 'very much'. In total seven ratings where completed throughout the study (base level plus 6 applications).

Then, in turn, the 6 applications were presented in a randomized order, replaying the video sequence of the selected event, enriched with the video/audio augmentation. To this purpose an interactive software emulator was implemented capable of recognizing in real time the cars in the video footage via a cascade classifier of Haar-like features (Viola & Jones, 2011) and of following such cars frame to frame by means of optical flow analysis. During the evaluation, an operator can chose what visual elements (e.g. smileys, green leaves) will be added to the scene, to which car(s) and when. The operator has also the possibility to choose among several background noises (light traffic, normal traffic, traffic congestion) and sound tracks (music, chatter) to be played in the headset. After participants evaluated all applications they received a reimbursement of AUD25 for their participation.

Table 1. Participants‘ evaluation of the 6 applications

Participant	Base			Eco-Driving Reward			Driving-Behaviour			Shared Emotions			Shared Music			Shared Chat			Shared Snapshot		
	A	I	C	A	I	C	A	I	C	A	I	C	A	I	C	A	I	C	A	I	C
1	4	6	2	0	-1	0	-1	-1	2	0	0	0	1	-1	1	0	1	-1	-1	-4	3
2	4	4	5	-1	-1	0	-1	-1	0	-1	-1	0	-1	-1	1	0	0	-2	0	1	-1
3	1	4	6	1	-1	0	1	-3	0	3	1	-3	1	0	0	3	1	-4	0	-1	-1
4	5	5	2	-1	-1	2	-1	-1	3	1	1	1	-2	-2	3	-2	-1	2	-2	-1	2
5	1	3	4	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0
6	1	5	2	2	0	1	0	0	0	0	1	0	0	-2	4	0	1	0	0	-2	1
7	1	1	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	2	5	3	1	-1	0	1	0	0	99	0	0	1	-1	0	2	0	0	1	0	0
9	2	2	6	-1	0	1	0	0	0	1	1	-1	-1	-1	1	-1	-1	1	0	0	0
Mean	2,33	3,89	4,11	0,11	-0,6	0,44	-0,1	-0,7	0,56	0,43	0,44	-0,4	-0,1	-0,9	1,11	0,22	0,11	-0,4	-0,2	-0,8	0,44

Note: Subjective ratings for emotional experience: A (Anger), I (Irritation), C (Calmness). Ratings are measured on a 7-point scale. Gain scores (D = Y2 − Y1) were calculated as indicator of mitigation of driver’s frustration. Lower values: Less anger, irritation and calmness; Higher values indicate more anger, irritation and calmness.

RESULTS

Qualitative and Quantitative analysis were conducted. First, audio recordings of the qualitative interview were transformed in written transcripts. Excerpts from the interviews, which are used describing the results, are at the level of whole sentences or paragraphs.

Second, quantitative data was prepared using Microsoft Excel. Subjective ratings of anger, irritation and calmness were analyzed as indicators of driver’s frustration. Gain scores (D = Y2 − Y1) were calculated. For example, Participant 1 (see Table 1) gave the following ratings for anger “4”, irritation “6” and calmness “2” (on a 7-point-likert scale). Differential values (gain scores) where then calculated for each application to have an indicator for the mitigating effect of driver’s frustration when confronted with each application prototype (Difference = Application-Rating – Base-Rating). For example, the sense of irritation was slightly lower for participant 1 when the simulation of Application 1 (Eco-Driving Reward) was applied to the frustrating driving scene (see Table 1 first column: I = -1); There was no change for anger and calmness for this participant evaluating the Eco-Driving Reward application. Thus, we follow that for this participant there isn’t a valuable mitigating effect for driving frustration when the Eco-Driving Reward-Application was applied. On the other hand, when the application Shared Snapshot was applied higher changes where observed (see Table 1). The ratings for each application were also measured on a 7-point-likert scale. For each participant the subjective evaluation of anger, irritation and calmness, with and without intervention, are reported in Table 1. Further on, the subjective evaluations are commented with the help of excerpts from the interviews.

Eco-Driving Reward and Driving-Behavior Badge

Of the other 4 applications, *Eco-Driving Rewards* and *Driver-behavior Badges* were given analogous evaluations: The effect on irritation was scored positive but modest by 5/9 Eco-Driving Rewards and by 4/9 participants respectively for the Driver-Behavior Badge. The effect on calmness was rated positive by 3/9 for Eco-Driving Rewards and by 2/9 participants for Driver-Behavior Badge. The effect on anger though was uncertain, with some participants feeling a minor improvement and other rating a smaller increase of anger. Analogously to the Snapshot, when commenting on the application *Driving-Behavior Badge* participants stressed on the risk of abuse, but also highlighted the advantages of giving and receiving feedback on driving behavior, with the dual advantage of expressing irritation and visualizing other driver’s reputation, allowing to take extra precautions.

P01: *“A lot of people would just use the bad and they'd use it for everybody. Or (...) use it too much and not enough would use the good. (...) I'd have my finger like over the red button all the time.”*

P06: *“If they do something wrong you can punish them, in a way, by voting down. (..) well, mostly it is kind of a way to make my frustration known and then it might take it off my chest.”*

P02: *“Yeah I think that one has some validity, in terms of… like you're getting continuous… like live feedback to someone's performance, which is really quite useful.”*

P09: *“I don't think it would change the way I felt (...) it would make me more careful depending on the rating that*

they had"

Shared Music

The application *Shared Music* emerges as the preferred one, in terms of its capacity to mitigate irritation (6 out of 9 participants) and improve calmness (5/9). In absolute terms, the presence of music has the higher (positive) impact on the sense of calmness (+1.11 on average) and on irritation (-0.89 on average). Even though the effect on anger is moderate 3 participants reported a small positive effect, 3 reported no effect at all, while 3 reported a slight increase of anger.

This effect resulted from the qualitative analysis, too: All participants agreed on the calming effect of music, and some agreed on the potential of such application of breaking the shield of anonymity.

P02: *"I like the fact that you'll be sharing a space with someone, rather than being isolated from them."*

P04: *"I think music is a great way to modify people's emotional experiences. (..) And it's a way of communicating where you know it doesn't have to get too personal about anyone who might be annoying you."*

P09: *"It's a way of associating a personality with some feature of that person."*

However, several participants were critical towards to risk of being *forced* to share music that they would not like, or anticipated a negative appraisal of other drivers based on their taste in music.

P01: *"I can just see myself getting really angry all the time listening to everyone else's choices, probably irritated as well [...] the odds are probably gonna get more people listening to stuff you don't like [...] I probably wouldn't cos' I like to listen to my stuff."*

P06: *"Since the other driver has a decent taste in music I guess I'm most likely to sympathize with him if we have something in common. If he was playing (...) then it might get on my nerves."*

Shared Emotions and Shared Chat

The applications *Shared Emotions* and *Shared Chat* on the other hand were reported to have a negative effect on all three dimensions, and hence to have potentially the effect of escalating anger and irritation.

This is represented in the interview comments, too. The attempts to support emotional reciprocity through a visualization of drivers' mood in *Shared Emotions* proved the more difficult and less effective. Specifically, the effect of *humanization* provided by the smileys superimposed on the scene doesn't appear to counteract, in general, the irritation caused by the driving conflict.

P01: "*He knows he's cranky and he's still driving like an idiot, which is pretty terrible.*"

P01 *"[I feel] Probably more irritated (...) I'm gonna really wanna avoid them because they've made an effort to show the world that they're cranky or upset or whatever"*

P02: *"If they're angry and they pull out in front of you, I think that would make me angry as well."*

P08: *"If I see something (...) that says that that driver is very irritated then I would be equally irritated."*

P09: *"I would assume that the motivation was because he was angry that he was paying less attention to other people, or taking advantage of them."*

Shared Snapshot

The application *Shared Snapshots* was rated to positively impact irritation by 4/9 and improved calmness by 3/9 participants. Again the effect on anger was moderate according to the subjective ratings.

Taking the qualitative data into account, the aspect of socially rating other drivers' behavior received (sometimes strongly) contrasting evaluations. The application *Shared Snapshots* was understood as a way of taking immediate revenge on another drivers' actions.

P01: *"That's actually brilliant, because it's [immediate] gratification and naming and shaming. That this guy is an idiot and everybody gets to see, which is fantastic!"*

P03: "*A bit like name and shame isn't it? I wouldn't like somebody doing it to me [...] seeing lots of that on my Facebook, with other people doing it."*

DISCUSSION AND CONCLUSION

The aim of this study was to identify potential design alternatives for future in-car communication systems aimed at mitigating driver aggression. Focusing on 6 design alternatives we have invited several drivers and potential users to evaluate if/how their behavior, mood and feelings towards other drivers could be influenced.

Our results indicate that a potential benefit could come from advanced IVIS in terms of reducing driver aggression. Confirming related research, music is an appropriate means to reduce driver's anger in frustrating driving events. Our contribution additionally shows that such channel can be further exploited to break the sense of isolation and territoriality, as was the aim of the Shared Music Application. Yet, several conflicting points of view emerged, especially from the qualitative analysis, which suggests that no one single application could be developed to *solve* the problem of aggressive driving. We summarize below some criticalities that were evidenced.

Behavior versus Feelings: it could be argued that an application aimed at reducing driver aggression should have the dual goal of *(i)* improving the mood of the driver, and *(ii)* encouraging a safer driving behavior. Shared Music had a largely positive impact on feelings, but the participants wouldn't change their driving behavior when using it. On the other end, Shared Emotion was rated to have a negative impact on feelings, causing the participant to reciprocate the other drivers' bad mood. In this case however participants consistently declared that they would give more space from another driver, were his/her bad mood visible. Answering whether the two goals (safety vs feeling) could be conciliated is out of scope for this study. Our contribution indicates that different users react differently to a same application, suggesting, once more, that no one application, but rather a choice or combination of applications, can positively address the problem.

Territoriality: breaking the sense of territoriality proved to be extremely difficult, and several participants resisted strongly to the idea of a possible intrusion. While sensible differences could exist between age groups in this respect, given e.g. the different attitude towards social networking sites manifested by younger users, yet most of the applications failed in leveraging a sense of being in a *public space* when driving.

User Acceptance: participants were often suspicious regarding the introduction of new devices in the car. Again, there may be expected important differences between age groups, that this study was not designed to capture. Yet table 2 evidences strong differences between drivers that reported sensible differences when, either in positive or in negative, when using the applications (e.g. P04) and drivers that starting from the base evaluation reported 'ideal' values of total calmness and no anger/irritation (e.g. P07) or simply didn't notice any significant chance when using the applications (P05). If both positions are legitimate and there's no reason to attribute them to an excess of enthusiasm or hostility, yet such contrast seems to reveal a methodological issue that designers are likely to face during evaluation.

Driver Distraction, together with speed, alcohol, fatigue and unfastened seat-belts, is a major issues, and the cause of a large proportion of on-road fatalities. Distraction was intentionally left out of the study, that aimed instead at assessing which ones (if any) of the 6 applications were worth investigating further and why. Evaluating the potential distraction of IVIS is of paramount importance, and the distraction caused by some of the applications was underlined by several participants. The problem was exacerbated by the imperfections of the prototype implementation. However, it is clear that any added information layer (such as the visual overlay given by a HUD) adds to the cognitive burden of the driver; yet, striking the balance between the cognitive cost and the safety benefit was out the scope this study.

This study has implications for companies interested in improving car safety. Recently, automotive industry has started to examine the influence of in-vehicle communication on aggressive driving. In particular, they are interested in understanding how individuals can be influenced by in-vehicle communication systems to decrease maladjusted driving behavior. The points raised above are open questions that this study helped to focus, and will be the aim of future work.

REFERENCES

AAAFoundation.org. (2009). *Aggressive Driving: Research Update*. AAAFoundation.org.

AAMI. (2011). Annual Road Safety Index. Melbourne, Australia: AAMI Ltd.

Altman, I., & Chemers, M. (1980). *Culture and environment*. Brooks/Cole, Monterey, CA.

Cai, H., Lin, Y., & Mourant, R. R. (2007). Study on driver emotion in driver-vehicle-environment systems using multiple networked driving simulators. In *Proceedings of the Driving Simulation Conference North America (DSC'07)*.

Cardoso, C., & Clarkson, P. (2012). Simulation in user-centred design: helping designers to empathise with atypical users. *Journal of Engineering Design*.

Cook, L. J., Knight, S., & Olson, L. M. (2005). A comparison of aggressive and DUI crashes. *Journal of Safety Research*, *36*(5), 491.

Deffenbacher, J. L., Lynch, R. S., Oetting, E. R., & Swaim, R. C. (2002). The Driving Anger Expression Inventory: a measure of how people express their anger on the road. *Behaviour Research and Therapy*, *40*(6), 717–737.

Diewald, S., Möller, A., Roalter, L., Stockinger, T., & Kranz, M. (2013). Gameful design in the automotive domain. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13* (pp. 262–265). New York, New York, USA: ACM Press. doi:10.1145/2516540.2516575

Dula, C. S., & Geller, E. S. (2003). Risky, aggressive, or emotional driving: Addressing the need for consistent communication in research. *Journal of Safety Research*, *34*(5), 559–566. doi:10.1016/j.jsr.2003.03.004

Fitz-Walter, Z., Tjondronegoro, D. W., & Wyeth, P. (2012, November 5). Encouraging learners to drive using game elements and smartphones. *Proceedings of the 24th Australian Computer-Human Interaction Conference (OzCHI 2012)*. ACM.

Fleiter, J., Lennon, A., & Watson, B. (2010). How do other people influence your driving speed? Exploring the "who"and the "how"of social influences on speeding from a qualitative perspective. *Transportation Research Part F: Traffic Psychology and ….*

Fraine, G., Smith, S. G., Zinkiewicz, L., Chapman, R. L., & Sheehan, M. C. (2007). At home on the road? Can drivers' relationships with their cars be associated with territoriality? *Journal of Environmental Psychology*, *27*(3), 204–214.

Gellatly, A., & Hansen, C. (2010). Journey: General Motors' move to incorporate contextual design into its next generation of automotive HMI designs. ... *Conference on Automotive ….*

Harris, H., & Nass, C. (2011). Emotion regulation for frustrating driving contexts. *Proceedings of the SIGCHI Conference on Human ….*

Jonsson, I. M., Nass, C., Harris, H., & Takayama, L. (2005). Matching In-Car Voice with Driver State : Impact on Attitude and Driving Performance. In *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 173–181).

Leckie, G. J., & Hopkins, J. (2002). The public place of central libraries: Findings from Toronto and Vancouver. *The Library Quarterly*, *72*(3), 326–372.

Lindgren, A., Chen, F., Amdahl, P., & Chaikiat, P. (2007). Using personas and scenarios as an interface design tool for advanced driver assistance systems. *Universal Access in Human- ….*

Lupton, D. (1999). Monsters in Metal Cocoons: `Road Rage' and Cyborg Bodies. *Body & Society*, *5*(1), 57–72. doi:10.1177/1357034X99005001005

Lupton, D. (2002). Road rage: drivers' understandings and experiences. *Journal of Sociology*, *38*(3), 275–290. doi:10.1177/144078302128756660

Marcus, A. (2004). The next revolution: Vehicle User Interfaces. *Interactions*, *11*(1), 40. doi:10.1145/962342.962358

Mitrevska, M., Castronovo, S., Mahr, A., & Müller, C. (2012). Physical and spiritual proximity: Linking Car2X Communication with Online Social Networks. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '12* (p. 249). New York, New York, USA: ACM Press. doi:10.1145/2390256.2390297

Nass, C., Jonsson, I.-M., Harris, H., Reaves, B., Endo, J., Brave, S., & Takayama, L. (2005). Improving automotive safety by pairing driver emotion and car voice emotion. In *CHI '05 extended abstracts on Human factors in computing systems* (pp. 1973–1976). New York, NY, USA: ACM. doi:10.1145/1056808.1057070

O'Brien, S. R., Tay, R. S., & Watson, B. C. (2004). Situational factors contributing to the expression of aggression on the roads. *IATSS Research*, *28*(1), 101–107.

Parkinson, B. (2001). Anger on and off the road. *British Journal of Psychology*, *92*(3), 507–526. doi:10.1348/000712601162310

Parkinson, B. (2008). Emotions in direct and remote social interaction: Getting through the spaces between us. *Computers in Human Behavior*, *24*(4), 1510–1529. doi:10.1016/j.chb.2007.05.006

Rakotonirainy, A., Feller, F., & Haworth, N. (2009). In-Vehicle Avatars to Elicit Social Response and Change Driving Behaviour. *International Journal of Technology and Human Interaction (IJTHI)*, *5*(4), 80–104.

Renge, K. (2000). Effect of driving experience on drivers' decoding process of roadway interpersonal communication. *Ergonomics*, *43*(1), 27–39. doi:10.1080/001401300184648

Schroeter, R., Rakotonirainy, A., & Foth, M. (2012). The social car : new interactive vehicular applications derived from social media and urban informatics. In *4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2012)* (pp. 107–110). Portsmouth, New Hampshire: ACM.

Schroeter, R., Soro, A., & Rakotonirainy, A. (2013). Social Cars : Sensing , Gathering , Sharing and Conveying Social Cues to Road Users. In B. Guo, D. Riboni, & H. Peizhao (Eds.), *Creating Personal, Social and Urban Awareness through Pervasive Computing*. IGI Global.

Soole, D., Lennon, A., Watson, B., & Bingham, C. (2011). Towards a comprehensive model of driver aggression: a review of the

literature and directions for the future. In C. N. Ferraro (Ed.), *Traffic Safety* (pp. 69–97). Nova Science Publishers, Inc.
Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, *47*(2), 143–148. doi:10.1016/0001-6918(81)90005-6
Viola, P., & Jones, M. (2011). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001* (Vol. 1, pp. 511–518). IEEE Comput. Soc. doi:10.1109/CVPR.2001.990517
Wollstädter, S., Pfister, H., & Höger, R. (2013). The role of music characteristics on drivers' anger regulation. *Arsrpe.acrs.org.au*.
Young, M. S., Birrell, S. A., & Stanton, N. A. (2011). Safe driving in a green world: a review of driver performance benchmarks and technologies to support "smart" driving. *Applied Ergonomics*, *42*(4), 533–9. doi:10.1016/j.apergo.2010.08.012

Experimental Studies on Driver's Expectations Regarding the Positions of Display and Control Elements in the Car Cockpit

Matti Schwalk, Stephan Hetzinger and Thomas Maier

Institute for Engineering Design and Industrial Design (IKTD)
Research and Teaching Department Industrial Design Engineering
University of Stuttgart, Germany

ABSTRACT

The number of functions within the car cockpit is increasing continuously independent from the class of vehicle. The driver's cognitive resources are mainly occupied by the traffic, while there are several other possible tasks which can be diverting. Driver's tasks include tasks of controlling the vehicle, checking relevant display information, adjusting navigation or infotainment systems etc. Despite these sources of diversion in safety relevant situations the driver must be able to find crucial display and control elements immediately and without confusion to avoid an accident. In these situations the driver is using evolutionary developed patterns of behaviour which are activated unconsciously and differing from learned system knowledge. The present work is meant to reveal driver's expectations regarding the positions of relevant interface elements in the car cockpit. Results are concluded on the basis of a standardised cockpit to give independent design recommendations. The effects and relevance of intuitively expected positions in comparison to unfavourable positions of different interface elements are verified by reaction time measurements.

Keywords: Driver's Expectations, Positions of Display and Control Elements, Car Cockpit, Interface Design

INTRODUCTION

Significance

An important example for a safety relevant traffic situation is the need of activating the hazard lights immediately if the car has stopped in case of damage or getting into a sudden traffic jam. Another scenario can be the urgent need of blowing the horn if being overseen by another driver. Beside these most important cases there are lots of further display and control elements which can influence the driver's attention and ability of focussing on the traffic. There are guidelines for ergonomic positioning and arrangement of display and control elements in the car cockpit, cf. primary and secondary control elements in DIN EN ISO 15005, but most of them are still not adapted to the current range of functions in modern car cockpits. In case of emergency it is absolutely necessary to find important interface elements intuitively to avoid an accident. Surely every driver should be aware of all important functions of his own car to be able to react adequately in critical situations. Though there is still a huge amount of vehicles which are used as rental cars or part of a vehicle fleet with almost daily changing drivers. In these cases it is even more important that most relevant interface elements can be found intuitively because drivers of these cars hardly check the whole system of the car before driving and cockpit designs vary between different manufacturers.

Modern Car Cockpit Design

Today's cars provide by far more functionality than necessary to accomplish the task of driving. While the first vehicles were limited to these basic functions, there have been integrated more and more additional elements like speedometer, tachometer, radio and heating followed by electric windows, air conditioning or phone- and navigation systems. Currently there is a huge progress in developing driver assistance systems as well as integrated infotainment including internet connectivity. The result of this technological development has been an increasing number of display and control elements in the car cockpit. To handle this complexity many manufacturers established a central rotary push control like the BMW iDrive or Audi MMI to perform many different operating tasks by using just one single control element. After recognising that this can cause some problems regarding intuitive operation leading to increased mental workload in some situations, most manufacturers started to add additional controls around the rotary push control again. Further technologies are head-up displays, touchpads, voice control and multifunctional TFT displays. In general there is a huge amount of functions in modern cars and there are also many possibilities of positioning displays and controls in the cockpit.

Standards and Guidelines

One of the most important tasks of the cockpit engineer is to make sure that all drivers of a corresponding profile are able to reach all crucial controls and operate safely while driving (Tiemann, 2005). There are several standards in designing car cockpits to achieve these demands as well as possible. After Tiemann (2005) a common procedure is the hand and foot reachability determination in accordance to guideline SAE J287 (Society of Automotive Engineers). Thereby mostly the digital human model of RAMSIS software by Human Solutions is used. After defining significance and frequency of use of crucial interface elements the virtual human model can be used to evaluate reachability in different situations (Tiemann, 2005).

Besides different SAE guidelines there are also some norms for the positioning of interface elements in the car cockpit, for example: BS ISO 4040 (2009): Road vehicles – Location of hand controls, indicators and tell-tales in motor vehicles. For hand reachability issues also ISO 3958 (1996) is used, while DIN EN ISO 15005 (2003) describes ergonomic standards regarding driver information and assistance systems. Gaze behaviour is addressed with DIN EN ISO 15007-1 (2002) and discussed below.

In general most of the described guidelines are out of date concerning recent technology standards and the huge amount of functions in modern car cockpits. This is why there are several differences regarding the positioning of displays and controls between various manufacturers and partially even within one brand due to different vehicle models and classes.

OBJECTIVE AND MOTIVATION

Gaze Behaviour

As far as there is no blind operation the driver's gaze always orientates towards the target respectively the control element before the operation itself. When reading a display first the area of the display is the target before the actual display information can be focused and recognized (DIN EN ISO 15007-1, 2002). Figure 1 shows an exemplary situation of reading a display while driving.

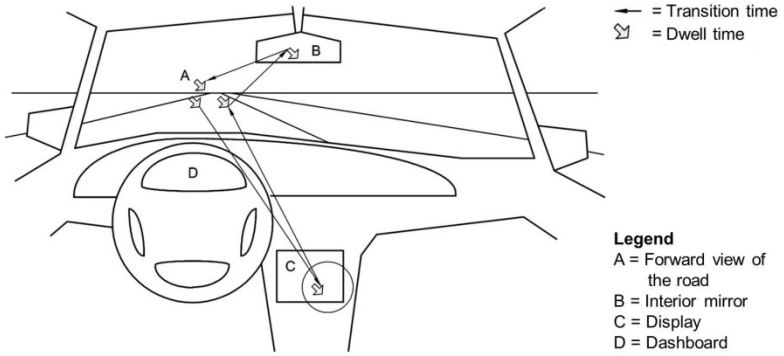


Figure 1: Gaze behaviour in car cockpit, based on DIN EN ISO 15007-1 (2002)

The duration of distraction consists of the transition times for finding an interface element (target C), getting back to the initial area (A) and the dwell time of the gaze on the respective display or control until the demanded information is encoded by the human information processing (figure 2).

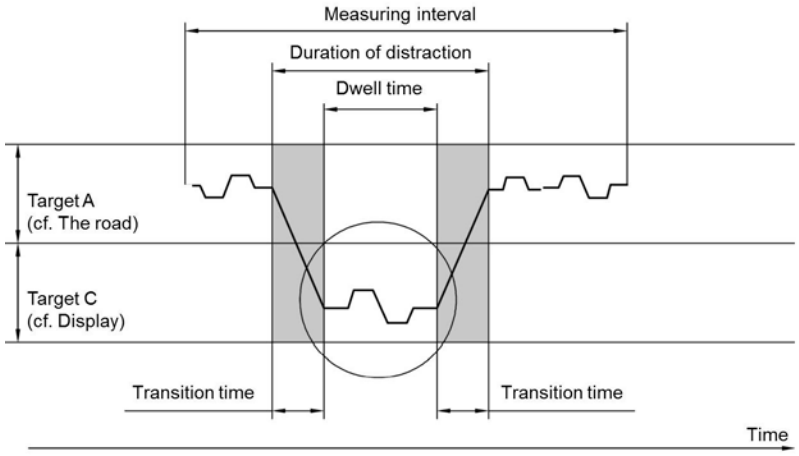


Figure 2: Duration of averting gaze off the road, based on DIN EN ISO 15007-1 (2002)

By ergonomic and intuitive cockpit design transition times as well as dwell times can be reduced to achieve a safer driving situation. Through the positioning of important and safety relevant displays and controls within the central visual field transition times can be minimized. Also a clear differentiation of interface elements from the vehicle interior by shape, colour and contrast can support short transition times.

Motivation

A survey by the German Automobile Club ADAC has shown that even since the late 1980s many drivers have stated, that "inconsistent, sometimes difficult usability" is bothering them the most about their cars (Färber & Färber 1987). Of course in safety relevant situations it is not acceptable that the driver is struggling with the usability of his vehicle while he should focus on the traffic situation and be able to act and react intuitively. Intuitive operation requires proper design and positioning of interface elements in a way, that persons who haven't been driving for a longer period of time or switching to another type of vehicle are able to manage the operating safely and clearly at all times (Färber & Färber 1987). This kind of intuitive design is the motivation for the present work. Besides many standards and guidelines regarding ergonomic cockpit design the aim is to reveal the driver's intuitive expectation of where crucial displays and controls are positioned, when faced with different cockpit designs. The driver's expectations are based on habits including practice, learning and frequent execution (Enzyklo, 2013a) or pure intuitive determination meaning the ability of the human mind to register a situation immediately (Enzyklo, 2013b).

EXPERIMENT PREPARATION

Approach

The presented experimental studies on driver's expectations regarding the positions of display and control elements in the car cockpit were prepared with a systematic selection of different vehicle classes and manufacturers. Relevant interface elements were classified in terms of frequency of use and importance respecting safety issues. The classification is segmented into primary, secondary and tertiary driving tasks and stationary tasks which are accomplished while the vehicle is not in motion. In total the positions of 21 interface elements were examined within a selection of 15 different virtual car cockpits recorded from driver's viewing perspective.

Classification of Driving Tasks

Based on the definition of primary and secondary driving tasks by Jürgensohn & Timpe (2001) the following classification has been established for the present work:

- **Primary driving tasks:** Actual vehicle control in longitudinal and transverse direction, divided into planning, manoeuvring and stabilising.
- **Secondary driving tasks:** Support of the driver in primary driving tasks, e.g. activating indicators, wipers or headlights.
- **Tertiary driving tasks:** Activities and functions which are not involved in the vehicle control (primary driving tasks) including air conditioning, infotainment or navigation system.
- **Stationary tasks:** Activities which are normally performed only when the car is not moving, e.g. adjusting the positions of seat and mirrors or activating the parking brake.

Basically the secondary and tertiary driving tasks must not affect the driver or distract her or him from the actual primary task of driving safely within the complex traffic situation. Therefore displays and controls regarding secondary and tertiary driving tasks should be found quickly in the driver's expected positions.

Systematic Selection of Relevant Interface Elements

The considerations above in addition with the frequency of use according to Löffler (2010) lead to the following 21 interface elements shown in Table 1, which are going to be examined. The list also includes two future-oriented functions (No 20 and No 21).

Table 1: Selection of interface elements for examination

No	Function / Interface element	No	Function / Interface element
1	Hazard lights	12	Speedometer
2	Ignition / Start-Stop	13	Cruise control
3	Parking brake	14	Navigation advice
4	Horn	15	Air circulation
5	Headlights (on / off)	16	Voice control
6	Window lifter	17	Adaptive cruise control (automatic)
7	Exterior mirror adjustment	18	Lane departure warning
8	Central locking	19	Parking assistant
9	Filler cap unlocking	20	Operating mode: Electric motor
10	Bonnet unlocking	21	Autonomous driving
11	Audio (+/-) (</>)		

Systematic Selection of Relevant Vehicles

The selection of crucial car cockpits for the examination is based on a systematic analysis of current vehicles (models from 2009 to 2014). As a result 15 car cockpits are selected including different vehicle classes, manufacturers and formal designs. Also important for the experimental studies is a comparable view and image section of the cockpits. To simulate a realistic driving situation the driver's view is selected for all 15 cockpit designs (cf. figure 3).

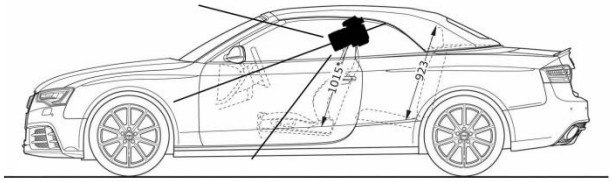


Figure 3: Defined view for the presentation of different cockpits (Automobilesreview, 2013)

The presented virtual car cockpits were graphically edited to remove all originally existing interface elements. Thus participants had the opportunity to mark their expected positions without being influenced (figure 4).

Figure 4: Graphically edited car cockpit with removed interface elements

EXPERIMENTAL STUDIES

Collective of Test Persons

Prior to the actual experiment the participants filled in a questionnaire about their driving behaviour and experience. For the present studies a collective of 33 test persons (12 female, 21 male) with an average age of 37.8 years have been interviewed. 37% of the participants are regularly using more than one car and the average driven distance within one year is 14062 kilometres. The average age of the cars is 7.8 years. 21% of the test persons stated that they have already driven more than 10 different cars for at least a few months, 49% stated 5 to 10 cars, about 30% stated less than 5 cars. 24% often or very often use an unfamiliar car which they are not used to, while just 21% of them extensively familiarise themselves with the new cockpit environment before they start driving. 45% state that the number of interface elements is too high while only 6% say it never occurs that they can't find a display or control element. In other words 94% of the test persons are rarely, often or very often not able to find some of the interface elements they want to use. Particularly hazard lights (stated by 30% of the participants) and headlights (stated by 21%) are mentioned most frequently in this context.

Procedure

After being informed about the subject of the experiment and filling in the questionnaire the participants were asked about their expectations regarding the positions of the 21 selected interface elements in the 15 edited car cockpits. The study was performed with a tablet computer and starting with an introduction page, where the test persons could try out the tablet pen. After that the experiment started. To simulate getting into the car there has always been an exterior view of the vehicle first, followed by the graphically edited cockpit in combination with an advice which interface element position should be marked with the pen. To avoid an extraordinary period of time for the participants not all of the 21 interface element positions were asked within all of the cockpits. Instead within each cockpit a varied selection of the 21 elements was enquired to also randomise the influence of respective cockpit designs.

RESULTS AND DESIGN RECOMMENDATIONS

Interpretation of Results

For each participant all marked positions were recorded and visualised, which is exemplary shown in figure 5 for the button of the hazard lights in four cockpits. Red dots indicate the expected positions of different test persons, green dots additionally indicate the desired positions and the blue ring shows the real position in the respective car cockpit. Original images are taken from Automobilesreview (2013).

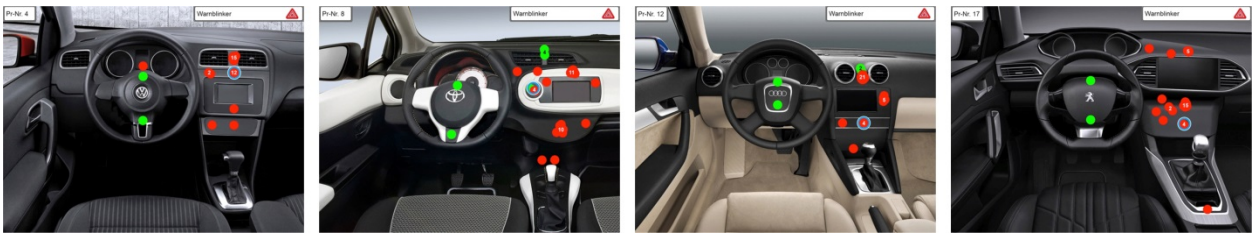


Figure 5: Car dependent results of expected position of hazard lights button

A significant discrepancy between expected and actual position could e.g. be detected in the Porsche 911 Carrera cockpit as shown in figure 6.

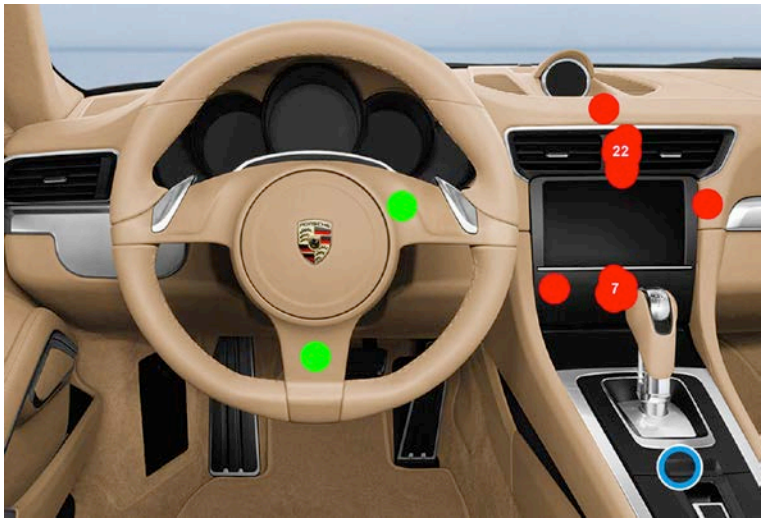


Figure 6: Expected position of hazard lights button in Porsche 911 Carrera (Original image from Automobilesreview, 2013)

Design Recommendations

To derive design guidelines from the results a standardised independent car cockpit with defined grid was generated to merge all data (figure 7).

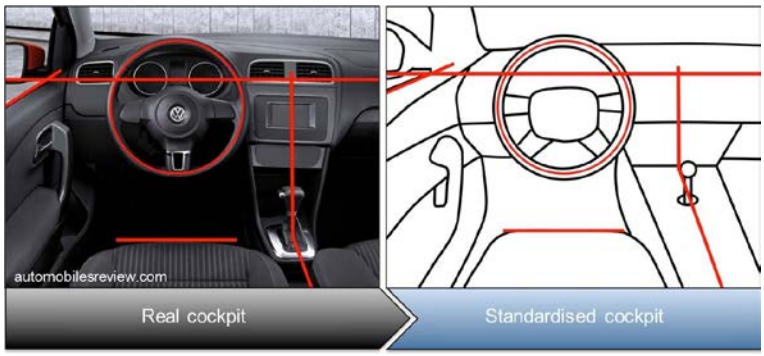


Figure 7: Original and standardised car cockpit

The standardised cockpit is used to gather the data of all test persons and cockpits for each interface element. The procedure is divided into the following four steps (exemplary shown for hazard lights in figure 8):

1) Superposition of all results for the expected position of a specific interface element
2) Representation of mentioning concentration by half transparent layers
3) Percentage of mentioning of expected car independent position in defined areas
4) Design recommendation for positioning of specific interface element

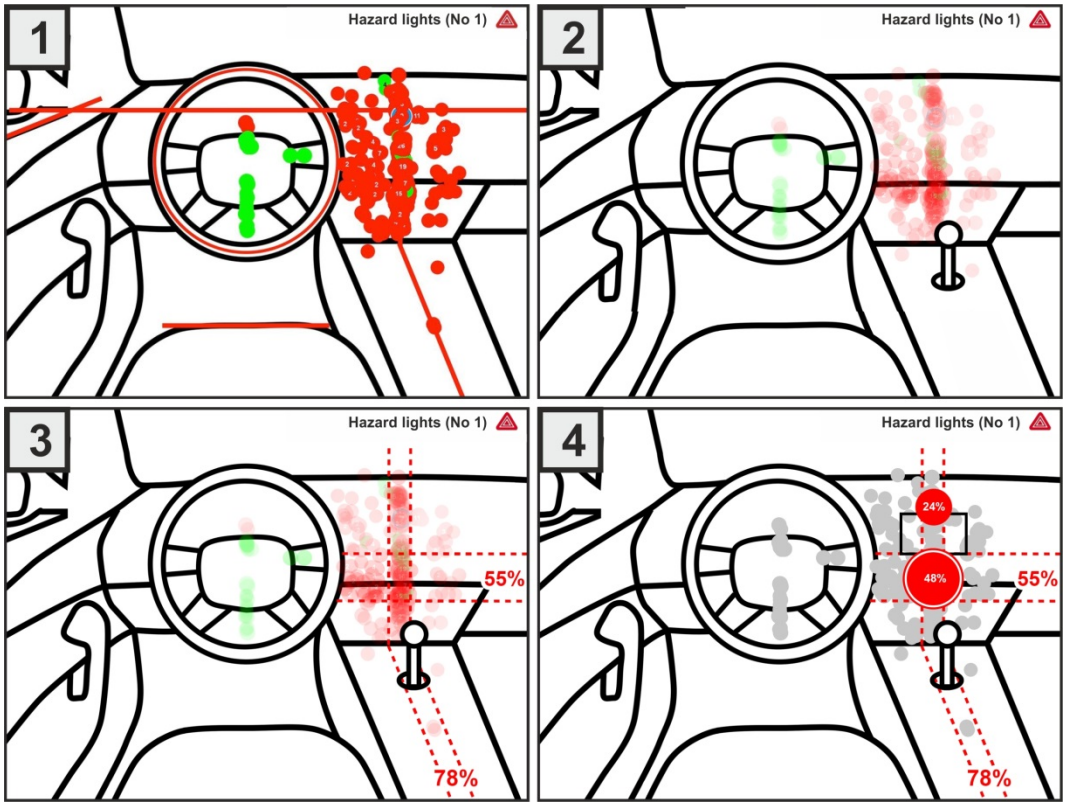


Figure 8: Process of deriving car independent design recommendations

Examination results are concluded as areas of expected positions for each interface element. All derived design recommendations for the selected 21 interface elements in the car cockpit are shown in the following figures 9, 10 and 11.

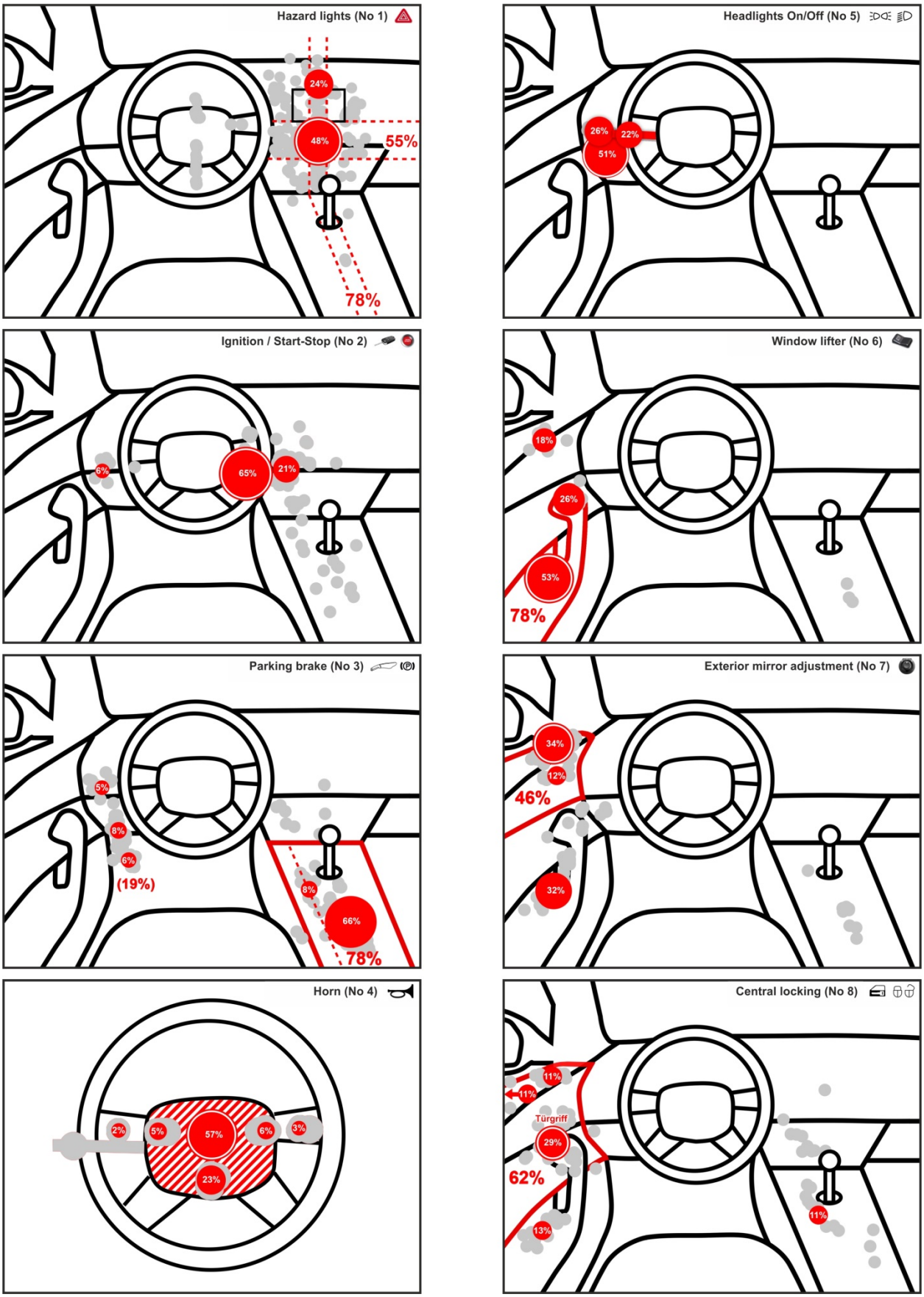


Figure 9: Design recommendations for interface elements 1 to 8

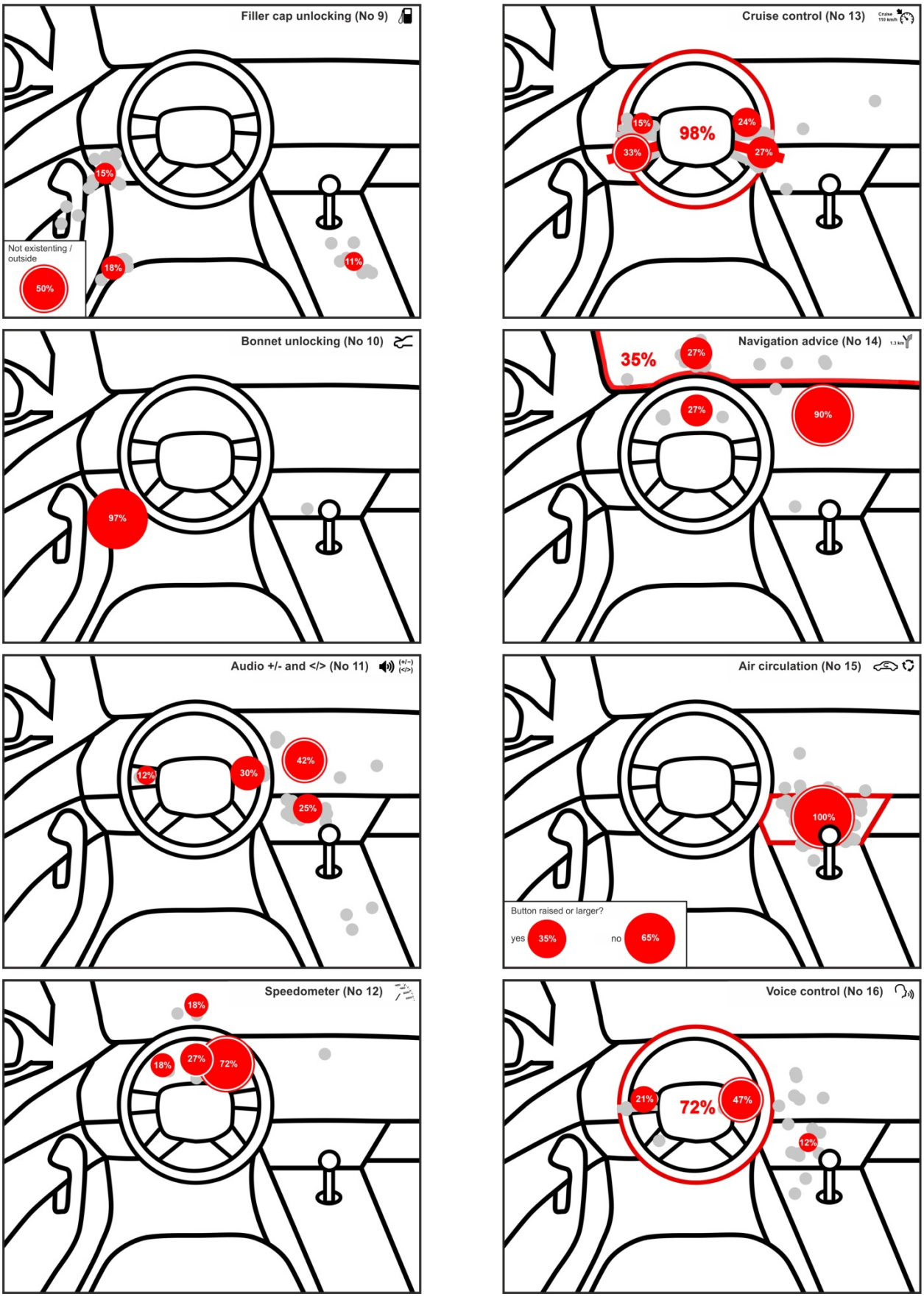


Figure 10: Design recommendations for interface elements 9 to 16

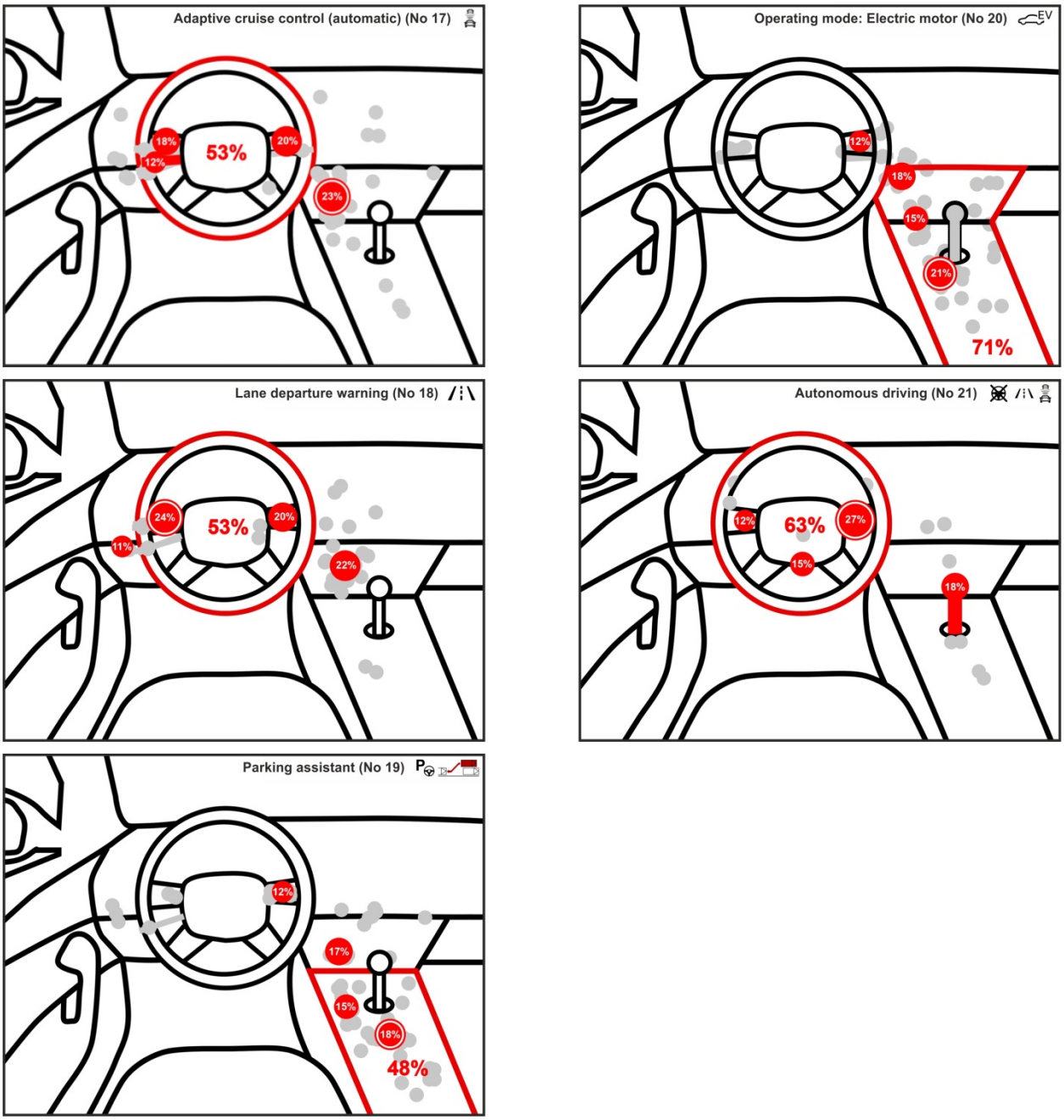


Figure 11: Design recommendations for interface elements 17 to 21

Verification by Reaction Time Measurement

In a further study with 6 test persons (2 female, 4 male, average age 36.7 years) the effects of good and bad positioned interface elements were examined to make a first step of verifying the derived design recommendations in a real car cockpit environment. In this study the participants were asked to find selected displays and controls during a simulated driving situation. 14 crucial and common interface elements were chosen and each once placed in a good position, once in a bad position regarding the design recommendations. Interface elements which are unfamiliar or always expected in the same area are excluded. For the experiment display and control elements are printed on photo paper and cut out to be placed in the cockpit of a Toyota RAV4 (from 2009). To generate a situation close to reality also distracting buttons are placed in the cockpit. Real interface elements are either included in the examination or covered respectively visually neutralized. In total three cockpit setups are generated, each

containing good and bad positioned interface elements. The order of query is defined in a way that the participant preferably doesn't recognise interface elements which are going to be asked for later on, while searching for the current one. To simulate the driving situation a video is shown on a tablet computer placed in front of the steering wheel. Thus the gaze of the test person is directed at the virtual road before being asked to search for a respective display or control. During the experiment reaction time is measured. The order of the three cockpit setups is randomised between all participants. Figure 12 shows the results of reaction time measurements.

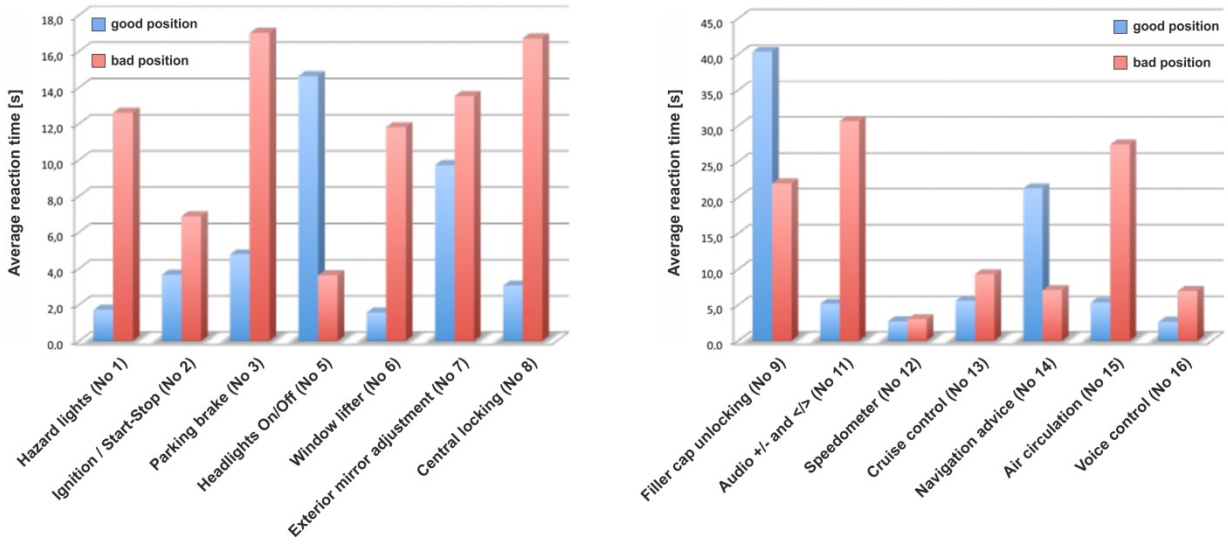


Figure 12: Results of reaction time measurements

CONCLUSIONS

Overall most of the results are hardly surprising. Many of the expected positions of displays and controls in the car cockpit have either evolved historically or are caused by functional aspects like the spatial proximity to the functional component. On the other hand there are still some current cockpit designs containing interface elements in quite unfavourable and unexpected positions. Regarding safety relevant functions like hazard lights or horn this can lead to dangerous traffic situations. Also long transition times are causing unnecessary distraction, if crucial interface elements cannot be found within a few seconds. Furthermore during the second study it was detected that test persons partially react stressed when they cannot find the respective display or control quickly. Thus an intuitive positioning also supports the driver's joy of use. Of course the expectations depend on the driver's individual experience with cars because there are several manufacturer-specific operating concepts. Despite of these facts it would be useful to standardise the positioning of crucial safety relevant interface elements more strictly. For example it is not acceptable that some manufacturers almost hide the button for activating hazard lights somewhere between other similar buttons or in other areas where it wouldn't be expected by the majority of drivers, e.g. if it is not centred in the cockpit. In general the number of driver assistance systems is increasing continuously and in addition with new drive concepts the complexity of car cockpits increases as well. Although the modern vehicle is able to undertake more and more tasks from the driver autonomously the knowledge about the positions of crucial displays and controls is still important and necessary, especially in case of failure. With the present work it has been accomplished to detect relatively small areas where specific interface elements are expected to be positioned by a majority of questioned drivers. If there are no important interface elements placed in unexpected positions the basis for safe and user-centred cockpit design is given. In this way also unfamiliar cars could be used without obstacles, e.g. rental cars or cars of a vehicle fleet which are additionally often used under time pressure without the opportunity of familiarising with the cockpit before driving. Further examinations should focus even more on the empirical analysis of the influence of structure and shaping of different car cockpits on the expected interface positions. Thereby the number and arrangement of displays and controls should be considered in detail as well.

REFERENCES

Automobilesreview (2013): "*High Resolution Car Picture Galleries*". Website: http://www.automobilesreview.com/pictures – accessed 19.05.2013

DIN EN ISO 15005 (2003): "*Straßenfahrzeuge – Ergonomische Aspekte von Fahrerinformations- und -assistenzsystemen*".

DIN EN ISO 15007-1 (2002): "*Messung des Blickverhaltens von Fahrern bei Fahrzeugen mit Fahrerinformations- und -assistenzsystemen – Teil 1: Begriffe und Parameter*".

Enzyklo (2013a), Online Enzyklopädie: "*Begriff Gewohnheit*". Website: http://www.enzyklo.de/ Begriff/gewohnheit – accessed 09.05.2013.

Enzyklo (2013b), Online Enzyklopädie: "*Begriff Intuition*". Website: http://www.enzyklo.de/ Begriff/intuition – accessed 09.05.2013.

Färber, Be. Färber, Br. (1987): "*Sicherheitsorientierte Bewertung von Anzeige- und Bedienelementen in Kraftfahrzeugen*". In: FAT Schriftenreihe Nr.64. Frankfurt am Main: Forschungsvereinigung Automobiltechnik e.V. (FAT)

ISO 3958 (1996): "*Straßenfahrzeuge, Personenkraftwagen – Handreichweiten des Fahrzeugführers*".

Jürgensohn, T. Timpe, K.P. (2001): "*Kraftfahrzeugführung*". Springer-Verlag, Berlin.

Löffler, S. (2010): "*Priorisierung von Interfacefunktionen in Fahrzeugcockpits der Kompaktklasse*". Student research project, University of Stuttgart: Institute for Engineering Design and Industrial Design (IKTD)

Tiemann, R. C. (2005), "*Cockpits im Automobil – von der Idee zum Insassen*". Expert-Verlag, Renningen.

Evaluation of Telematic Applications for Information Provision in Public Transport

Andrzej Szarata[a], Olga Kokkinou[b], Socrates Basbas[b] and Ioannis Politis[b]

[a]Kracow University of Technology
ul. Warszawska 24
Krakow, 31-155, POLAND

[b]Aristotle University of Thessaloniki
University Campus
Thessaloniki, 54-124, GREECE

ABSTRACT

This paper investigates various parameters that are related with the information provided by telematics applications in public transport. The telematics applications examined are a) a web site portal that provides real time information to PT users, b) a smartphone application that utilizes the advantages of location identification and c) an interactive voice response (IVR) system. The examination was performed through a revealed and stated preference questionnaire-survey that was addressed to public transport users in the city of Thessaloniki, Greece. The parameters investigated are related with the socioeconomic background of the public transport users, their mobility patterns and attitudes as well as with stated opinions about changes that telematics applications may have on their current travel behavior. Finally, a linear multiple regression model was developed, indicating that the parameters which may have an impact on the perceived cost of these application (as a percentage of the total fare level) are related with the application itself, the age of the PT user as well as with his/her educational level.

Keywords: Telematics Applications, Public Transport, Evaluation, Linear Multiple Regression

INTRODUCTION

The effects and the impacts of the provision of (real-time) information to public transport users, has been examined by numerous studies. Katrin & Kottenhoff, 2007, identified 7 main effects that real-time information may have to the Public Transport customers; on their study, they found that a 20% reduction of the perceived waiting time can be achieved through the installation of dynamic Variable Message Signs at stops. Furthermore, they found that these installations may have impact on adjusted walking speeds. The effect of passengers' waiting time on Public Transport assignment to lines and routes and the value of information to schedule Public Transport were extensively discussed on a paper by Odd & Sunde, 2008. The benefits of Real Time Passenger Information Systems and the way these systems affect the perceived quality of Public Transport have been recently explored by Monzon, Hernadez & Cascajo, 2013. Finally, in a paper by Guo, 2011, the relationship between the conventional transit maps and the Advanced Transit Information Systems and their effect on passengers' travel decision are examined in a comparative manner.

Interactive Voice Response (IVR) systems have been traditionally used by public transport operators to support

customers on their trip Casey et al., 1998. These systems tend to have limited use the last decade since information provision is gradually replaced by new developed telematics applications and therefore few research studies which analyze the impact of IVR systems on travel behavior can be identified in the literature (Nelson & Mulley, 2013).

Web based public transport information systems are the pioneer applications for real time information provision. In most cases, an advance Web Geographic Information System (web-G.I.S.) is developed and among other services, these applications are primarily used for path finding and optimization (Peng & Ruihong, 2000 and Ziliaskopoulos & Travis, 2000). From simple one-mode applications, these applications are now providing multimodal path services with various real cases implementations worldwide (Su & Chih-hung, 2010). The price sensitivity as well as the willingness to pay for web-enable Public Transport Information Services was discussed at papers by Molin & Harry, 2006 and Zheng et al, 2010.

The continuous increase of mobile usage is also considered as a new entry key player on the information provision. The ability of smartphones to locate the position of the user anytime, as well as the development of high technology applications that can be downloaded for free by smartphone users, gave the opportunity for provision of customized/personalized accurate information with numerous positive effects. Watkins et al, 2011 found that mobile real time information provision reduces not only the perceived waiting time at bus stops, but also the actual waiting time. New implications of smartphone usage, such as the Crow Sourcing, and how these can affect the routing and the path change of Public Transport users are being examined (Velaga et al, 2013).

Hickman & Wilson, 1995 explored the impact that real time information provision may have on path choice between public transport users. On a simulation study, various origin-destination paths were examined and a behavioral model of Public Transport path choice was developed. Based on their analysis, which was published in 1995 where real time information provision was limited and given only through traditional methods, they did not found valuable quantitative benefits that can be gained by the user.

Another important factor, which will not be discussed in detail in this paper, is the content of the information. One study by Abdel-aty, 2001 showed that frequency of service, number of transfers, seat availability, walking time to the transit stop and fare information are among significant information types that commuters desire. Almost most of these services are provided nowadays from the variety of the telematics applications

The paper examines the impact of three telematics applications installed by the Organization of Urban Transport of Thessaloniki, Greece (OASTH) on travelers' behavior. The applications are providing real time travel information on issues related with the routing and scheduling of the buses, aiming to increase in this way the attractiveness of Public Transport (P.T.) against private cars. The applications being examined are a) the website portal of the organization b) a smartphone application and c) an interactive voice response (IVR) system. Three revealed preference questionnaire surveys were conducted respectively in order to assess the level of awareness and acceptance of these applications. The research corresponds to preliminary descriptive and inferential statistic analyses and investigates a number of parameters that are related both with the passenger (e.g. age, income etc) and with the application itself and examines their potential for forcing travelers to change their current travel behavior e.g. by selecting more sustainable ways of travelling. Furthermore, a linear multiple regression model was developed in order to assess the cost of these applications as it is perceived by the customers.

THE TELEMATICS APPLICATIONS

As it is mentioned in the previous section, the telematics applications examined in this paper are a website portal, a smartphone application and an interactive voice response system. All these applications have been developed, operated and regularly updated by the Organization of Urban Transport of Thessaloniki (OASTH) which is the sole Public Bus Transport Operator in the conurbation of Thessaloniki. The objective of this section is to briefly present the three applications and their respective services.

The *website* portal, directly accessible through the address http://www.oasth.gr, is a dynamic webpage which provides information about: bus routes, optimal route search, maps of bus lines (though Google maps), real time information about bus arrivals at a selected bus stop, real time depiction of buses in Google maps, points of interests, duration of a route etc. According to OASTH, in December 2010, almost 250,000 unique visitors had visited this webpage (OASTH, 2010).

Similar information is provided also by the *smartphone* application which can be easily downloaded by android or iphone users. The services are directly accessible through the address http://m.oasth.gr. Figure 1 shows screenshots from the telematics application of the website portal and the smartphone respectively.

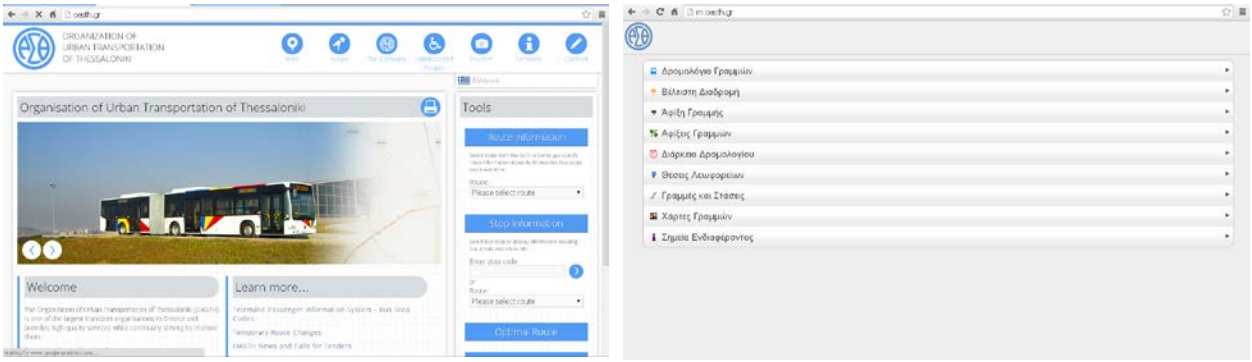


Figure 1:. The telematics application of the website portal and the smartphone application of OASTH respectively

Finally, the *interactive voice response* system is accessible though the telephone center of OASTH. A voice record menu is available and guides the listener to real-time services such as arrivals of a bus line to a specific stop, information about the bus stops of a specific bus line, which bus lines are passing through a specific bus stop etc. According to OASTH, more than 2,500 people had been served by the IVR system, in January 2011 (OASTH, 2011).

RESEARCH METHODOLOGY

The Questionnaire Survey

In order to evaluate the effects and the impacts of these applications on peoples' mobility and in order to examine in what extent these applications are perceived by the customers as useful and important, a questionnaire survey was carried out addressed to the public transport users of the city of Thessaloniki (Kokkinou, 2011).

The questionnaire survey was conducted at three bus stops which are located at the city center. The bus stops were selected in order to primarily avoid bias at the sample population; the city center is daily attracting high number of people with various socioeconomic, travel and behavioral profiles. Additionally, these bus stops serve large number of people and therefore it was easier to collect the necessary number of questionnaires. The sample size was calculated by using the following traditional sample size calculation formula (ASTM, 2013):

$$N = \frac{Z^2 * (p) * (p-1)}{c^2}$$

where:

Z = Z value (e.g. 1.96 for 95% confidence level)
p = response distribution (percentage picking a choice expressed as decimal, usual 0.5 used for normally distributed responses)
c = the amount of uncertainty (confidence interval, expressed as decimal, for example 0 .04 represent a tolerance of ±4)

At the survey design phase, the assumptions made to predefine the sample size where: confidence interval (c) equal to 0.06, response distribution (p) equal to 0.5, confidence level equal to 95%. The targeted population was estimated to 250,000 since the modal split for Public Transport at the Greater Agglomeration of the City of Thessaloniki (populated over 1,000,000 inhabitants) is estimated to 24-25%. With these assumptions the sample size was calculated to 267 cases. Finally, 319 valid questionnaires were gathered, returning a margin of error (confidence interval) equal to approximately 5%.

The questionnaire itself, consisted of 18 questions and it was divided into five discrete sections; at the first section, the socioeconomic profile of the respondents was identified (sex, age, education etc), at the second section the travel characteristics of the respondents were determined (e.g. whether the respondent is using his car for daily trips or not, the type of the ticket is usually used), at the third section, the experience of the customer related to the three telematics applications was examined (e.g. how often they are using these applications, whether they find the information provided useful and reliable, what type of information usually want), at the fourth section , the impact that these applications may have on traveler behavior is explored (e.g. how possible is to shift from car to bus due to the applications and how possible is to make additional trips due to the existence of these applications). Finally, at the last section, questions regarding the willingness to pay (WTP) were included so as to identify the level of acceptance for the examined applications.

It was decided to collect equal responses (questionnaires) for each one of the three applications. Therefore, from the 319 collected questionnaires, 103 questionnaires were for the mobile application, 105 questionnaires for the IVR application and 111 questionnaires for the website application.

For the purposes of the study, the questions of the questionnaire were coded into specific variables. Table 1 show the 13 variables which were used in this study, together with a short description and their measurement scale.

Table 1: Variables used in the study

ID	Variable Code	Description	Measurement Scale
1	**PLACE**	Bus Stop of the Questionnaire Survey	1: Bus Stop 1, 2: Bus Stop 2, 3: Bus Stop 3
2	**TYPE**	Telematic Application Type Being Examined	1: Mobile, 2:IVR, 3:Web
3	**GENDER**	Gender of respondent	1: Man, 2: Women
4	**AGE**	Age of respondent	1:<24, 2:25-39, 3:40-54, 4:55-64, 5:>65
5	**EDUC**	Educational Level of the Responder	1: Primary, 2:Secondary, 3:Technological, 4:University, 5:Master, 6: PhD
6	**CAR_US**	Do you use your car for daily trips?	1: yes, 2: no
7	**TICKET_T**	Type Ticket Usually Used	1: Single, 2:Double, 3:Multiple, 4:Discount
8	**TELE_US**	Have you ever used the Telematic Application?	1: yes, 2: no
9	**EASINESS**	Was it easy to get the information that you want?	1: Yes, 2: Do not Know, 3:No
10	**RELIABLE**	Did you found this information Reliable?	1: Yes, 2: Do not Know, 3:No
11	**CAR2BUS**	Is this application a reason to shift from car to bus?	1: yes, 2: no
12	**AD_TRIPS**	Is this application a reason to make additional trips by bus?	1: yes, 2: no
13	**TICKET_P**	Percentage of Ticket Price correspond to this application	Original response

From the Table above, it can be concluded that only the variable TICKET_P is considered as continuous. All other variables are coded as nominal with the exception of the variables AGE, EDUC, EASINESS and RELIABLE which are considered as ordinal.

Descriptive and Inferential Statistics

The variables of the study were manipulated through the statistical programs of SPSS and LISREL. Table 2, presents the descriptive statistics of the 13 variables together with the normality tests of Skewness and Kurtosis. Although the table indicates a kurtosis for the variables of additional trips (AD_TRIPS) and percentage of ticket that corresponds to the cost of the application (TICKET_P), it was decided not to apply any traditional transformation of these variables since the values are not high enough and the Q-Q plots indeed did not indicated any serious violation of the normality. Additionally, the mean values of the variables that are related with the socioeconomic characteristics of the respondents (e.g. 1.6 value for gender, 2.37 value for age, 3.18 value for educational level) confirms the initial willingness to equally represent various population segments in the study.

Table 2: Descriptive statistics and normality tests

Variable	N	Min	Max	Mean	S.D.	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
PLACE	319	1	3	1,91	0,823	0,176	0,137	-1,501	0,272
TYPE	319	1	3	2,03	0,82	-0,046	0,137	-1,511	0,272
GENDER	319	1	2	1,6	0,49	-0,418	0,137	-1,837	0,272
AGE	319	1	5	2,37	0,975	0,656	0,137	0,074	0,272
EDUC	317	1	6	3,18	1,056	0,194	0,137	-0,646	0,273
CAR_US	319	1	2	1,53	0,5	-0,107	0,137	-2,001	0,272
TICKET_T	319	1	4	1,92	1,235	0,737	0,137	-1,237	0,272
TELE_US	319	1	2	1,71	0,454	-0,939	0,137	-1,126	0,272
EASINESS	319	1	3	1,46	0,632	1,061	0,137	0,035	0,272
RELIABLE	319	1	3	1,31	0,515	1,369	0,137	0,9	0,272
CAR2BUS	253	1	2	1,87	0,342	-2,157	0,153	2,673	0,305
AD_TRIPS	318	1	2	1,9	0,297	-2,727	0,137	5,471	0,273
TICKET_P	295	0	0,640	0,082	0,080	2,899	0,142	12,422	0,283

Table 3, presents the various correlations (pearson product-moment, tetrachoric, polychoric, contingency etc) between the 13 variables, together with the statistical significance for two tailed probabilities.

Table 3: Correlation matrix and two tailed probabilities

Variable/ID	1	2	3	4	5	6	7	8	9	10	11	12	13
PLACE	1												
TYPE	,03	1											
GENDER	,00	-,14[c]	1										
AGE	-,06	-,01	-,24[a]	1									
EDUC	,06	-,09	,14[c]	-,44[a]	1								
CAR_US	-,20[a]	-,11[c]	,53[a]	-,08	-,18[b]	1							
TICKET_T	-,04	,06	,35[a]	-,37[a]	-,05	,56[a]	1						
TELE_US	-,09	-,41[a]	,09	,26[a]	-,37[a]	,27[a]	-,02	1					
EASINESS	-,21[a]	-,08	,04	,49[a]	-,35[a]	,17[b]	-,09	,78[a]	1				
RELIABLE	,03	-,14[c]	-,08	,25[a]	-,03	-,08	-,15[b]	,55[a]	,57[a]	1			
CAR2BUS	-,18[a]	-,10	,01	,08	-,37[a]	,22[a]	-,18[b]	,57[a]	,61[a]	,94[a]	1		
AD_TRIPS	-,19[a]	,07	-,29[a]	-,01	-,32[a]	,05	-,15[b]	,43[a]	,43[a]	,55[a]	,82[a]	1	
TICKET_P	-,04	,20[c]	,04	-,01	-,06	,04	,20[a]	-,01	-,05	-,11	-,02	-,12[c]	1

a) p <0.001, b) p<0.01, c) p<0.05

The presentation of the correlations between the variables, can lead to numerous useful outcomes. Regarding the selection of the bus stops, it seems that there is a statistically significant difference at the responses related to the car usage (r = -.20, p<0.001), the way people are perceiving the given information (r = -.21, p<0.001) and the potential to change their current travel behavior (r = -.18, p<0.001 for replacing car trips with the bus and r = -.19, p<0.001 for making additional trips with the bus to the existence of the telematics applications). No violation on the assumption of equal representation of the socioeconomic background between the respondents is detected within the bus stops. The above indicate that indeed no differences can be observed regarding the socioeconomic profile between bus stops that are located very close each other and in the city center; however strong variation is observed on the travel characteristics of the respondents and the impact that telematics application can have on their travel behaviour.

Regarding the telematics applications, differences have been observed on the gender (r = -.14, p<0.05) and this outcome is obviously related with the sampling distribution between the three types of the examined applications, the usage of the car for daily trips (r = -.11, p<0.05), the experience of the respondent to the specific application type (r = -.41, p<0.001), the level of perceiving the information provided as reliable or not (r = -.14, p<0.05) and the level of the perceived cost needed to develop and operate each one of the three applications (r = .20, p<0.05). Table 4 presents the distribution of the responses for the four variables that was found to have a statistically significant difference between the telematics applications. From this Table it can be concluded that the people who use their car for their daily trips are those who detected most on the website questionnaires. Furthermore, as expected, this application has the broadest usage between the users (almost 51% has an experience on using the website portal) and it is considered as the telematics application with the most reliable information (more than 78% believe that the information is reliable). However, respondents believe that the cost of the website application is greater compared to the other two applications (mean perceived cost as a percentage to the total fare which is 0.8 euros).

Table 4: Statistically different responses related to the telematics applications

Variable Code	Value	Application Type					
		mobile		ivr		web	
		Value	%	Value	%	Value	%
Car_Us	Yes	49	47,6%	48	45,7%	54	48,6%
	No	54	52,4%	57	54,3%	57	51,4%
Tele_Us	Yes	18	17,5%	18	17,1%	56	50,5%
	No	85	82,5%	87	82,9%	55	49,5%
Reliable	Yes	63	61,2%	78	74,3%	87	78,4%
	Dont Know	40	38,8%	26	24,8%	17	15,3%
	No	0	0,0%	1	1,0%	7	6,3%
Ticket_price	Mean/%of Ticket	,079	9,9%	,064	8,0%	,103	12,8%

Regarding gender differences, Table 3 indicates a strong differentiation in the age levels (r = -.24, p<0.001) with the males at the sample to be older compared to females, with a lower educational level (r = .14, p<0.05). As expected, females stated that they do not use their car for daily trips so often compared to males (r = .53, p<0.001) and prefer tickets for multiple trips (r = .35, p<0.001). Finally, males stated more frequently compared to females that the existence of the telematics applications will not motivate them to make additional trips (r = -.29, p<0.001).

Regarding statistically significant differences between age segments, the correlation matrix confirms the expected outcome that older users do not use very often the telematics applications (r = .26, p<0.001), have problems in easily understanding the information provided by the applications (r = .49, p<0.001) and do not trust the content of this information (r = .25, p<0.001), compared to younger users.

Amongst other valuable outcomes delivered from the correlation matrix is that users that do not often use the telematics applications do not intent to change their current travel behavior (using more often bus instead of cars - r = .57, p<0.001- and make additional trips by bus - r = .43, p<0.001) as well as the fact this change is not likely to happen when users meet difficulties to understand the given information (r = .61, p<0.001) or believe that this information is not accurate (r = .43, p<0.001)

THE LINEAR MULTIPLE REGRESSION MODEL

Model Specification

Following the outcomes that have been derived from the previous sections, it was decided to develop a regression model, in order to estimate which parameters - and in which manner - can have an impact on users' perception regarding the cost of this telematics application. As already discussed, the cost of these applications was identified as a perceived percentage of the passenger fare that corresponds to the development and operation of these

applications (total ticket fee for a single ticket equals to 0,8 euros).

Therefore, the proposed model has as dependent variable, the variable TICKET_P and as independent variables the total of the variables presented in Table 1. Assuming a linear relationship between the dependent and the independent variables, a linear multiple regression model was developed and tested.

Since all independent variables are non-continuous scale (ordinal and nominal), they all have been recorded as dichotomous with a reference category each time. Table 5 gives the reference category for each ordinal scale variable

Table 5: Recoding of Ordinal Variables

Variable Code	Description	Reference Category
TYPE	Telematics Application Type Being Examined	IVR
AGE	Age of respondent	<24
EDUC	Educational Level of the Responder	Primary
TICKET_T	Type Ticket Usually Used	Single
EASINESS	Was it easy to get the information that you want?	Yes
RELIABLE	Did you found this information Reliable?	Yes

Model Estimation

For the model estimation, a stepwise forward procedure was applied in order to estimate which variables have a statistically significant impact on the dependent variable TICKET_P. Table 6, presents the beta estimates (standardized and unstandardized) of these variables which are a) the web application (compared to the reference category of IVR), b) the people aged more than 65 (compare to the reference category of people aged less than 24 years old), and c) the users holding a Master Degree (compared to those have completed Primary education). Further, the calculation of Variance Inflation Factor (VIF) indicates no outliers and influential cases (values close to 1).

Table 6: Multiple Regression Model Results

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	,073	,006		12,431	,000		
Web Application	,031	,010	,186	3,283	,001	,997	1,003
Aged more 65	,068	,030	,129	2,269	,024	,997	1,003
Master Education	-,031	,015	-,116	-2,039	,042	,995	1,005

a. Dependent Variable: Ticket_price

The ANOVA statistical test, indicated that the prediction model was statistically significant, $F(3, 291) = 6.809$, $p < 0.001$, and accounted for approximately 6% of the variance of ticket price. This means that although the model includes variables that can predict the dependent variable, there are also other important predictors that are contributing on the estimation of the cost of the telematics applications and are not included in the proposed model.

The regression formula that derives after the results of Table 6 is the following:

$$(\text{TICKET_P})_i = \hat{\beta}_0 + \hat{\beta}_1 \cdot (\text{TYPE})_i + \hat{\beta}_2 \cdot (\text{AGE})_i + \hat{\beta}_3 \cdot (\text{EDU})_i + \varepsilon_i$$

$$= 0{,}073 + 0{,}031 \cdot (\text{TYPE})_i + 0{,}068 \cdot (\text{AGE})_i - 0{,}031 \cdot (\text{EDU})_i + \varepsilon_i$$

Where:

i=1,,,,,,295, the valid cases to predict the variable TICKET_P for each individual i
TYPE= 0 for IVR application and 1 for the web application
AGE= 0 for users aged less than 24 years old and 1 for users aged more than 65 years old
EDU=0 for users with Primarily Education and 1 for users with Master Degree.

Normality and Homogeneity of Variance

A main test that is performed in linear regression analyses is that of error variance assumption. Figure 2 presents the histogram for the residuals where a normally distribution can be observed. Thus, based on these results, the normality of residual assumption is satisfied.

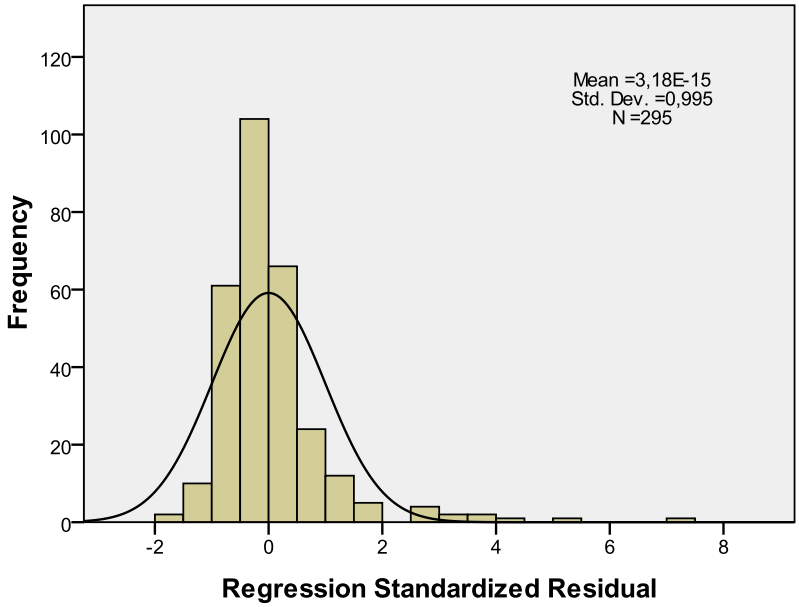


Figure 2. Frequency histogram of the standardized residuals of the dependent variable

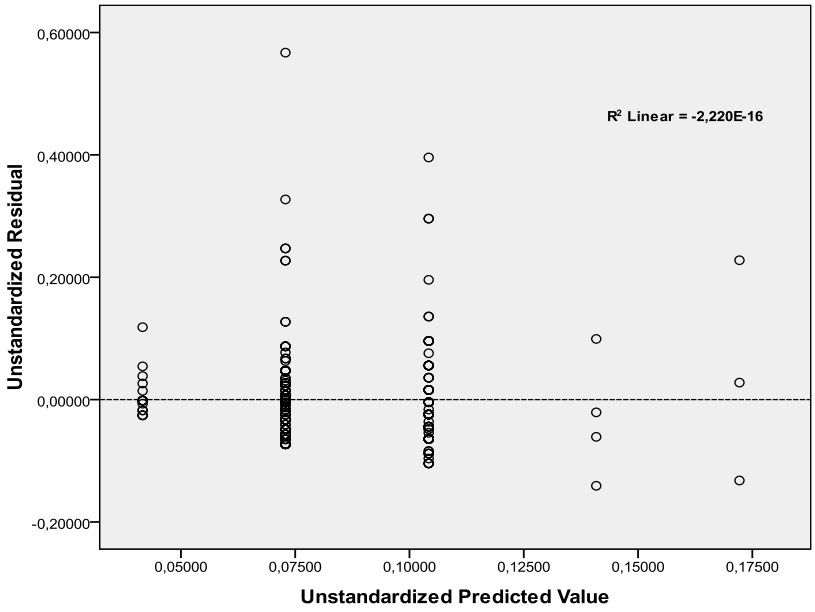


Figure 3. Scatter Plot of Unstandardized Predicted Variable against Unstandardized Residuals

In order to evaluate the homogeneity of the variance, a Scatter Plot of Unstandardized Predicted Variable against Unstandardized Residuals is presented in Figure 3. From the scatter plot it can concluded that no serious heteroscedasticity exists, e.g. no violation of the homogeneity of variance assumption is observed.

Finally, DfBetas values of the proposed model are examined in order to identify outliers and influential cases. Table 7 shows the descriptive statistics for the DfBetas of the statistical significant variables of the proposed model. Minimum and Maxinum values do not exit the cutting off criterion of ±2 therefore it can conclude that the dataset does not include outliers or influential cases.

Table 7: Outliers and Influential Cases Examination

	N	Minimum	Maximum	Mean	Std. Deviation
Standardized DFBETA Intercept	295	-0,13541	0,60331	0,0002347	0,05884591
Standardized DFBETA Type_Web	295	-0,33746	0,42833	0,0000676	0,06321904
Standardized DFBETA Aged_more_65	295	-0,79275	1,28787	-0,0000009	0,10637444
Standardized DFBETA Ed_Master	295	-0,16782	0,28068	-0,0001066	0,03410833

Model Interpretation

The interpretation of the model is done through the unstandardized coefficients of Table 6. The *constant* value of 0.073, corresponds to the ticket price that is used to developed and operate the IVR telematics application (Type=0), as it is perceived by young users aged less than 24 years old (AGE=0) with Primary education (EDUC=0).

Useful outcomes can be also derived from the interpretation of the coefficients for the independent variables that found to have a statistical significant prediction. For example, the coefficient value of 0.031 means that regarding the type of the application, the website portal is perceived as more costly compared to the IVR application by 0.031 units (euro) per ticket.

However, the most important parameter regarding the perceived cost of the telematics applications is the age of the user. As expected, older users (aged over 65 years old) perceive the examined applications as more costly compared to younger users (below 24 years old) by 0.068 units (euro) per ticket. This is a plausible outcome since older users tend to be impressed more about the capabilities of the new technologies and the services that can offer and therefore they intend to overestimate the budget that is needed to develop and operate them.

Finally, statistical significant differences were also observed regarding the educational level of the responders and how they are estimating the cost of the telematics applications. The regression model indicates that people, who hold a master degree, perceive the cost of telematics applications lower by a unit of 0.031euros per ticket, compared to people who have completed just the primary education. This is also a plausible outcome, since people who are highly educated can give better estimations about the technological background needed in order to develop such applications and therefore they tend to give lower cost values compared to lower educated users.

CONCLUSIONS

Information and Communication Technologies (ICT) are considered as one of the most promising supporting tools to tackle urban congestion and manage with an optimum manner, the travel demand and mobility. Ongoing research and new applications, such as social media and crowdsourcing, are utilizing the benefits of ICT and provide more accurate, reliable and user-demand services.

However, these ICT are subject to users' perception, level of acceptance and willingness to pay. Like other services, these applications can differ among the end users (e.g. public transport users) in various ways. An in-depth analysis of the factors that may have an impact on the acceptance of these applications by the end users as well as an impact on influencing their travel behavior is essential.

In this study, the various parameters that may have an effect on peoples' mobility and acceptance of ICT applications were investigated. Three telematics applications were tested, already installed and operating at the city of Thessaloniki Greece. The analysis showed that various differentiations can be observed between the population clusters and these differentiations can have an impact on the travel behaviour of the public transport user. Furthermore, the study found statistical significant differences related to age, application examined and educational level as well as the perceived cost required developing and operating the three case study applications.

All the above, indicateg that an appropriate categorization and analysis of the targeted population must be done ex ante, before the installation of any telematics application, so as to have better estimations about the impacts that are expected to be observed afterwards.

REFERENCES

American Society for Testing and Materials (ASTM), E122-07: Standard Practice for Calculating Sample Size to Estimate, With Specified Precision, the Average for a Characteristic of a Lot or Process, 2013

Abdel-aty, Mohamed A. 2001. "Using Ordered Probit Modeling to Study the Effect of ATIS on Transit Ridership." *Transportation Research Part C: Emerging Technologies* 9: 265–77.

Casey, Robert F, Lawrence N Labell, Elisabeth J Carpenter, Joseph A Lovecchio, Jackson W Royal, Judith C Schwenk, and Carol L Schweiger. 1998. "Transportation Systems : The State of the Art.", US Department of Transportation, Federal Transit Association.

Dziekan, Katrin, and Karl Kottenhoff. 2007. "Dynamic at-Stop Real-Time Information Displays for Public Transport: Effects on Customers." *Transportation Research Part A: Policy and Practice* 41(6): 489–501.

Guo, Zhan. 2011. "Mind the Map! The Impact of Transit Maps on Path Choice in Public Transit." *Transportation Research Part A: Policy and Practice* 45(7): 625–39.

Hickman, Mark D., and Nigel H.M. Wilson. 1995. "Passenger Travel Time and Path Choice Implications of Real-Time Transit Information." *Transportation Research Part C: Emerging Technologies* 3(4): 211–26.

Kokkinou Olga. 2011, *"Evaluation of Traveler Information Systems at the Public Transport Users"*, Master Thesis, October 2011 (In Greek).

Larsen, Odd I., and Øyvind Sunde. 2008. "Waiting Time and the Role and Value of Information in Scheduled Transport." *Research in Transportation Economics* 23(1): 41–52.

Li, Zheng, David a. Hensher, and John M. Rose. 2010. "Willingness to Pay for Travel Time Reliability in Passenger Transport: A Review and Some New Empirical Evidence." *Transportation Research Part E: Logistics and Transportation Review* 46(3): 384–403.

Molin, Eric J.E., and Harry J.P. Timmermans. 2006. "Traveler Expectations and Willingness-to-Pay for Web-Enabled Public Transport Information Services." *Transportation Research Part C: Emerging Technologies* 14(2): 57–67.

Monzon, Andres, Sara Hernandez, and Rocio Cascajo. 2013. "Quality of Bus Services Performance: Benefits of Real Time Passenger Information Systems." *Transport and Telecommunication* 14(2): 155–66.

Nelson, John D, and Corinne Mulley. 2013. "The Impact of the Application of New Technology on Public Transport Service Provision and the Passenger Experience: A Focus on Implementation in Australia.", *Research in Transportation Economics*, 39: 300–308.

Organization of Urban Transportation of Thessaloniki (OASTH), 2011, Various Unpublished Statistical Data.

Peng, Zhong-Ren, and Ruihong Huang. 2000. "Design and Development of Interactive Trip Planning for Web-Based Transit Information Systems." *Transportation Research Part C: Emerging Technologies* 8(1-6): 409–25.

Su, Jau-ming, and Chih-hung Chang. 2010. "The Multimodal Trip Planning System of Intercity Transportation in Taiwan." *Expert Systems with Applications* 37(10): 6850–61.

Velaga, N., Nelson, J., Edwards, P., Corsar, D., Sripada, S., Sharma, N., and Beecroft, M. (2013). "Development of a Map-Matching Algorithm for Rural Passenger Information Systems through Mobile Phones and Crowd Sourcing." Journal of Computational Civil Engineering, 27(6), 732–742.

Watkins, Kari Edison, Brian Ferris, Alan Borning, G. Scott Rutherford, and David Layton. 2011. "Where Is My Bus? Impact of Mobile Real-Time Information on the Perceived and Actual Wait Time of Transit Riders." *Transportation Research Part A: Policy and Practice* 45(8): 839–48.

Ziliaskopoulos, Athanasios K, and S.Travis Waller. 2000. "An Internet-Based Geographic Information System That Integrates Data, Models and Users for Transportation Applications." *Transportation Research Part C: Emerging Technologies* 8(1-6): 427–44.

A Driving Simulator Study on the Effect of Transversal Rumble Strips Located Nearby Dangerous Curves Under Repeated Exposure

Caroline Ariën[1], Kris Brijs[1,2], Wesley Ceulemans[2], Giovanni Vanroelen[2], Ellen M.M. Jongen[1], Stijn Daniels[1], Tom Brijs[1] and Geert Wets[1]

[1]Transportation Research Institute
Hasselt University
Wetenschapspark 5 box 6
3590 Diepenbeek, BELGIUM

[2]Faculty of Applied Engineering Sciences
Hasselt University
Agoralaan, Building H
3590 Diepenbeek, BELGIUM

ABSTRACT

This study examined drivers' behavior nearby dangerous curves when they are repeatedly exposed to the same transversal rumble strips which were located on the tangent before the curve. During a period of five successive days, sixteen participants completed a 17km test-drive in a driving simulator with four dangerous curves (two without and two with transversal rumble strips) in a within-subjects design. The selection of these curves was based on the official Belgian accident database and both curves were replicated in the driving simulator as detailed and realistic as possible. Results indicated that the transversal rumble strips induced a speed reduction of 2.3 to 5.9 kph on the tangent before the curve. This speed reduction effect sustained over the experimental period of five days. Taking the speed reduction effect into account, we can conclude that transversal rumble strips have a potential positive traffic safety effect because the reduced speed on the tangent provides more time to the drivers to make a good evaluation of the curve characteristics and environment and adapt their driving behavior in an appropriate way. Notwithstanding, we advise policy makers to make a good selection of potential dangerous curves to avoid excessive implementation of transversal rumble strips.

Keywords: Horizontal curves, pavement markings, transversal rumble strips, driving simulator, road safety engineering

INTRODUCTION

Accident analyses show that curves are typically prone accident locations on the road network: accident rates are 1.5 to 4 times higher than on tangents (i.e. straight road sections) and 25 to 30% of all fatal accidents occur in curves. In addition, 60 to 70% of all fatal curve-related crashes are single-vehicle run-of-road accidents, whereas head-on collisions occur in 11% of the fatal accidents (Safetynet, 2009a; Torbic et al., 2004).

According to Charlton (2007), inappropriate speed monitoring, failure to maintain a proper lateral position and inability to meet increased attentional demands are the three main behavioral causative factors for accidents in curves. These factors are also related to an inadequate evaluation of the degree of hazard associated with a given curve (Staplin, Lococo, Byington, & Harkey, 2001).

Experimental research on road design and human factors showed that geometric curve properties often relate to these behavioral problems (Brenac, 1996; Khan, Bill, Chitturi, & Noyce, 2014; Safetynet, 2009a). Low curve radii (<200m), inappropriate superelevation, too narrow road lanes and too long curve lengths are most frequently mentioned curve design elements which have adverse effects on traffic safety (Bonneson, Pratt, Miles, & Carlson, 2007; Khan et al., 2014; Safetynet, 2009a). In addition, a long preceding tangent length and a deviant sharp curve design of a single curve within a succession of gently designed curves are related to the extent to which the individual curve geometry fits within the surrounding road environment and showed to increase accident risks (R. Elvik, Hoye, Vaa, & Sorensen, 2009; Findley, Hummer, Rasdorf, Zegeer, & Fowler, 2012; Safetynet, 2009a).

Several studies proposed a wide variety of pavement markings (i.e., directional arrows, centerline or shoulder rumble strips and (peripheral) transversal strips) and signs (i.e., (dynamic) warning signs, advisory speed signs, (chevron) alignment signs and delineators) in order to induce appropriate speed and lateral control in curves (Charlton, 2004, 2007; Comte & Jamson, 2000; Federal Highway Administration, 2009; Hallmark, Smadi, & Hawkins, 2014; Katz, 2004; McGee & Hanscom, 2006). Since this study focuses on pavement markings, and more specifically on transversal rumble strips, we briefly elaborate on the main working mechanism behind this perceptual countermeasure. Transversal rumble strips (TRS) consists of a sequence of transverse colored lines with a raised profile at decreasing distance apart in the travel direction (see figure 1c). They manipulate the visual driving scene and the raised profile generates auditory and tactile feedback. These sensory inputs are meant to create an illusionary impression of increased motion which should result in a decrease in driving speed. Besides assisting drivers in more optimally speed monitoring, TRS have an important alerting function (Godley, 1999; Merat & Jamson, 2013).

Although a wide variety of patterns, colors and spacings are implemented, several field and driving simulator studies have demonstrated the potential speed reduction effect of transversal (rumble) strips in combination with intersections (Godley, 1999; Jamson & Lai, 2011; Montella et al., 2011), rural-urban transitions (Jamson, Lai, & Jamson, 2010), work zones (Bryden, Corkran, Hubbs, Chandra, & Jeannotte, 2013; Meyer, 2004) and curves (Ariën et al., 2012; Comte & Jamson, 2000; Gates, Qin, & Noyce, 2008; Godley, 1999). Elliot, McColl and Kennedy (2003; in Charlton & Baas, 2006) reported localised speed reductions between no effect up to 9,6 kph for transverse groupings of rumble strips. Godley (1999) established speed reductions between 8 and 11 kph near intersections and curves equipped with transverse lines. These results are in line with the speed reduction effects near intersections reported by Montella et al. (2011) (i.e., between 3 and 15 kph). Nevertheless, Rossi et al. (2013) found only moderate speed reductions for optical transversal speed bars near roundabouts (i.e., up to 2 kph). According to Elvik, Høye, Vaa and Sørensen (2009), rumble strips have a positive effect on road safety near junctions: injury accidents are reduced by 33% and the number of property-damage-only accidents decreased with 25%. Although these auspicious results, there is some doubt about the durability (both in time and distance) of the speed reduction effects (Comte & Jamson, 2000; Gates et al., 2008). The literature review of Martens et al. (1997) described that some experiments found that effects remained stable after a year (Zaidel, Hakkert, & Barkan, 1986), while others report that the effects lessen after some weeks or days (Maroney & Dewar, 1987).

Related to these inconclusive effects under repeated exposure (i.e., effect over time), Ariën et al. (2014) performed a literature review concerning the potential influence of novelty effects related to traffic calming measures on driving performance data in driving simulator research as described by Jamson and Lai (2010). Besides the various advantages related to driving simulator research (e.g., total control over various driving conditions, safe, cost efficient, collection of a variety of continuous high rate driving performance data), simulator validation, participant self-selection, simulator sickness and novelty effects should be taken into account (Jamson et al., 2010; Nilsson, 1993; Rudin-Brown, Williamson, & Lenné, 2009). These novelty effects can be related to the simulator system itself, but can also apply for the specific treatment being tested (for instance traffic calming measures or perceptual

countermeasures). Ariën et al. (2014) subdivided driving simulator experiments during which subjects were repeatedly exposed to an identical treatment into two groups: (1) participants were exposed several times to the same treatment during one single simulator session (e.g., Brown, 2001; Jamson et al., 2010; Lewis-Evans & Charlton, 2006; Rossi, Gastaldi, Biondi, & Mulatti, 2013a, 2013b) and (2) participants were exposed several times to the same treatment during multiple simulator sessions spread over different days (e.g., Åkerstedt et al., 2010; Charlton & Starkey, 2011; Domeyer, Cassavaugh, & Backs, 2013; Jenssen, Bjoerkli, Sakshaug, & Moen, 2007; Lenné, Triggs, & Redman, 1997; Manser & Creaser, 2011; Martens & Fox, 2007).

However, the literature available is rather scarce when it comes to examining the impact of technological and/or infrastructural treatments under conditions of repeated exposure on driving behavior. The studies of Jamson and Lai (2010) and Rossi et al. (2013a, 2013b) are the only references we are knowledgeable of which test the impact of infrastructural perceptual countermeasures under repeated exposure specifically. In both studies subjects participated during one single simulator session during which each participant passed four (Jamson et al., 2010) and ten times (Rossi et al., 2013a, 2013b) the same infrastructural measurements. Rossi et al. (2013a, 2013b) averaged the driving performance parameters over the ten trials and did not analyze the effect of the repeated exposure. Jamson and Lai (2010), on the other hand, observed three types of behavioral effects within their range of tested treatments: initial behavior shows a stronger / weaker than future behavior and future behavior can be predicted by initial behavior. Based on this literature review, the main objectives and more specific research questions are formulated.

OBJECTIVES AND RESEARCH QUESTIONS

This study will investigate the impact of transversal rumble strips (TRS) located on the tangent before two dangerous curves on the driving behavior of a sample of participants who will be repeatedly exposed to this specific perceptual countermeasure. The main research questions are formulated as follows:

1. Do TRS nearby dangerous curves influence mean speed?
2. How far does the influence of TRS nearby dangerous curves reach?
3. Does the effect of TRS nearby dangerous curves change when the same subjects are repeatedly exposed?

METHODOLOGY

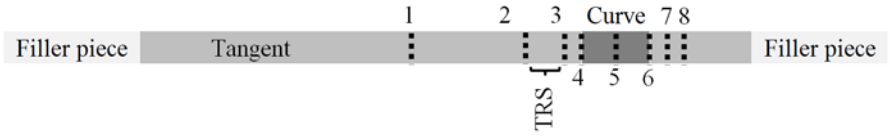
Participants

Participants were recruited via e-mail at Hasselt University. Twelve of the twenty-nine volunteers were excluded: three participants suffered from simulator sickness, six participants could not complete the total experimental period of five successive days due to technical problems and two participants were identified as outlier. A participant was defined as an outlier when he/she drove faster or slower than three inter-quartile distances from the group's mean during 25% of the analysis section. Thus, the remaining sample consists of 18 participants (8 men; mean age: 27.7; SD age: 11.5). All participants had (corrected to) normal vision and gave informed consent. Gender and age were not taken into account as between-subject factors in the statistical analysis.

Apparatus

The experiment was conducted on a medium-fidelity driving simulator (STISIM M400; Systems Technology Incorporated). It is a fixed-based (drivers do not get kinesthetic feedback) driving simulator with a force-feedback steering wheel, brake pedal, and accelerator. The simulation includes vehicle dynamics, visual and auditory feedback and a performance measurement system. The visual virtual environment was presented on a large 180° field of view seamless curved screen, with rear view and side-view mirror images and depiction of the speedometer. Three projectors offer a resolution of 1024 x 768 pixels and a 60 Hz refresh rate. The sounds of traffic in the environment and of the participant's car were presented. Data were collected at a 60 Hz frame rate.

(a)



1	500m before curve	5	curve middle
2	166m before curve	6	curve end
3	50m before curve	7	50m after curve
4	curve entry	8	100m after curve

Location A **Location B**

(b)

(c)

Figure 1: (a) Scenario overview; (b) satellite image and real world versus simulator images nearby curves at location A (left) and B (right); (c) simulator images of TRS at location A (left) and B (right)

Simulation scenario

During the 17 km driving scenario four curves, alternated with filler pieces (see figure 1a), were presented to the participants. The curves were programmed according to geo-specific database modelling method. Yan et al. (2008) defined this method as "replicating a real-world driving environment in a simulated virtual world" and is to be differentiated from the fictive driving scenarios. The real-world curves which were replicated in this driving simulator scenario were picked from the existing Flemish road network by means of extensive selection procedure using for instance the official Belgian accident database (Federal Government Statistics Belgium, n.d.), cross-sectional streets views and detailed accident maneuver diagrams. The detailed selection procedure is described in Ariën et al. (2012) and resulted in two dangerous left-oriented compound curves (i.e., combination of different curve radii in one curve) which were both preceded by a long tangent. Table 1 shows more detailed information on the curve characteristics.

Table 1: Curve properties of Location A and B

	Loc A	***Loc B***		***Loc A***	***Loc B***		***Loc A***	***Loc B***
Radius 1	170m	169m	Length 1	17m	51m	Total curve length	130m	116m
Radius 2	94m	92m	Length 2	29m	19m	Speed limit	90kph	70kph
Radius 3	161m	97m	Length 3	46m	21m	Road lane width	3.2m	2.8m
Radius 4	219m	688m	Length 4	38m	25m	Bicycle facilities	Yes	No

Two of the four presented curves were located at location A, while the other two curves had the road and environmental characteristics of location B. At both locations, one curve was equipped with TRS (see figure 1c), while no additional countermeasures were implemented at the other two curves. TRS were located between 155 and 66 m before the curve entrance (Vanduyver & Depestele, 2002) and each passage over a strip was accompanied by both auditory and tactile feedback provided by the sound equipment and the steering wheel of the driving simulator.

The road sections which were used for the statistical analyses consisted of a tangent (1200 m) followed by the compound curve and ended again with a tangent (300 to 375 m). The filler pieces, which connected these curve sections, were meant to provide some variation in the driving scenario and consisted of road segments with a variety of speed limits (e.g., 30, 50, 70 and 90 kph), surrounding environment (e.g., rural or urban) and daily changing interactions with other road users. The last 700 m before the first analysis point (i.e., 500 m before the curve entry) was standardized in order to prevent interference from these small day-to-day variations. Weather conditions were sunny and dry.

Procedure

Participants agreed to take part for a period of five consecutive weekdays. On the first day, participants were asked to fill out their personal data (e.g., gender, date of birth) and to give their informed consent. The general introduction in the driving simulator was followed by a practice session (first scenario of 4 km rural road with some slight curves and second scenario of 7 km with successively a motorway, a 70 kph rural road with a dangerous curve in right direction and an urban area) in order to get acquainted with the simulator. During the subsequent test trip of 17 km participants passed four dangerous curves (i.e., two curves at location A and two curves at location B and at each location once with TRS and once without TRS) in a counterbalanced order. The order of the four curves (location and TRS present or absent) did not change during the whole experiment for a particular participant because we were specifically interested in the driving behavior of participants who were repeatedly exposed to the TRS in the same configuration. The necessary guidance instructions were provided by a GPS voice. Subjects were instructed to apply the traffic laws as they would (or would not) do in reality and to drive as they normally would in their own car.

Data collection and analysis

The main purpose of the TRS under investigation is to improve road safety. Because of the positive relationship between driving speed on the one hand and crash risk and severity on the other hand (Safetynet, 2009b), mean speed is analyzed at eight analysis point along the driving scenario (see figure 1a).

A 2 (marking: no TRS, TRS) × 5 (day) × 8 (analysis points) within-subjects analysis of variance (ANOVA) was conducted on mean speed [kph] for each location separately. Based on Kolomgorov-Smirnov tests of normality and Mauchly's test of sphericity we corrected for deviation from normality (Bonferroni correction) and sphericity (Greenhouse-Geisser epsilon correction). *P*-value was set at 0.05 to determine statistical significance. F- and partial eta squared values are mentioned.

RESULTS

Table 2: ANOVA statistics for location A and B (significant p-values are indicated in bold)

Greenhouse-Geisser	***Location A***			***Location B***		
	F (dfs)	*p*	Parial eta squared	F (dfs)	*p*	Parial eta squared
Marking	5.9 (1, 17)	**0.027**	0.257	4.6 (1, 17)	**0.047**	0.212
Day	1.0 (3, 51)	0.593	0.036	8.1 (3, 45)	**<0.001**	0.322
Point	121.5 (2, 39)	**<0.001**	0.877	43.1 (2, 38)	**<0.001**	0.717
Marking × Day	1.5 (3, 48))	0.228	0.081	1.4 (3, 47)	0.244	0.078
Marking × Point	5.8 (2, 36)	**0.006**	0.255	3.6 (3, 48)	**0.023**	0.174
Day × Point	3.3 (8, 129)	**0.002**	0.164	3.5 (7, 115)	**0.002**	0.171
Marking × Day × Point	1.0 (7, 124)	0.415	0.057	1.8 (5, 93)	0.106	0.098

Location A

The daily values for mean speed on the 8 analysis points separated for the condition without (left graph) and with (right graph) TRS are shown in figure 2. At 500 m before the curve mean speed was highest and close to the speed limit of 90 kph. During the first four days, participants decelerated to a minimal mean speed near the curve middle. On the fifth day, mean speed was already minimal at the curve entry. Once participants passed the curve middle, they started to accelerate again at a continuous level, but there is some indication that they accelerated more as the days progressed. Overall, mean speed seems to be lower at 166 m and 50 m before the curve entry when TRS were present. In addition, at the curve entry mean speed was very constant during the whole experimental period when TRS was present, compared to the slightly higher mean speeds during the first two days of the experiment and the larger spread over the different days when TRS was absent.

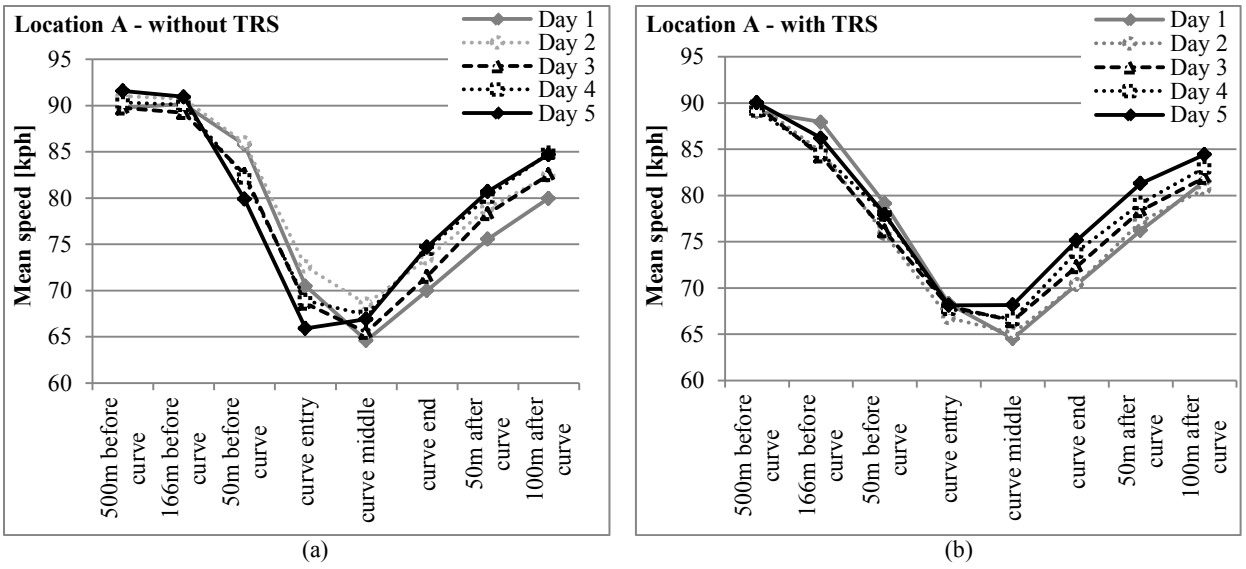


Figure 2: Mean speed as a function of Marking × Day × Point: (a) TRS absent and (b) TRS present at location A (TRS were located between 150 en 66 m before the curve entry)

The ANOVA for location A showed a significant main effect of *Marking* and *Point*. In addition, there was a significant interaction effect of *Marking* × *Point* and *Day* × *Point*. Since the combination of the factors *Marking* and *Day* were not significant in a two- or three-way interaction, we can conclude from the significant interaction of *Marking* × *Point* that mean speed varied across the different analysis points in function of the presence or absence of TRS, but not in function of the day. This means that the effects generated by the TRS on a certain day did not significantly differ from the other four days. Figure 3a shows the mean speed values on the 8 analysis points, separated for the condition with or without TRS but irrespective of the day. Post-hoc analysis for the interaction effect of *Marking* × *Point* showed that mean speed was 4.7 to 5.9 kph lower at respectively 166 and 50 m before the curve entry when TRS was present (166 m before curve entry: $F_{(1, 17)} = 8.4$, $p = 0.010$, $\eta_p^2 = 0.330$; 50 m before curve entry: $F_{(1, 17)} = 12.6$, $p = 0.002$, $\eta_p^2 = 0.426$). At the other six analysis points, there were no significant differences in mean speed between the condition with or without TRS.

Mean speed values on the 8 analysis point, separated for the 5 days but irrespective of the presence or absence of TRS are shown in figure 3b. Post-hoc analysis for the interaction effect of *Day* × *Point* showed that some mean speed values significantly varied across the different analysis points on the different days. Interestingly, during the first four days minimal mean speeds were reached at the curve middle, while on the last day participants reached a minimal speed already at the curve entry and continued this speed until the curve middle. It is however important to note that there was no significant difference in mean speed between the curve entry and middle at day 4. Although figure 2 and 3b gave some indication that mean speed increased at the curve end and at 50 and 100 m after the curve end as the days progressed, this was not confirmed by the pairewise comparisons.

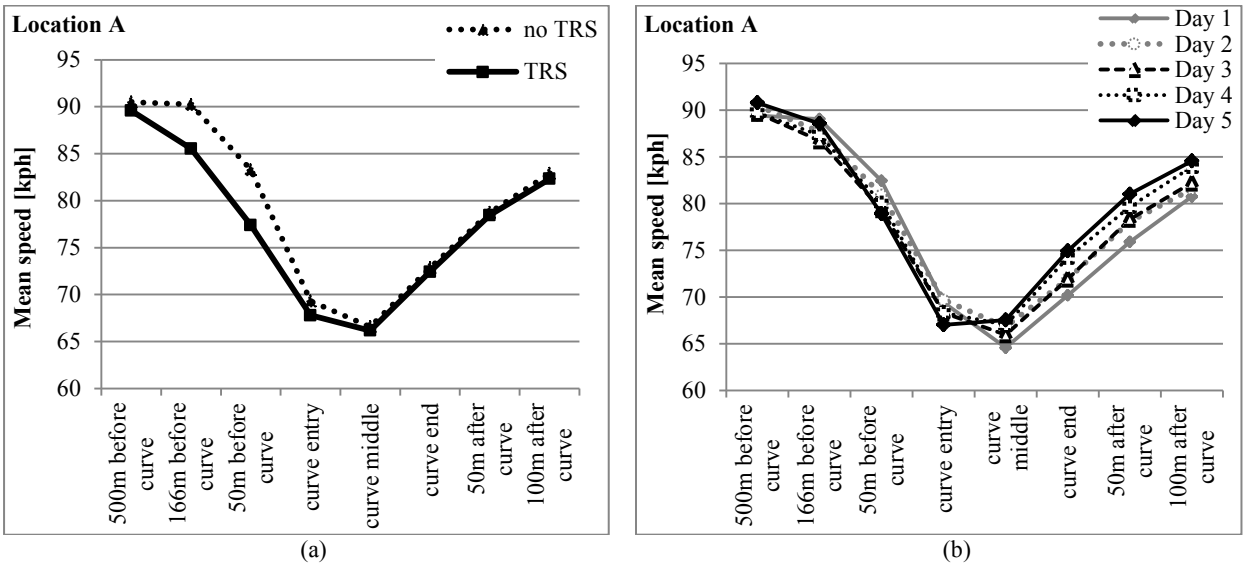


Figure 3: Mean speed for the interaction of (a) *Marking* × *Point* and (b) *Day* × *Point* at location A (TRS were located between 150 en 66 m before the curve entry)

Location B

Figure 4 shows the daily mean speed values for location B on the 8 analysis points, separated for the condition (a) without and (b) with TRS. Starting from a mean speed slightly about the speed limit of 70 kph at 500 m before the curve, participants decelerated to a minimal speed at the curve middle during the five experimental days when TRS was absent. When TRS was present, the same deceleration behavior was present during the first three days, but at day 4 and 5 participants reached their minimal speed already at the curve entry. In addition, both in the condition with and without TRS there seem to be an indication that mean speed increased as the days passed by. Overall, mean speed seems to be lower at 166 m and 50 m before the curve entry when TRS were present. At the curve entry, there seems to be some indication that mean speed was lower during the first two days when TRS was present.

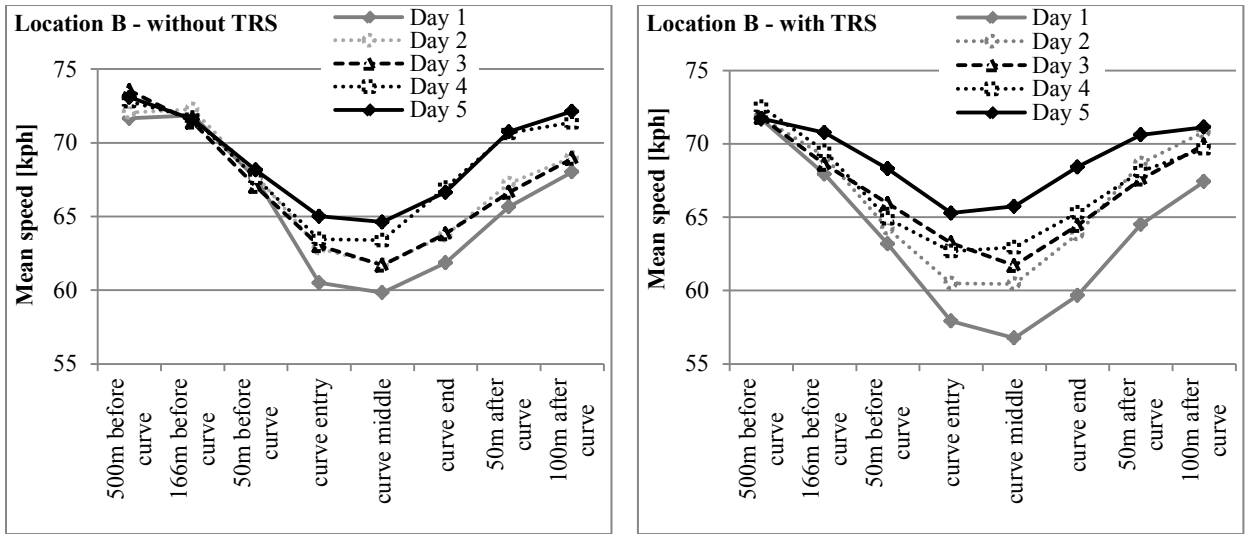


Figure 4: Mean speed as a function of Marking × Day × Point: (a) TRS absent and (b) TRS present at location B
(TRS were located between 150 en 66 m before the curve entry)

The ANOVA for location B showed comparable significant main and interaction effects as at location A, but the main effect of *Days* was also significant at location B. Since the interaction between *Marking* and *Day* or between the three factors was not significant, the interaction effect of *Marking* × *Point* indicated that mean speed might have varied across the different analysis points in function of the presence or absence of TRS, but this interaction was irrespective of the day. This means that the effects generated by the TRS on a certain day did not significantly differ from the other four days.

Mean speed values at the 8 analysis points separated for the condition with and without TRS, but irrespective of the day are shown in figure 5a. Post-hoc analysis for the interaction effect of *Marking* × *Point* showed that mean speed was 2.6 to 2.3 kph lower at respectively 166 and 50 m before the curve entry when TRS was present (166 m before curve entry: $F_{(1, 17)} = 12.0$, $p = 0.003$, $\eta_p^2 = 0.414$; 50 m before curve entry: $F_{(1, 17)} = 8.2$, $p = 0.011$, $\eta_p^2 = 0.325$). In addition, TRS generated a marginally significant speed reduction of 1.0 kph at the curve entry ($F_{(1, 17)} = 3.5$, $p = 0.077$, $\eta_p^2 = 0.173$).

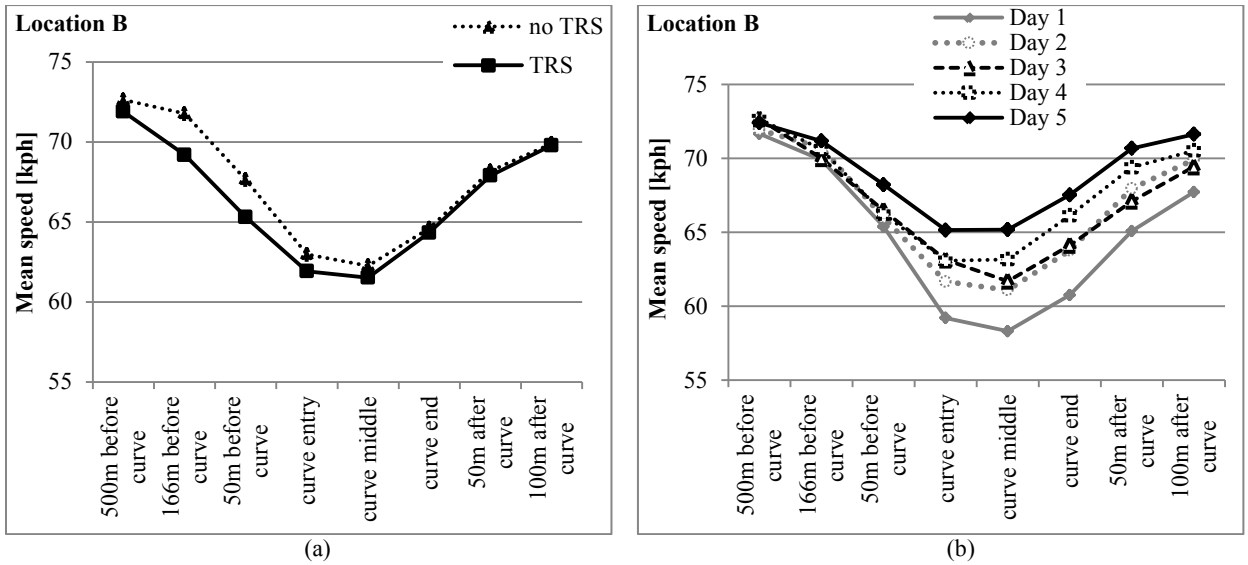


Figure 5: Mean speed for the interaction of (a) *Marking* × *Point* and (b) *Day* × *Point* at location B
(TRS were located between 150 en 66 m before the curve entry)

Figure 5b shows the mean speed values at the 8 analysis points, separated for the 5 days but irrespective of the presence or absence of TRS. Although there were significant speed differences between the curve entry and middle during the total experimental period, mean speeds were slightly (but not significant) lower at the curve middle during the first three days. On the two last days, minimal speed was reached at the curve entry. Comparing the mean speed values at each analysis point between the different days shows that mean speed was significant lower at the first two days compared to the last day from the curve entry until the curve end. In addition, mean speed was also significant lower from the curve end until 100 m after the curve at the first day compared to day 4 and 5. Finally, mean speed at day 3 was significant lower than at day 5 at 50 m after the curve.

DISCUSSION

In this driving simulator study we analyzed mean speed to find out (1) whether TRS located on the tangent before dangerous curves influences mean speed; (2) how far the influence reaches and (3) whether the effect would change when the same participant is repeatedly exposed during a period of 5 successive days. In addition, we try to relate the established results with the geometric curve characteristics of the two dangerous curves under investigation.

Besides some main effects, the ANOVA showed a significant interaction effect of *Marking* × *Point* and *Day* × *Point* for both location A and B. The absence of a significant interaction between the factors *Marking* and *Day* or between the three factors *Marking*, *Day* and *Point* reveal that the potential influence of TRS on mean speed is independent of the day. This means that the effects generated by TRS on a certain day were not significant different from the other four days.

During the five successive days of the experiment period and at both curve locations (A and B), TRS generated a significant speed reduction on the tangent in the direct vicinity of the TRS, more specifically at 166 and 50 m before the curve (TRS was located between 150 and 66 m before the curve). At location A, significant speed reductions between 4.7 and 5.9 kph were measured. The size of the speed reduction effect of the TRS at location B was smaller (i.e., 2.3 to 2.6 kph), but there was also a marginally significant speed reduction effect at the curve entry of 1.0 kph. These speed reductions are in line with the results of for instance (Elliot et al., 2003; Montella et al., 2011; Rossi et al., 2013b). According to the Elvik's power model for rural roads (Rune Elvik, 2009, p. 58), speed reductions of that size might induce a decrease in fatal accidents and injury accidents on the tangent at location A up to 35% and 12% respectively and up to 16% and 6% at location B.

Several studies established that drivers start to explore curves between 100 m and 30 m before the curve entry (Milleville-Pennel, Jean-Michel, & Elise, 2007; Tsimhoni & Green, 1999). The lower driving speed on a tangent equipped with TRS gives drivers thus more time to satisfy the increased need for visual information on curved roads, to make an adequate evaluation of the degree of risk and to meet increased attentional demands associated with the curve. In addition, due to their lower speed on the tangent, drivers are less forced to suddenly adapt their driving behavior just before they enter the curve or along the curve itself, compared to a situation in which their approaching speed was higher. This is an important issue because accidents occur primarily at both the curve entry or the curve end (PIARC, 2003).

The results of the interaction effect of *Day* × *Point* showed that a minimal mean speed was reached at the curve middle at the first three days and at the curve entry at the last two days. These results are somewhat in line with Mintsis (1988) who observed lowest speed in the middle of the curve and Taragin (1954) who suggested that drivers adjust their speed before entering a curve and continue at a contact speed throughout the curve. Another element related to this interaction effect is the (indication of an) increase in mean speed beginning at the curve entry as the days pass by. This evolution over the days might be related to Wilde's theory of risk homeostasis (Milleville-Pennel et al., 2007) or the driving behavior model of Weller and colleagues (Weller, Schlag, Friedel, & Rammin, 2008) in which drivers adjust their driving behavior as a result of an appraisal of the perceived risks with an acceptable risk threshold. During the first days, participants seem to 'overestimate' the perceived risk of the curves and lower their speed. However, the successive exposure during the following days might adjust their risk perception and participants feel confident to increase their mean speed which does not benefit road safety.

When we compare driving behavior in the curve at location A and B and determine the V_{85} speed differential between the tangent (166 m before the curve) and the curve entry, we see that V_{85} reduced with 12 kph at location A and 5 kph at location B. According to Lamm et al. (1999) the curve at location B has a good design quality, whereas

the design quality of the curve at location A is acceptable. In addition, Anderson et al. (in PIARC, 2003) established that the accident rate at a curve section with a speed differential of 10 kph to 20 kph is twice as high as a curve with a speed differential of less than 10 kph. The established differences in driving behavior at location A and B can be attributed to the differences in their geometric design characteristics and the different curve radii and curve lengths of the individual curve segments of the compound curves (see table 1). Odhams and Cole (2004) found for instance a positive relationship between speed choice and both lane width and curve radius. The lower speed limit of 70 kph, the absence of bicycle facilities and the smaller lane width at location B might be the main reasons why mean speed at location B was lower than at location A.

LIMITATIONS AND FUTURE RESEARCH

The reporting of results of driving simulator studies often goes together with discussions about external validity. Although moving base simulators generated a greater degree of realism (Bella, 2009), several studies showed indications that fixed-base driving simulators can examine geometric design issues in a perfectly adequate way (e.g., Bella, 2007, 2008; Benedetto, Calvi, & Messina, 2012; Calvi, Benedetto, & De Blasiis, 2012; Charlton, 2004; Federal Highway Administration, 2007). In addition, the 180° seamless curved screen used in this study satisfies the prescribed minimum field of view of 120° in order to make correct estimation of longitudinal speed (Kemeny & Panerai, 2003).

Future research on TRS can focus on additional driving parameters related to longitudinal and lateral speed (e.g. acceleration/deceleration or lateral position) or on different geometric design configurations to improve road safety nearby dangerous curves. Although we tried to anticipate the potential influence of novelty effects of TRS on mean speed by this quite unique experimental setup (besides the studies of Ariën et al., 2014; Jamson & Lai, 2011; Rossi et al., 2013a, 2013b) where participants were repeatedly exposed during 5 successive days, we are unable to pronounce upon the long term effect of these TRS as in a before-after field experiment. Future research can thus focus on longer term naturalistic driving studies and a before-after field experiment.

CONCLUSION AND POLICY RECOMMENDATIONS

The paper has investigated the effect of transversal rumble strips (TRS) located near dangerous curves on mean speed of a sample of participants who were repeatedly exposed to this specific perceptual countermeasure. The driving simulator study has established that TRS generated a significant reduction of mean speed (i.e., between 2.3 kph and 5.9 kph) on the tangent proceeding to the curve and that these effects on mean speed are irrespective of the day. The speed reduction effect sustained thus over the five-day experiment period. Although the speed reduction effect did not proceed until the curve entry and further along the curve, the lower speed on the tangent gives drivers more time to make an adequate evaluation of the degree of risk with the curve and to adapt their driving behavior in an appropriate way. Besides this potential positive effect on mean speed, TRS work also as an alerting device (Merat & Jamson, 2013). Despite these favorable effects, some studies warn for the produced noise when a vehicle passes by the TRS (Dewar & Olson, 2007; Martens et al., 1997). Based on these results, we can conclude that TRS is a low-cost perceptual countermeasure that has the potential to improve road safety near dangerous curves. Notwithstanding, we advise policy makers to make a good selection of potential dangerous curves to avoid excessive implementation of transversal rumble strips.

REFERENCES

Åkerstedt, T., Ingre, M., Kecklund, G., Anund, A., Sandberg, D., Wahde, M., … Kronberg, P. (2010). Reaction of sleepiness indicators to partial sleep deprivation, time of day and time on task in a driving simulator – the DROWSI project. *Journal of Sleep Research*, *19*(2), 298–309. doi:10.1111/j.1365-2869.2009.00796.x

Ariën, C., Brijs, K., Brijs, T., Ceulemans, W., Vanroelen, G., Jongen, E. M. M., … Wets, G. (2014). Does the effect of traffic calming measures endure over time? – A simulator study on the influence of gates. *Transportation Research Part F: Traffic Psychology and Behaviour*, *22*, 63–75. doi:10.1016/j.trf.2013.10.010

Ariën, C., Brijs, K., Ceulemans, W., Jongen, E. M. M., Daniels, S., Brijs, T., & Wets, G. (2012). The effect of pavement markings on driving behavior in curves: A driving simulator study. Presented at the Transportation Research Board, Washington, D.C.: Transportation Research Board.

Bella, F. (2007). Parameters for Evaluation of Speed Differential: Contribution Using Driving Simulator. *Transportation Research Record: Journal of the Transportation Research Board*, *2023*, 37–43. doi:10.3141/2023-05

Bella, F. (2008). Driving simulator for speed research on two-lane rural roads. *Accident Analysis and Prevention*, *40*, 1078–1087.

Bella, F. (2009). Can the driving simulators contribute to solving the critical issues in geometric design? Transportation Research Board Annual Meeting 2009. Retrieved from http://pubsindex.trb.org/view.aspx?id=880535

Benedetto, A., Calvi, A., & Messina, M. (2012). Potentialities of driving simulator for engineering applications to Formula 1. *Advances in Transportation Sciences*, 127–138. doi:10.4399/978885484657912

Bonneson, J., Pratt, M., Miles, J., & Carlson, P. (2007). *Horizontal curve signing handbook* (No. FHWA/TX-07/0-5439-P1) (p. 56). Texas: Texas Transportation Institute.

Brenac, T. (1996). Safety at Curves and Road Geometry Standards in Some European Countries. *Transportation Research Record: Journal of the Transportation Research Board*, *1523*(-1), 99–106. doi:10.3141/1523-12

Brown, C. M. (2001). New in-vehicle technologies: are lane departure warnings a good thing? (Vol. 2, pp. 70–74). Presented at the SELF-ACE 2001 Conference - Ergonomics for changing work.

Bryden, J. E., Corkran, M. O., Hubbs, C. W., Chandra, A. K., & Jeannotte, K. L. (2013). *Traffic enforcement strategies for work zones* (p. 44). Washington, D.C.: Transportation Research Board - National Cooperative Highway Research Program.

Calvi, A., Benedetto, A., & De Blasiis, M. R. (2012). A driving simulator study of driver performance on deceleration lanes. *Accident Analysis & Prevention*, *45*, 195–203. doi:10.1016/j.aap.2011.06.010

Charlton, S. G. (2004). Perceptual and attentional effects on drivers' speed selection at curves. *Accident Analysis & Prevention*, *36*(5), 877–884. doi:10.1016/j.aap.2003.09.003

Charlton, S. G. (2007). The role of attention in horizontal curves: A comparison of advance warning, delineation, and road marking treatments. *Accident Analysis & Prevention*, *39*(5), 873–885. doi:10.1016/j.aap.2006.12.007

Charlton, S. G., & Baas, P. H. (2006). *Speed change management for New Zealand roads* (Research report No. 300). New Zealand: Land Transport New Zealand.

Charlton, S. G., & Starkey, N. J. (2011). Driving without awareness: The effects of practice and automaticity on attention and driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*(6), 456–471. doi:10.1016/j.trf.2011.04.010

Comte, S. L., & Jamson, A. H. (2000). Traditional and innovative speed-reducing measures for curves: an investigation of driver behaviour using a driving simulator. *Safety Science*, *36*(3), 137–150. doi:10.1016/S0925-7535(00)00037-0

Dewar, R., & Olson, P. (2007). *Human factors in traffic safety*. Tucson: Lawyers & Judges Publishing Company, Inc.

Domeyer, J. E., Cassavaugh, N. D., & Backs, R. W. (2013). The use of adaptation to reduce simulator sickness in driving assessment and research. *Accident Analysis & Prevention*, *53*, 127–132. doi:10.1016/j.aap.2012.12.039

Elliot, M. A., McColl, V. A., & Kennedy, J. V. (2003). *Road design measures to reduce drivers' speed via "psychological" processes: a literature review.* (No. TRL564). Crowthorne, Berkshire, UK: Transport Research Laboratory.

Elvik, R. (2009). *The Power Model of the relationship between speed and road safety - Update and new analyses* (No. 1034/2009). Oslo: Institute of Transport Economics.

Elvik, R., Hoye, A., Vaa, T., & Sorensen, M. (2009). *The handbook of road safety measures* (second.). UK: Emerald.

Federal Government Statistics Belgium. (n.d.). Official Belgian accident databank (1997-2007).

Federal Highway Administration. (2007). *Drivers' evaluation of the diverging diamond interchange* (TechBrief No. FHWA-HRT-07-048). FHWA Federal Highway Administration.

Federal Highway Administration. (2009). Manual on uniform traffic control devices for streets and highways. FHWA Federal Highway Administration.

Findley, D. J., Hummer, J. E., Rasdorf, W., Zegeer, C. V., & Fowler, T. J. (2012). Modeling the impact of spatial relationships on horizontal curve safety. *Accident Analysis & Prevention*, *45*(0), 296–304. doi:10.1016/j.aap.2011.07.018

Gates, T., Qin, X., & Noyce, D. (2008). Effectiveness of Experimental Transverse-Bar Pavement Marking as Speed-Reduction Treatment on Freeway Curves. *Transportation Research Record*, *2056*(1), 95–103.

Godley, S. T. (1999). *A driving simulator investigation of perceptual countermeasures to speeding.* Monash University.

Hallmark, S. L., Smadi, O., & Hawkins, N. (2014). Speed reduction impacts of dynamic speed feedback signs on rural two lane curves. Presented at the Transportation Research Board 93rd Annual Meeting, Washington, D.C.

Jamson, S., & Lai, F. (2011). Are novelty effects of road safety treatments observable in simulator experiments? Presented at the TRB 2011 Annual Meeting, Washington DC: Transportation Research Board.

Jamson, S., Lai, F., & Jamson, H. (2010). Driving simulators for robust comparisons: A case study evaluating road safety engineering treatments. *Accident Analysis & Prevention*, *42*(3), 961–971. doi:10.1016/j.aap.2009.04.014

Jenssen, G. D., Bjoerkli, C. A., Sakshaug, K., & Moen, T. (2007). Behavioural adaptation to adaptive front lighting systems

(AFS): a six day driving simulator study. Presented at the 14th World Conference on Intelligent Transport Systems, Beijing. Retrieved from http://trid.trb.org/view.aspx?id=916173
Katz, B. J. (2004). Pavement markings for speed reduction. Science Applications International Corporation - Turner-Fairbank Highway Research Center.
Kemeny, A., & Panerai, F. (2003). Evaluating perception in driving simulation experiments. *Trends in Cognitive Sciences*, *7*(1), 31–37. doi:10.1016/S1364-6613(02)00011-6
Khan, G., Bill, A. R., Chitturi, M., & Noyce, D. A. (2014). Safety evaluation of horizontal curves on rural undivided roads. Presented at the 92nd Annual Meeting of the Transportation Research Board, Washington DC.
Lamm, R., Mailaender, T., & Psarianos, B. (1999). *Highway design and traffic safety engineering handbook*. McGraw-Hill.
Lenné, M. G., Triggs, T. J., & Redman, J. R. (1997). Time of day variations in driving performance. *Accident Analysis & Prevention*, *29*(4), 431–437. doi:10.1016/S0001-4575(97)00022-5
Lewis-Evans, B., & Charlton, S. G. (2006). Explicit and implicit processes in behavioural adaptation to road width. *Accident Analysis & Prevention*, *38*(3), 610–617. doi:10.1016/j.aap.2005.12.005
Manser, M., & Creaser, J. (2011). Assessing driver behavioral adaptation to rural intersection driver support system. Presented at the 6th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Lake Tahoe, California.
Maroney, S., & Dewar, R. (1987). Alternatives to enforcement in modifying the speeding behavior of drivers. *Transportation Research Record*, *1111*, 121–126.
Martens, M. H., Comte, S., & Kaptein, N. A. (1997). *The effect of road design on speed behavior: A literature review* (Deliverable D1 No. RO-96-SC.202). Finland: VTT Communities & Infrastructure.
Martens, M. H., & Fox, M. R. J. (2007). Do familiarity and expectations change perception? Drivers' glances and response to changes. *Transportation Research Part F: Traffic Psychology and Behaviour*, *10*(6), 476–492. doi:10.1016/j.trf.2007.05.003
McGee, H. W., & Hanscom, F. R. (2006). *Low-cost treatments for horizontal curve safety* (No. FHWA-SA-07-002). Washington DC: Federal Highway Administration.
Merat, N., & Jamson, A. H. (2013). The effect of three low-cost engineering treatments on driver fatigue: A driving simulator study. *Accident Analysis & Prevention*, *50*(0), 8–15. doi:10.1016/j.aap.2012.09.017
Meyer, E. (2004). *Evaluation of data from test application of optical speed bars to highway work zones* (Final report No. KTRAN: KU-00-4). Kansas: University of Kansas.
Milleville-Pennel, I., Jean-Michel, H., & Elise, J. (2007). The use of hazard road signs to improve the perception of severe bends. *Accident Analysis & Prevention*, *39*(4), 721–730.
Mintsis, G. (1988). Speed distribution on road curves. *Traffic Engeneering and Control*, *29*, 21–27.
Montella, A., Aria, M., D'Ambrosio, A., Galante, F., Mauriello, F., & Pernetti, M. (2011). Simulator evaluation of drivers' speed, deceleration and lateral position at rural intersections in relation to different perceptual cues. *Accident Analysis & Prevention*, *43*(6), 2072–2084. doi:16/j.aap.2011.05.030
Nilsson, L. (1993). Behavioural research in an advanced driving simulator - Experiences of the VTI system. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, *37*(1), 612–616.
Odhams, A. M., & Cole, D. J. (2004). Models of driver speed choice in curves. Cambridge University Engineering Department.
PIARC. (2003). *Road safety manual - Recommendations from the World Road Association (PIARC)*. Route 2 Market.
Rossi, R., Gastaldi, M., Biondi, F., & Mulatti, C. (2013a). Effects on lateral position of perceptual measures in affecting driver's perceived speed. Presented at the 4th Internation Conference on Road Safety and Simulation, Rome, Italy.
Rossi, R., Gastaldi, M., Biondi, F., & Mulatti, C. (2013b). Oppel-kundt illusion and lateral optic flow manipulation in affecting perceived speed in approaching roundabouts: experiments with a driving simulator. Presented at the 92nd Transportation Research Board Meeting, Washington DC.
Rudin-Brown, C. M., Williamson, A., & Lenné, M. G. (2009). *Can driving simulation be used to predict changes in real-world crash risk?* (No. 299). Monash University Accident Research Centre.
Safetynet. (2009a). Roads.
Safetynet. (2009b). Speeding.
Staplin, L., Lococo, K., Byington, S., & Harkey, D. (2001). *Guidelines and recommendations to accomodate older drivers and pedestrians* (Final report No. FHWA-RD-01-051) (p. 92). Federal Highway Administration.
Taragin, A. (1954). Driver performance on horizontal curves. *Proc. Ann. Meeting*, *33*, 446–466.
Torbic, D. J., Harwood, D. W., Gilmore, D. K., Pfefer, R., Neuman, T. R., Slack, K. L., & Hardy, K. . (2004). *Guidance for implementation of the AASHTO strategic highway safety plan. Volume 7: A guide for reducing collisions on horizontal curves* (No. 500, volume 7). Washington DC: Transportation Research Board - National Cooperative Highway Research Program.
Tsimhoni, O., & Green, P. (1999). Visual demand of driving curves determined by visual occlusion. Presented at the Vision in Vehicles VIII Conference, Boston.
Vanduyver, A., & Depestele, R. (2002). *Verkeerssignalisatie* (second.). Brugge: Uitgeverij Vanden Broele.
Weller, G., Schlag, B., Friedel, T., & Rammin, C. (2008). Behaviourally relevant road categorisation: A step towards self-explaining rural roads. *Accident Analysis & Prevention*, *40*(4), 1581–1588. doi:10.1016/j.aap.2008.04.009
Yan, X., Abdel-Aty, M., Radwan, E., Wang, X., & Chilakapati, P. (2008). Validating a driving simulator using surrogate safety measures. *Accident Analysis and Prevention*, *40*, 274–288.
Zaidel, D., Hakkert, A. S., & Barkan, R. (1986). Rumble strips and paint stripes at a rural intersection. *Transportation Research Record*, *1069*, 7–12.

Development of a Concept for Ergonomic Design and Work Organization of Remote Lock Operation Centers

J. Bützler[1], F. Motz[2], A. Mertens[1], H. Horoufchin[2], C. Bröhl[1], N. Jochems[1] and C. M. Schlick[1, 2]

[1]Chair and Institute of Industrial Engineering and Ergonomics of RWTH Aachen University, Bergdriesch 27, 52062 Aachen, Germany
[2]Fraunhofer Institute for Communication, Information Processing and Ergonomics, Neuenahrer Str. 20, 53343 Wachtberg, Germany

ABSTRACT

A trend can be observed towards grouping and shifting local operations to centralized remote lock operation centers (RLOCs) due to costs and limitations in personnel resources and the increasing automation level of technical systems. However, safety-critical activities still require active monitoring of processes by a human operator with help of adequate technical systems and interfaces. Therefore, many video images and abstract representations must be shown simultaneously at one workplace to guarantee an all-inclusive overview. Given this situation, it has to be determined from an ergonomic point of view how many systems (e.g. lock chambers) can be controlled by one operator. Therefore, the challenge is to develop a new efficient and ergonomic control concept for RLOCs that prevents straining work conditions in terms of mental over- and underload. This control concept should enable the short-term and situation-dependent allocation of systems to operators. Moreover, prospective personnel resource planning and the design of the control panels have to be adapted to the control concept and human performance factors.

The following paper presents a study covering the analysis of tasks and processes in remote lock operation centers around Germany including soft- and hardware design components in use. Based on the results, a new concept integrating ergonomic requirements for remote control and monitoring has been developed. This concept enables the assignment of optimal workload to one workplace by determining the potential straining condition within one shift. For the purpose of integrating prospective and retrospective evaluation, a practical guide for personnel management and scheduling has been developed. It includes a questionnaire and calculation specifications that can be generically adapted to other remote operation centers. In addition to the strain which is evoked by the main tasks of lock control and monitoring, it also considers the impact of additional tasks. Based on the developed concept for remote control and monitoring and in accordance with the assessed ergonomic design requirements, the control panel layout has been determined with regard to anthropometric dimensions.

Keywords: Ergonomic Design, Remote Operation Centers, Lock Operation, Control Panel, Personnel Management, Usability

INTRODUCTION

Since advances in process control engineering enable to operate even safety-critical systems remotely, an automation of systems and centralized control and monitoring processes in RLOC can be observed in Germany. This trend affects the process control level in terms of higher degree of automation and digitalization of the work system (e.g. Ivergard and Hunt, 2008; Stanton et al., 2010) and consequently evokes various changes in the sociotechnical work system (DIN EN ISO 6385, 2004). Due to the higher automation, steps which prior to this trend were executed manually by the operator are now omitted. However, for safety-critical processes, there is still a need to actively initiate and monitor a process by a human operator. Therefore, grouping and shifting local operations to centralized RLOCs represents a solution that leads to more efficient use of personnel resources but is also accompanied by new challenges in work space design and work organization.

At the beginning of the 2000s, the organization of lock operation changed from on-site control to remote control and monitoring in Germany. Several technical systems are connected with the RLOCs and can be remote controlled by the operators via a computer interface. These systems include lock chambers, weirs, pumping stations, flood barrages, light-signal systems and bridges. Until 2012, 27 RLOCs were built across Germany with 113 connected lock chambers plus other systems. In the next few years, it is planned to establish 33 RLOCs with 172 connected lock chambers. However, no standard for ergonomic design and work organization has been established yet, rather a series of regional solutions can be found.

This paper focuses on the consequence of this trend regarding ergonomic issues such as work environment, work place and human-system interaction on the one hand and organizational aspects as allocation of system(s) to operator and personnel resources on the other hand. One challenge in this context, for example, is the simultaneous depiction of a lot of information (video images and abstract schemas of the control interface) at one workplace. This results in a high number of monitors and is often insufficiently implemented regarding ergonomic criteria. Hence, the aim of this paper is to develop solutions for the ergonomic challenges evoked by remote lock control. As a result, a concept for an ergonomic design and work organization is presented.

METHODOLOGICAL APPROACH

The aim of this study was to develop an ergonomic design and work organization concept for RLOCs in Germany. Therefore, a two-stage methodology was developed consisting of an analysis of the current state and the development of recommendations in terms of a concept for RLOCs (Figure 1).

Within the first stage – the analysis of the current state – several aspects were differentiated: analysis of the work process, an ergonomic analysis of physical and environmental conditions and an ergonomic analysis of human-system interaction. The analysis of the work process also included factors that influence the work process in a technical and organizational sense. The analysis of the current state was conducted for a sample of 14 RLOCs in Germany and covers process monitoring for different shifts (morning, evening, night) and interviews with 42 operators. The analysis of the current state is described in the section analysis of the current state of this paper (1.work process, 2.physical and environmental ergonomics, 3.ergonomics of the human-system interaction).

The analysis of the current state allowed the identification of deficits and with respect to the ergonomic design enabled to derive design recommendations for RLOCs. As a part of these recommendations we propose a comprehensive concept for ergonomic design and work organization of RLOCs. According to the derived work process and the determining factors, first, an ergonomic concept for remote control and monitoring was developed consisting of four cases depending on the traffic volume. Second, this concept served as a basis for the suggestion of a guide for personnel management and scheduling. This guide is in turn used to determine which case of the developed concept for remote control and monitoring has to be applied. Third, an ergonomic control panel was designed based on the ergonomic recommendations for the work place and the ergonomic concept for remote control and monitoring. Furthermore, design recommendations were derived regarding the work environment and

the design of the Graphical User Interface (GUI). Some examples of these recommendations are given in the design recommendation section.

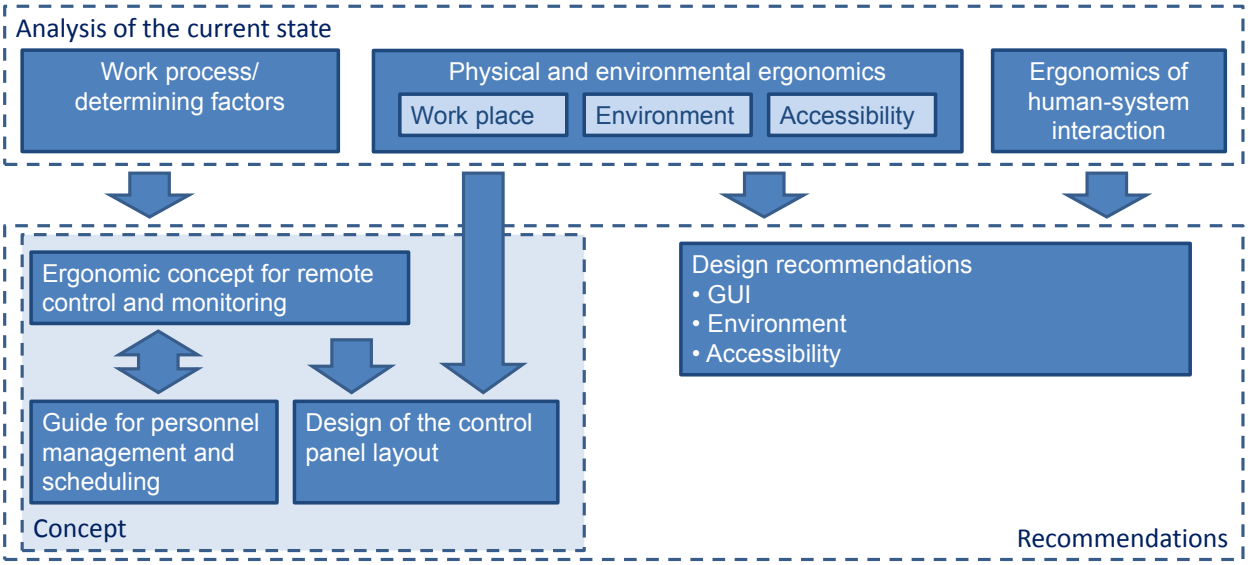


Figure 1: Methodological approach

ANALYSIS OF THE CURRENT STATE

The first step in the development process of a concept for ergonomic design and work organization of RLOCs was an investigation of the current state. This investigation involved an analysis of the work process including the specification of determining factors and the classification of main and additional tasks. Furthermore, relevant ergonomic criteria were derived from respective ISO standards and guidelines and were adapted to the practical context.

Analysis of the work process and determining factors

(a) Method

First, a workshop was conducted with representatives of all stakeholders to collect and map all occurring processes and classify them into main and additional tasks. Furthermore, information about determining factors, restrictions and special characteristics of the systems was collected. An iterative procedure based on participants' feedback in different RLOCs was chosen for modeling all main tasks. For additional tasks a similar approach was taken as all occurring tasks were collected and classified. Information about frequency, duration and the possibility to conduct the task in parallel to a main task was added.

Second, a process observation was carried out to verify the results of the workshop. Additionally, the Rating Scale Mental Effort (RSME) (Zijlstra, 1993) was used to analyze the short-term mental effort which is evoked by executing occurring tasks. The operators were instructed to rate their mental effort every 15 minutes during one shift on a scale between 0 and 150. The results were matched with the tasks that were carried out by the operator in the corresponding 15 minutes.

(b) Results

Main tasks that occur within remote lock operation centers in Germany include all operations and tasks that are directly related to the lock operation or monitoring process, e.g. gate control, light-signal control, monitoring of the gate movements and the water level adjustment in the lock chamber. For the main tasks a meta-process-model was derived integrating several variants for safety systems, light-signal and gate control as the main processes are not identical in all remote lock operation centers and depend highly on the given technical conditions. The lock operation is executed remotely via control software using a mouse as input device. The implementation of the GUI

of the control software also differs between the investigated RLOCs. Regarding the allocation of system(s) to operator, different variants exist but currently no consistent solution has been implemented for RLOCs in Germany. The monitoring process is supported in average by seven video images per lock chamber. These video images depict the current state of the lock chamber in a sequence from the upper to the lower side of the lock chamber. When a ship that accesses the chamber is monitored the operator has to observe adjacent monitors.

Additional tasks include for example telephone or radio information and guidance, administrative tasks such as traffic management and additional tasks that have to be executed at high watermark.

Regarding the short-term mental effort, the analysis revealed that long-running processes lead to subjectively higher mental effort for the operators than short processes which can be interrupted. In addition, for processes that occur in parallel or in blocks a higher mental-effort was recorded.

Physical and Environmental Ergonomics

(a) Method

For the ergonomic evaluation of the physical and environmental factors an approach consisting of two parts – an expert review with a checklist and a user interview – was adopted. The checklist was developed based on established standards and guidelines (DIN EN ISO 9241-6, 2001; DIN EN ISO 7726, 2002; DIN EN ISO 11064-5, 2008; DIN EN ISO 9241-303, 2011; DGUV: BGR 131-1, 2008; VBG: BGI 7003, 2008; VBG: BGI 5128, 2008; Probst, 2003) and adopted to RLOCs. It is divided into three main categories: (1) “work environment”, (2) “workplace and posture” and (3) “accessibility”. The first category includes items regarding the environment e.g. climate or illumination in RLOCs. Examples of items in this category are:

- Minimal illuminance is ensured (500 lx for work area, 300 lx for surrounding areas)
- Brightness distribution is adjustable
- Direct and reflected glare is limited
- Interior temperature is adjustable and between 23°C and 26°C in summer and 19°C and 24°C in winter
- Interior humidity lies between 40 and 65%
- Outdoor air supply is provided
- Noise emissions are reduced to a maximum of 55dB(A)

The second category focuses on the design of the workplace, e.g. screens, control panels and work seats. Example items are:

- Viewing distance to video monitors is between 500 mm and 1500 mm
- Viewing distance to several monitors is as constant as possible
- Minimum size of video monitors (minimum 15’’) is ensured
- Monitors are arranged within the vertical ergonomic field of view (0°-30° underneath the horizontal)
- Monitors are arranged within the horizontal field of view (0°-45° to each direction)
- Glare and reflection on monitors are prevented
- Minimal depth of the working desk (800 mm) is ensured

The last category deals with accessibility of the operation center for physically disabled people, e.g.:

- Work place is accessible for physically disabled people
- Bathroom for handicapped is installed

In order to receive an overall view of the ergonomic situation the subjective perception of the operators was recorded with user interviews. The interviews focused for example on ergonomic aspects that vary over the year (climate, illuminance) and aspects which were difficult to observe objectively (e.g. difficulties to match acoustical signals).

(b) Results

Regarding the ergonomic design, the investigated RLOCs show mainly positive results. Within the category “work environment” 74%, in category “workplace” 66% and in category “accessibility” 44% of the items were fulfilled in average. Most of the ergonomic deficits are due to the establishment of RLOCs in old lock buildings as for example lock towers. These buildings often have large windows and a multi-story architecture without elevators. For on-site lock operation a sufficient view on the lock chamber is guaranteed by this architecture but computer work is affected negatively by large windows (e.g. contrast ration between screen and background, glare).

Ergonomics of the Human-System Interaction

The work system in RLOCs consists of a software system that enables to control several technical subsystems remotely. The interface of this control software is realized via a GUI. Aim of this part of the investigation was to analyze and evaluate the usability of the different GUIs which are applied in the investigated RLOCs.

(a) Method

For the evaluation of the GUI of the control software an expert-review with a checklist based on guidelines, directives and standards (DIN EN ISO 9241-110, 2006; DIN EN ISO 9241-303, 2011; DIN EN ISO 11064-5, 2008; MIL-STD-1472D, 1994; O´Hara et al., 1994; Smith and Mosier, 1986; Widdel and Motz, 2002) was conducted. The checklist is divided into seven categories, (1) the “representation of information”, (2) “text and input formats”, (3) “display elements (characters, icons, scales)”, (4) “codes”, (5) “disorders”, (6) “process, redundancy and information requirements” and (7) “system support/learnability and safety”. Each category includes 5 to 28 items. Examples of items are:

- The required minimum font size is at least 16 arc minutes
- The GUI gives a schematic representation of the reality
- All information needed for a process is depicted on the GUI at that time
- Icons are similar to the objects/processes they represent
- Alarm and error-messages are in the front layer
- Mouse-over effect for functions gives detailed information

(b) Results

For the categories “representation of information”, “text and input formats”, “display elements” and “process, redundancy and information” 80% of the items were fulfilled in average. The items in the categories “disorders” and “system support/learnability” were only fulfilled to 64% and 51% in the investigated RLOCs. The category with the highest amount of fulfilled items is the category “codes” (94%). Altogether, none of the investigated GUIs had strong ergonomic deficits but neither fulfilled all investigated criteria.

CONCEPT

Ergonomic Concept for Remote Control and Monitoring

In order to enhance efficient work processes and develop systems that are robust against errors an ergonomic concept for remote control and monitoring was developed. This concept aimed at preventing mental over- and under load of the operators and is based on the traffic volume/condition of each system (e.g. lock chamber). Therefore, a categorization depending on the traffic volume into low-, average- and high-traffic volume systems was conducted. For high-traffic volume systems a 2:1 allocation of systems to operator is needed whereas for average-traffic volume systems a 4:1 allocation and for low-traffic volume systems a n:1 allocation of systems to operator is recommended. In addition for systems where high-risk operations occur in a high-frequency a 1:1 allocation of system to operator is suggested. It is furthermore assumed that each operator is responsible for all incoming operations at one workplace at a time. The allocation concept can vary for different workplaces within one RLOC.

(a) High-traffic volume systems (2:1 system to operator allocation)

For systems that are frequented by high-traffic a two-to-one allocation in which two systems are controlled/monitored by one operator is recommended. In this case the concept should ensure a permanent visualization of all relevant information in a symmetrical spatial arrangement to the left and to the right of the operator. Due to the stationary permanent visualization of all components of the two systems the error probability caused by confounding can be minimized. The operator is informed about the current status at any time. This supports the prospective planning of process operations. Furthermore, strain can be reduced actively because all relevant Information that is important for the decision process is directly accessible (Pashler, 1994).

(b) Average-traffic volume systems (4:1 system to operator allocation)

Average-traffic volume systems have a normal traffic volume that is (relatively) uniformly distributed within one shift and allow the operator to be responsible for up to four systems including the workload of incoming additional tasks. In this case we recommend a hybrid control and monitoring concept which consists of two control panels that are symmetrically arranged to the left and to the right of the operator. Each control panel displays one system at a time and allows switching between two systems. This concept enables the operator to control two systems in parallel which can be selected via a corresponding switching element and presents an upper limit for concurrent operations (Müller-Jung, 2010). Two systems are assigned permanently to one control panel and allow no switching between the two panels, which in turn reduces possible confusion. Furthermore, switching between the systems is only allowed via the switching element on the corresponding control panel. This minimizes the “costs of switching” (Monsell, 2003). Due to only binary transitions between systems, the deficits that occur for free assignment (conformity with user expectations, suitability for learning) can be reduced.

Error tolerant controlling/monitoring is supported by the permanent assignment of a process to whether the right or left side of the workplace within one shift. This hybrid concept guarantees that one operator can only control two processes in parallel. Ongoing processes that are switched to the background and are not visible any more are automatically brought to a save state by the system (corresponding to the emergency stop described in DIN EN ISO 13850 [2008]) as far as a continuous monitoring is necessary.

(c) Low-traffic volume systems (n:1 system to operator allocation)

Low-traffic volume systems require only little time for controlling/monitoring within a shift, for example during night shift when the traffic volume is at a low level. In this case one operator can be responsible for more than four systems. For operation hours with low workload of main and additional tasks a concept is recommended that allows allocating any number of systems to one operator. However, this concept involves certain limitations compared to the other concepts in order to ensure a process that prevents errors. To avoid confusion each operator is only allowed controlling one process at a time. Therefore the workplace consists of one control panel which visualizes exactly one system at a time. Concentration on solely one work process guarantees a sufficient conformity with user expectations during the control process despite multiple systems (Trouvain, 2008). Basic requirement for this solution is that the control and monitor operations are uniformly distributed within one shift and do not occur simultaneously. The application of this concept is not recommended when the workload is low but occurs simultaneously in blocks. In this case, the limitation of controlling no more than one system at a time would lead to a reduced efficiency and a temporally high strain of the operator.

(d) High-risk operations that occur in a high-frequency (1:1 system to operator allocation)

For high-risk operations, which allow no parallel processing due to high workload or the case that events emerge unpredictably and in blocks, a one-to-one allocation of system to operator is required. In this case a concept that allows only visualizing one system at one workplace is recommended. As a result, all required information is visible at any time and the attention does not have to be divided between several operations (Wickens, 1984). This case is only being listed for reasons of completeness and should be considered as an option when designing similar systems.

Design of the Control Panel

Based on assessed requirements and with regard to anthropometric dimensions a concept for the control panel layout was determined. Regarding the ergonomic requirements (e.g. maximum viewing distances, viewing angle) it is advisable to keep the number of monitors which are arranged at the control panel as low as possible. Based on the

developed recommendations for the remote control and monitoring concept the representation of one system per control panel is preferred. Therefore, two control panels are needed for high-traffic volume systems: one operator controls two systems) and average-traffic volume systems: one operator is in charge of three to four systems as a maximum). For high-traffic volume systems (a one-to-one assignment between systems and control panels is needed whereas in the case of average-traffic volume systems (two systems are assigned to each of the two control panels. Low-traffic volume systems (one operator is in charge of more than four systems but does not control these in parallel) require only one control panel per operator. This control panel enables to operate any number of systems sequentially by remote control but only allows controlling of one system at a time. Switching between the systems is realized via a separate switching element. Therefore, a touch screen solution seems beneficial that allows switching between the systems as well as a visualization of the active systems. Furthermore, this switching element can be used to display the state of the respective lock chamber (e.g. via arrows) and to integrate a zoom-function for the camera.

In general, vertical adjustability of the control panel by the operator has to be guaranteed to enable the alternation of sitting and standing postures. 800 mm were chosen for the depth of the control panel in order to comply with the requirements. As the majority of considered lock chambers require displaying seven video images per system, this results in a high amount of monitors arranged at a control panel (7 video monitors, 1 control monitor, 1 traffic management monitor) and contradicts the realization of all ergonomic criteria. However, since a reduction of the video images that are displayed permanently is accompanied by an information deficit, the number of the displayed video images per lock chamber and thus per control panel was kept at seven video images. For each of the seven video monitors a minimum size of 15'' was chosen, whereas for the control monitor and the traffic management monitor a larger size (e.g. 22''-Widescreen) is recommended in order to depict more detailed information and to meet the required minimum of angular font size.

Regarding the monitor arrangement several options were considered. One option is to arrange all nine monitors in a row with a centered control monitor and a traffic management monitor. An argument against this arrangement is the violation of the maximum viewing distance for the external monitors. Furthermore, the arrangement of the control monitor and the traffic management monitor in between the video monitors would lead to a spatial separation of the video images. Another option is to position the control monitor and the traffic management monitor in front of the video monitors, however, for this arrangement the video monitors would be masked by the control monitor and traffic management monitor. A further option is a monitor arrangement in two rows, but since it is not possible to arrange all monitors within the vertical ergonomic field of view (0°-30° underneath the horizontal) (Schlick et al., 2010) it does also not provide a sufficient solution. As an alternative we recommend to arrange all 7 video monitors in a row underneath the horizontal visual axis and to position the control monitor as well as the traffic management monitor on a 20°-sloped board in the control panel (Figure 2). The two monitors arranged on the sloped board are adjustable regarding the inclination in order to prevent glare. The vertical angle of view on the control monitor measures for this arrangement between 19.4° (5. percentile women) and 29.7° (95. percentile man) and therefore fulfills the ergonomic requirements. The vertical angle of view on the video monitors measures between 4.3° (5. percentile women, regarding the external video monitors with the highest viewing distance) and 17.2° (95. percentile man, regarding the centre video monitor with the lowest viewing distance). It has to be mentioned that, since a movement of the operator is assumed, there is always a certain variability of the viewing angle.

Due to the high amount of monitors that need to be observed, it is not possible to arrange all video monitors of a control panel within the required maximum viewing distance in the horizontal field of view (45°). Therefore, the extended field of view (Schlick et al., 2010) with an angle of 70° served as a constraint. The viewing distance to the video monitors amounts between 980 mm (due to the required depth of the panel) and 1400 mm (in compliance with the maximum viewing distance). This results in a curvature radius of the table plate of R=2267.04 mm. Unfortunately, larger differences in the viewing distances between the different video monitors have to be accepted, but there is no solution that meets all ergonomic criteria for the requested length of the monitor row. However, since only adjacent monitors are observed and a moveable chair is required, this criterion can be neglected for the design.

Control panel case (a) and (d) **Control panel case (b) and (c)**

1 – Control Monitor
2 – Video monitor
3 – Nautical information radio
4 – Touch screen element for switching
5 – Safety-stop
6 – Traffic management monitor

Figure 2: Three-dimensional model of the workplace for cases (a) and (d) [left] and (b) and (c) [right].

The control monitor and the traffic management monitor are arranged on the sloped board with a distance of 601.7 mm to the front edge of the panel. In doing so, a viewing distance larger than the required minimum viewing distance of 500 mm is also ensured for these monitors. The control monitor that depicts the GUI is needed for the main task of lock control and therefore arranged in the center of the panel. The traffic management monitor is needed for additional tasks (e.g. register ships that have been locked) and arranged in the right part of the panel. The board also serves to arrange the device for safety-stop (emergency stop and lock stop), the nautical information radio and the touch screen element. These devices usually require just grasping or reaching movements or only a short switching and thus can be arranged on the board in a control panel area distal from the body. The sloped board facilitates the grasping motion. For the lock control which is conducted via the control monitor no keyboard input is required and the mouse serves as sole input-device. For the traffic management text-input is required and therefore a mouse and a keyboard are needed. The input devices are arranged in the primary grasping area close to the respective monitors (Figure 2).

For high-traffic volume systems and average-traffic volume systems the workplace for an operator consists of two control panels which are arranged next to each other with preferably low distance in between (Figure 2). In the case of low-traffic volume systems and high-risks operations that occur in a high frequency, however, one control panel is sufficient. The relevant measures for the ergonomic design of the control panel are specified in figure 3.

An important component of the human systems integration plan should be a verification and validation process that provides a clear way to evaluate the success of human systems integration. The human systems integration team should develop a test plan that can easily be incorporated into the systems engineering test plan. The effectiveness and performance of the human in the system needs to be validated as part of the overall system. It may seem more attractive to have stand-alone testing for human systems integration to show how the user interacts with controls or displays, how the user performs on a specific task. This methodology can address the performance of the human operator or maintainer with respect to the overall system. The most important thing is to develop a close relationship between human systems integration and systems engineering.

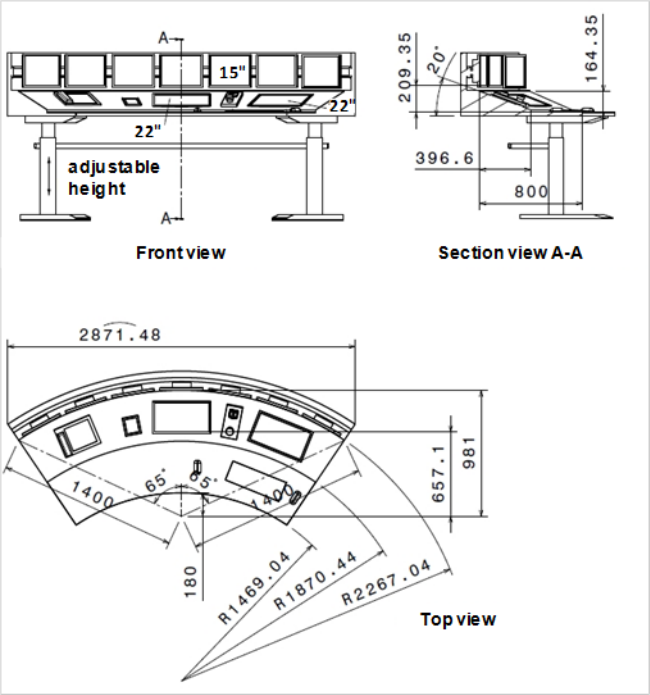


Figure 3: Measurements for the ergonomic design of the control panel

Guide for personnel management and scheduling

According to the introduced control and monitoring concept an approach for personnel resource planning and scheduling is proposed that not only integrates the strain evoked by main tasks but also considers the impact of additional tasks.

In order to obtain an objective quantification of the workload, a method based on statistical analysis of traffic data and standardized questionnaires was applied. Based on operation journals, system logs or statistics, the workload due to the main tasks has to be determined. Furthermore, the workload evoked by additional tasks has to be identified via information about the frequency of occurrence and duration of each individual additional task. This ensures a valid basis for the required prospective personnel management and scheduling. The determination of the cumulative workload is always conducted for a specified period within which a continuous or homogenous workload is expected (e.g. for one shift). This allows an application of the results to all periods with similar conditions. Aim of these determinations is to model the impact of different concepts on the workload as well as the efficiency of the remote lock operation center. Based on the results the most suitable constellation can be chosen

(a) Main tasks

First, the personnel which are required to manage the workload of the main tasks have to be estimated. Therefore, the time for the main tasks is evaluated and accumulated for each operator considering the respective control concept. This allows monitoring the workload for the entire RLOC regarding the temporal resources and monitoring the workload of each operator according to the individual conditions. It has to be considered that parallel processing is allowed for the 2:1 allocation and the 4:1 allocation and therefore a correction factor has to be integrated for the calculation for each operator to avoid a worst-case-estimation.

This correction factor (percentage of parallel tasks) can be calculated retrospectively based on statistics or determined via estimation and is multiplied by the total time and subtracted from it. To estimate the probability of temporal parallelism, the accumulated duration for each task has to be detected and put into relation to the investigated time period in order to determine the percentage of the respective time period. Subsequently, these values are multiplied for all tasks one operator is responsible for in order to receive the probability of parallel activities (Bortz, 1993. The multiplication of the probability of parallel tasks with the total duration of the main tasks results in the corrected value for the required time for main tasks.

(b) Additional tasks

Furthermore, the required personnel due to additional tasks have to be considered. Therefore, it has to be differentiated between additional tasks that can be carried out in parallel to main tasks and additional tasks that do not allow parallel processing. Tasks which can be carried out in parallel need to be taken into account in terms of higher strain for the operator, but do normally not require additional personnel resources as the required time for the additional tasks per shift is normally smaller than the total working time minus the calculated time for the main tasks. Resources for additional tasks that do not allow parallel processing need to be included and added to the required personnel resources calculated for the main tasks. Additional tasks that occur in blocks afford a separate determination of the required personnel for the respective time period. For the quantitative analysis of the additional tasks for a specified time period for one RLOC a questionnaire was developed. The questionnaire addresses issues such as frequency of additional tasks, duration of tasks, and parallel operation to main tasks.

Based on the identified task profiles, the workload and amount of operators can be modeled mathematically for each control concept. Furthermore, the straining condition that occurs for each operator can be estimated for each constellation. As a result, the most suitable constellation which leads to an ergonomic, optimal loaded and economic work for the operator can be established.

DESIGN RECOMMENDATIONS

Based on the investigations we were able to derive generic design recommendations for RLOCs. As it is not possible to provide a comprehensive overview in this paper, only some of the most important recommendations are summarized exemplarily.

Concerning the environmental ergonomic design and under consideration of the detected deficits in the operation centers, we among others recommend:

- Ventilation with external air
- Equal distributed illumination
- Glare shields
- Better accessibility for physically disabled people
- Ergonomic seats

Regarding the GUI of the control software we recommend for instance:

- Standardization of the GUI to facilitate the learning process and reduce errors when rotating between different RLOCs (especially important for safety-critical operations)
- An automatic prioritization of error-reports by the software system in order to facilitate the decision process of the operator and reduce strain.

In figure 4 the GUI for the lock control is shown exemplarily. Essential elements that need to be depicted are the lock chamber, gates, paddles and signals. These elements should give a schematic illustration of the respective lock system. This example depicts a lock chamber with two mitre gates and ground paddles (culverts). All elements that move (gates and paddles) need to be highlighted when moving as well as the arrow which indicates that a locking is operating and the direction of the locking (e.g. blinking). All values that need to be monitored are labeled and arranged close to the respective elements to prevent confusion. If relative values need to be monitored as for example the difference between the water level in the chamber and the water level at the upper level these relative values are given directly by the system (e.g. labeled with level difference). The operating mode indicates if the system runs in automatic mode (start is initialized by the operator and the sub-operations are run automatically) or in manual-remote mode (each sub-operation has to be initialized by the operator manually). For some locks different locking categories, depending on the size of the ships or amount of ships that need to be locked, are furthermore differentiated.

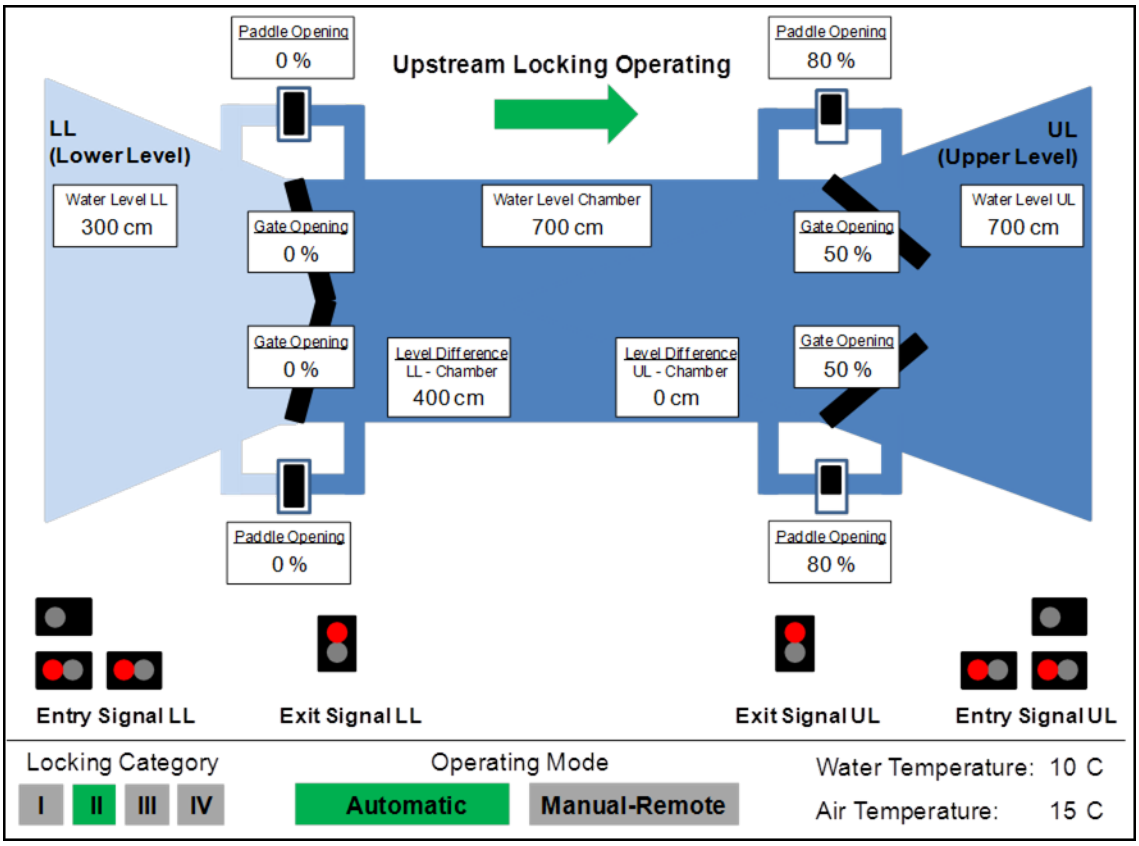


Figure 4: Example depiction of the GUI of the control monitor

CONCLUSION AND OUTLOOK

In this paper a concept for the ergonomic design and work organization in RLOCs is introduced. An approach consisting of two parts was used. The current state of 14 RLOCs in Germany was investigated regarding work process, physical and environmental ergonomics and ergonomics of the human-system interaction. Based on the results of the evaluation, recommendations for improvement were given and a new concept for the ergonomic design and work organization in RLOCs was developed. This concept includes an ergonomic concept for remote control and monitoring, an accordingly designed control panel and guidance for personnel management and scheduling. It differentiates four cases depending on the traffic volume. For each case a specific allocation of system(s) to operator is proposed and the respective ergonomic realization of the control panel. The decision on the case to be applied is based on the calculation via the guide for personnel management and scheduling.

The next step is to implement this concept to RLOCs and to investigate it in a long-term study. Therefore, a prototype control panel was constructed and is currently being discussed with the operators to validate the theoretical derived concept in praxis.

ACKNOWLEDGEMENTS

The paper refers to a project that was conducted on behalf of the Federal Ministry of Transport, Building and Urban Development (BMVBS) and the Waterways and Shipping Administration of the Federal Government (WSV). We are grateful for the close cooperation with the WSV and BMVBS and thank all participants for their contribution, especially the operators in the RLOCs for their support (contact: BMVBS, Referat WS11, Thomas Rosenstein).

REFERENCES

Bortz, J. (1993). *"Statistik für Sozialwissenschaftler*" (4. Aufl.). Berlin: Springer.

DGUV: BGR 131-1. (2008). 'Natürliche und künstliche Beleuchtung von Arbeitsstätten Teil 1: Handlungshilfen für den Unternehmer', (Berufgenossenschaftliche Regel Nr. BGR131-1). Berlin : DGUV.

DIN EN ISO 9241-6. (2001). 'Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten - Teil 6: Leitsätze für die Arbeitsumgebung'.

DIN EN ISO 7726. (2002). 'Umgebungsklima - Instrumente zur Messung physikalischer Größen'.

DIN EN ISO 11064-5. (2008). 'Ergonomische Gestaltung von Leitzentralen - Teil 5: Anzeigen und Stellteile'.

DIN EN ISO 9241-303. (2011). 'Ergonomie der Mensch-System-Interaktion - Teil 303: Anforderungen an elektronische optische Anzeigen'.

DIN EN ISO 13850, (2008). 'Sicherheit von Maschinen - Not-Halt – Gestaltungsleitsätze'.

DIN EN ISO 6385, (2004). 'Grundsätze der Ergonomie für die Gestaltung von Arbeitssystemen'.

DIN EN ISO 9241-110, (2006). 'Ergonomie der Mensch-System-Interaktion - Teil 110: Grundsätze der Dialoggestaltung' .

Ivergard, T., Hunt, B. (2008). "*Handbook of Control Room Design and Ergonomics*", 2. Aufl. Boca: CRC Press Inc.

MIL-STD-1472D. (1994) 'Human Engineering Design Criteria for Military Systems and Facilities', Washington, DC: US Department of Defense.

Monsell, S. (2003). "*Task switching*", Trends in Cognitive Science Volume 7, pp. 134-140.

Müller-Jung, J. (2010). *"Resultat der Hirnforschung - Multitasking ist ungesund*", Frankfurter Allgemeine Zeitung, Onlineveröffentlichung (Stand: November 2012).

O´Hara, J.M., Brown, W.S., Baker, C.C., Welch, D.L., Granda, T.M., Vangelis, P.J. (1994). "*Advanced Human-System Interface Design Review*", Guideline. Document NUREG/CR-5908, BNL-NUREG-52333, Vol. 2. Washington, DC: US Nuclear Regulatory Commission.

Pashler, H. (1994). *"Dual-task interference in simple tasks: Data and theory"*, Psychological Bulletin Volume 116, pp. 220-244.

Probst, W. (2003). "*AWE 123: Bildschirmarbeit -Lärmminderung in kleinen Büros, Arbeitswissenschaftliche Erkenntnisse*", Dortmund : BAuA.

Schlick, C.M., Luczak, H., Bruder, R. (2010). "*Arbeitswissenschaft*" (3. Auflage). Berlin: Springer.

Smith, S.L., Mosier, J.N. (1986). "*Guidelines for Designing User Interface Software"*, (Report-No. ESD-TR-86-278). Bedford, MA: MITRE.

Stanton, N.A., Salmon, P.M.; Jenkins, D.P., Walker, G. H. (2010). "*Human Factors in the Design and Evaluation of Central Control Room Operations*", 1. Aufl. Boca Raton (FL): Crc Press Inc.

Trouvain, B. (2008). "*Theoretische und experimentelle Untersuchungen zur Gestaltung multimodaler Mensch-Roboter-Schnittstellen für mobile Mehrrobotersysteme"*, Schlick. C.M. (Hrsg.), Industrial Engineering und Ergonomics (Band: 5). Aachen: Shaker Verlag.

VBG: BGI 7003. (2008). 'Beurteilung des Raumklimas, VBG Fachinformationen', Hamburg : VBG, BAuA.

VBG: BGI 5128, (2008). 'Arbeitsstätten sicher planen und gestalten, Berufsgenossenschaftliche Informationen', Hamburg : VBG.

Wickens, C.D. (1984). "*Processing resources in attention"*, Parasuramen. R., & Davies, D. R. (Hrsg.), Varieties of Attention. Orlando: Academic Press, pp. 63-101.

Widdel, H., Motz, F. (2002). "*SEBEW - Software-ergonomisches Bewertungsverfahren*", FKIE-Bericht Nr. 40, Wachtberg.

Zijlstra, F.R.H. (1993). "*Efficiency in work behavior: A design approach for modern tools*", Delft: Delft University Press.

Section 19

Road and Rail:

Driver Distraction and Inattention

Perceived and Real Impacts of Digital Advertising Billboards on Driving Performance

Virginia P. Sisiopiku[1], Md Mozahidul Islam[1], Shannon Wittig[2], Sharon C. Welburn[2], Despina Stavrinos[2]

[1] Civil, Construction, and Environmental Engineering; [2] Psychology
University of Alabama at Birmingham
Birmingham, AL, 35294, USA

ABSTRACT

This paper explores perceived and real impacts of digital advertising billboards on driving performance of individuals from representative samples across the lifespan. Perceived impacts were assessed through an online driver questionnaire that documented perceptions and attitudes of Alabama motorists as they relate to roadside billboards. The questionnaire solicited demographic information, inputs on driving patterns and attitudes, drivers' experiences with advertising billboards, and perceptions regarding presence and type of advertising billboard and traffic safety. Actual impacts of digital advertising billboards on driving performance were examined using a data-informed, empirical driving simulator study. Driving performance variables of interest were electronically coded by the simulator (i.e., lane deviations, speed exceedances, and crashes and statistical analyses were performed to test whether digital billboards presented driving performance decrements. The paper provides detailed description of the study approach and major findings and discusses conclusions and recommendations. The study sheds light on the impact of roadside advertisements on driving performance across the lifespan which, in turn, provides useful insights on the potential links between digital outdoor advertising and traffic safety.

Keywords: Digital advertising billboards, driving performance, traffic safety, driver perceptions, driver simulation.

INTRODUCTION

Roadside advertising billboards are used for advertisement of various products and services and are meant to attract drivers' attention to the message or information conveyed by the billboards. According to the Outdoor Advertising Association of America (OAAA), there were over 365,000 unique billboard faces in the United States in 2013 (Outdoor Advertising Association of America [OAAA], 2013). Roadside advertising billboards can be either static or digital. Static billboards show the same message for an extended period of time (typically days). They are the traditional type of outdoor advertising and the most commonly used type of advertising billboards in the United States. The digital billboards (DBBs) were introduced in the recent years and utilize light-emitting diode (LED) technology to show multiple messages one after another that are updated using computer input. Because DBBs flash images every four to ten seconds (Copeland, 2010), a single board can advertise to far more clients than a traditional board, making them an attractive advertisement option. Thus, despite the fact that DBBs are initially more expensive to build compared to their static counterparts, over time they prove to be cost-effective. While static billboards are still dominant, digital billboards are a fast growing sector of the outdoor advertising market (OAAA, 2013).

The increased number and sophistication of DBBs raises questions about their potential impact on traffic safety. As an advertising medium, DBBs purposely encouraging drivers to shift their attention away from the driving task.

Moreover, DBBs brightness may be especially problematic at night and may affect the driver's ability to observe changes in the surrounding environment such as brake lights or signal changes. Also, frequently changing images may compel more glances, and sequential messages may hold drivers' gazes longer until the entire message is read. Lastly, targeted messages that promote interactivity with the driver are particularly troublesome as they are hypothesized to be distracting to the driver (Sisiopiku et al, 2013).

Several studies have been performed worldwide to document the relationship between roadside advertising billboards, driver distraction, and traffic safety. These include a) crash studies analyzing historical crash records, b) laboratory studies using driving simulators, and c) naturalistic studies observing driver behaviors on-road using instrumented vehicles. A comprehensive synthesis of the literature is available at Sisiopiku et al. (2013) while Wallace (2003), Coetzee (2003), Molino et al. (2009), and Wachtel (2009, and 2011) provide useful insights on the subject of outdoor advertising and driver distraction through literature reviews and meta-analyses.

Several earlier studies provide some evidence for a correlation between advertising billboards and increased driver distraction. However, local conditions, experimental settings, and other factors may play a role in the impact of driver distraction due to advertising billboards on traffic safety. Overall, the outcomes of earlier research studies have resulted in somewhat contradictory conclusions, indicating a need for further research. Furthermore, earlier studies have looked at the distractive effects of billboards in the general population; however, little research has been done to examine how these distractive effects differ across the lifespan, namely young drivers, middle aged drivers, and older adults (65 years and older).

To bridge existing knowledge gaps, this study investigated the potential relationship between DBBs and traffic safety based on perceived and real impacts of DBBs on individual drivers across the lifespan. In order to determine the perceived impacts of roadside advertisement, an online driver questionnaire survey was conducted and used to document perceptions and attitudes of Alabama motorists as they relate to roadside billboards. The real impacts of DBBs on driver performance where obtained from the driving simulator experimental study. This study focused on the effect of external distractions from billboards on driving performance.

SURVEY OF ROAD USERS

Methods

One straightforward approach toward understanding transportation users' choices and behaviors is through questionnaires. In the present research, an online questionnaire instrument was developed and used in 2013 to gather and analyze data from Alabama road user's perceptions and attitudes related to roadside advertising billboards. The questionnaire included a total of 22 questions that assessed several variables of interest including demographic information (e.g., age, ethnicity, and gender), exposure (driving patterns and experience, frequency of billboard encounters), driver's behaviors, attitudes, and perceptions toward billboards with respect to safety and efficiency, and respondents' stated preferences regarding placement, frequency and regulation of roadway advertising billboard. To ensure random sampling, a company specialized in web based surveys was hired to recruit a diverse group of survey participants. In order to be eligible to participate in the survey, subjects had to possess a valid driver's license and reside in Alabama.

Analysis

In aggregate, 295 respondents from Alabama participated in this survey. Incomplete questionnaire responses were omitted in order to maintain consistency for analysis. Eventually, responses from 231 participants across the lifespan were used in the analysis. The questionnaire extracted information related to driver demographics, driving experience level, perception towards billboards, in general, and digital billboards, in particular, attitudes related to use of information billboards, and perceptions on traffic safety with respect to billboards and digital billboards. Participants' questionnaire responses were collected and then processed using 'Microsoft Excel' to make usable for analysis.

Results

Out of 231 questionnaire respondents, 133 (57.58%) were male and 98 (42.32%) were female drivers. Aggregate responses from the questionnaire are summarized in Table 1.The findings reveal that 45.89% of respondents find billboards distracting in general, and an overwhelming 67.53% perceive DBBs as more distracting than static billboards. Moreover, the majority responded that they are more likely to read a message on a DBB rather than a static billboard. The majority (58.88%) also admitted that they stare at a DBB long enough to read the entire message but they rarely slow down (87.88%) when doing so. Interestingly, while responders admit that the messages posted on DBBs capture their attention, three fourths of them (74.46%) state that they rarely use the information.

Table 1: Aggregate response from online questionnaire survey

Question or Information	Response	% of total respondents
Are billboards distracting in general?	Yes	45.89
	No	31.60
	Not sure	22.51
Do you think that DBBs are more distracting than static billboards?	Strongly agree	22.08
	Agree	45.45
	Neither agree nor disagree	20.35
	Disagree	11.26
	Strongly disagree	0.87
Are you more likely to read a message on a digital billboard than a static one?	Yes	48.92
	No	38.10
	Not sure	12.99
Do you glance long enough at a DBB to read the entire message?	Rarely	25.54
	Sometimes	42.86
	Often	16.02
	It depends on message	15.58
How often do you slow down to read a DBB message?	Rarely	87.88
	Sometimes	10.82
	Often	1.30
How often do you use the information from DBBs?	Rarely	74.46
	Sometimes	23.81
	All the time	1.73

For further analysis, the drivers were categorized into 7 age classes as summarized in Figure 1. Approximately 13% of responders were under 20 years of age and 11.26% were older than 55. The responses were then stratified according to the age the participants. When asked about their perception as it related to billboard distraction, 106 respondents (45.89%) reported that billboards cause 'distraction.' The respondents in the 56-65 year old bracket had maximum rate of agreement on the issue of distraction from presence of billboard (65%). The younger driver population, i.e., drivers of ≤20 years and 21-25 years of age also had a high percentage on the agreement that the billboards cause distraction (53.33% and 46.34%, respectively). The findings are summarized in Figure 2.

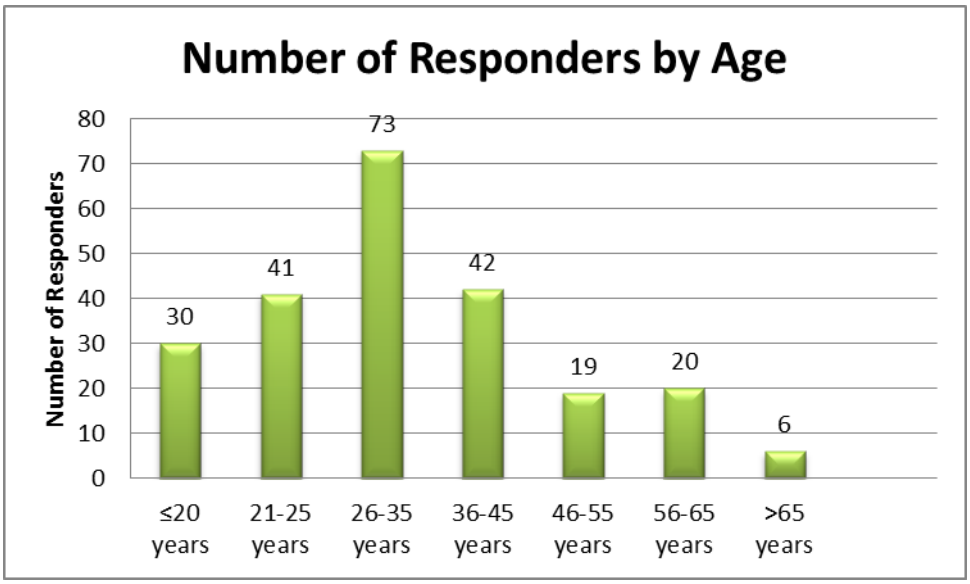


Figure 1: Number of respondents with age class

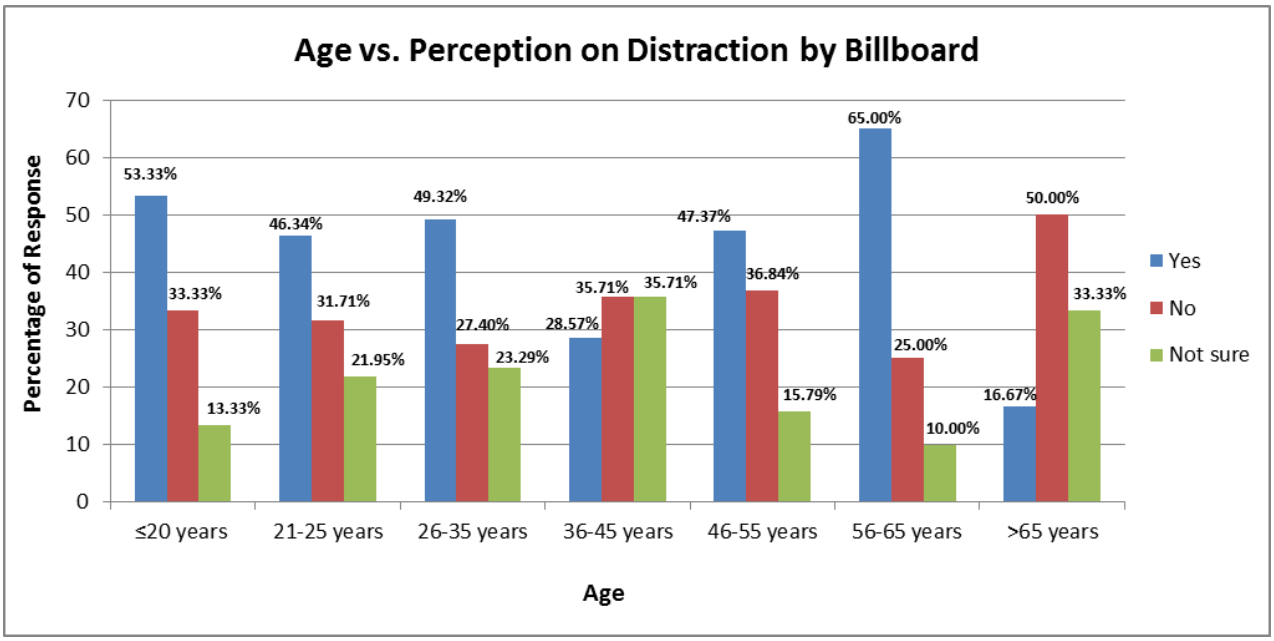


Figure 2: Age vs. perception on distraction by billboard

When asked if DBBs are more distracting than static billboards, nearly half of the respondents (45.45%) agreed on the greater distracting power of the digital billboards compared to static billboards. Also, as shown in Figure 3, approximately 56%of those 21-25 years of age agreed that digital billboards are more distracting. The percentage is higher than any measures from other age classes. Their immediate juniors (≤20 years) were not far behind (53.33%) to accept the notion of potentially more distracting power of digital billboards. So, it can be inferred that, the rate of acceptance of potential distraction by digital billboards in this study was higher among young drivers.

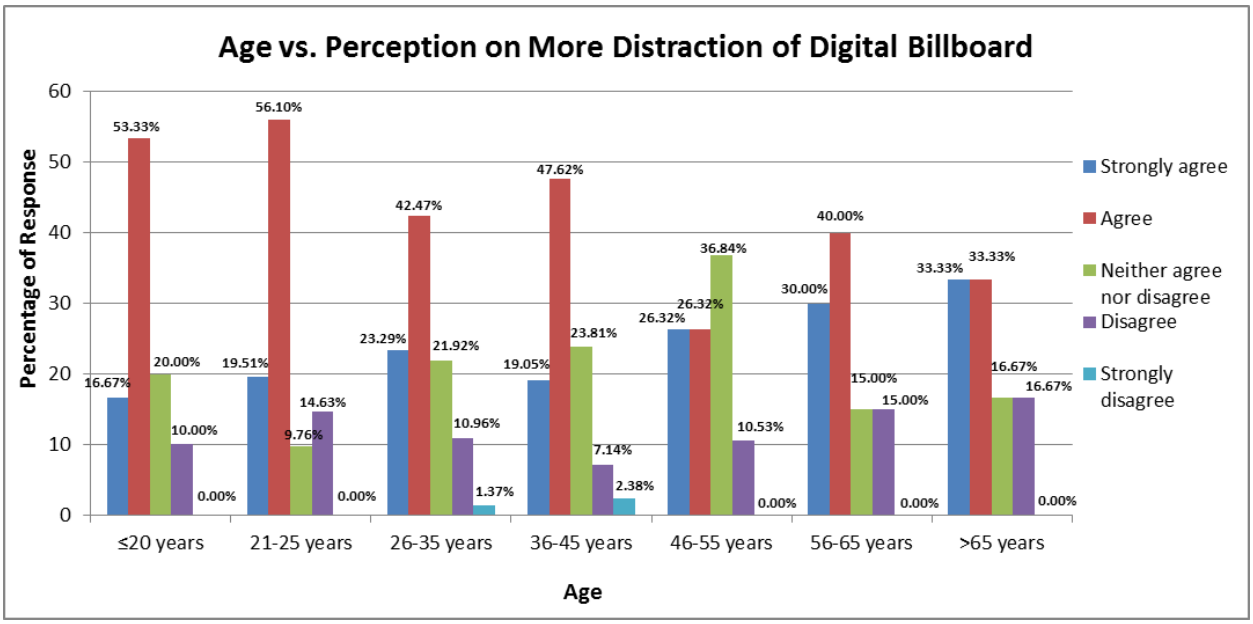


Figure 3: Age vs. perception on more distraction of digital billboard

Almost half of the respondents also mentioned that they are more likely to read messages from digital billboards (48.92%). This is a clear intention of the road users to be tempted by messages from digital billboards. Taking gender into consideration, the tendency was greater among male drivers (52.63%) than their female counterparts (43.88%). Interestingly, as depicted in Figure 4, this response was fairly consistent across all age groups, including the elderly.

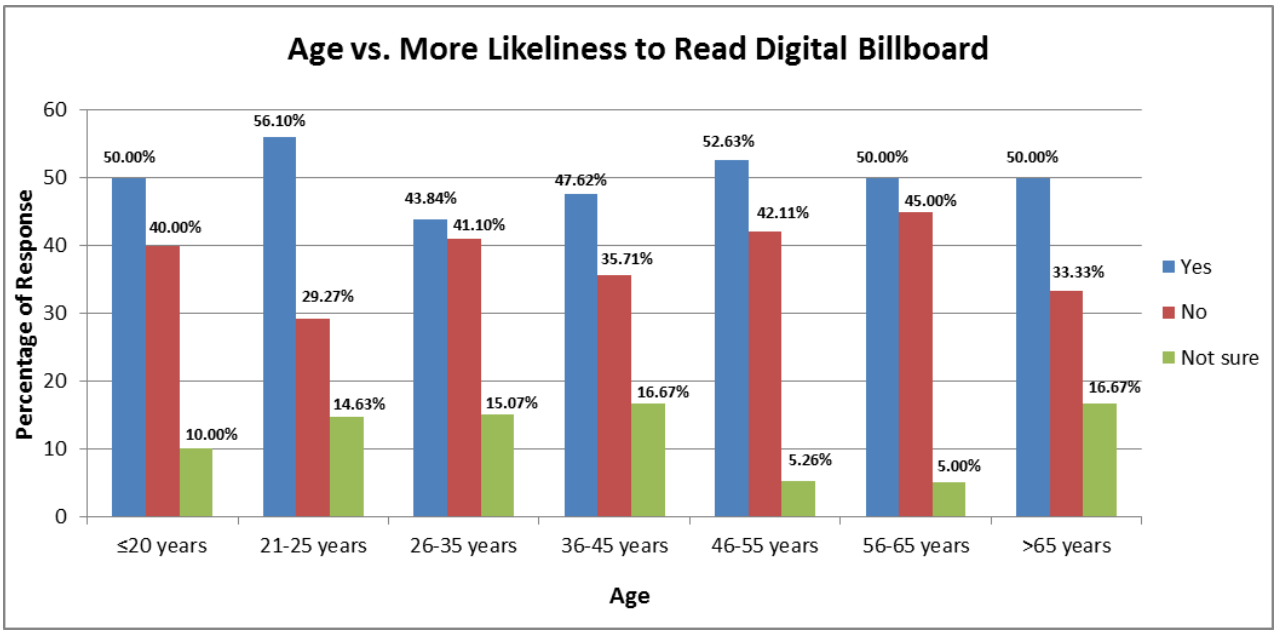


Figure 4: Age vs. more likeliness to read digital billboard

The analysis also revealed that over 42% of the road users sometimes glance at the digital billboard for significantly long time. Although the exact time was not described, the term 'long' may be akin to several seconds. This rate was highest among the participants when asked about long glance, meaning that the digital billboards can make people to look at them for a significantly 'long' time. This scenario (long glance at digital billboard) was further broken down by age class and the results are shown in Figure 5. More than half (56.67%) of the young drivers (≤20 years of age) 'sometimes' looked at the digital billboard for a long time, which is quite natural because the respondents of this age might have a curiosity to the appearance and messages of digital billboards. Though they sometimes glance for a long time, a small percent of drivers across the lifespan reported doing it 'often'.

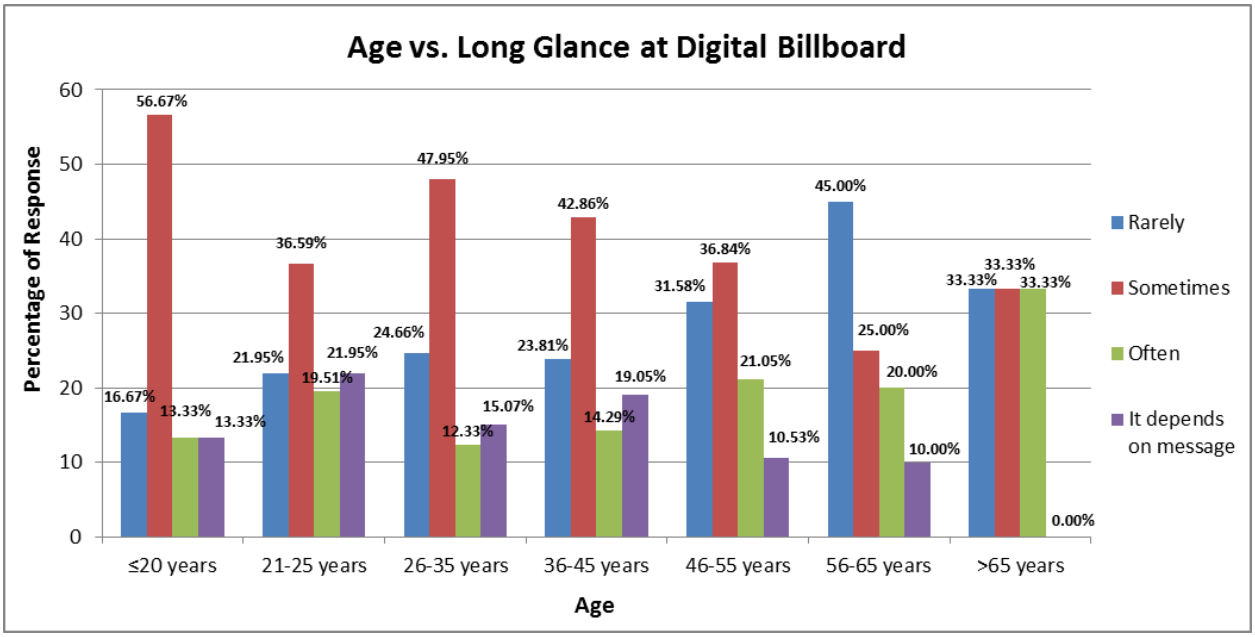


Figure 5: Age vs. long glance at digital billboard

It can be deduced from the analysis of the responses that the overwhelming majority of the questionnaire participants (87.88%) had a rare tendency to slow down near digital billboards. Very small percentage of the drivers 'sometimes' reduced their speed (10.82%). Figure 6 shows the result of 'slow down at digital billboard' scenario based on age. The youngest driver group (≤20 years) rarely reduced their vehicle speed disregarding the presence of digital billboard.

Interestingly, most of the participants stated that they rarely used information from digital billboards, and just over one-fifth of them (23.81%) used the information sometimes. The rate was highest (36.84%) for participants between 46 and 55 years of age. As can be seen in Figure 7, the youngest population group and the older population (>65 years) showed almost no intention to use digital billboard's information.

Survey participants were asked also their opinion regarding the need for stricter regulation of billboards. Nearly 60% of responders suggested that there should be restrictions on all billboard locations for the purpose of traffic safety and agreed on the need for better regulation of the size and number of DBBs.

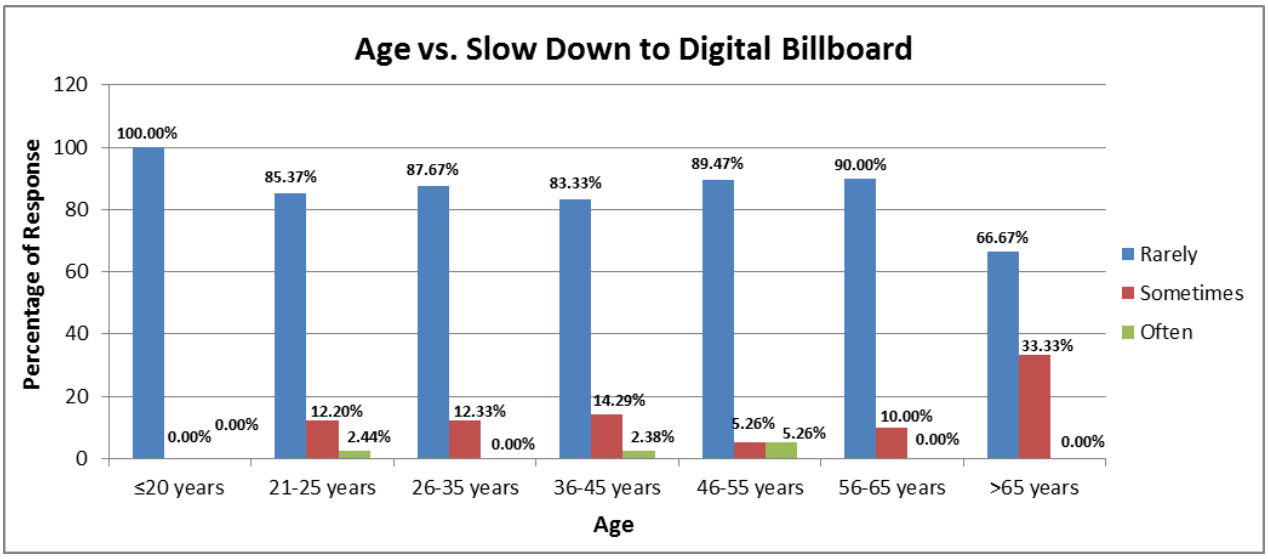


Figure 6: Age vs. slow down to digital billboard

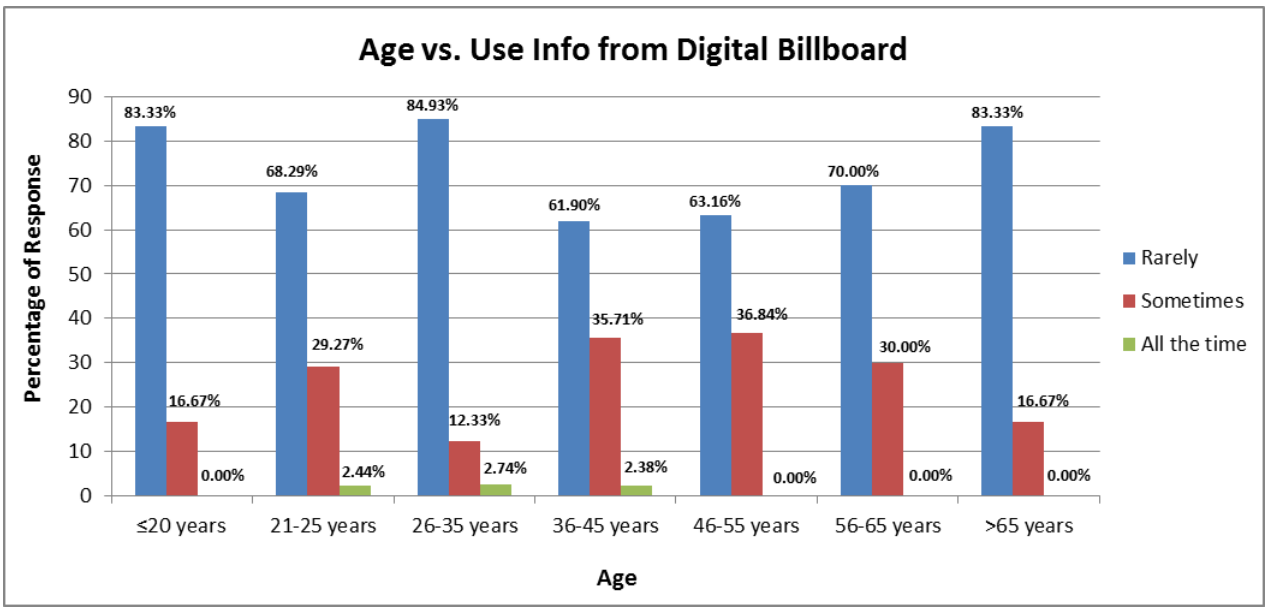


Figure 7: Age vs. use of information from digital billboard

Apart from the general analysis of the responses between genders and age groups, chi-square test has been performed across age groups and gender separately. The observed values for the chi-square test have been found from the survey itself and the expected values have been determined. The result of this test has been presented in Tables 2 and 3.

The p-values from Table 2 suggest that there is no significant difference among responses across different age groups of drivers when asked for their perception and/or actions. From Table 3 it can be implied that, there is no significant differences between the responses of male and female drivers when asked for their perception and/or actions.

Table 2: Chi-square test result for age groups

Notion/Information/Query	p-value	Comments
Are billboards distracting in general?	0.2855	Not significant
Do you think that DBBs are more distracting than static billboards?	0.9323	Not significant
Are you more likely to read a message on a digital billboard than a static one?	0.9488	Not significant
Do you glance long enough at a DBB to read the entire message?	0.7010	Not significant
How often do you slow down to read a DBB message?	0.3677	Not significant
How often do you use the information from DBBs?	0.2768	Not significant

Table 3: Chi-square test result for male and female

Notion/Information/Query	p-value	Comments
Are billboards distracting in general?	0.6431	Not significant
Do you think that DBBs are more distracting than static billboards?	0.6610	Not significant
Are you more likely to read a message on a digital billboard than a static one?	0.2938	Not significant
Do you glance long enough at a DBB to read the entire message?	0.3410	Not significant
How often do you slow down to read a DBB message?	0.6763	Not significant
How often do you use the information from DBBs?	0.3405	Not significant

Discussion

The analysis of the driver questionnaire has produced some interesting findings regarding the perceptions and attitudes of Alabamians with respect to DBBs. Younger drivers (≤20 years and 21-25 years) have a significant agreement rate on the distraction caused by the billboards in general. They also provided similar opinions when asked if digital billboards are more distracting than static billboards. So, this younger driver group actually perceived the distraction caused by billboards and also the higher distraction level when the billboards are digitized. It has also been revealed that young drivers usually have a long glance at digital billboards but very rarely slow down. This behavior might be a matter of concern as it could lead to potential risk for traffic crash occurrence.

Another interesting finding is that the older population group (>65 years) was more likely to use information from digital billboards but barely used the information conveyed by the billboards. In fact, the effectiveness of billboards, in general, and DBBs in particular to convey a message to the drivers is found to be questionable since the vast majority of respondents confirmed that they rarely use information from outdoor advertising billboards.

The analysis of aggregate responses of the drivers showed that almost half of the participants agreed that billboards distract drivers while 22.51% 'were not sure'. A similar percentage of drivers perceived that the digital billboards are more distracting than their static counterparts. The online survey also suggests that more than 40 percent of the drivers looked at the digital billboards for a sufficiently long time, but most of the drivers barely slowed down. This behavior is a matter of concern as the combination of speed and inattention is found to increase the risk for a crash. Last but not least, survey responders emphasized the need for stricter regulation of DBBs and restriction of size and frequency of placement for the benefit of traffic safety.

At the end, chi-square test has been performed across age groups and genders discretely. There was no significant difference in the responses among the drivers groups. No significant change was also found between the responses of male and female drivers when asked about their perception or action while driving.

DRIVING SIMULATOR STUDY

Method

The current study aimed to evaluate the distractive effects of roadside billboards through the use of a driving simulator, providing a safe environment for imposing driver distractions. Participants were recruited using advertisements on social networking websites, flyers, and letters and were asked to drive through a simulated scenario embedded with a variety of billboards (static and digital).Sixty-six participants were recruited and divided into three groups: 16–19 years old for teens (n = 20), 35–55 years old for middle adults (n = 21), and 65 and older for older adults (n = 25).

Participants were familiarized with the simulator during a brief, 2.84 mile, standardized four lane highway calibration scenario to assure that they met a minimum standard proficiency with basic driving tasks. Participants then engaged in a driving task comprised of driving on a 16-mile simulated four lane bi-directional highway with a median during daytime. A variety of billboards were programmed to appear at predetermined distances within the scenario, and participants were instructed to drive as they normally would on a real interstate. A posted speed limit of 65 mph was displayed periodically throughout the scenario

More specifically, the simulation was displayed on three, 20" LCD computer monitors, providing a 135° field of view. Participants sat within the simulator's passenger compartment, which provided a view of the roadway and dashboard instruments, including a speedometer. The vehicle was controlled by moving a steering wheel in a typical driving manner and depressing accelerator and brake pedals accordingly. An on-board stereo sound system provided naturalistic engine sounds, external road noise, and sounds of passing traffic (see Figure 8).

The driving simulation displayed a mixture of digital and static billboards interspersed throughout the drive and always appeared on the right side of the road. A total of 16 billboards were presented in the simulation drive. Each billboard was presented once per simulation, thus the billboard order was fixed across participants. Transition times for the digital billboards varied to mimic naturalistic digital billboards, which transition at different points in time while a driver passes. Two transition time points were established at 250 feet and 500 feet away from the billboard to ensure clear visibility of both first (initial) and second (changed) advertisements. Therefore, if the billboard was digital, the first advertisement would change to another advertisement once the participant passed the predetermined marker (i.e., 250 or 500 feet from the billboard) while driving.

The 16-mile drive was further broken into four equal parts for development purposes. Each part consisted of the following: 1) a billboard that transitioned (i.e., changed from one advertisement to another) when the driver was 500 feet away from it, 2) a billboard that transitioned (i.e., changed from one advertisement to another) when the driver was 250 feet away from it, 3) a billboard that was static and therefore did not transition, and 4) a segment that did not include a billboard at all. Each of these 4 parts spanned one mile each and were populated in a randomized order according to a Latin square design.

Analysis

Key driving variables were coded electronically by the simulator and analyzed across four conditions: 250-foot billboard transition, 500-foot billboard transition, static billboard, or no billboard present. These include a) the number of speed limit exceedances, i.e., situations in which the participant's speed exceeded 69 miles per hour (mph) as recorded by the simulator, b) the number of road edge excursions, i.e., situations in which the right tire touched or crossed the right line of the road, and c) the total number of motor vehicle collisions, i.e., situations in which the driver made contact with another vehicle or structure within the scenario.

Figure 8: Photo of driving simulator

To examine the impact of billboards on various components of driver distraction, primary analyses involved a series of Repeated Measures ANOVAs (RM ANOVA) where the between subjects factor was age group (teen, middle, older) and the within-subjects factor was billboard type (static vs. digital, not present vs. static vs. 250-ft transition vs. 500-ft transition). Interactions between age group and billboard type were tested using RM ANOVAs. Degraded driving performance (e.g., speed exceedances, motor vehicle collisions, and road edge excursions) served as dependent variables.

Results

Consideration of participants' demographic information showed no significant differences of gender and ethnicity across age groups. As Figure 9 shows, a marginally significant difference among age groups for speed exceedances was observed among age groups ($F(2, 54) = 2.85$, $p = 0.066$). In a pairwise comparison of age groups, teens had significantly more speed exceedances than older drivers ($p = 0.042$). The results show no significant interaction between billboard type and age group ($p = 0.635$). There was not a significant main effect of road edge excursions ($F(2, 54) = 0.551$, $p = 0.580$). However, there was a marginally significant effect of billboard type on road edge excursions ($p = 0.071$). A closer pairwise RM ANOVA showed that there was statistically significant more road edge excursions in the no billboard condition compared to all other billboard conditions (no billboard vs. static billboard, $p = 0.045$; no billboard vs. 250 transition, $p = 0.037$; no billboard vs. 500 transition, $p = 0.045$) (Figure 10).Furthermore, there was no interaction of billboard type and age group ($p = 0.141$). Finally, there was only one motor vehicle collision across all age groups, which is too few to produce significant findings ($F(2, 54) = 1.000$, $p = 0.375$).

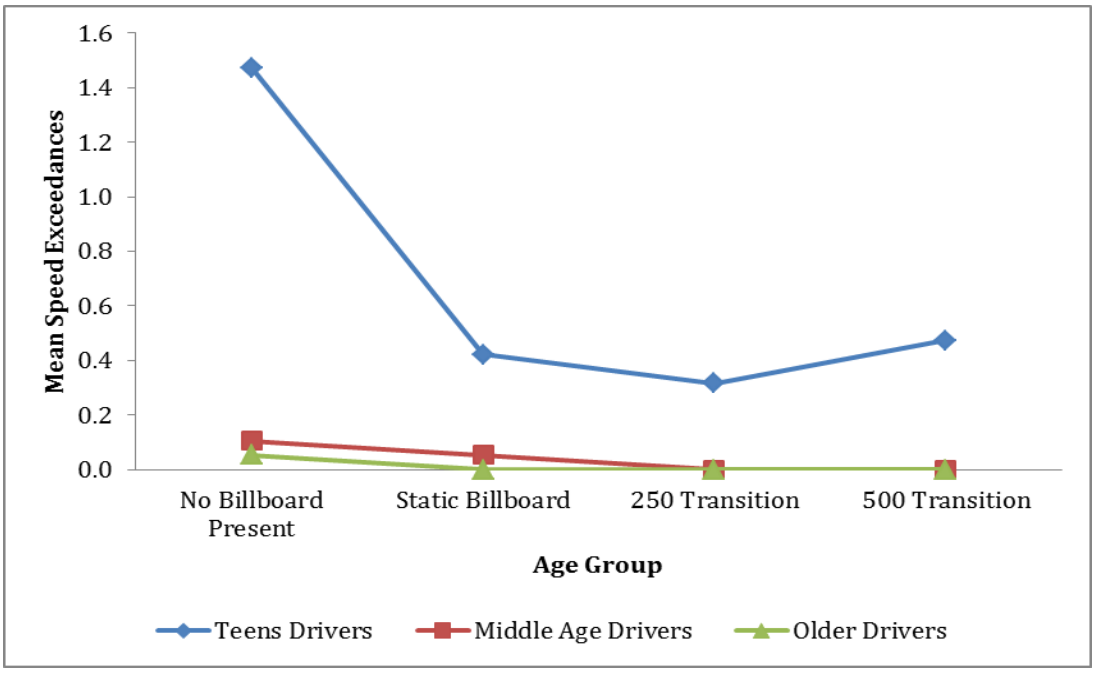


Figure 9: Mean frequency of speed exceedances

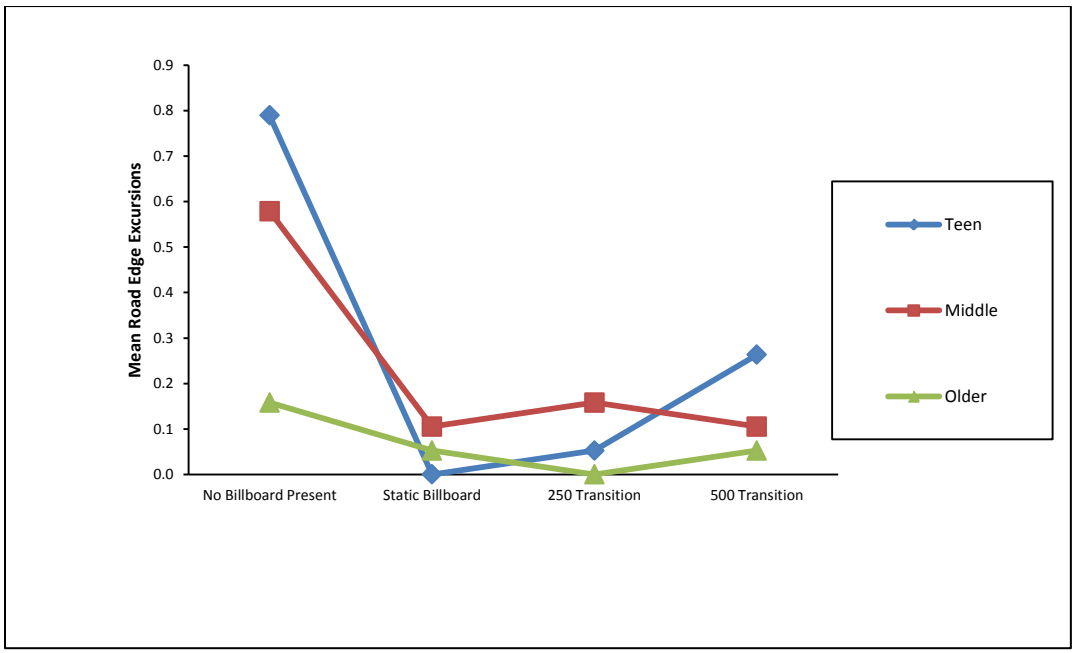


Figure 10: Mean road edge excursions

Discussion

This study is among the first to look at billboard distraction across different age groups, namely teens and older drivers, who have the highest rates of motor vehicle collisions. Overall, participants in the driving simulator study had fewer speed exceedances when there was a billboard present, indicating that their attention may have been captured by the billboards resulting in a slowed speed. Teens, as expected, had more speed exceedances than middle aged and older drivers. Also, in the presence of billboards, the number of speed exceedances across all age groups was significantly reduced. Finally, fewer road edge excursions were recorded in the presence of billboards of any type as compared to the no billboard condition. These findings suggest that there is an impact of billboard distraction on driving performance, which could be attributed to drivers paying more attention to the billboard.

CONCLUSION AND FUTURE STEPS

This paper explored perceived and real impacts of digital advertising billboards on driving performance of individuals in Alabama from representative samples across the lifespan. This work presents a contribution to the traffic safety research as it provides some insights that can help inform future public policy relating to driver distraction and billboards, especially in regards to regulations for billboard use. The findings offer support to future interventions such as incorporating billboard awareness into driver education courses for teen drivers, since younger drivers' performance was most affected out of all age groups.

The online survey suggested that more than 40 percent of the drivers admitted looking at the digital billboards for a sufficiently long time, but most of the drivers rarely slow down. However, some inconsistencies in responses were observed, indicating that some respondents might indicate one perception in general but that perception might not necessarily reflect on their action when put in a real situation. This is further supported by the findings of the simulation study that showed participants having fewer speed exceedances when there was a billboard present, indicating that their attention may have been captured by the billboards resulting in a slowed speed.

A close observation of the drivers' actual behavior in a naturalistic study environment would be more suitable to capture the real reaction of drivers as they approach a static or digital advertising billboard, and it is recommended for future study. Moreover, an expansion of the questionnaire survey sample to include responses from other states is planned for the near future. Statistical comparisons will be performed to establish if regional differences affect the findings of the study. Also, the scope of the driving simulation study will be expanded to investigate the effect of external distractions from billboards on driver's visual and cognitive attention.

A limitation of the current work is the lack of significant results in performance degradation. This may be due to the fact that hazards were not presented throughout the simulation - there were not ample opportunities to crash. Rather, the simulation was a mundane driving situation. Speed exceedances seemed to have the largest impact in this study, perhaps because the scenario was a straight road, which could have been interpreted as boring by participants. In addition, there was not a significant increase in road edge excursions. Consideration of within lane deviation may pick up subtleties in swerving that went undetected. Future studies should look at lane deviation as a measure of driving performance.

ACKNOWLEDGMENT

Funding for this study is provided by US DOT/RITA through the National Center for Transportation Systems Productivity and Management (NCTSPM) and the Alabama Department of Transportation (ALDOT). The support of the study sponsors is greatly appreciated.

REFERENCES

Coetzee, J. (2003). The Evaluation of Content on Outdoor Advertisements. Presented at the Southern African Transport Conference, 14-16 July 2003.

Copeland, L. (2010). More Communities Banning 'Television on a Stick.' USA Today; 23 Mar 2010.

Molino, J.,Wachtel, J.,Farbry, J., Hermosillo, M., Granda, T. (2009). The Effects of Commercial Electronic Variable Message Signs (CEVMS) on Driver Attention and Distraction: An Update. Publication FHWA-HRT-09-018.FHWA, Feb 2009.

Outdoor Advertising Association of America [OAAA]. (2013). Out of home media formats: Number of out of home displays (2013), October 14, 2013, from http://www.oaaa.org/OutofHomeAdvertising/OOHMediaFormats/OOHMediaFormats.aspx

Sisiopiku, V.P., Hester, D., Gan, A., Stavrinos, D., Sullivan, A. (2013). "Roadside Advertising and Traffic Safety", 3rd Annual International Conference on Civil Engineering, Athens, Greece.

Wachtel, J. (2009). Safety Impacts of the Emerging Digital Display Technology for Outdoor Advertising Signs.Project 20-7 (256), Final Report.NCHRP, Apr 2009.

Wachtel, J. (2011). Digital Billboards, Distracted Drivers. Planning, Vol. 77, Issue 3, Mar 2011, pp. 25-27.

Wallace, B. (2003). Driver Distraction by Advertising: Genuine Risk or Urban Myth? Proceedings of the Institution of Civil Engineers: Municipal Engineer, Vol. 156, No. 3, Sept 2003, pp. 185-190.

Recording and Evaluation Procedure of Drivers' Distraction in Accordance with Driver's Characteristics in High Speed Arterials

Eleni Misokefalou and Nikolaos Eliou

University of Thessaly
Volos, Greece

ABSTRACT

Over the last years, distracted driving constitutes a considerably increasing road safety problem with disastrous results and it possesses a leading position among the accidents causes. The present study deals with driver's distraction due to out of the vehicle factors as well as factors related to the driver such as age, gender, driving experience etc. Considering exterior factors as the most significant, we can group them in four categories: built roadway, situational entities, the natural environment, and the built environment. Regarding the fourth category, it is related to the wide variety of civil infrastructure, the commercial land use combined with the high vehicle speeds. All these contribute to the setup of a very dangerous environment by increasing driver's distraction and inattention. This research is based on a medium scale experimental procedure which took place in three urban freeways in Greece, using a sample of 77 drivers. The distraction of the driver's attention is evaluated via a continuous recording of his gaze which acts as the main indicator regarding driver's performance. The main objective of this paper is to assess the side effects of roadside advertising and overloaded informational signs to driver's distraction and inattention. The results of this type of research procedures are very useful as a tool to prevent the forthcoming pressure for more and more billboards and trademarks on the roads as well as to encourage the adaptation of more precise regulations with regard to the road infrastructure, the placement of roadside elements, etc.

Keywords: Driving, Distraction, Advertising, Billboards, Naturalistic, Research

INTRODUCTION

The distraction of driver's attention during the implementation of the driving task is not simply a theory. It is a procedure which is activated and developed depending on many factors. It is detected in all drivers, with varying extent and frequency of appearance, but, in every case, the results of this distraction are intense for the driving task, the driver's safety and, finally, for the rest of road users. Distraction at all forms, has become object of research recently, with distraction from a secondary task concentrating most of the research on the subject, particularly after the widespread use of mobile phones and the integration of driver assistance systems in modern vehicles. Naturally, priority is given to drivers of passenger cars without overlooking the other road users' categories such as truck drivers, motorcyclists, bicyclists etc (Misokefalou et al. 2010).

Definition and Characteristics of Driver Distraction

The first step to a proper approach is to understand the basic characteristics of distraction as it appears in general. Distraction may be visual, cognitive, biomechanical and auditory (Ranney et al. 2001). Many researchers have tried to define driver distraction and as a result the related literature contains a significant number of those definitions. In the first International Conference on Distracted Driving (Hedlund et al. 2005:2) the scientific community agreed on a definition for distracted driving:

"Distraction involves a diversion of attention from driving because the driver is temporarily focusing on an object, person, task, or event not related to driving, which reduces the driver's awareness, decision-making, and/or performance, leading to an increased risk of corrective actions, near-crashes, or crashes" [3].

The main causes of distraction are classified into two categories: Those coming from the interior of the vehicle and those from the external environment. In the second category, one finds some very important potential sources of driver distraction which can be grouped in four major categories: built roadway, situational entities, the natural environment, and the built environment (Horberry et al. 2009). According to the writers one component of built environment which consists the fourth category of the previous categorization is the billboards and any other kind of road advertising. In the case of causes related to advertising, it should be particularly emphasized that the purpose of their presence at some point at the roadside, or even in a moving vehicle in the road, is to capture the driver's gaze in order for him/her to devote the required time so as to assimilate the information obtained. Roadside advertising billboards are designed by their very nature to attract attention. Crucially, though, the related potential threat to road safety is generally not acknowledged by the industry (Crundall et al. 2006). This is the reason why distraction by advertising is a very significant road safety phenomenon which must be elucidated by using all the available means and methods.

Frequency of Driver Distraction

The importance of this issue emerges from data which shows distraction from a secondary task as a cause of serious accidents as well as crashes. A characteristic research was carried out by the Virginia Tech Transportation Institute (VTTI) for NHTSA, the "100- Car Naturalistic Driving Study" (Klauer et al. 2006). During the 100-Car Naturalistic Driving Study, driver involvement in secondary tasks contributed to over 22 % of all crashes and near-crashes recorded during the study period. At this point it should be noted that, as near crash is defined the subjective judgment of any circumstance that requires, but is not limited to, a rapid, evasive maneuver by the subject vehicle, or any other vehicle, pedestrian, cyclist, or animal to avoid a crash (Klauer et al. 2006). These secondary tasks, which can distract the driver from the primary task of driving (steering, accelerating, braking, speed choice, lane choice, manoeuvring in traffic, navigation to destination, and scanning for hazards), are manifold and include such things as eating/drinking, grooming, reading billboards, using and adjusting in-vehicle entertainment devices, conversation with passenger(s), viewing the scenery, tending to children and pets, smoking, cell phone use and related conversation, use of other wireless communication devices, and note-taking, to name a few (Hendlud et al. 2006). But, not all of the above distracters involve secondary tasks initiated by driver –they can be events, objects, activities or people both inside/outside the vehicle (Tasca 2005). The statistics of the above research are confirmed by the data from accidents in many countries (e.g. accident data from United States in 2008 (NHTSA, 2009) and Greek Police for 2009 and 2010 (Greek Police 2011)) as well as by a large number of other studies in the field.

Particularly for billboards, Crundall in his study (Crundall et al. 2006) supports that though it is acknowledged that research into advertisement distraction has been extremely limited (Beijer et al. 2004), the few studies that have been conducted have demonstrated that drivers do look at and process roadside advertisements (Hughes et al. 1986), and that fixations upon advertisements can be made at short headways or in other unsafe circumstances (Smiley et al. 2004). Previous studies of accident statistics have also identified external distractors, including advertisements, as a significant self-reported cause of traffic accidents (Stutts et al. 2001). Particularly, for roadside distractors, evidence is mounting that roadside distractions (and advertising in particular) present a 'small but significant' risk to driving safety (Lay2004). Conservative estimates collated from a review of several accident databases put external distractors responsible for up to 10% of all accidents (Wallace 2003). This is confirmed also by a recent simulator study (Young et al. 2009) in which there is a tentative suggestion that more crashes occur when billboards are present.

METHOD

Methods of Evaluating Driver Distraction

The only certain way for the researcher to detect driver's distraction is via the results that distraction produces. The use of standardized methods gives the researchers the possibility to exchange data, conclusions and best practices. Therefore, it is important to detect the most suitable method of data collection (Rockwell. 1998). This aim can be achieved via a comparative study between the allocated methods, examining the advantages and disadvantages of every method separately as well as the usefulness and necessity of the results that every one of them produces. An analysis of this kind was made by the Univerity of the Thessaly (Eliou et al. 2009). During this study all the available methods were examined. The most popular among the available methods are based on elements of accidents, on experiments, on observation and surveys. Furthermore, there are some kinds of methods that are not included in any of the previous categories like Peripheral Detection Task and Visual Occlusion.

Selection of the Appropriate Method

The method considered the most appropriate is an observational-naturalistic study, which takes place in the field, using specially equipped vehicles with regard to record the driver's eye movements in order to measure the frequency and the duration of the glances at every object considered potential source of visual distraction. The available equipment (Facelab machine) is capable of making continuous data recording. The main advantage is that with this method, in contrast with all the others categories, driving comes as close to the real thing as possible which is important for the research when we study human reactions. Naturally, there are some limitations both in designing and carrying out the experiment. The most important of these is the limited number of participants in comparison with other methods like questionnaires study, the unfamiliar vehicle which causes stress to the driver, the anxiety because of the sense of being monitored as the vehicle is equipped with cameras and, finally, the subjective discretion of the analyst-observer at the data processing.

Captiv software, which is compatible with FaceLab L2100, was used for the analysis of the results. This software gives the opportunity to analyse the data in detail by recording the number of glances at every billboard as well as the total time that the billboard captured driver's gaze during driving. At this point it should be noted that as distraction, in this study, is considered the continuous or intermittent but repeated capture of the gaze from a theme for longer than a total of two seconds as glances that last more than this time are related to driving errors (Rockwell, 1998). Smaller time intervals (0-0.7 seconds, 0.7-1.0 seconds, 1.0-1,6 seconds, 1,6-2.0 seconds >2 seconds) were also studied in order the results to be comparable with the results of other related studies. Highly experienced analysers carefully analysed the produced by Facelab machine videos and all the data were statistically analysed.

Experimental Site

The research took place from June 2011 to February 20102, in 3 suburban greek roads. The first road is Attiki Odos, a suburban highway in the city of Athens. There were three routes under observation with a length of 19, 16.8 and 15 km respectively. The second road is Leoforos Kifisou which is also located in the city of Athens and the route under observation has a total length of 10 km. The third road is the national road between the city of Thessaloniki and the city of Giannitsa. The studied section is 7.7 km.At the last two roads, drivers drove both directions. The flow of vehicles is continual without being interrupted by traffic lights. The speed limit of the road is 90km/h. The total number of the points at which the data were collected is 136, of which the first 69 are at the first route, the following 40 at the second route and the last 27 at the third route. In this study other objects besides the advertisement has been selected to participate in order to act as comparison points to the advertising. The general categorization and the percentages of its participating category are shown in the following chart (Figure 1).

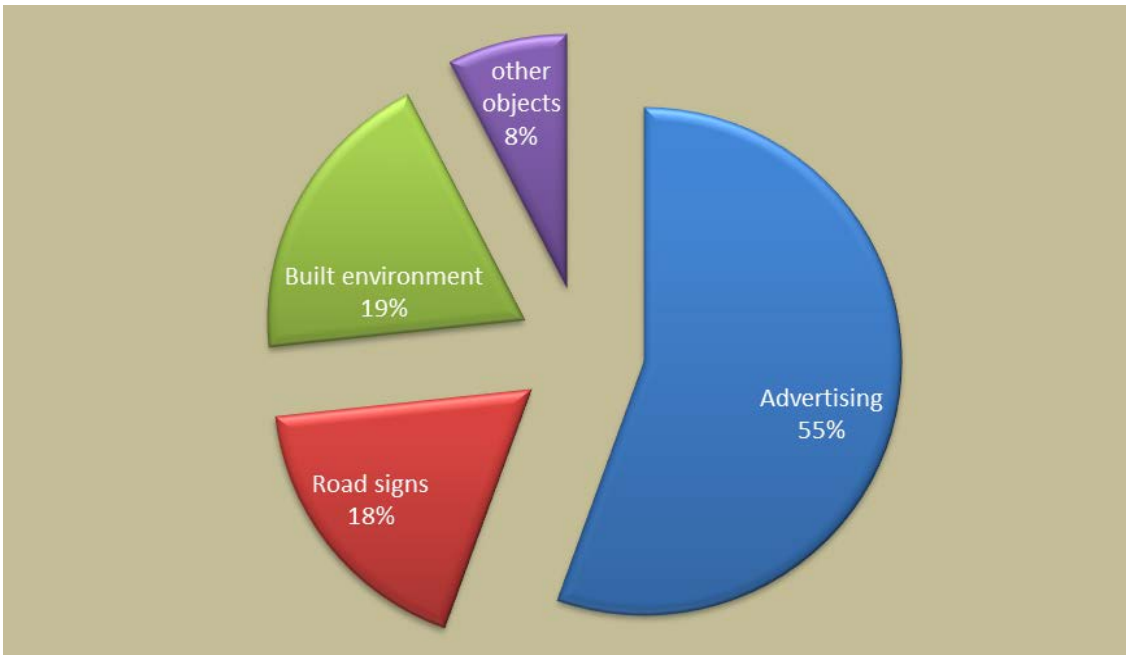


Figure 1. Categorization of the 136 studied points.

Participants

Using volunteer drivers, who are required to drive a car on the three selected roads, under the supervision of the researcher, who was always in the passenger seat checking the proper function of the system, the obtained results are characterized by a high degree of reliability and validity. Seventy seven drivers participated in the survey. Their sex distribution can be seen in Figure 2.

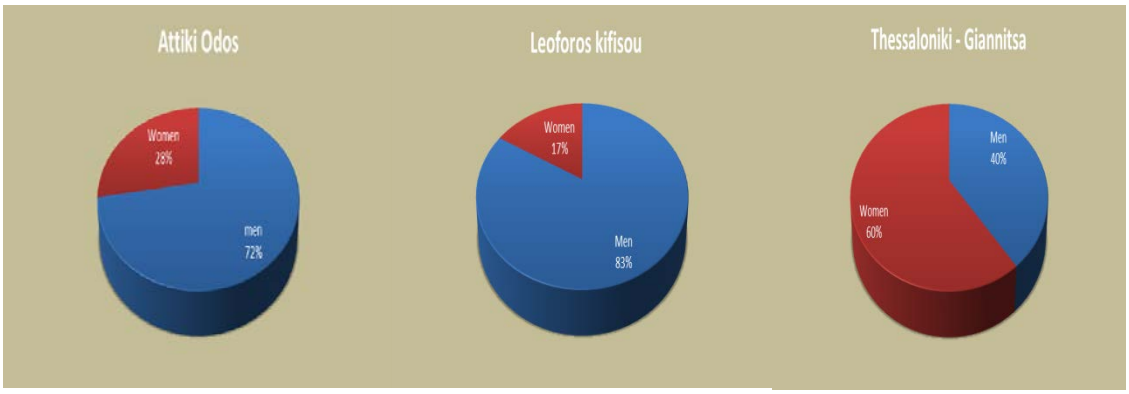


Figure 2. Participants' sex distribution.

The drivers were selected with criterion their familiarity with the selected roads in order to eliminate the stress which an unfamiliar route induces. All drivers were familiar with the road, as they use it in a daily basis, but the subject of the study was completely unknown to them. Regarding the limitation of the unfamiliar vehicle, each one of them, in order to become familiar the vehicle drove the selected route 2 times before the third run which was the one that we focused our attention at during the analysis process.

The drivers are between 26 and 45 years old and the vast majority of them, in all three traffic routes, possesses a driving license and drives systematically for more than ten years.

Material - Data Collection

The equipment used in the survey was very carefully chosen in order to produce the optimal quality, completeness and integrity of results. It includes a passenger vehicle and a monitoring and recording system (Facelab, Seeing

machines), which detects and records every single movement of the driver's gaze and the driver's head. It is composed of two cameras for the recording of the above and an external camera for the recording of the road scene. All measurements for the experiment took place during the day, under normal traffic conditions as well as normal weather and lighting conditions.

RESULTS

In this study, the information isolated and analysed in depth, is related to the external impulses that cause driver distraction and concentrates interest mainly on billboards near the road and their role in driver's distraction of attention. For this purpose, billboards along the road were identified and mapped for both directions of the route. Among them there are billboards, banners, soundpannels with graffiti and posters on them and gas station signs. Additionally, other elements on the road, such as road signs and buildings were under study. The analysis included an examination of driver behaviour, as far as concern the reactions of drivers' pupils of the eyes while driving under the existence of these potentially evocative distraction elements of the road environment.

The following 3 figures (Figure 3,4 and 5) show the percentages from the analysis of the gaze direction to the observed points of each route. Each driver drove the selected route 3 times but we decided to focus our attention at the third one because of the familiarization of the driver with the vehicle which was analysed at the method section. As it shows, distraction there are certain points on the route which attract attention more than the allowed, from the aspect of safety, time.

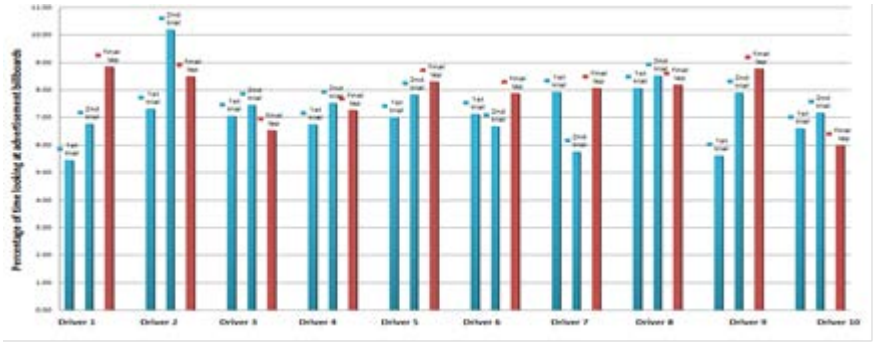


Figure 3. Percentage of driving time looking at advertisement billboards.

The detailed analysis of the data came from the eye gaze, in terms of glance duration and frequency, led us to the following conclusions:

- All roadside billboards of the route distract the majority of the drivers (gaze captured for more than 2 seconds).
- At the points of the route where many billboards are placed in a short distance (e.g. advertisement billboard number 9 and 21), the majority of drivers are distracted as their gaze is captured by more than one billboard. At these points, it is observed intermittent but repeated capture of the drivers' gaze.
- Billboards attracts women's gaze more than men's. The average of percentages of the total time that women look at advertisement billboards is 8.7%, while for men is 7.5%.
- The billboards found in the center or near the central field of vision are more likely to attract the driver's gaze.
- At the section which contains the high gather of posters placed in disorder, the large number of illegal posters attracts multiple glances from drivers so that the visual disturbance leads to confusion.
- During the third route, 50 percent of drivers' gaze is captured by more advertisement billboards than during the first route.
- At Fig. 3 is obvious that there are certain advertisement billboards that capture drivers' gaze during all

three rides. The percentage of billboards that capture drivers' gaze mostly at the second and the third route is limited.

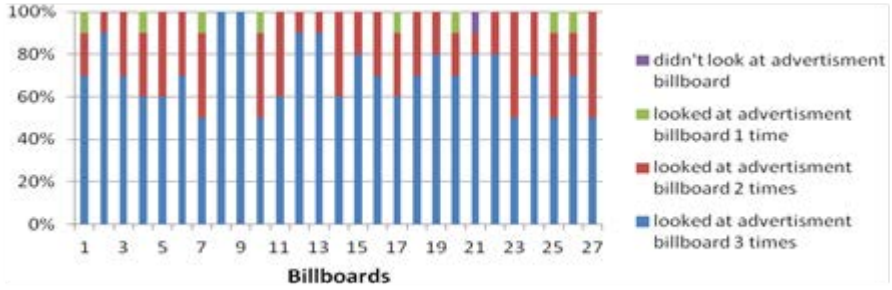


Figure 4. Percentages of drivers distracted from billboards at all three, at two, at one and at none of the routes.

From the survey results arises that the presence of distraction, and more specifically the kind caused due to billboards, is common in drivers and depends largely on the characteristics of the billboards and their position in the field of vision.

CONCLUSIONS

Distraction of driver's attention during driving is a major road safety problem, which threatens not only the driver's safety but also the safety of other drivers and road users. The focus of the research on drivers of passenger vehicles is due to the fact that those drivers consist the largest category of road users with growing involvement in accidents, which are caused by the distraction of driver's attention. The goal of the research is to identify and clarify the causes, the frequency of appearance and the way that certain factors influence the distraction of attention of each driver, focusing on the role played by roadside advertising in Greece as a parameter of the distraction of the driver's attention.

The methods commonly used in a study of driver distraction aren't all feasible or effective to the same extent. The chosen method allows the continuous data recording with its main advantage being the fact that driving is as close to the real thing as possible. Thus, the results are characterized by a high degree of reliability and validity. It, also, gives the opportunity to the participant to have an adjustment period with the vehicle in order to obtain a normal driving behaviour. The small possibility of the researcher to control the situations and create desirable driving scenarios is among the disadvantages of this method. The environmental conditions, also, cannot be controlled. Another disadvantage is the increased cost of the method due to the eye tracker. Finally, as disadvantage of the eye tracker we could mention the difficulty of the installation in the car as well as its sensitivity to changes (e.g. lightness conditions).

This research concluded that all roadside billboards of the route distract the majority of the with signs in the raw causing the greater distraction. Also, the more centrally positioned in the field of vision the signs are placed, the more eye-catching they are. There is a need to relate the drivers' distraction to specific aspects of advertising signs (size, message content, position by the road).

The results of this type of research procedures are very useful as a tool to prevent the forthcoming pressure for more and more billboards and trademarks on the roads as well as to encourage the adaptation of more precise regulations with regard to the road infrastructure, the placement of roadside elements, etc.

Much of the data analysis requires collaboration with experts such as psychologists and doctors in order to provide an integrated approach. Furthermore, a comprehensive policy to reduce the visual pollution near the roads, such as billboards, can help not only to improve the road aesthetic but also to significantly improve road safety by eliminating driver's visual distraction of attention.

To sum up, it is a fact that driver distraction is a major cause of accidents; therefore, the responsibility over the issue translates into efforts to reduce the number of injured and dead drivers. In order to achieve that, the solution of the distraction phenomenon could be found in the combination: "research-education-design- legislation".

This research is under further analysis, as a number of 30 variables - from the road, the conditions, the points and the driver - is available to the researchers and is being statistically analysed in order to conclude about contribution of those variables in driver distraction.

ACKNOWLEDGEMENTS

We acknowledge the contribution of the company Attikes Diadromes S.A. who granted the use of the Facelab equipment for the purposes of this research. We also acknowledge the valuable contribution of all the participants who voluntarily devoted their time and energy to our research.

REFERENCES

Beijer, D.D., Smiley, A. & Eizenman, M. (2004), "Observed driver glance behavior at roadside advertising signs", Transport Research Record, 1899, 96–10.

Crundall, D., Van Loon, E. & Underwood, G. (2006), "Attraction and distraction of attention with roadside advertisements", Accident Analysis & Prevention, 38, 671–677.

Eliou, N. & Misokefalou, E. (2009), "Comparative analysis of drivers' distraction assessment methods", 22nd ICTCT Workshop, Towards and Beyond the 2010 Road Safety Targets-Identifying the Stubborn Issues and their Solutions, Leeds, UK

Greek Police. (2011), Accident Statistics for the Year 2011, Ministry of Citizen Protection, Greece.

Hedlund, J., Simpson, H.M. & Mayhew, D.R. (2006), "International conference on distracted driving-summary of proceedings and recommendations", International Conference on Distracted Driving, Toronto (2–5October 2005).

Hughes, P.K. & Cole, B.L. (2010), "What attracts attention when driving?", Ergonomics, 29, 311–391, 1986.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D &, Ramsey, D.J. (2006), The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic driving study data, Report No. DOT HS 810 594, National Highway Traffic Safety Administration, Washington, D.C.

Lay, M.G.: Design of traffic signs. In C. Castro & T. Horberry (Eds.). (2004), "The human factors of transport signs", 25–48, Boca Raton, FL: CRC Press.

Misokefalou, E., Eliou, N. & Galanis, A. (2010), "Driving distraction of attention audit of motorcycle and bicycle users in urban areas, Case study in the city of Volos", Greece, 5th international congress on transportation research, Volos, Greece.

Ranney, T.A., Garrott, W.R. & Goodman, M.J. (2001), "National Highway Traffic Safety Administration driver distraction research: past, present and future", 17th International Technical Conference on Enhanced Safety of Vehicles, Amsterdam.

Rockwell, T.H. (1998), "Spare visual capacity in driving- revisited", In A.G. Gale, H.M. Freeman, C.M. Haslegrave, P. Smith & S.P. Taylor (Eds.) Vision in Vehicles II, 317-324. Amsterdam: Elsevier.

Smiley, A., Smahel, T. & Eizenman, M. (2004), "Impact of video advertising on driver fixation patterns", Transport Research Record, 1899, 76–83.

Stutts, J.C., Reinfurt, D.W., Staplin, L.W. & Rodgman, E.A. (2011), "The Role of Driver Distraction in Traffic Crashes", AAA Foundation for Traffic Safety, Washington, DC.

Tasca, L. (2005), "Driver distraction: towards a working definition", International Conference on Distracted Driving, Toronto (2–5October 2005).

Young, K., Mahfoud, J.M., Stanton, N. A., Salmon, P.M., Jenkins, D.P. & Walker, G.H. (2009), "Conflicts of interest: The implications of roadside advertising for driver attention while driving: Skill and awareness during inspection of the scene", Transportation Research Part F 12, 381–388.

Wallace, B.: "Driver distraction by advertising: Genuine risk or urban myth?", (2003), Municipal Engineer, 156, 185–190.

Driving Maneuver Prediction Based on Driver Behavior Observation

Frederik Diederichs[a] *and Gloria Pöhler*[b]

[a]*Human Factors and Vehicle Interaction department*
Fraunhofer Institute for Industrial Engineering (IAO)
Nobelstr. 12, 70569 Stuttgart, Germany

[b]*Institute of Human and Industrial Engineering (ifab)*
Karlsruhe Institute of Technology (KIT)
Kaiserstr. 12, 76131 Karlsruhe, Germany

ABSTRACT

With respect to an increasing amount of driver assistance systems and automated driving functions, a higher chance of unappreciated action and intervention of these systems can be registered, which in turn lowers the acceptance by drivers and passengers. A reduction of unnecessary warnings and interventions can be achieved by making them adaptive to driver's intentions and maneuvers planning. In order to learn which driver behavior indicates certain maneuver intentions, a rater-based method using video recordings is proposed in this paper. Three driving maneuvers, namely turning, changing lane and braking for a pedestrian who intends to cross the road, were chosen for analyzing their predictability due to behavior observation. As a first step, a driving simulator study was conducted in order to collect behavior data of 24 drivers. Subsequently, clearly distinguishable behavior classes for each maneuver were extracted from video data, resulting in five superior behavior categories with 29 behavioral classes. Based on these classes four human observers were trained to detect at the earliest convenience maneuver intentions. Overall in 97 % of all cases the observers could predict the maneuvers. Inter-rater reliabilities showed to be between κ= 0.30 and κ = 1.00.

Keywords: Maneuver Prediction, Driver Intention, Driver Model, Automated Driving, Driver Assistance

INTRODUCTION

Modern driving assistance systems have to deal with a warning dilemma between timely intervention and false alarms. Rigid thresholds have proven to fail in this dilemma which calls for thresholds that are driver state adaptive. In addition future assistance systems and Car2x applications leading towards automated driving come up with the demand for anticipating planned driving maneuvers and conclude on driver's intentions when the system needs to take a decision whether to take over control. In such situations it is of upmost importance to guess the drivers intentions in time and correctly in order to avoid system interventions that are counteracting the driver's planning.

In this context Blaschke (2007) comments that a reduction of false warnings and interventions decreases driver stress and leads to a gain in security and acceptance. Typical examples of the need to guess drivers intentions are autonomous emergency breaking interventions. These systems fail too often when drivers plan close distance passing of vehicles or pedestrians. In a future which is introducing more and more automated driving maneuvers it will be also necessary to guess drivers intentions i.e. for lane changes and turning maneuvers in order to suppress corrective safety interventions and initiate supporting assistance (Kopf, 2005). Further Car2x based applications will

broadcast information on driver and maneuvers intentions in order to organize a collaborative traffic behavior (TEAM-project, 2012).

Although there are many definitions of driver intention (see Kobiela, 2012), a good summary of the relevant aspects is provided by Kopf (2005). He describes driver intention as one of the influencing factors on driving next to emotion or attention, changing on a short-term basis (minutes or seconds), compared to long-term factors such as personality or driving experience and daily rhythm or tiredness, which are changing on middle-term. Intentions include the travel or arrival time, the planned route and especially the driving manoeuvres or manoeuvre order planned for the imminent future. Or more simply put driving intention refers to the manoeuvre planned next.

The remaining question however is how intentions can be detected. Ajzen (2002) acknowledges that intentions are defined not to be directly measurable, but rather detectable through observable behavior. If this behavior leads to manipulation of a machine (pedal or steering wheel) it can be observed indirectly via these actuators. A direct observation is assumed when human behavior is observed via video analysis. This approach is well known from social science (König, Heintz, & Scheuch, 1962) and is characterized by a systematic observation with a variety of specific methods ranging from simple annotation to classification methods and even magnitude estimation approaches (Schnell, Hill, & Esser, 2011).

Following Ajzen's advice the authors of this paper planned and conducted an observation study with direct behavior observation in order to determine driver's behavior which indicates intentions. The results can be used to detect intentions with sensors that would measure behavior directly, such as the emerging video camera sensors (InCarIn-Project 2014). Systematic direct behavior analysis also provides good reference for indirect measures as well in correctness as in timing matters. The research questions of this approach were

- to describe and classify the behavior that precedes driving maneuvers,
- create a set of behavioral classes and magnitudes that allow anticipation of most driving maneuvers,
- and to develop a procedure to predict maneuvers based on human rater observation.

An exemplary subset of maneuvers were selected from a set of basic maneuvers provided by Dambier (Dambier, 2010 and Manstetten et al., 2011) with a consciously broad range of behavior variance, including longitudinal and perpendicular maneuvering, maneuvers with high potential for assistance and – for practical reasons – to be simulatable in our immersive driving simulator. This resulted in three maneuvers:

- lane change,
- left-turn and
- close bypassing a pedestrian.

With respect to the state of the art on drivers intention recognition predicting lane change maneuver was subject to quite a view studies (Pentland & Liu (1999), Kuge et al. (2000), Oliver & Pentland (2000), Salvucci (2004), Gerdes (2006), Feyer et al. (2007), Henning et al. (2009), Berndt et al. (2008), Schroven & Giebel (2008)).

Turn prediction was investigated by Pentland & Liu (1999), Oliver & Pentland (2000), Takagi et al. (2000), Färber (2005), Gerdes (2006), Blaschke et al. (2007), Berndt et al. (2008) and Schroven & Giebel (2008).

A potential braking intention when passing a pedestrian has so far not yet been investigated according to our systematic literature review (planned for publication). The most related maneuver would be stopping at an intersection which was part of the extensive investigation in the studies of Pentland & Lui (1999) and Oliver & Pentland (2000). Those authors also investigated on turn prediction and lane change.

By means of velocity, steering data and acceleration Pentland and Liu (1999) were able to predict maneuvers with an accuracy of 95%. In their study dynamic Markov models were used to detect lane change, turning, following, overtaking maneuvers and stopping at the next section. 2 Seconds after the instruction for the maneuver or 1.2 seconds after the beginning of action, the prediction accuracy reached 90%. The maneuver to be executed was displayed via text in a 20-minute driving simulation with surrounding traffic, houses and road markings. For the included 72 stop maneuvers this happened approximately 70 m before the stopping point and for the 24 analyzed overtaking maneuvers (including lane change) with a speed of 30-35 mph, ca. 30 m before the preceding vehicle. After the driving simulator study Oliver and Pentland (2000) run a study with 70 participants who drove a test vehicle on urban and highway routes around Boston, during which driving instructors judged and instructed the maneuvers. Here they used dynamical graphical, hidden Markov and coupled hidden Markov models for the

prediction of five different maneuvers. With just vehicle related data (brake, steering angle, acceleration and gear) their hidden Markov model predicted overtaking, start and stop with 100% accuracy. A combination of vehicle data with GPS and headway distance information, together with head and viewing direction was effective for turn and lane change. Here precisions of 6% (lane change right) to 86% (right turn) could be achieved.

The prediction probabilities and times achieved by these authors are all based on indirectly measured behavior. It was hence time to introduce also direct behavioral observation and apply human observer based methods from social sciences to this field of application. Our approach uses extensive data recording in a driving simulator, a video based analysis of the drivers' behavior, a behavior classification and the development of an observer based intention analysis method. Intra and inter-rater correlations are reported for the likelihood of detecting a maneuver intention.

DATA RECORDING

The data recoding was carried out in the Fraunhofer IAO driving simulator lab. The simulator is equipped with video recording technology and delivers synchronous vehicle sensor data. The test participants where balanced with respect to driving style, age and gender.

Driving Simulator and Scenarios

The driving simulator, in which the data recording was performed, is a real vehicle with automatic transmission surrounded by seven screens to attain maximum immersion into the simulation. The chosen setting, which is roughly displayed in figure 1, enabled almost 360° view and included a projection for looking backwards behind the shoulder (right picture of figure 1). Road noises as well as surface irregularities could be displayed in the driving simulator.

Figure 1. Driving simulator at Fraunhofer IAO.

The simulated urban driving environment itself was implemented by using the pc-based simulation software SILAB 3.0 (WIVW, 2013). The three maneuvers where settled into a city with random curves, houses, trees, parks, pedestrians and traffic. None of the scenario scenes contained dynamic objects to avoid distracting the driver from the maneuver itself. For each maneuver a baseline scenario was included which was identical to the test scenario but without introduction to carry out the maneuver. Each maneuver and each baseline was included 5 times and cued up in random order resulting in 30 maneuver recordings per participant. Further road strips where included in-between and sum up the trip to 36 km / 45min driving. A screenshot of all three maneuver's can be seen in figure 2.

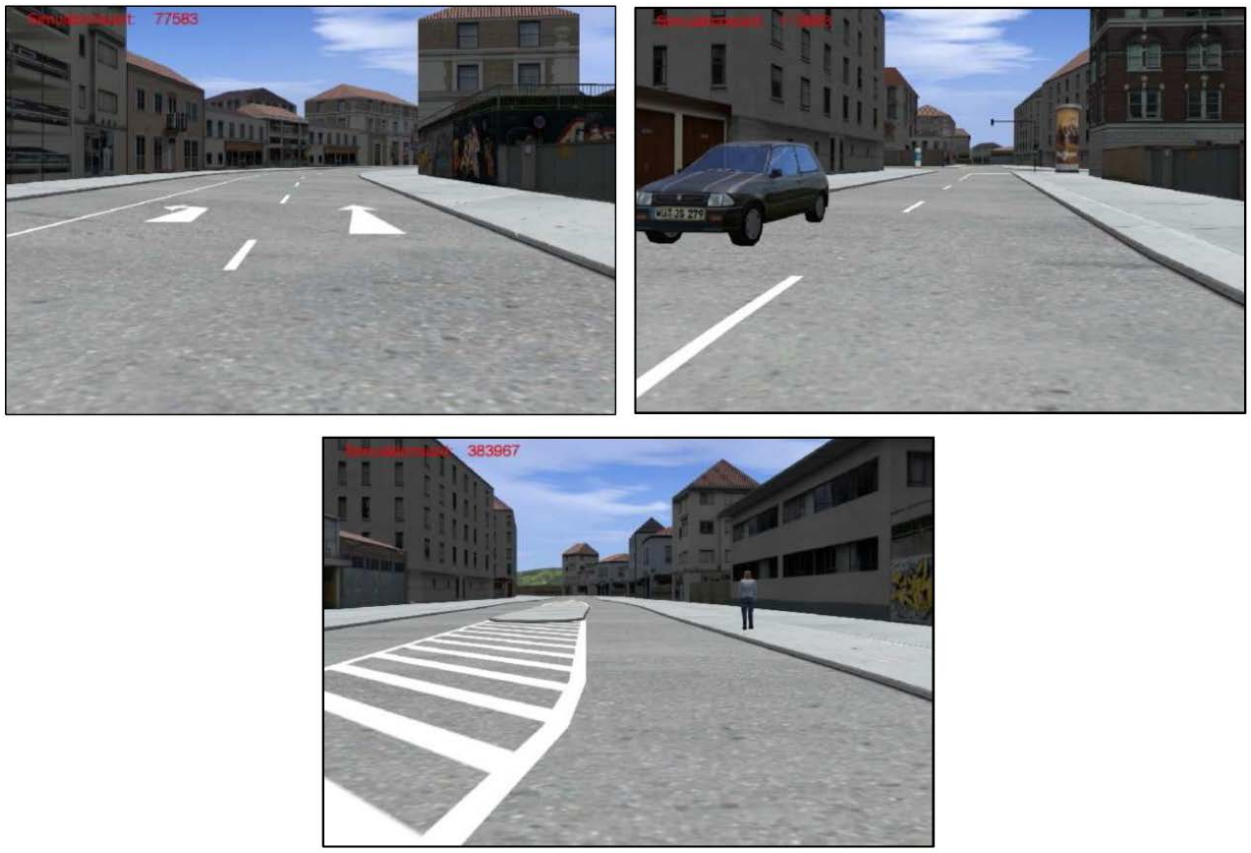


Figure 2. The three driving maneuvers "lane change" (top left), "left turn" (top right) and "close bypassing a pedestrian" (bottom) in the simulation software SILAB.

The first maneuver, turning left, was implemented as an intersection turn. The traffic signal turned green at viewing distance. At the same time, the driver was instructed to turn left via an acoustic navigation command. This command was missing in the respective baseline condition. This setup allowed to discover different behavior between baseline and maneuver condition.

With regards to the second driving maneuver, changing lane to the left, a gauge widening has been realized in SILAB. Two seconds after passing the gauge the lane change command was given (no command in the baseline). At this time a car which was following the ego vehicle changed lane to the left and accelerated slightly in order to make backwards checking necessary.

The third and last driving maneuver, bypassing a pedestrian, differed in terms of the baseline condition. A pedestrian was walking on a pavement close to the street with a traffic island on the opposite site. In a similar scenario participants had experienced a road crossing here. This indicated a probability that the pedestrian may cross the road without looking. The objective here was to detect at an early stage in time whether the driver showed any intentions of braking – or not. Experiences collected in former experiments where used to design a scenario which would stimulate a braking readiness in many but not all drivers (Diederichs et al., 2011 and Ganzhorn et al., 2011)

Recording Procedure

To gather video recordings, three IP cameras were built into the vehicle to record driver behavior from different perspectives. The first camera was installed leftwards in front of the driver to log gaze and head movements. The second camera with a fisheye lens was fixed on the ceiling to record the driver's body and hands. The third camera in the pedal room records foot movements. The exact positioning of each camera and their corresponding shots can be seen in figure 3.

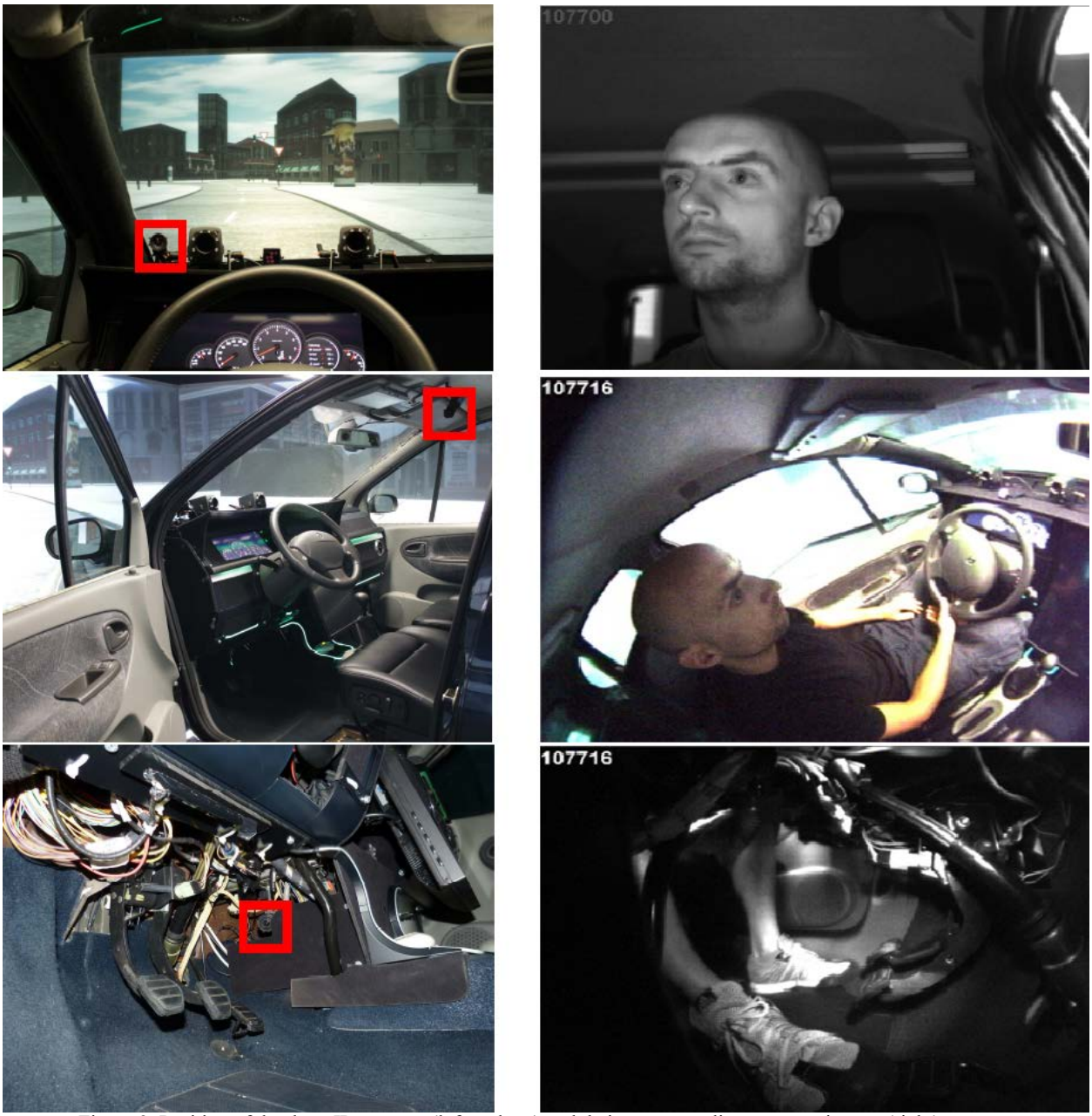


Figure 3. Position of the three IP cameras (left, red tag) and their corresponding camera pictures (right).

The three IP cameras recorded with a frequency of 30 Hz and the SILAB scenario itself recorded with 60 Hz. Time stamps were used for synchronization.

Test Participants

Altogether, 40 participants took part in the data gathering, of which 16 participants had to be excluded due to simulator sickness although the participants had been pre-selected for driving simulator robustness. Especially urban scenarios seem to cause a comparable high number of simulator sickness related data loss. For future studies we recommend to use curves with wide radius, reduce turning situations and reduce braking until full stop.

Consequently, 24 participants provided basis for further analysis with eight people classified to a driving style. As reported by Färber (2005) the individual driving style has an important influence on driver's intentions, therefore a standardized questionnaire was utilized. This questionnaire allowed classifying three different driver styles, namely "less-experienced, indecisive driver", "ordinary driver" and "sporty and ambitious driver". Participants were selected with an equal representation of these driving styles, having within the driving style groups a balanced

number of young and old and female and male drivers.

The experiment itself started with reading various documents such as a declaration of consent and information about the project itself as well as filling in a demographic questionnaire. A training ride was included to familiarize the participants with the driving simulator and the simulation environment. The actual test ride took around 45 minutes. Subsequent to the preceding steps a final questionnaire was presented to the test subjects, asking them about the recent simulator ride. The whole procedure took approximately 90 minutes.

DRIVER BEHAVIOR OBSERVATION AND CLASSIFICATION

Driver behavior was then categorized based on the video recordings. In a first step behavior categories and their classes where defined. Secondly the videos where analyzed for each maneuver and the observed behavioral differences where specified. Finally a procedure for human raters was developed which allows standardized behavior observation and prediction of the maneuvers.

Behavior Categories and Their Associated Behavior Classes

The data gathering led to 720 videos in total, 19 videos had to be excluded from further analysis in consequence of insufficient quality. Due to the effort encoding the single videos, both in an organizational and timely manner, only a representative part of 67 videos of this sample was used to develop the rater-based observation method. The analysis of the videos yielded five superior behavior categories with 29 behavioral classes collectively, which are listed in table 1.

Table 1: Superior behavior categories and their assigned behavioral classes.

Gaze Behavior	Front, Side-view mirror left, Pedestrian, Speedometer, Left glance over shoulder, Orientation glance right, Orientation glance left, Rear-view mirror
Hand posture on the steering wheel	Relaxed posture, Encompass, Unfasten the grip, Tighten the grip, Moving hand to the indicator, Using the indicator
Steering	No Movement, Left, Right
Pedalry	Accelerate, Moving the foot from the gas pedal, Foot is located on the gas pedal / no movement, Changing from gas to brake pedal, Changing from brake to gas pedal, Brake, Moving the foot from the brake pedal, Foot is located on the brake pedal / no movement
Torso movements	Head movement / rotation, Torso movement / rotation, Relaxed posture, Increased body tension

The behavior classification was designed for time windows of 500 ms. Thus every 500 ms one class per category is marked which results in a behavior description composed of five behavior classes per 500 ms. The raters were instructed to rate the current video as long as their subjective assurance changed from either “very unsure”, “unsure” or “sure” to “very sure”, with respect to maneuver / no maneuver. After rating four times in a row “very sure”, raters were told to stop coding the video and go on for another video.

Specific Behavior for Baseline and Maneuver Conditions

Using the introduced coding scheme, specific behavior classes for the three driving maneuvers were defined, aiming for distinguishable behavior between one of the maneuvers and their baseline. Table 2 provides an overview of specific differences in driving behavior, forming behavioral patterns for the three selected driving maneuvers.

Table 2: Overview of behavioral classes, which appear (almost) exclusively in one of both possible conditions.

	Turning left		**Changing lane**		**Bypassing a pedestrian**	
	Maneuver	**Baseline**	**Maneuver**	**Baseline**	**Maneuver**	**Baseline**
Gaze Behavior	Orientation glance right, Orientation glance left	Speedometer, Rear-view mirror	Side-view mirror left, Left glance over shoulder, Rear-view mirror	Speedo-meter	-	-
Hand posture on the steering wheel	Moving hand to the indicator, Using the indicator	Relaxed posture	Moving hand to the indicator, Unfasten the grip, Using the indicator	Relaxed posture	Encompass, Tighten the grip	Unfasten the grip
Steering	-	-	Left	-	-	-
Pedalry	Changing from gas to brake pedal, Changing from brake to gas pedal, Brake, Moving the foot from the brake pedal, Foot is located on the brake pedal / no movement	Foot is located on the gas pedal / no movement	Moving the foot from the gas pedal, Changing from gas to brake pedal, Changing from brake to gas pedal	Foot is located on the gas pedal / no movement	Moving the foot from the gas pedal, Changing from gas to brake pedal, Changing from brake to gas pedal, Brake	Foot is located on the gas pedal / no movement
Torso move-ments	Torso movement / rotation, Increased body tension	Relaxed posture	Head movement / rotation, Torso movement / rotation	Relaxed posture	Head movement / rotation, Increased body tension	Relaxed posture

Concerning the driving maneuver turning to the left, the executed analysis indicates that drivers show more dynamic sequences of movements in the maneuver condition in comparison to the baseline condition. This observation applies to all five behavioral categories. It can be said that in the maneuver condition visual exploration of the environment was obvious from the driver's gaze behavior. On the other hand, in the baseline condition, the drivers seemed to focus on controlling their speed rather than exploring their environment. Moving the hand to the indicator and using it has naturally only been detected in the maneuver condition.

A similar observation has been found for the second driving maneuver, changing lane. It has also been found that dynamic movement appears in the maneuver rather than the baseline condition, especially regarding behavior for exploring the environment. For this driving maneuver it can also be concluded that the baseline condition was rather spent on controlling speed parameters, whereas in the maneuver condition drivers focused on ensuring a safe lane changing. Again, moving the hand to the indicator and using it has only been detected in the maneuver condition.

Detecting behavior differences for bypassing a pedestrian was more difficult than for the other two driving maneuvers. No differences could be found in gaze behavior, which however might be attributed to an uncentralized camera position. Also steering behavior showed no difference. The main differences resulted from the behavioral categories torso movements and hand posture on the steeling wheel, in a kind that drivers who were about to show a braking behavior are more likely to show a tensed body language accompanied with hand movements than drivers who do not intend to use their brake pedal. Clearly distinguishable behavior could be found in the pedalry behavior category, as dynamic movements in interaction with the gas or brake pedal only occurred in situations where the

driver showed an intention to brake, whereas no movements could be registered when the driver did not intend to use the brake pedal.

Observer Training and Rater-Based Observation Method

To instruct the observers with the above-mentioned findings, an observer training has been performed which lasted five days and included four student raters in total. The aim of the training was to familiarize the raters with both: the behavior classes rating in 500ms scenes and the coding program Mangold Interact Version 9. The observers were trained in the behavior classification for each baseline and maneuver condition. They were instructed to rate the behavior classes and additionally to indicate at each interval if they consider intention for a baseline or maneuver condition. Raters were informed about the potential driving maneuver, e.g. turning left, changing lane or bypassing a pedestrian; their overall objective was to decide whether a baseline or maneuver condition was to be performed.

Due to timely limitations so far the raters coded only a subset of all recorded videos, consequently every observer decoded 27 videos whilst the first three and the last three videos were redundant in order to compute intra-rater-reliabilities. The videos to be coded were randomly chosen, the only constraint was to have one video of every of the 24 participants in the simulator data recording to ensure an even appearance of driving style. This random selection of videos also avoided any anticipation of the driving maneuvers condition. All of the raters decoded the same videos but in a randomly assigned order.

The rating itself lasted for three days, with a daily video decoding of four hours maximum to decrease any effects of fatigue.

VALIDATION OF THE OBSERVER METHOD

The above-mentioned rating with four student raters yielded 108 codings altogether of which 97 % of all codings proved to be correct with only seven videos coded wrongly. The majority of all correct codings, namely 72 codings, were correct at the first attempt, whilst in the other cases the participants decided on the wrong condition at first but corrected later to chose the right condition.

To reveal conclusions about the accuracy of the developed rating, three different methods were executed.

Extraction of Crucial Behavioral Classes

It was of interest to understand which behavioral classes supported an observer most. No chronological sequence of single behavioral classes could be extracted from the observers' codings. However, there were apparently certain behavioral classes, which seem to be crucial and led to a change in the observers' choice between baseline or maneuver condition. They are summarized in table 3.

It can be noted that the specific behavior for each condition does not completely correspond to the behavior classes which led to the change in their subjective assurance. Additionally, some behavior classes occur at a very short time before the maneuver is started or is indicated with the indicator.

Table 3: Overview of behavioral classes, which had the most influence on changing observers' choice for a certain condition. The first behavioral class in each cell indicates the driver's behavior in the 500ms time frame before the observers' decided for maneuver or baseline. The second behavioral class indicates the behavior which led to the change.

	Turning left		**Changing lane**		**Bypassing a pedestrian**	
	Maneuver	**Baseline**	**Maneuver**	**Baseline**	**Maneuver**	**Baseline**
Gaze Behavior	Orientation glance right Orientation glance right	Front Front	Side-view mirror left Side-view mirror left	Front Front - Speedometer	Pedestrian Pedestrian	Pedestrian Pedestrian
Hand posture on the steering wheel	Relaxed posture - Moving hand to the indicator Relaxed posture - Encompass	Relaxed posture Encompass - Encompass	Relaxed posture Moving hand to the indicator	Relaxed posture Relaxed posture	Unfasten the grip Unfasten the grip - Tighten the grip	Relaxed posture Relaxed posture - Encompass
Steering	Right – Left Right - No movement	Left Left	No movement No movement	Right - No movement No movement - Right	Right Right	Right Left
Pedalry	Changing from gas to brake pedal – Brake Foot is located on the gas pedal / no movement - Moving the foot from the gas pedal	Changing from brake to gas pedal Accelerate	Foot is located on the gas pedal / no movement Foot is located on the gas pedal / no movement	Moving the foot from the gas pedal Moving the foot from the gas pedal	Foot is located on the gas pedal / no movement - Moving the foot from the gas pedal Moving the foot from the gas pedal - Moving the foot from the gas pedal	Foot is located on the gas pedal / no movement Foot is located on the gas pedal / no movement
Torso movements	Relaxed posture Relaxed posture	Relaxed posture Relaxed posture	Head movement / rotation Head movement / rotation	Relaxed posture Relaxed posture	Relaxed posture Relaxed posture	Relaxed posture Relaxed posture

Inter- and Intra-Rater-Reliability

Inter- and intra-rater measurements were performed using Cohen's kappa coefficient for categorical items of two raters. The translation of values into categories according to Altman (1991) is displayed in table 4.

Table 4: Altman's kappa values.

slight agreement	fair agreement	moderate agreement	substantial agreement	(almost) perfect agreement
$\kappa \leq 0.20$	$0.21 \leq \kappa \leq 0.40$	$0.41 \leq \kappa \leq 0.60$	$0.61 \leq \kappa \leq 0.80$	$0.81 \leq \kappa \leq 1.00$

Inter-rater reliability calculations revealed κ - values between $\kappa = 0.30$ and $\kappa = 1.00$. Table 5 provides an overview of the inter-rater-reliability, arranged according to the different driving maneuvers and the distinct behavior categories.

Table 5: κ overview per behavioral class, which appear (almost) exclusively in one of both possible conditions.

Driving Scenario	Gaze Behavior	Hand posture on the steering wheel	Steering	Pedalry	Torso movements	Mean
Turning Left - Maneuver	0.44	0.37	0.66	0.64	0.92	**0.61**
Turning Left - Baseline	0.49	0.30	0.70	0.56	1.00	**0.61**
Changing Lane - Maneuver	0.64	0.59	0.62	0.67	0.60	**0.62**
Changing Lane - Baseline	0.58	1.00	0.63	0.60	0.69	**0.70**
Bypassing pedestrian - Maneuver	0.50	0.58	0.52	0.55	1.00	**0.63**
Bypassing pedestrian - Baseline	0.33	0.56	0.58	0.39	1.00	**0.57**
Mean	**0.50**	**0.57**	**0.62**	**0.57**	**0.87**	

Torso movements ($\kappa = 0.87$) and steering ($\kappa = 0.62$) showed substantial inter-rater reliability, while hand position and pedal position (both $\kappa = 0.57$) showed moderate correlation. Gaze behavior resulted in the smallest kappa value ($\kappa = 0.50$), which was probably due to a non-optimal camera position. In order to interpret gaze direction adequately from video images the camera should be positioned centrally in front of the driver's face.

With respect to the maneuvers the inter-rater reliability showed for all maneuvers a substantial correlation. Changing lane yielded the highest inter-rater-values ($\kappa = 0.70$ and $\kappa = 0.62$) whereas the choice in behavioral classes of the scenario "bypassing a pedestrian" varied most ($\kappa = 0.57$ and $\kappa = 0.63$).

The intra-rater-reliability amounted to be in a similar range between $\kappa = 0.50$ (Hand posture on the steering wheel) and $\kappa = 0.80$ (Torso movements).

Observers' Subjective Feedback

The questionnaire, which was presented to the raters on the last day of their rating, was supposed to deliver ideas for forthcoming edited versions of the observation method. Overall, the raters were satisfied with the developed coding scheme, although they thought the scheme to be insufficient in some situations during the rating. Therefore some behavior categories, in particular pedalry and hand posture on the steering wheel, should be described by more classes. They commented that in some scenarios it was impossible to attribute a behavior either to any intention or to a situational context, specifically curves result in corrective steering which could be confounded with preparation of a lane change or turn.

Another important part of the final questionnaire was dealing with the predictability of the single driving maneuvers. All four raters indicated that they found turning left and changing lane easily detectable whereas again all of them thought that an intention to use the brake pedal when passing a pedestrian was rather difficult to foresee.

CONCLUSIONS

We found in our observation study that driving maneuvers are preceded by observable behavior. The classification of driver's behavior allows identifying the differences between a maneuver preparation (driver intention) and continuation of the current maneuver, following the lane in our case. Moreover it was possible to specify the behavior classes so objectively, that even different raters showed an acceptable correlation.

Hence, we assume that an observation approach is suitable to classify driver behavior and to predict driving maneuvers. The behavior classes need to be specified for each maneuver separately. It seems that for each maneuver, or even variation of maneuver, a specific set of behavior classes is needed. Only generally speaking, longitudinal maneuvers are mainly predicted by foot movements, perpendicular maneuvers are predicted by hand and steering movements as well as shoulder glances and side mirror views. The study however showed that each maneuver is different and needs specific investigation.

Our raters indicated a high influence of the situation on the behavior. On curved road strips and even shortly after curves steering wheel movements could ether be attributed to the curve, or to a lane change maneuver preparation. In those cases a correct maneuver prediction was not possible - fact that should be even more relevant in real driving. Typical situational contexts influence the behavior and interfere with the behavior reflecting intention.

Another important aspect are inter- and intra-individual behavior differences of drivers. Simulator studies allow, in contrast to real life driving, a high level of scenario and maneuver standardization. And even here we discovered a wide range of individual behavior. Some drivers raise the foot over the brake pedal when they prepare to brake, others only lift the gas pedal. Some drivers change hand position before turning, others keep their hands in a suitable position all the time. Left glance over shoulder may happen when chaining lane, but may also not.

In conclusion maneuver prediction by driver behavior classification is influenced by driver's intention, the situational context and personal states and traits.

AKNOWLEDGEMENTS

The work presented in this paper is part of our activities in the UR:BAN project "Human in Traffic" (Urban-project, 2012-2016) which is co-funded by the Federal Ministry for Economic Affairs and Energy according to a resolution of the German Bundestag. The authors express their special thanks for the fruitful discussions within the consortium and especially thank Paul Czech, Matthias Beggiato and Dietrich Manstetten for their stimulation and support of this study. At Fraunhofer IAO the technical staff around the vehicle interaction lab and its driving simulator receives a special thanks for their support and Prof. Dieter Spath for his advice and support of this publication.

REFERENCES

Ajzen, I. (2002). *Constructing A Theory of Planned Behaviour Questionnaire*. Website of University of Massachusetts http://people.umass.edu / aizen / pdf / tpb.measurement.pdf

Altman, D. G. (1991). Practical Statistics for Medical Research. Boca Raton: Chapman & Hall / CRC.

Berndt, H.; Emmert, J.; Dietmayer, K. (2008): *Continuous Driver Intention Recognition with Hidden Markov Models*. 11th International IEEE Conference on Intelligent Transportation Systems. Institute of Measurement, Control, and Microtechnology University Ulm. Beijing, China, 12.10.2008.

Blaschke, C.; Schmitt, J.; Färber, B. (2007): *Fahrmanöver-Prädiktion über CAN-BUS Daten*. In: VDI-Bericht (2015), S. 165–172.

Dambier, M. (2010): *Adaptive Information Flow Control Recognition and Prediction of Factors Contributing to Driver's Stress*. Dissertation. Universität der Bundeswehr, München.

Diederichs, F.; Weber, B.; Ganzhorn, M. (2011): *Experimental evaluation of ITST based school bus stop notifications for car drivers*. In: 11th International Conference on Intelligent Transport Systems Telecommunications, ITST 2011. Proceedings. Piscataway, NJ: IEEE, S. 345–350.

Färber, B. (2005): *Erhöhter Fahrernutzen durch Integration von Fahrerassistenz- und Fahrerinformationssystemen*. In: Markus Maurer und Christoph Stiller (Hg.): Fahrerassistenzsysteme mit maschineller Wahrnehmung. [basiert auf ausgewählten Vorträgen eines Workshops in Walting (Altmühltal)]. Berlin: Springer, S. 141–160.

Freyer, J.; Deml, B.; Maurer, M.; Färber, B. (2007): *ACC with enhanced situation awareness to reduce behavior adaptations in lane change situations*. IEEE Intelligent Vehicles Symposium. Istanbul, Turkey, 13.06.2007.

Ganzhorn, M.; Diederichs, F.; Widlroither, H. (2011): *Enhancing the reproduction of road signs in automotive HMIs to protect children.* In: 11th International Conference on Intelligent Transport Systems Telecommunications, ITST 2011. Proceedings. Piscataway, NJ: IEEE, S. 338–344.

Gerdes, A. (2006): *Driving Manoeuvre Recognition*. German Aerospace Center. Braunschweig, Germany.

Henning, M.; Georgeon, O.; Krems, J. (2007): *The Quality of Behavioral and Evironmental Indicators Used to Infer the Intention to Change Lanes*. Vision in Vehicles 11. Dublin, Ireland, 2007.

InCarIn-project: Project website: http://www.incarin.de 2014-2016

Kobiela, F. (2012): *Fahrerintentionserkennung für autonome Notbremssysteme*. Dissertation. Technische Universität Dresden, Dresden.

König, R., Heintz, P. R., & Scheuch, E. K. (1962). *Beobachtung und Experiment in der Sozialforschung*. Köln: Kiepenheuer & Witsch.

Kopf, M. (2005): *Was nützt es dem Fahrer, wenn Fahrerinformations- und -assistenzsysteme etwas über ihn wissen?* In: Markus Maurer und Christoph Stiller (Hg.): Fahrerassistenzsysteme mit maschineller Wahrnehmung. [basiert auf ausgewählten Vorträgen eines Workshops in Walting (Altmühltal)]. Berlin: Springer, S. 117–139.

Kuge, N.i; Yamamura, T.; Shimoyama, O.; Liu, A. (2000): *A driver behavior recognition method based on a driver model framework*. In: Intelligent vehicle systems, S. 47–54.

Manstetten, D.; Dambier, M.; Krautter, W.; Quander, M. & Rimini-Döring, M. (2011): *Adaptive und kooperative Technologien für den intelligenten Verkehr*: Forschungsinitiative Aktiv. Schlussbericht. Beitrag des Zuwendungsempfängers: Robert Bosch GmbH zum Teilprojekt: Aktiv-Verkehrsmanagement Teil 4. Stuttgart

Oliver, N.; Pentland, A. (2000): *Driver Behavior Recognition and Prediction in Smart Car*. Proceedings of IEEE Conference and Intelligent Vehicles. Dearborn (MI), USA, 2000.

Pentland, A.; Liu, A. (1999): *Modeling and prediction of human behavior.* In: Neural Computation 11, 1999, S. 229–242.

Salvucci, D. (Hg.) (2004): *Inferring driver intent: A case study in lane-change detection*. In: Proceedings of the Human Factors Ergonomics Society 48th Annual Meeting. Santa Monica, USA.

Schnell, R., Hill, P. B., & Esser, E. (2011). *Methoden der empirischen Sozialforschung*. München [u.a.]: Oldenbourg Wissenschaftsverlag.

Schroven, F.; Giebel, T. (2008): *Fahrerintentionserkennung für Fahrerassistenzsysteme. Driver Intent Recognition for Advanced Driver Assistance Systems*. In: VDI-Bericht (2048), Wolfsburg. Düsseldorf: VDI-Verlag.

Takagi, T.; Nishi, T.; Yasuda, D. (2000): *Computer Assisted Driving Support based on Intention Reasoning*. Dept. Of Computer Science, Meiji Univ. Kanagawa-ken, Japan.

TEAM-project: Project website: http://www.collaborative-team.eu 2012-2016,

UR:BAN project: Project website: http://www.urban-online.org 2012-2016

WIVW. (2013). *Produkte und Dienstleistungen: SILAB*. Website of Würzburger Institut für Verkehrswissenschaften: http://www.wivw.de / Produkte Dienstleistungen / SILAB / index.php.de

Impact of Music Tempo on Simulated Driving Performance

Jordan Navarro and Emanuelle Reynaud

University of Lyon
Laboratoire d'étude des mécanismes cognitifs
5 avenue Pierre Mendès France, Bron, France

ABSTRACT

It is known that drivers often listen to music when they are in their vehicle. Listening to music while driving can impact drivers' behaviours in several ways, through emotion induction or arousal changes for instance. It has already been shown that various musical tracks with different tempos (i.e. the music pace) can impair drivers' performances (Brodsky, 2002). The current study focused on the impact of music tempo on vehicular control by using the same music track played at different tempos. Four experimental conditions associated with four different versions of the same track, differing only by their tempos, have been used while participants had to follow a lead vehicle and then drive with no particular constraints. Participants had to drive in a simulated environment without music, with the music track of their choice, with the music of their choice but with a modified tempo (+10% of the regular tempo and -10% of the regular tempo). Results indicated that the music tempo slightly changed the level of arousal of participants as observed through heart rate and subjective questionnaires data. However these changes in the arousal level did not translate into driving behaviours modifications (i.e. inter-vehicle time while following a lead vehicle and speed in free driving context).

Keywords: Music, Driving, Tempo, Driving Behaviour, Arousal.

INTRODUCTION

Playing and listening to music have always been an essential part of human culture. It is known that people often listen to music while driving their vehicle either for professional or recreational purposes. It is even the first self-declared music listening place before listening to music alone at home, while exercising or hanging out with friends (Rentfrow & Gosling, 2003). About 75% of drivers listen to music or to the radio while driving (Dibben & Williamson, 2007). Because in our modern societies driving is a daily activity for most people, it comes with no surprise that several authors intended to study how these two activities (listening to music and driving a vehicle) can affect each other. More precisely, these studies focused on the potential impact of music listening on driving performances (e.g. Bellinger, Budde, Machida, Richardson, & Berg, 2009; Brodsky & Slor, 2013; Dalton, Behm, & Kibele, 2007; Dibben & Williamson, 2007; Nelson & Nilsson, 1990; Pecher, Lemercier, & Cellier, 2009; Ünal, Steg, & Epstude, 2012; Ünal, de Waard, Epstude, & Steg, 2013; Van der Zwaag & al., 2012).

Music can be a subject of study on its own but it can also be studied as "background music" that is to say listening to music while performing another task in parallel. In many cases, such as travelling on public transport, shopping, eating at a restaurant or driving for instance, music listening is carried out as a background activity. The influence of background music on the main activity can be very different depending on both the music and the activity consid-

ered. While driving, background music has been shown neutral or positive in some cases but negative in some other cases. On the one hand, it was found no music effects on lateral position, speed and reactions to hazardous incidents (Hatfield & Chamberlain, 2008). Furthermore, it was also found that music listening did not impair driving performance and could even help drivers in some situations requiring sustained attention such as car following (Ünal & al., 2012) for instance. On the other hand when drivers were exposed to high volume background music, driving behaviours were impaired (Dalton, Behm and Kibele, 2007). Consequently it seems that the effects of background music on driving is a tricky question, presumably sensitive to several factors related to both driving and music and probably interacting with each other.

Music itself is a complex combination of interacting properties such as tempo, rhythm, mode, pitch and harmony played at a given intensity. Therefore listening to music behind the wheel can affect driving behaviour in several ways. Besides the distraction of drivers interacting with the radio player for instance (Young & Salmon, 2012), music is known to be an emotional inductor that might impair driving performances (Pêcher, Lemercier & Cellier, 2009).

In the present experiment we decided to focus on one of the dimensions of background music: the tempo. The tempo of a music track can be defined as its speed or pace (expressed in beats per minute or BPM). The tempo is considered as a crucial element of most musical compositions. Previous results on the effect of different music tracks associated with different tempos have been reported (Brodsky, 2002). Music tracks with faster tempos led to an increased spontaneous speed, and raised both the number of lane crossings and disregarded red traffic lights in a simulated driving environment. Brodsky classified music tempo in three categories: slow tempo for music tracks ranging from 40 to 70 BPM, moderate tempo for tracks ranging from 85 to110 BPM, and fast tempo for music ranging from 120 to 140 BPM. The main criticism that can be made to Brodsky's experiment is that the manipulation of tempo was obtained through the use of different music tracks. Indeed, Brodsky manipulated tempo by selecting different music tracks from the three categories defined earlier. As a consequence, the observed effects attributed to tempo can also be partly attributed to other music properties and to emotions induced by the different tracks.

The current study was designed to assess the impact of music tempo on vehicular control, using the same music track under different tempo conditions. Four experimental conditions associated with four different versions of the same track, differing only by their tempos, have been used while participants had to follow a lead vehicle at a given distance and then drive with no particular constraints. Participants had to drive without music, with the music track of their choice, with the music of their choice but with a modified tempo (+10% of the regular tempo and -10% of the regular tempo).

Following Kahneman (1973), it was hypothesized that, with an increased tempo, drivers' arousal will increase, leading drivers to be more reactive and therefore more able to maintain the distance with the leading vehicle. Therefore in the case of this simple following task, or free driving, an increase of the tempo is expected to improve performance. The opposite is expected with a decreased musical tempo.

METHOD

Participants

Twenty-four participants volunteered for the experiment. Participants were aged from 18 to 40 years old (22 years old ± 4.4 years) and received their driving licence from 0.2 to 22 years (3.7 years ± 4.5) before the experiment. They drove between 1000 and 27 000 kilometers per year (9075 km ± 8566). The experiment received authorisation from the local ethics committee.

Musical background

Prior to the experiment participants were asked to bring their current preferred music track. The software Audacity® (http://audacity.sourceforge.net/) has been used to play the music track for the duration of each driving session and to modify the tempo of the music. Musical tempo modification has been obtained without any pitch modification, as allowed by the SoundTouch library from Audacity®. Participants faced four auditory backgrounds in a randomized order: (1) No Music - NM- (2) Music without modification -M- (3) Music with +10 % of the regular tempo -M +10- (4) Music with -10 % of the regular tempo -M -10-.

Because music intensity is capable of modifying driving performances (Dalton & Behm, 2007), the intensity of the music was fixed at 75 dB for all auditory backgrounds except for the NM condition.

Procedure

After completion of the consent form, participants took place in front of a driving simulator. The driving scene was played on a 19 inches screen, steering wheel and pedals Logitech G27® were used to control the simulated vehicle. The open-source OpenSD2S software (Institut Image – Le2i/CNRS, ENSAM & Renault) was used for driving simulation. The same driving scenario was used for all four auditory conditions plus a training scenario prior to the proper start of the experiment. The driving scenario was composed of two phases both happening in a similar environment, consisting in a two-lane road with no traffic. During the first phase, participants had to follow a lead vehicle at a 40 meters distance. This lead vehicle (a blue Subaru Impreza) decelerated and accelerated once and then parked on the right side of the road after two and a half-minute of driving. At that time participants had to overtake the lead vehicle and to complete phase two of the scenario by driving with no constraints for about thirty seconds. Prior to the experiment participants were told to respect a speed limit set at 90 Kilometers Per Hour or KPH (i.e. the usual speed limit on two-lane roads in France). Each participant carried out the four driving conditions (NM, M, M+10, M-10).

Data analysis

Three different types of data were collected and analyzed. First of all, the heart rate average and standard deviation were computed for each experimental condition (Biopac® equipment and Acqknowledge® software). Then, the subjective emotional state was collected after each experimental condition using the Brief Mood Introspection Scale (BMIS scale, Mayer and Gaschke, 1988). Finally, driving data were recorded. The Inter-Vehicle Time (ITV) between the lead vehicle and the driven vehicle in phase 1 and the driven vehicle mean speed in phase 2, were then calculated from driving data.

Within-participants ANOVAs were used for statistical analysis. The level of significance of $p < .05$ was used for all statistics. p values between .05 and .1 were considered as a trend. For post-hoc analysis the LSD of Fisher was used in all cases.

RESULTS

Heart rate

Due to recording problems, the data from five participants were unusable and therefore not included in the following heart rate analyses.

Participants mean heart rate was significantly affected by the musical background $F(3, 51)=4.3$, $p< .01$; see Figure 1. The mean heart rate in the NM condition was significantly different from the mean heart rate in the M +10, M -10 conditions ($p< .05$) and tended to differ from the M condition ($p< .1$). Besides, still for the mean heart rate, the M condition tended to differ from the M +10 condition ($p< .1$). To sum up, the presence of music increased the mean heart rate compared to the no-music condition. In addition, the tempo sped up by 10% tended to increase the mean heart rate compared to music at its regular pace.

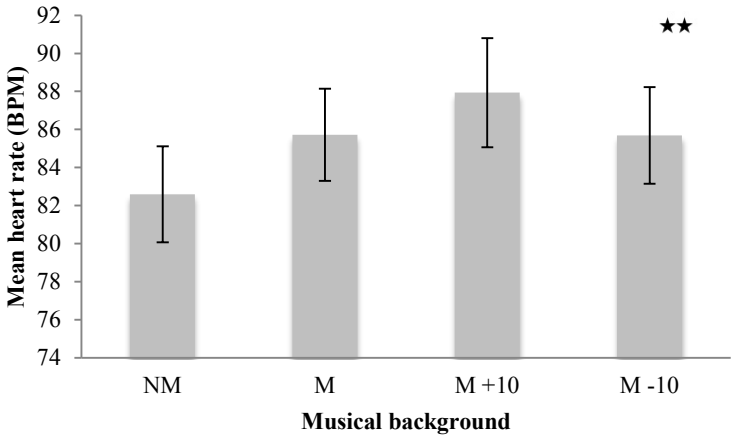


Figure 1. Mean heart rate in the four musical backgrounds in beats per minute.
Error bars represent standard errors.

The heart rate standard deviation was not significantly modified by the musical background F(3, 51)=0.9, NS; see Figure 2.

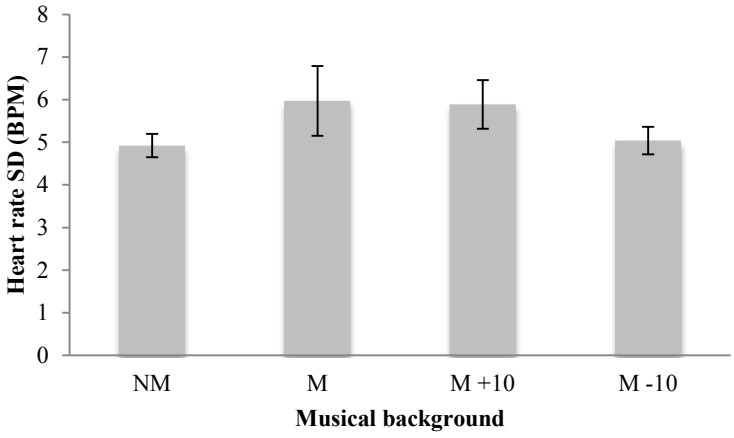


Figure 2. Heart rate standard deviation in the four musical backgrounds in beats per minute.
Error bars represent standard errors.

BMIS

The arousal-calm (Figure 3) and the pleasant-unpleasant (Figure 4) dimensions scores, as assessed by the Brief Mood Introspection Scale (BMIS), were analysed.

Arousal-calm dimension

The musical background significantly affected the level of arousal F(3, 69)=7.7, $p<.001$. Post-hoc analyses revealed that NM background differed from all other musical backgrounds ($p< .05$). In addition M +10 background tended to differ from both M and M -10 ($p< .1$) backgrounds.

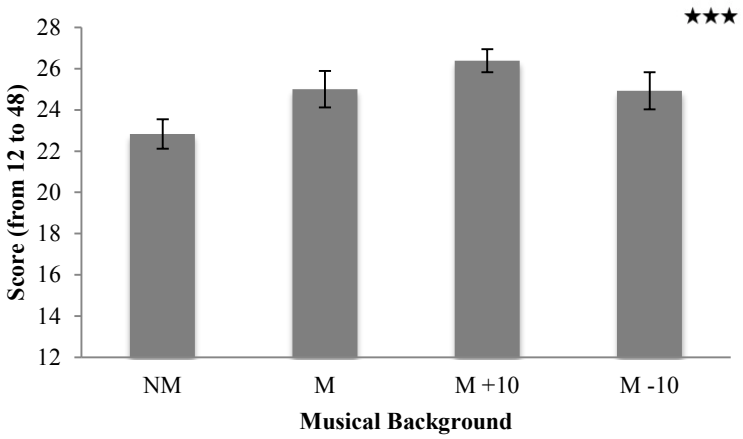


Figure 3. Arousal-calm dimension score from 12 to 48. The higher the score is, the higher is the arousal. Error bars represent standard errors.

Pleasant-unpleasant dimension

The musical background significantly affected the level of pleasantness $F(3, 69)=2.8$, $p<.05$. Post-hoc analyses revealed that NM background differed from M and M +10 backgrounds ($p< .05$).

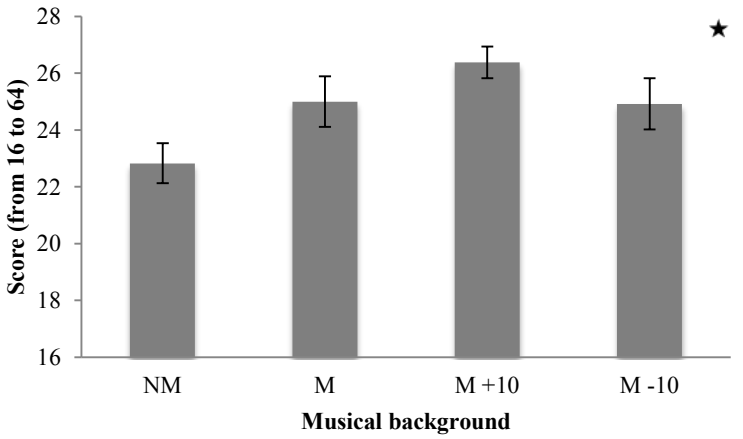


Figure 4. Pleasant-unpleasant dimension score from 16 to 64. The higher the score is, the more pleasant the participant felt. Error bars represent standard errors.

Driving data

Global analyses

Statistical analyses did not reveal any significant effect of the musical background neither for IVT ($F(3, 69)=1.9$, NS; Figure 5), nor for the mean spontaneous speed ($F(3, 69)=0.6$, NS; Figure 6).

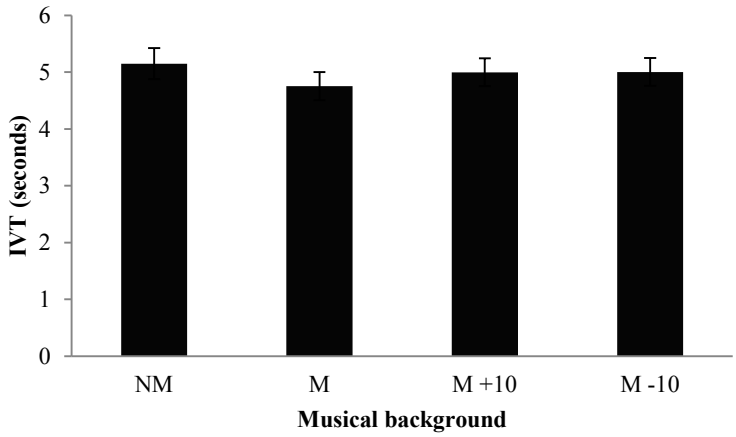


Figure 5. Mean Inter-vehicle Time (IVT) per musical background in seconds.
Error bars represent standard errors.

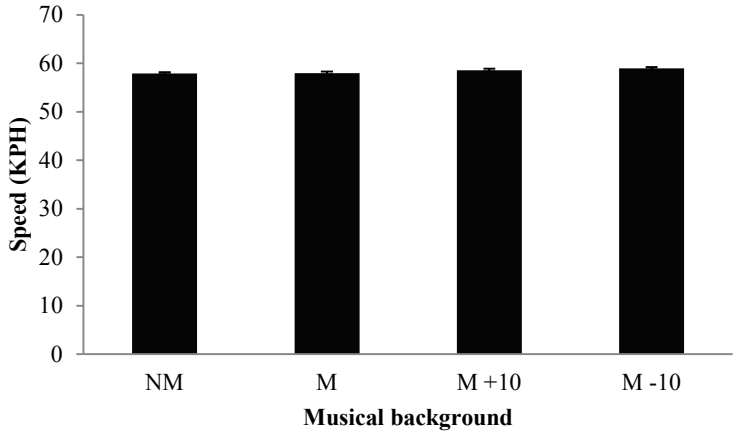


Figure 6. Mean speed per musical background in KPH.
Error bars represent standard errors.

Detailed analyses

To get a clearer picture of the potential impact of musical background on driving performances, analyses were carried out specifically during the acceleration and deceleration phases of the lead vehicle. The IVT and the number of steering wheel reversals (i.e. number of steering wheel direction changes) were computed. Statistical analysis did not reveal any significant effect neither on the acceleration nor on the deceleration phase, for every variable (IVT or steering wheel reversals) considered.

DISCUSSION

Consistent results from heart rate averages and the arousal-calm dimension scores of the BMIS indicated that the presence of background music modified the participants' level of arousal. Indeed the mean heart rate and the level of subjective arousal without music almost always significantly differed from all musical backgrounds. However, the slight changes of music tempo (plus or minus 10% from the original music track tempo) only tended to modify arousal. This trend was observed for music without any tempo modification compared to music sped up by 10%,

when looking at the mean heart rates and BMIS arousal scale. Only the BMIS arousal scale tended to indicate that music sped up by 10% increased the level of arousal more than music slowed down by 10%. The observed changes in terms of arousal level did not transfer into any impairments or improvements of driving performances. Both Inter-Vehicle Time (IVT) during the lead vehicle following task, and spontaneous speed in the free driving section were not significantly modified across the four musical backgrounds.

It is undeniable that music might impact people arousal and emotions whatever the country and the culture. Nevertheless, not all individuals are affected the same way by music. Moreover, not all individuals respond the same way to a given music track. Correlations between musical tastes and several personality dimensions have even been shown (Rentfrow & Gosling, 2003). As a consequence, each participant should elect his own musical background in order to ensure a positive impact of the music. This was the choice made in this experiment, where participants brought their own current preferred music track. BMIS questionnaires pointed out that when the musical tracks of the participants were played, they felt more pleased compared to the no-music background. Music track tempo modifications did not change significantly the level of pleasantness. Therefore, it seems that offering participants the opportunity to elect their own musical track was a good choice because it affected their mood consistently in all musical conditions.

Giving participants the opportunity to select their own musical track resulted in tempos ranging from 76 to 172 beats per minute (mean 115 BPM ± 23) before tempo modifications. Because the original tempo was not the same depending on music tracks, tempo modifications (plus or minus 10%) might not have the exact same effects on arousal and driving. In the music plus 10%, background tempos ranged from 83 to 190 (mean 127 BPM ± 26) and in the music minus 10%, background tempos ranged from 68 to 155 (mean 104 BPM ± 21). For comparison Brodsky (2002) used slow tempo tracks ranging from 40 to 70 BPM, moderate tempo tracks ranging from 85 to 110 BPM, and fast tracks ranging from 120 to 140 BPM. In Brodsky's experiment participants had to choose between a few pre-selected musical tracks. When averaged for all participants, all musical backgrounds used in the current experiment were included in a 104-127 BPM interval. These tempos only cover a small range of the large variety of tempos (from 40 to 140 BPM) used by Brodsky (2002).

Even if Brodsky did not observe any significant changes in the heart rate across the different musical condition, contrary to what was found in the current experiment, increases in spontaneous speed, number of lane crossings and disregarded red traffic lights were reported. The current experiment focused on the impact of music on vehicle control (mostly longitudinal through IVT and spontaneous speed). Possibly because of the much smaller tempo variation across conditions as compared to the ones used by Brodsky, the results on spontaneous speed reported were not reproduced in the current study. Brodsky found an almost linear increase in spontaneous speed from slow tempo tracks (mean speed about 132 KPH) through moderate tempo tracks (mean speed about 136 KPH) to fast tempo tracks (mean speed about 140 KPH). It must be noted that these speeds are more than twice the ones recorded in the current study (slightly below 60 KPH). The average spontaneous speeds reported by Brodsky are not realistic in a combined urban and rural environment (the Chicago ring simulated environment) with speed limits set at 50 KPH for municipal road sections and at about 105 KPH for interstate highway. The very high spontaneous speeds obtained by Brodsky compared to speed limits should invite to consider with caution the variation across musical conditions reported.

Based on the data from our experiment it is not possible to conclude that the music tempo (of a same track) is not impacting driving performances because the differences between musical backgrounds were possibly not big enough. It is likely that music tempo is affecting the level of arousal of drivers: the faster the tempo is, the higher the arousal level is. However, the differences observed in the current study are quite small and elicited only a statistical tendency in the results. This is possibly why the auditory background has not significantly affected drivers' performances. The next step is to collect data with larger differences in terms of tempo between musical background conditions.

CONCLUSION

A slight tempo modification (±10%) of the participants' own music track caused slight changes in arousal as observed through questionnaires and heart rate analysis. However these changes did not translate into significantly different driving behaviours in the current experiment. Further studies are needed before a definitive conclusion can

be drawn on the impact of music tempo on driving performances. In addition, driving context, music rhythm, music intensity, and emotions inducted through music are some of the main factors that probably combine and impact driving behaviours. Further research is also required to investigate how these different variables influence driving behaviours, for a full understanding of such a popular activity as listening to music while driving.

AKNOWLEDGMENTS

This work was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon. The authors are grateful to Manon Perdrix and Geoffrey Seccia for their help and to the developers of the OpenSD2S software for their support.

REFERENCES

Beh, H. C., & Hirst, R. (1999). Performance on driving-related tasks during music. *Ergonomics, 42*, 1087–1098.

Bellinger, D. B., Budde, B. M., Machida, M., Richardson, G. B., & Berg, W. P. (2009). The effect of cellular telephone conversation and music listening on response time in braking. *Transportation Research Part F: Traffic Psychology and Behaviour, 12*, 441–451.

Brodsky, W. (2002). The effects of music tempo on simulated driving performance and vehicular control. *Transportation Research Part F: Traffic Psychology and Behaviour, 4,* 219-241.

Brodsky, W., & Slor, Z. (2013). Background music as a risk factor for distraction among young-novice drivers. *Accident Analysis & Prevention, 59*, 382-393.

Dalton, B.H., & Behm, D.G. (2007). Effects of noise and music on human and task performance: A systematic review. *Occupational Ergonomics, 7*, 143-152.

Dalton, B.H., Behm, D.G., & Kibele, A. (2007). Effects of sound types and volumes on simulated driving, vigilance tasks and heart rate. *Occupational Ergonomics, 7,* 153-168.

Dibben, N., & Williamson, V. J. (2007). An exploratory survey of in-vehicle music listening. *Psychology of Music, 35*, 571–589.

Hatfield, J. & Chamberlain, T. (2008). The effect of audio materials from a rear-seat audiovisual entertainments system or from radio on simulated driving. *Transportation Research Part F, 11*, 52-60

Kahneman, D. (1973). Attention and effort. New Jersey, USA: Prentice Hall.

Mayer, J.D., & Gaschke, Y.N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology, 55,* 102-111.

Nelson, T. M., & Nilsson, T. H. (1990). Comparing headphone and speaker effects on simulated driving. *Accident Analysis & Prevention, 22*, 523–529.

Pêcher, L., Lemercier, C., & Cellier, J.M. (2009). Emotions drive attention: Effects on driver's behavior. *Safety Science, 47*, 1254-1259.

Rentfrow, P.J & Gosling, S.D (2003). The Do Re Mi's of Everyday Life: The Structure and Personality Correlates of Music Preferences. *Journal of Personality and Social Psychology, 84*, 1236–1256.

Ünal, A.B., Steg, L., & Epstude., K. (2012). The influence of music on mental effort and driving performance. *Accident Analysis and Prevention, 48*, 271-278.

Ünal, A. B., de Waard, D., Epstude, K., & Steg, L. (2013). Driving with music: Effects on arousal and performance. *Transportation Research Part F: Traffic Psychology and Behaviour, 21*, 52-65.

Van der Zwaag, M.D., Dijksterhuis, C., De Waard, D., Bulder, B.L.J.M., Westerink, J.H.D.M., & Brookhuis, K.A. (2012). The influence of music on mood and performance while driving. *Ergonomics, 55,* 12-22.

Young, K. & Salmon, P. (2012). Examining the relationship between driver distraction and driving errors: a discussion of theory, studies and methods. *Safety Science, 50,* 165-174.

Section 20

Aviation:

Human Factors Issues in Air Traffic Management II

Development and Application of a Hybrid Control Theory Model to Quantify Human-Machine Interaction Problems on the Flight Deck

Caitlin Anne Surakitbanharn [a] *and Steven J. Landry* [b]

[a]*Department of Industrial Engineering*
Purdue University
Hiyoshi, Yokohama-shi, 223-0061, Japan

[b]*Department of Industrial Engineering*
Purdue University
West Lafayette, IN, 46709, USA

ABSTRACT

Previous work has been conducted by McRuer [McRuer, 1973] to model the gain and delay of the pilot-aircraft system using classic control theory methods. However, a continuous extraction of gain and delay further enhances the ability to gather valuable, quantitative data to make inferences about human-machine interaction problems that occur on the flight deck. A parameter tracking system, based on McRuer's crossover model, has been developed to simultaneously and continuously track the gains and delays at the pilot-machine interface with which pilots track particular axes. The values extracted for gain and delay may be indicative of particular human factors issues that are prevalent in the cockpit, such as inattention, complacency, low situational awareness, and high workload. This model has been developed for use in an artificial tracking task, where it has been observed that if gain and delay are allowed to change freely, the model is overfit and not realistic, as it is believed that gain and delay change rarely, though not necessarily slowly. To combat this, one gain/one delay is applied to the system to find time periods where the fit is poor, and these frames of time are then re-evaluated to find a gain/delay value that fits, and this method is applied across all poorly fit time periods to create a well-fit model. This method of modeling is then applied to a realistic flight simulation task for pilots and is intended to evaluate the levels of gain and delay for specific human-machine tasks incurred on the flight deck, and which axis are affected the most in flight.

Keywords: Aviation Safety, Human Factors, Workload, Human-Automation Interaction

INTRODUCTION

The pilot-aircraft system as defined by McRuer states that the pilot and aircraft must be analyzed as one system, wherein the pilot reacts to displayed errors (e), normal and desired command inputs (i), and vehicle output motions (m) in order to produce control actions (e). This closed-control loop is illustrated below in Figure 1.

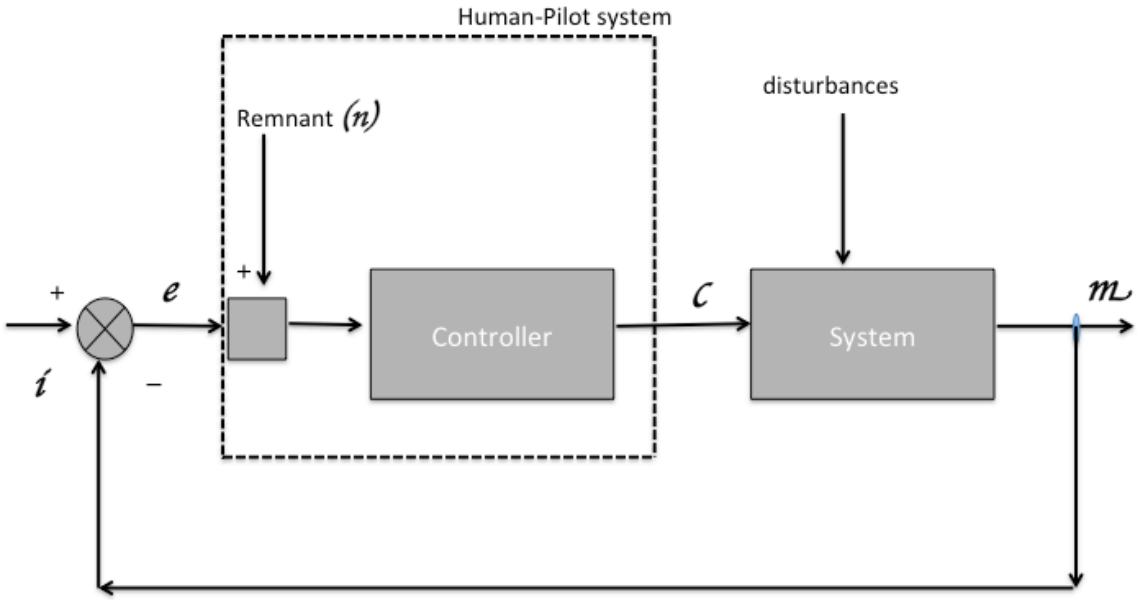


Figure 1. Pilot/Aircraft Closed Loop Control Circuit (Adapted from McRuer, 1973)

The remnant (n) and controller transfer characteristics are functions of operator, procedural, and environmental variables. That is, the decision making process for control actions is affected by the operator, procedures, and the environment. Once a control action is decided, it is performed by the Human-Pilot system and input into the system, which could be the rudder, the yoke, the elevator, a display etc. Disturbances also feed into the system occasionally. Once a control action is performed in the system, an output is recorded (m). If disturbances were present, the system output may not be desirable, resulting in an error, which will loop, back to the beginning of the system [McRuer, 1973] [McRuer, Krendel 1974].

This closed-loop control circuit is repeated several times throughout a given flight, and several loops may be present in the system at once [4]. In order to accomplish control and guidance of the aircraft, the pilot sets up many of the loops around the airplane, which could not perform these tasks by itself. For example, in order to increase the altitude of the aircraft, the pilot-human system may set up a closed loop around the aircraft elevator, around the aircraft yoke in the cockpit, and perhaps around radio as well. In this case, remnant transfer characteristics and control characteristics could be prior experience and knowledge of the operating pilot, physical condition of the pilot, time into flight (take-off, climb, cruise, descent, etc), standard operating procedures of the aircraft, governing aviation body rules and regulations, and/or weather conditions (Instrument or Visual Flight conditions).

Continuing the example of the control circuit around the aircraft elevator, the desired command input would be to increase altitude by 1,000 ft. The human-pilot system would decide on a course of action, given the transfer function characteristics, inside of the human-pilot system loop, and would then produce a control action, c, which feeds into the system. Here, the pilot would decide to pull the yoke of the aircraft to move the elevators of the aircraft upward. The system then processes the control action, and produces an outcome, or movement, m. The aircraft would, given the control action, tilt upward and climb in elevation. If the outcome is not the exact desired outcome, such as the aircraft climbed to a higher elevation than desire, and error would read back into the system (altimeter would show the altitude is too high), and the control loop would continue onward.

Of particular interest in these closed-loops circuits are the effective gains and delays incurred while these inputs and errors are occurring. Delays can be defined as the amount of time it takes for the human-pilot system to recognize the error, and Gains can be defined as the aggressiveness of the response; that is, how quickly the operator attempts to null the error.. Understanding the gains and delays of the human-pilot system are expected to be indicative of particular human-machine interface problems, such as situational awareness deficiency and reduced vigilance. These metrics may also help to identify periods of high or low workload in the cockpit [Ramadge, Wonham, 1987].

APPROACH TO FIND HUMAN-MACHINE INTERFACE ISSUES (METHODS)

In order to attempt to correlate variances in errors and the changes in the control parameters (gain and delay) in a single closed-loop event, it is valuable to determine several components for evaluation. First, the event must be defined, such as increasing aircraft altitude. Secondly, the event states must be defined, such as aircraft not climbing, aircraft climbing, etc. Next, the transitions between states must be identified. Finally, it must be determined which states are desirable and which states are undesirable. From this information, it is possible to evaluate what causes or influences transitions into undesirable states. Following the previous example of attaining desired altitude; a state model is defined in Figure 2. This state model is based on the assumption that the states of the system are closed over m, where m is defined as m=1 (¬ desired altitude) or m=2 (desired altitude) [Landry, 2010]. It is also assumed that the states of the system are closed over m.

This kind of state model can be applied to any closed-loop circuit that the human-pilot system enacts onto an aircraft. In the given example, the output of the function is m=2, the desired altitude. 1a is the current state of the system, or the current action of the airplane, and state 1a1' are the future states of the system. The importance of these state models is to help determine which states are undesirable, and then these states can be further evaluated to potentially determine what factors cause the transition from a desirable state into an undesirable state. In this example, an undesirable state is one which does not directly lead to reaching the output goal state, m=2. These states would be { 1a1, 1a3 1a1', 1a'2a, 1a'4a} as these states do not sufficiently lead to from m=1 to m=2 without transitioning to another state [Landry, 2010].

From this perspective, it is interesting to look at what behaviors or factors cause the system to be in undesirable states, and what behaviors or factors cause transitions from desirable states to undesirable states. Regarding the given example, it would be valuable to look at what causes the system to go from 1a2, the aircraft is descending or climbing to reach its desired altitude, to 1a1, the aircraft is not climbing or descending but has not reached its desired altitude.

The McRuer crossover model, which was previously developed, extracts continuous delays and gains, where the pilot-aircraft system is modeled as an integer of error with gains and delays. However, in the last 30 years, flight characteristics have changed, and large portions of all flights are flown using autopilot. Autopilot behavior is mostly uninteresting, as changes to the aircraft behavior are typically small and fairly routine throughout the flight, unless initiated by the pilot. However, there are parameters associated with the pilot's discrete control behavior, such as the pilot's selection of autopilot settings or modes, which is interesting, and defining the behaviors and factors which result in undesirable autopilot settings or mode selections could provide valuable insight into human-machine interface issues. Tracking this discrete behavior, however, is challenging, as there is no McRuer crossover equivalent for tracking discrete control systems.

DEVELOPMENT OF DISCRETE CONTROL

A discrete event system is a dynamic system that evolves by way of abrupt occurrences of physical events at unknown intervals. The large goal of discrete control attempts to disable events that cause undesirable transitions, and enable events that cause desirable transitions [Wickens, 2008]. That is, try to disable the human-pilot system from performing the undesirable transition from state 1a2 to 1a1, and put controls in place to help enable the desirable transition from 1a1 to 1a2. A discrete control model can be theorized as follows:

$$G = (Q, \Sigma, \delta, q_0, Q_M)$$

$where$:
$Q = set\ of\ states$
$\Sigma = set\ of\ events$
$\delta = transition\ function\ between\ states\ of\ given\ events$
$q_o = initial\ state$
$Q_M = subset\ of$ marker $states\ of\ interest$

In this model, G can be thought of as a "device that starts it its initial state," or in this case, G is the aircraft or the part of the aircraft in which control is being applied to achieve a desired state [Wickens, 2008].

It is also of value, from a human factors perspective, to track the delay in the event occurrence. That is, it is interesting to track how long it takes for the human-pilot system to recognize that there is an error in the system (delay, τ), and it is also worthwhile to note how long it takes the human-pilot system to apply control to the system to correct the error (gain, k).

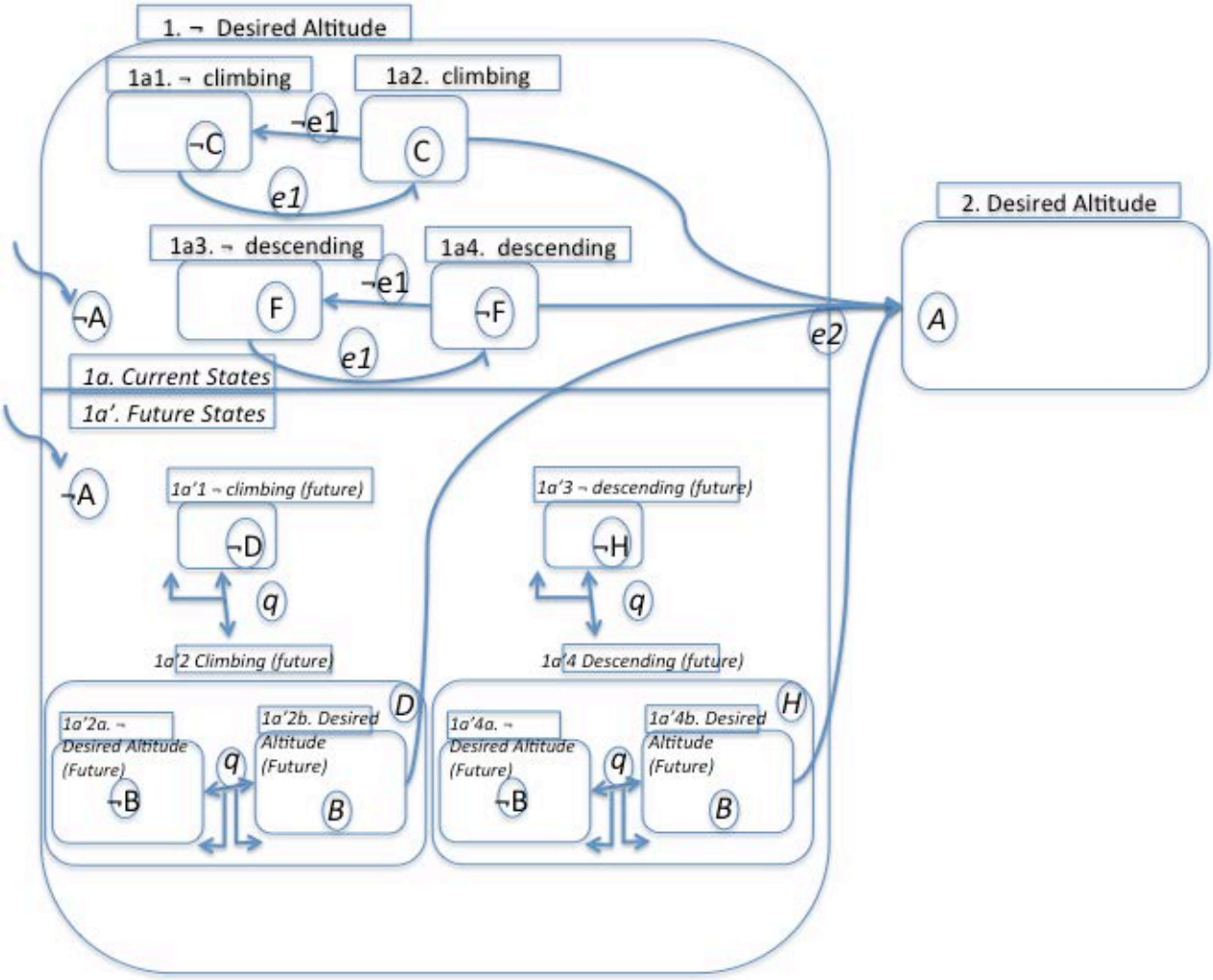


Figure 2 : State Model for Desired Altitude

$A: Altitude = Desired\ Altitude$
$B: Altitude(t_{future}) = Desired\ Altitude$
$C: Rate\ of\ Climb > 0$
$D: Altitude(t_{future}) < Desired\ Altitude$
$F: Rate\ of\ Climb < 0$
$H: Altitude(t_{future}) < Desired\ Altitude$

$e1 = t_{\substack{climb \\ descend}^{+}} < t < t_{\substack{climb \\ descend}^{-}}$
$e2: t = t_{desired\ altitude}$

q : 4D Trajectory change or "control applied"

APPLICATION OF THEORIES TO DETECT FAULTS

In order to determine if discrete control theory will help to answer the human-factors issues that arise in transitioning between desirable and undesirable states, several questions will attempt to be answered.

The first question is to determine if error variance correlates with workload. In order to test this theory, a human-subject experiment will be necessary. Human subjects (pilots) will be placed in a simulated flight environment where flight conditions will be presented under various conditions of workload. Workload conditions will vary according to multiple resource theory, where some tasks will come up unexpectedly when the pilot has additional resources available, and some tasks will come up unexpectedly where the demand for resources will exceed what is available [Wickens, 2008] [McRuer 1980]. Error variances will be then compared with NASA-TLX models to determine if they are correlated.

The next question is to determine whether control mode can be reliably identified. This will also require a human experiment where subjects are placed in a simulated environment where flight is presented under various conditions of workload, where workload expectations continue to be established by multiple resource theory. In this experiment, under various unexpected situations, the subject will have the method of control applied to correct a system error recorded, as well as have eye movements tracked. After the experiment is over, subjects will be interviewed and asked to report their control strategy. Although this data will be subjective, it will help to determine if the control method reported by the human subject and the apparent control method applied (determined by observing actions taken and eye movements) seem to match up appropriately.

Thirdly, it is invaluable to determine if delay and gains can predict human-machine interaction problems. A preliminary experiment to potentially prove or disprove this theory has been developed, where the subject is given a simple tracking task (a blue bar moves away from the center of the screen at random intervals) and the subject must use the mouse or tracking device to correct it back to the center. The amount of time it takes for the subject to initially begin moving the bar back to center is recorded, as well as the amount of time it takes for the bar to reach the center once the initial control is applied. This model is the first step in building a system that could track the gain and delay of a pilot correcting minor errors that present throughout a flight. While the current version of the experiment does not have human factors controls, such as the subject being subjected to a high workload or a low workload while tracing the bar, these may be implemented in the future to determine if there is any correlation between the values of gain and delay and the amount of workload. This tracking task is continues to be in-development, and key hurdles include an over-fit model which is generally not realistic due to these parameters being able to change freely, where in reality, gain and delay may change infrequently, but not necessarily slowly. In order to attempt to combat this, the model is being fit using one gain/one delay for a small portion where the fit is poor, and another gain/delay value is being used to fit the next portion of the model, etc. This method is being employed in hopes of creating a pieced-together but overall well-fit model for tracking gain and delay.

Finally, it is important to answer if the delays and control discrepancies can be clearly identified. While this question's answer is the least developed, it is essentially the final step in determining if human-machine interface issues and human-factors problems can explain gain and delays in the human-pilot-aircraft control loop, and determine if paths from desired states to undesired states can be disabled. Methodology for answering this question will contain modeling open-loop flights first, and then closed-loop flights to compare what kind delays are incurred and which control measures are applied, but full determination of testing theories is beyond the scope of this paper.

CONCLUSION

When using a modified version of the McRuer crossover method to model the human-pilot and aircraft as a closed-loop system, it becomes potentially possible to determine various desired and undesired states in a given physical event during flight. When these states are evaluated, it may be possible to enable desired states and disable undesired states through discrete control theory. Unfortunately, no discrete control theory equivalent of the McRuer model exists, but development and capture of the gains and delays which occur in the system during an event when an error is presented could help to show human-machine interface issues which are prevalent. The understanding of this could lead to measures to disable undesired events and also help disable transitions from desired events to undesired events. In order to do this, a tracking task is being developed which is intended to be applied to a human-

subject (pilots) simulation to evaluate the levels of gain and delay. These values may help understand human factors issues which occur on the flight deck and how this affects the flight itself.

REFERENCES

Landry, S.J. , Lagu, A. Kinnari, J. (2010). "*State-based modeling of continuous human integrated systems: An application to air traffic separation assurance."* Reliability Engineering and System Safety. Vol 95. Issue 4. April 2010. Pg 345-353.

McRuer, D. (1973). "Development of Pilot-in-the-Loop Analysis." AIAA Guidance and Control Conference, Stanford, CA, 1973.

McRuer, D. (1980). "*Human Dynamics in Man-Machine Systems."* Automatica. Vol. 16, Issue 3. May 1980, pg. 237-253

McRuer, D. Krendel, ES. (1974}. "*Mathematical Models of Human Pilot Behavior."* Advisory Group for Aerospace Research, France. Pg 25-39.

Ramadge, P. J., & Wonham, W. M. (1987). Supevisory control of a class of discrete event processes. *SIAM Journal of Control and Optimization, 25*, 206-230.

Wickens, CD. (2008). "*Multiple Resources and Mental Workload."* Human Factors: The Journal of Human Factors and Ergonomics Society. June, 2008, vol. 50, no.3. pg 449-455.

How Can Procedures and Checklists Help Pilots in Abnormal Flight Situations?

Andreas Haslbeck[1], Patrick Gontar[1], and Ekkehart Schubert[2]

[1]Institute of Ergonomics
Technische Universität München
Boltzmannstrasse 15, 85747 Garching, Germany

[2]Section Flight Guidance and Air Transportation
Technische Universität Berlin
Marchstrasse 12-14, 10587 Berlin, Germany

ABSTRACT

This paper introduces a flight simulator study to evaluate procedures and checklists for use in abnormal situations. These utilities are meant to support pilots in hazardous situations, but for extreme situations, they might be inappropriate. Sixty crews (A340, A320) were required to fly in an abnormal flight scenario with a technical defect (loss of one main hydraulic system), which leads to several subsequent events. The aircrews had to mitigate these situations in order to safely land. The decreasing remaining fuel intensified time pressure while overloaded procedures had to be performed. In this study, the provided procedures were tested in this simulated crisis scenario.

Keywords: flight simulator scenario, abnormals, standard operation procedures

INTRODUCTION

Many checklists (digital and paper-based) and also standard operation procedures (SOPs) have been established to help pilots act in abnormal, often critical situations (simply called abnormals in aviation). These tools have to be as accurate and comprehensive as possible but also short and easy to handle. Thus, aircraft manufacturers and airlines have to manage a tradeoff between these opposing goals. As procedures are typically developed by expert pilots and technicians, and also by lawyers, a crucial question is whether these checklist tools are partially polluted by items whose sole purpose is to relieve the manufacturers and airlines of legal responsibility in case of an accident. Another issue is the context a procedure was made. It might happen that these checklists are not suitable for all practical situations and contexts, for example during fuel emergencies or adverse weather conditions. Furthermore, recommendations stated by accident reports led to stretched procedures. These additional guidelines are important to avoid future incidents, however, these items frequently lengthen the affected checklists.

An abnormal situation in aviation is frequently defined as the opposite of normal operation and more serious than non-normal situations. When analyzing such abnormals, they frequently turn out as highly complex and unfamiliar to the pilots. These circumstances normally lead to frequent human/operator errors (VDI, 2003) and sometimes time pressure often emerges as an additional hazard, especially in the approach and landing phase. A well-known accident is US Airways flight 1549, which ditched into the Hudson River. Cpt. Sullenberger and his crew did not have much time and were losing thrust from both engines. After the birdstrike, which resulted in the rapid loss of thrust, the crew had about three minutes to reach the ground – in this case, the Hudson River off midtown Manhattan. In these short moments, they were not able to properly complete procedures and checklists, because the flight crew also had to communicate with air traffic control (ATC), control the airplane, try to restart the engines and to find a spot to land. This situation can be described as an abnormal with high time pressure. Another abnormal with very different circumstances was an uncontained engine failure of an Airbus A380 known as Qantas Flight 32 on the 4th November 2010. In this incident, a second aircrew was present and aided the primary aircrew. Additionally, the aircraft had a

sufficient amount of fuel remaining and therefore no time pressure was involved. Despite these factors, the pilots also decided to not follow through with all the required procedures.

This leads to one problem of recent recurrent training in aviation: There is a strong focus on scenarios which are partly known to pilots beforehand and which offer a rather standardized way to accomplish (Casner, Geven, & Williams, 2013). This approach enables one to learn how to solve such a scenario and delivers objective and standardized evaluation criteria for instructors and check-pilots. In contrast, this kind of scenario-based training does not promote the ability to handle unknown hazardous situations. Therefore, this paper introduces a study to evaluate the use of abnormal checklists and procedures in an unknown and time critical scenario.

PREVIOUS STUDIES EVALUATING PILOTS' BEHAVIOR IN ABNORMAL SITUATIONS

Nikolic and Sarter (2007) conducted a flight simulator study to observe and evaluate pilots' strategies for diagnosing and recovering from abnormals. Their twelve Boeing 747 pilots were required to handle three different abnormal events: an unexpected and difficult climbing clearance, and two different issues of mode awareness. The authors did not find any significant differences in the captains' performance compared to the results of the first officers. Nevertheless, all pilots struggled with the given events at some point in the scenario. Further results showed a broad variety of recovery methods and nearly all pilots failed to diagnose the situation properly. Ebbatson, Harris, Huddlestone, and Sears (2010) combined a manual flying task with an engine failure (single engine ILS with backing crosswind) to evaluate pilots' performance. Some might not consider flying an aircraft manually as an abnormal, but recent accidents (e.g. Colgan Air 3407, Air France 447, Asiana 214) have powered a debate about the influence of automation on manual flying skills, which led to official attention (Federal Aviation Administration, 2013). Under manual flying circumstances, which are also associated with higher workload, an evaluation of manual handling skills is a valid approach for abnormals. Ebbatson et al. (2010) found an influence of short-term flight experience on manual flying skills and control input strategies. Casner et al. (2013) faced 18 Boeing 747 pilots with different abnormal situations (stall, wind shear and engine failure) under two different experimental conditions: under known conditions like airline training and under unexpected conditions. The results seen in this study imply that pilots can cope with known and frequently trained abnormals, but they have difficulty in unknown or unforeseen events. The outcome of the flown scenarios showed a wider spread of actions, decisions, and flight performance than someone might anticipate in recurrent training. Their conclusions on this study represent argue for improving pilots' training with surprising abnormal scenarios.

These studies describe the different problems today's pilots are faced with. In sum, aircrews show weaknesses in diagnosing unexpected events and such surprising scenarios are very rare in pilot's training. Another issue is the use of automation and consequential skill degradation. One important guideline for operators in safety critical domains is that under normal conditions they must adhere to the given SOPs. They should only deviate from SOPs when this behavior seems to provide a safer consequence to their actions. This principle is also valid for pilots. Thus, an experiment evaluating procedures and pilots should consider both situations: the need to adhere to SOPs and the need to dismiss SOPs. One approach to create a situation where a pilot needs to deviate from SOPs is to significantly reduce the time budget. In aviation, the time budget and the remaining fuel strongly correlate with each other.

METHOD

Research questions

The main research question (1) of this study is whether the provided procedures and checklists adequately support pilots in abnormal situations. This question addresses two different issues: whether these procedures are appropriate for safe flight conduction in relevant situations and whether the provided material (paper-based or electronic checklists) is adequately usable for use in abnormal situations? The first aspect is addressed in this study and the latter aspect is not directly addressed, but corresponding observations are documented. The authors hypothesize that the provided checklists and procedures are not suitable for time-critical situations with abnormal events like technical problems of an aircraft.

A corresponding research question (2) is whether the level of practice and training or the experience of pilots has an effect on pilots' capability to successfully manage a time-critical scenario during an abnormal event. It can be as-

sumed that experienced pilots make better decisions than pilots with less experience. The opposite may be true for manual flying skills where pilots with a high level of practice and training are assumed to perform better than pilots with low levels. Pilots' level of practice and training is negatively correlated with their experience (Haslbeck, Schubert et al., 2012). While first officers on the short-haul only have little experience, they are well practiced and trained because of their frequent flight operations. On the other hand, captains on the long-haul are very experienced, but since they conduct less flight operations in the same time period, they suffer from a rather low level of practice and training.

Operationalization and measurement

In the following section, this paper introduces an experimental design to evaluate the usefulness of some abnormal procedures and checklists. This experiment was performed in cooperation with a major European airline within two qualified full-flight simulators (JAR-FSTD A) and professional airline pilots as participants.

According to the stated research question addressing the handling of checklists (1), eye-tracking data are expected to show the direct interaction with these lists and represent the data acquisition of the relevant information. After perceiving and processing environmental information, communication is the next step in pilots' information processing to maintain a shared or team mental model. Helmreich and Sexton (2004) found differences in the use of first person plural between captains and first officers. Other indicators are the amount of spoken words and density of spoken words of the two crew members and the distribution between them. Furthermore, the content of the spoken speech acts are of concern. Frequently asked questions or speech acts of dissent for instance may indicate an ambiguous implementation of checklists in cockpit. Therefore, all audio channels have to be recorded separately and synchronously. This allows a separate analyzing for both pilots later on. Non-verbal communication has to be recorded by video cameras as well. Also here, when analyzing shared team mental models, the gaze behavior of both pilots has to be recorded. When performing the different checklist actions, pilots on one hand interact with each other but just as well with the aircraft. This interaction can be observed on video data basis and even more precisely when recording the aircraft parameter like certain switches' position or control inputs – to name only a few. The flight parameters additionally help to answer the further research questions such as manual flying qualities, which might become a concern later on.

According to the described operationalization, the following measurement should be incorporated: both pilots' gaze behavior is tracked by two synchronously recording eye tracking devices, three audio channels are recorded (CPT, FO, ATC) simultaneously, and flight parameters are recorded by the flight simulator data recorder. To ease the understanding of the pilots' interaction later on, cameras record from two different view angles. The level of practice and training can be easily assessed by the help of questionnaires.

Further requirements

To measure operators' performance in professional domains, Haslbeck, Schubert et al. (2012) have stated some requirements to be taken into account for finding variance in the experimental results. An important recommendation is the need for non-voluntary (randomized sample) participants to avoid a self-selection bias. One lesson learned from the abovementioned study is a distinct self-selection bias that occurs when selecting only volunteer participants, especially when there is a chance to fail the experiment. In the first announcement of this earlier experiment (Haslbeck, Schubert et al., 2012), pilots were asked to participate on a voluntary base. Afterwards, additional pilots were randomly chosen and mandated to participate by the respective airline. Some of the participants had also answered to the announcement for the volunteer participation and after a review of the results, it turned out that the original volunteer pilots were amongst the best rated participants. In addition, a realistic setting with pilots in uniform, real recorded communication from ATC and foreign aircraft were created for this study. This was done to maintain the appearance and feel of a real flight to evoke a pilot's typical behavior in flight.

Technical flight scenario

To incorporate the aspect of different levels of practice and training, the experiment has to be conducted on two different fleets: long and short haul. As usual for the fleets but also a foreign destination, the long haul crews had to perform a non-precision landing scenario at New York's John F. Kennedy (KJFK) airport in the USA in an Airbus A340-600 aircraft (with four engines); the short-haul crews (Airbus A320-200, with two engines) had to perform a similar non-precision approach at Nice Côte d'Azur (LFMN) airport in France (see Figure 1).

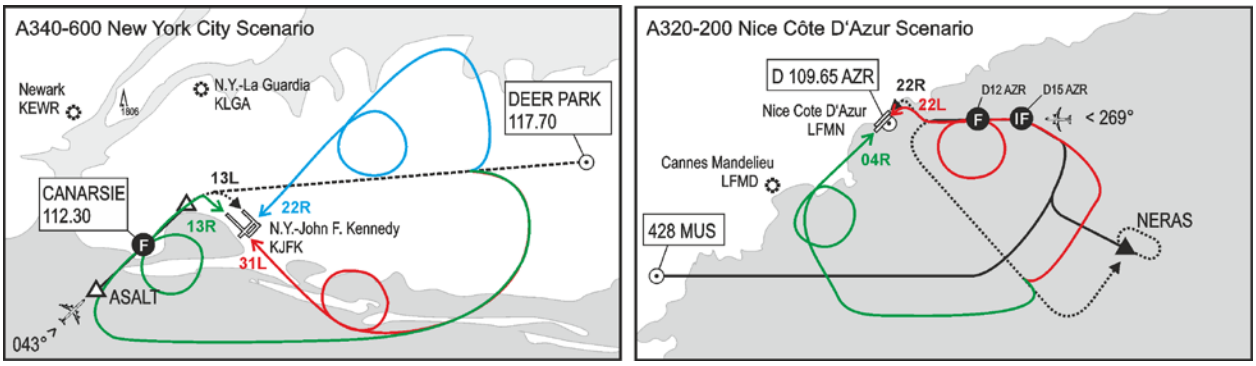


Figure 1. New York City (left) and Nice Côte D'Azur (right) flight scenario (with exemplary flight tracks)

The weather conditions in both scenarios (A320, A340) were moderate with steady wind from south (180°/10kt) and light rain (runways *slippery when wet*). The crew entered the simulator, which was already prepared for the approach (flight management system programmed for a non-precision approach, autopilot setup in *managed* mode and *altitude hold* mode, auto thrust in *speed* mode, navigational and radio setup prepared). The aircraft was already flying with activated position freeze (like a helicopter mode) and realistic ATC in the background. When the crew had taken place and completed their individual settings and adjustments, the scenario began with position freeze released. Fuel on board lead to a remaining flight time of approximately one hour (A340=7,500kg and A320=2,500kg); both scenarios started in 3,000ft altitude with the flaps in position two and appropriate speed (A340=190kt and A320=170kt). The crew then reduced speed for the approach and, with the gear lever selected down by the pilot monitoring, a malfunction of the green hydraulic system was invoked. The green hydraulic fluid was lost within ten seconds and the nose gear remains unlocked as a first consequence. The crew receives cautions from the affected hydraulic system and a master warning because of the unsafe landing gear (main gear deployed and nose gear retracted) after 30 seconds. A go-around had to be performed and the gear remained down because green hydraulic system was needed for retraction. In New York the A340 crew followed the standard missed approach procedure climbing 4,000ft to DEER PARK (see Figure 1 dotted line on left side); in Nice the A320 crew has turned left climbing 2,000ft for a hold at NERAS (see Figure 1 dotted line on right side). With gear down, the fuel consumption was two times higher than normal, which halves the remaining flight time (see Table. 1) and caused additional time pressure for the crew.

Table 1: Approximate time schedule for both scenarios (A340/A320)

Event	Time	Remaining Flight Time (from Fuel)
start of the scenario	0 min.	60 min. left
gear down, loss of green hydraulic	3 - 10 min.	57 - 50 min., with gear down shorten to 25 min. (<50%)
status cleared, ready for second approach	15 min.	20 - 15 min. left
slats/flaps jammed	17 - 19 min.	12 - 10 min. left
touch down, end of the scenario	30 min.	approx.. 5 min. left

With the well-known aviation principle *Aviate, Navigate, Communicate than Manage the Systems* the electronic centralized aircraft monitoring system (ECAM) for managing the systems came now into the crews' focus. They had to execute the necessary ECAM actions but also get deeper into the time consuming emergency checklist in the onboard *Quick Reference Handbook*. Under time pressure, because of the unsafe and also noisy gear with high fuel consumption, the crew had to decide which airport is appropriate for the upcoming emergency landing. Therefore, the inflight landing distance procedure (ILD) had to be performed. With only 20 minutes of fuel remaining, the crew had to declare an emergency at the latest with minimum diversion fuel with a commitment to stay (in New York a diversion to Newark is possible but also time consuming with no further improvement, same situation in Nice with Marseille or Genua). With four runways in KJFK (13R/31L, 13L/31R, 04L/22R, 04R/22L) and two in LFMN (04L/22R, 04R/22L) the crew had also to choose a safe kind of approach (precision with tail wind or non-precision with possible second go-around) and proper runway length, which fits best with the current situation (see Figure 1 for possible flight tracks and Figure 2 for airport map and available runway lengths). ATC assisted the crew with available information about weather and runways. The cabin crew asked for an announcement from the cockpit

to the passengers to improve the feeling of reality in such a situation. According to their airport and runway decision, a second approach was planned. Consider automatic terminal information service (ATIS), 13L and 22R in KJFK were available runways under the current wind and airport situation. Because of a displaced threshold on 22R, only 2,638m runway was available for landing but additional 821m could have been used if the crew would have required full runway length (they had to undershoot the glideslope; see Figure 2 left side). Runway 31L with 3,428m (also with a displaced threshold and additional 994m) was offered by ATC with a slight tailwind component as well. Runway 13R with 3,800m was also offered with a second longer non-precision approach (CANARSIE). For Nice Cote D'Azur airport the crew could choose between a second non-precision approach on runway 22L (which is longer than 22R; see Figure 2) or a longer but precision approach on 04R with also a slight tailwind component.

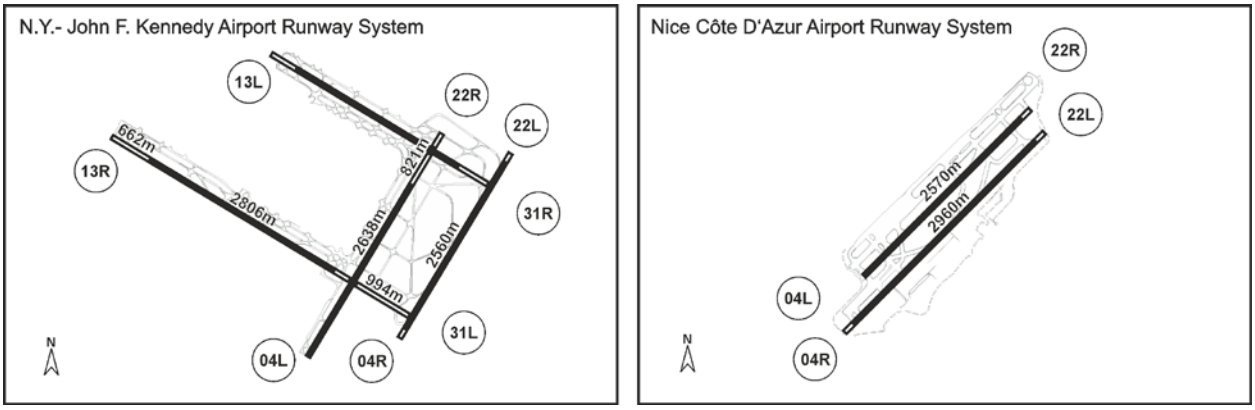


Figure 2. New York City (left) and Nice Côte D'Azur (right) airport map with runway system

A second sequenced malfunction was invoked when the crew was setting slats (or flaps if not in clean configuration): The missing green hydraulic system also slowed down slats and flaps movement so that an additional subsequent error occurred. With dissymmetrical slats and flaps movement, the wing tip brake came active and stopped the slats or flaps in a jammed position. This subsequent error with an occurrence probability of about 10^{-7} is irrevocable and can only be resolved by the maintenance on the ground. A second missed approach or, under increasing time pressure from the remaining fuel, a circling (as shown in Figure 1) had to be flown by the crew. After informing ATC, the crew had again to complete (clear) the ECAM system displays. In this situation, the crews decided whether to skip some steps from the QRH procedure and to shorten the second ILD checking or to try to sequence these procedures fully. The scenario ended after approximately 30 minutes (about 5 min remaining flight time; see Table 1) with touching down and aircraft full stop on the runway. If the crew ran out of fuel, the instructor interrupted the scenario shortly before, in order to prevent the participants leaving the experiment with a negative experience.

DISCUSSION AND OUTLOOK

The introduced experimental design is able to show the appropriateness of currently used checklists and potential improving possibilities. One question to address is whether the A340 and A320 scenarios can be compared to each other. When considering the usefulness of abnormal procedures and checklists on both fleets, the study can be regarded as two separate experiments, delivering results for both aircraft types. Both types have different technical procedures in part, but the level of difficulty of the scenario is very equal for both. The crews of both fleets are faced with the same problems: loss of green hydraulic and go-around, ECAM actions and ILD procedure under time pressure, slats/flaps problem during second approach, ECAM actions and ILD procedure under severe time pressure and landing with an abnormal high speed. The only significant difference between both scenarios is given by the two airports: Nice Cote D'Azur has only two runways and there is no additional difficulty by displayed thresholds. Therefore, the results from both fleets can also be compared to each other when thinking about pilots' performance (e.g. in communication) under different levels of practice and training (A340 crews vs. A320 crews). The authors of this paper believe that abnormal procedures and checklist provided will emerge as inappropriate in critical situations with time pressure. Either these utilities cannot be accomplished on time or won't be even started. The provided ILD procedure, supplied as a paper-copy will also lead to errors in use because of font sizes, attention errors or misinterpretations.

Further research questions refer to the influence of communication on the crews' mental models and overall performance. Regarding communication, the Crew Resource Management can be observed and evaluated by the instructor

or the pilots themselves (Gontar, Hoermann, Deischl, & Haslbeck, 2014). Besides the above mentioned main research questions, some further aspects of this scenario are to be recorded and analyzed afterwards. One important field of pilots' performance is manual flying. In the presented scenario pilots will be faced with manual flying in the normal manner at the end of the approach. Under abnormal conditions, it's the captain's duty to take over control and serve as the pilot flying. This can lead to a comparison of captains' manual flying skills between both fleets with some limitation, considering the different airports and aircraft types.

ACKNOWLEDGEMENTS

This work was funded by the Federal Ministry of Economics and Technology via the Project Management Agency for Aeronautics Research within the Federal Aeronautical Research Program. The authors want to thank CPT Manfred Müller, the retired Training CPTs Manfred Binder and Peter Cröniger as well as CPT Carsten Schmidt-Moll and Verena Porstner for their contribution to the project and their support during the experimental conduction.

REFERENCES

Casner, S. M., Geven, R. W., & Williams, K. T. (2013). The Effectiveness of Airline Pilot Training for Abnormal Events. Human Factors, 55(3), 477–485. doi:10.1177/0018720812466893

Ebbatson, M., Harris, D., Huddlestone, J., & Sears, R. (2010). The relationship between manual handling performance and recent flying experience in air transport pilots. Ergonomics, 53(2), 268–277.

Federal Aviation Administration. (2013). Safety Alert for Operators (No. SAFO 13002). Washington, DC.

Gontar, P., Hoermann, H.-J., Deischl, J., Haslbeck, A. (2014, in press). How pilots assess their non-technical performance – a flight simulator study. In *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics*.

Haslbeck, A., Schubert, E., Onnasch, L., Hüttig, G., Bubb, H., & Bengler, K. (2012). Manual flying skills under the influence of performance shaping factors. Work: A Journal of Prevention, Assessment and Rehabilitation, 41(Supplement 1/2012), 178–183. Retrieved from http://dx.doi.org/10.3233/WOR-2012-0153-178

Helmreich, R. L., & Sexton, J. B. (2004). Group interaction under threat and high workload. In T. M. Childress & R. Dietrich (Eds.), Group interaction in high risk environments (pp. 9–23). Aldershot: Ashgate.

Nikolic, M. I., & Sarter, N. B. (2007). Flight Deck Disturbance Management: A Simulator Study of Diagnosis and Recovery From Breakdowns in Pilot-Automation Coordination. Human Factors, 49(4), 553–563. doi:10.1518/001872007X215647

VDI guideline VDI 4006-2 (2003). Human reliability - Methods for quantitative assessment of human reliability. Berlin: Beuth Verlag.

Using System-wide Trust Theory to Analyze Passenger Loss of Trust in Aircraft Automation

Scott R. Winter, Stephen Rice, and Katie M. Reid

College of Aeronautics
Florida Institute of Technology
Melbourne, FL 32901-6975, USA

ABSTRACT

The current study focuses on airline passengers and how they might be affected by the contagion of an unreliable automated device. Participants took part in an online survey where they were presented with a scenario in which they were on an airline flight and oxygen masks erroneously dropped from the compartment above. They were then asked to rate their trust in that automated device, along with four other unrelated devices (auto-pilot system, flaps, landing gear, and video screen). In a control condition, no error occurred with the oxygen masks, and trust was high for all automated devices. In the experimental condition, trust was lower for all automated devices. In fact, in most cases, trust in the other automated devices dropped nearly as much as it did for the unreliable device. These results show that system-wide trust theory can make accurate predictions about the contagious effects of unreliable automated devices on otherwise unknown or reliable devices for non-expert airline passengers. This provides evidence that people tend to: a) view different automated components of an aircraft as part of a system even when they are logically independent of each other; and b) distrust other components of that "system" when one component fails.

Keywords: System-wide trust, trust in automation, consumer perceptions, airline passengers, reliability

INTRODUCTION

Background

Previous research has been conducted looking at operator's trust in automated devices while completing a task testing the system-wide trust (SWT) theory. The research looked primarily at how the operator interacted with one automated aid known as a single-device paradigm (Dixon & Wickens, 2006; Dixon, Wickens, & Chang, 2005, Dixon, Wickens, & McCarley, 2007; Geels et al., 2011, Keller & Rice, 2010; Lee & Moray, 1994; Parasuraman, Sheridan, & Wickens, 2000; Rice, 2009; Rice & Geels, 2010; Rice et al., 2008; Wickens & Colcombe, 2007; Wiegmann, Rich, & Zang, 2001). In more recent studies (Geels-Blair, Rice, & Schwark, 2013; Keller & Rice, 2010; Rice & Geels, 2010), there was a shift from looking at single-device paradigms to encompass multiple devices, and it revealed a system-wide trust effect. In other words, it was concluded that operators group multiple aids together into a single system, and when one device in the system is unreliable, the operator treats the entire system as unreliable.

This study seeks to examine consumer views toward automation trust using European participants via an online survey. The adverse effect found with multiple aids and operators is hypothesized to occur with *consumers* when

examining commercial airplanes. Instead of real-life experiences, participants were given scenarios to reveal a SWT effect in consumers. In one scenario, there was a failure of an automation system and the experimental group was asked to rate their trust in other, non-relevant automated aids. The control group received the same scenario, but it did not include the failure of an automation system.

A discussion on trust in automation will be examined in the following sections, and an explanation of system-wide trust will be provided. These sections will illustrate the need for the current study and offer more information about the implication of system-wide trust theory on consumer perceptions.

Automation

The definition of automation (Wickens & Hollands, 2000) is a mechanical or electrical task of work that would usually be completed by a human operator. The four stages of automation include synthesis, diagnosis, response selection, and response execution (Parasurman, Sheridan, & Wickens, 2000). When a task is unsuitable for a human to undertake, automation can be a valuable asset in assisting a human operator to complete a task. For example, when monitoring multiple gauges for an extended period of time, automation can help relieve some of the workload (Bainbridge, 1983). As convenient as automation can be, it is by no means a perfect system. When false alarms (FA) and misses occur, it can lead to mistrust in the automation system, or worse, completely ignoring the system.

When an automation system alerts an operator to a condition that does not actually exist, that is referred to as a false alarm. Alternatively, misses occur when the automation fails to notify the operator of a condition when there is an event. One or the other of these situations can impact the operator's trust in automation (Geels-Blair, Rice, & Schwark, 2013; Parasuraman & Riley, 1997; Rice, 2009; Rice & Geels, 2010). In regards to previous research, misses and false alarms have been seen to drag down operator trust in the device with the condition present, but also in unrelated automated devices.

Trust

In social psychology, trust is defined as the predictability of another person (Deutsch, 1958; Eckel & Wilson, 2004; Ergeneli, Saglam, & Metin, 2007), and through research it has been determined that this concept can be related to automation (Reeves & Nass, 1996; Rice, 2009; Parasuraman & Riley, 1997). Humans interact on a daily basis and with trust, humans can predict what a person is expected to do. A high level of trust equates to how a person believes that another person will do what is expected, and generally it will be a positive interaction (Lee & See, 2004). Human interactions can be extrapolated to include how humans interact with machines. In other words, trust can be interrelated to human-machine interactions (Geels-Blair, Rice, & Schwark, 2013). The current study was interested in how trust varies when devices fail or do not operate as anticipated by the consumer.

In the past few decades, there has been technological advancements and along with that, an increased use of automation. However, there is often an added level of complexity with the introduction of automation. Individuals use trust to help adjust to the complexity by augmenting supervision. This helps reduce uncertainty and therefore, reliance is guided by trust in automation. It is impractical to assume that a complete understanding in a situation is possible, but using trust and reliance, individuals can adapt to the situation and deal accordingly with automation (Lee & See, 2004). Automation failure is a key feature that researchers are eager to understand, especially in the context of multiple automated aids (Keller & Rice, 2010; Rice & Geels, 2010).

Previous research has examined how operator's trust level changes depending on the reliability of the automated aid. A positive relationship has been observed, which means that if the operator has a high level of trust, they will rely on the aid (Rice, 2009). Additionally, operators appear to view unrelated automated aids as one system (Geels-Blair, Rice, & Schwark, 2013; Keller & Rice, 2010; Rice & Geels, 2010) instead of treating an unreliable device on its own. This demonstrates that operators treat the entire system as unreliable and provides support of system-wide trust theory.

System-Wide Trust Theory

In earlier research, single-aid paradigms where the primary focus of automation trust (Maltz & Shinar, 2003; Meyer, 2001, 2004; Rice, 2009). This meant that operators were only paired with a single automated device. Keller and Rice (2010) embarked on a study to look at whether participants had component-specific or system-wide tendencies of trust when using multiple aids. In a component-specific perception, operators would differentially place their trust in different devices. To elaborate, an operator would be able to discern the reliability differences between two aids and

treat only the faulty aid as such and continue to trust the other reliable aid. The system-wide view occurs when an operator is unable to treat various aids as independent entities. In other words, if one of the aids is unreliable, then it adversely affects the level of trust the operator has towards the other unrelated, reliable aids.

To test these two theories, a simulated flight test was used. In the experiment, the operator was assigned to fly an unmanned aerial system (UAS) and check for system errors in two gauges that were positioned below the screen (Keller & Rice, 2010). In each condition, the percentage of reliability varied by 100%, 85%, or 75% reliable. It was solely up to the operator to agree or ignore the aid. It was apparent that by the end of the study, the data pointed to a system-wide trust of automation aids. The research showed that one unreliable aid could adversely affect the trust in the reliable aids (Keller & Rice, 2010). The limitation in this research study was the use of only two gauges, which caused the unreliable aid to represent 50% of the system. Another limitation included that there was not a large pool of participants, and all were from the subject university.

A later study included four gauges, which increased the ratio of perfect to imperfect aids from 1:1 to 3:1 (Rice & Geels, 2010). A single-task paradigm was used to help avoid confounding dual-task variables. Misses were examined in this study instead of false alarms, to see if the same effect would occur. By the end of the experiment, the results coincided with the earlier study. Even when the participants were given accurate and detailed information about the reliability of each automated aid, they still succumbed to the effects of system-wide trust. The same demographic limitation was present with a low participant pool as the previous experiment.

Geels, Rice, and Schwark (2013) continued research along these lines, but instead utilized eight gauges and compared the adverse effect that false alarms and misses had over the operator's trust. Of the eight gauges, one was either 100% or 70% reliable, and included either false alarms or misses. The data shows that regardless of the number of automated aids used in the experiment, participants still grouped aids together. These findings support system-wide trust theory. When comparing false alarms and misses, false alarms correlated with a stronger system-wide trust effect. This study had a larger sample pool, but the demographics of the participants were still limited to the subject university.

METHODS

Participants. Ninety-eight (30 females) participants from European countries took part in the study via an online survey. The mean age was 30.42 (*SD* = 9.57).

Materials and Recruitment. The study was presented online using FluidSurveys ®. Participants were recruited via Amazon's ® Mechanical Turk ® (MTurk). MTurk is a global online service that enables participants (Turkers) to participate in Human Intelligence Tasks (HITs) in exchange for monetary compensation. Participation in any HIT is voluntary and anonymous.

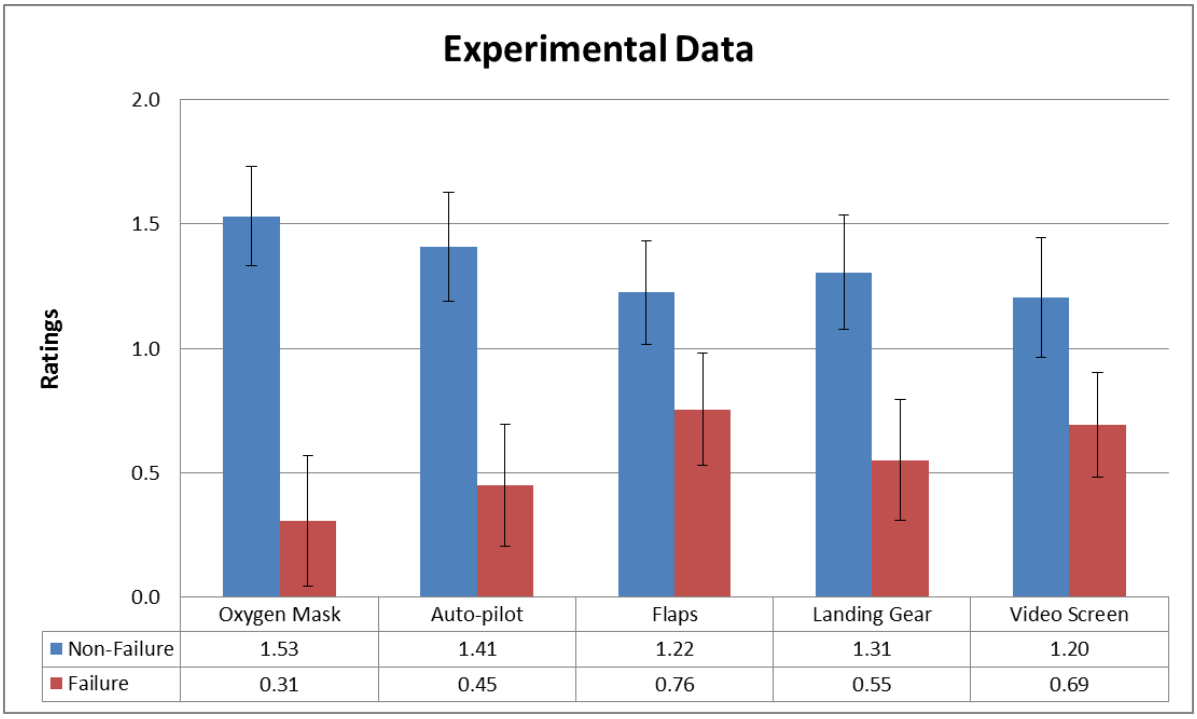
Procedure. Participants first signed an electronic consent form. They were then presented with the following scenario: "*Imagine that you are flying on a 4-hour commercial airplane flight from one major city to another. Sometime during the flight, an alarm goes off throughout the cabin and oxygen masks fall from the compartments above passenger seats. Following this, the pilot comes on the intercom and says that there was a mistake and the automation that operated the oxygen masks failed. He says that there is no actual emergency and not to worry. The pilot then tells you the altitude of flight and how long it will be before you land.*" In a separate control condition, participants were given the same scenario without the failed oxygen mask situation.

Participants were then asked to rate their trust in the automation that operates the oxygen masks, auto-pilot system, airplane flaps, landing gear, and seat video monitor on a Likert-type scale from -3 (extremely distrust) to +3 (extremely trust) with a neutral option of zero (neither trust nor distrust). Following this, participants were asked to provide demographic information and were then dismissed.

Design. A mixed design was employed with different participants in the experimental and control conditions, and all participants providing ratings for all the different automated devices.

RESULTS

Figure 1 presents the data from the study. A two-way 2 x 5 analysis of variance (ANOVA) was performed on the data, with FailureNonfailure being the between-participants condition and TypeofAutomation being the within-participants condition. There was a marginally significant interaction in the data, $F(4, 384) = 2.22$, $p = .066$, $\eta p^2 = .02$. The main effect of FailureNonfailure was significant, $F(1, 96) = 9.21$, $p = .003$, $\eta p^2 = .09$, indicating that trust in the automated devices were lower in the failure condition compared to the non-failure condition. The main effect of TypeofAutomation was not significant, $F(4, 384) = 0.07$, $p = .99$, $\eta p^2 = .001$.



	Oxygen Mask	Auto-pilot	Flaps	Landing Gear	Video Screen
Non-Failure	1.53	1.41	1.22	1.31	1.20
Failure	0.31	0.45	0.76	0.55	0.69

GENERAL DISCUSSION

The purpose of this study was to examine the effects of SWT theory on consumers in a commercial aviation setting. The unique piece of this experiment was the investigation of consumers instead of operators. Participants were from European countries, placed in either a control or experimental condition, and then were asked to rate their trust level in automated aids. In the experimental group, a failure of an automated aid was present, while this was restricted with the control group. The findings of the study support the hypothesis set forth by the authors that, like operators, consumers would follow a system-wide trust process.

The current study found that participants rated non-related automated aids as less reliable when paired with a failed automated device, and there was evidence that the SWT effect caused lower trust ratings across all automated devices. The main effect of FailureNonfailure was significant, which indicates that consumer's trust in automated devices was lower in the failure condition compared to the non-failure condition. This supports the finding that consumers gage their trust in the system as a whole. It would seem that consumers have a difficult time differentiating between separate, independent components when it comes to failures. Also, it is, to some extent, illogical for consumers to presume that the failed component is related to the other components. These results are uniform with the findings from previous research, which had focused on operators (Geels-Blair, Rice, & Schwark, 2013; Keller & Rice, 2010; Rice & Geels, 2010). These results are interesting since consumers do not usually have as much information as operators.

Practical Implications, Limitations, and Recommendations

The findings from this study pose some interesting practical implications. The data demonstrates that there was a reduction in consumer's trust level of automated devices. This concept of system-wide trust could greatly influence participants' views of the safety of commercial aircraft. If analyzed from a business perspective, then it may translate to consumers seeking the service of a different airline entirely. Every day the news is full of stories surrounding issues like the batteries found in Boeing's 787 Dreamliner. Does this make consumers less trusting of the other functions of the Dreamliner or of Boeing as a company? This is an area that requires more research to identify the impact that these issues may have on consumers.

Severity of the incident could likely influence this decision as well. The main limitation of the current study is that it was unable to be executed in a real-world setting. Given the findings from previous research concerning operators, it can be assumed that this particular limitation would not pose a serious threat to the findings of this study. Further research would be needed to check the validity of this assumption and the findings of this study.

CONCLUSIONS

System-wide trust theory suggests that individuals are unable to accurately differentiate between trust in failed and non-failed automated components. Earlier studies have examined trust in automation using both single and multiple aids. Initially, research focused on the response of operators or those trained in the functionality of the systems. These studies all found that a system-wide trust view was present, and operators rated all automated devices as less trustworthy despite little logical explanation. In the current study, *consumers* were asked to rate their trust in automated components during a scenario where they were a passenger on a commercial aircraft flight. The findings of this study seem to replicate those of earlier studies. Passengers in the experimental group that received the failure of an automated device rated their trust in other unrelated automated aids as less than a control group that did not receive a failure. As with earlier research, these findings support that consumers are also likely to use a system-wide trust theory in their interpretation and trust of automated commercial aircraft components.

REFERENCES

Bainbridge, L. (1983). Ironies of automation. *Automatica, 19*, 775-779.

Dixon, S. R. & Wickens, C. D. (2006). Automation reliability in unmanned aerial vehicle control: A reliance-compliance model of automation dependence in high workload. *Human Factors, 48,* 474-486.

Dixon, S. R., Wickens, C. D., & Chang, D. (2005). Mission control of multiple unmanned aerial vehicles: A workload analysis. *Human Factors, 47*, 479-487.

Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2007). On the independence of compliance and reliance: Are automation false alarms worse than misses? *Human Factors, 49,* 564-572.

Deutsch, M. (1958). Trust and suspicion. *The Journal of Conflict Resolution, 2,* 265-279.

Eckel, C. C. & Wilson, R. K. (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization, 55,* 447-465.

Ergeneli, A., Saglam, G., & Metin, S. (2007). Psychological empowerment and its relationship to trust in immediate managers. *Journal of Business Research, 60,* 41-49.

Geels-Blair, K., Rice, S., & Schwark, J. (2013). Using system-wide trust theory to reveal the contagion effects of automation false alarms and misses on compliance and reliance in a simulated aviation task. *The International Journal of Aviation Psychology, 23*(3), 245-266, DOI: 10.1080/10508414.2013.799355

Keller, D. & Rice, S. (2010). System-wide versus component-specific trust using multiple aids. *The Journal of General Psychology, 137*(1), 114-128.

Lee, J. D. & Moray, N. (1994). Trust, self-confidence and operators' adaptation to automation. *International Journal of Human-Computer Studies, 40,* 153-184.

Lee, J. D. & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*, 50-80.

Maltz, M. & Shinar, D. (2003). New alternative methods in analyzing human behavior in cued target acquisition. *Human Factors, 45*(2), 281-295.

Meyer, J. (2001). Effects of warning validity and proximity on responses to warnings. *Human Factors, 43*, 563-572.

Parasuraman, R. & Riley, V. (1997). Humans and automation: Use, misuse, abuse. *Human Factors, 39*, 230-297.

Parasuraman, R., Sheridan, T., & Wickens, D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 30,* 286-297.

Reeves, B. & Nass, C. (1996). The media equation: How people treat computers, television, and new media like real people and places. *International Journal of Instructional Media, 33,* 19-36.

Rice, S. (2009). Examining single and multiple-process theories of trust in automation. *Journal of General Psychology, 136*(3), 303-319.

Rice, S. & Geels, K. (2010). Using system-wide trust theory to make predictions about dependence on four diagnostic aids. *The Journal of General Psychology, 137*(4)*,* 362-375.

Wickens, C. D. & Colcombe, A. (2007). Dual-task performance consequences of imperfect alerting associated with a cockpit display of traffic information. *Human Factors, 49,* 839-850.

Wickens, C. D. & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Wiegmann, D. A., Rich, A., & Zhang, H. (2001). Automated diagnostic aids: The effects of aid reliability on users' trust and reliance. *Theoretical Issues in Ergonomics Science, 2*(4), 352-367.

Fuzzy Model of Air Traffic Controller Attitude to the Risk During Decision Making

Alexii M. Reva and Sergii P. Borsuk

National Aviation University
Kiev, prosp.Kosmonavta Komarova 1
Ukraine 03058

ABSTRACT

Air traffic controllers (as "first line" operators) professional activity may be described as a continuous chain of decisions, being developed and implemented under influence of different natural factors, including risks of stochastic and deterministic type. Thus, taking into account air traffic control (ATC) influence upon flight safety, problem solution of ATC danger estimation that may appear under conditions of aircraft flight levels separation failure is permanently urgent. For that purpose term set of linguistic variable "risk level" was formed. This set includes terms "very high", "high", "above common", "common", "below common", "low" "very low". Using modified Cooper-Harper scale and information gather method that may be described as "point on research interval" along with membership function construction procedure based on "supplementary matrices", fuzzy generic models of 132 ATC students attitude to 10-kilometer interval between aircrafts (for horizontal en-route flight in AC (Approach Control) TMA (Terminal Control Area) with ATC AS (automated system)) violation was received. This allows to place in correspondence quantitative and qualitative factors of flight conditions and to define human factor influence upon acceptance of distances between aircrafts. Since formation of initial term set includes modifier "very", use of procedures reverse to concentration and dilation may reduce term set to five terms in it. That is exact number of ICAO recommended quantity of danger identifiers. This allowed to transform correspondent membership functions. Received connections between quantitative and qualitative estimates of distances between aircrafts was basics for solving ICAO's "risk triangle" in response to "event danger" parameter.

Keywords: Flight Safety, Human Factor, Air Traffic Controllers, Operator Attitude, Risk Levels, Fuzzy Modeling

.

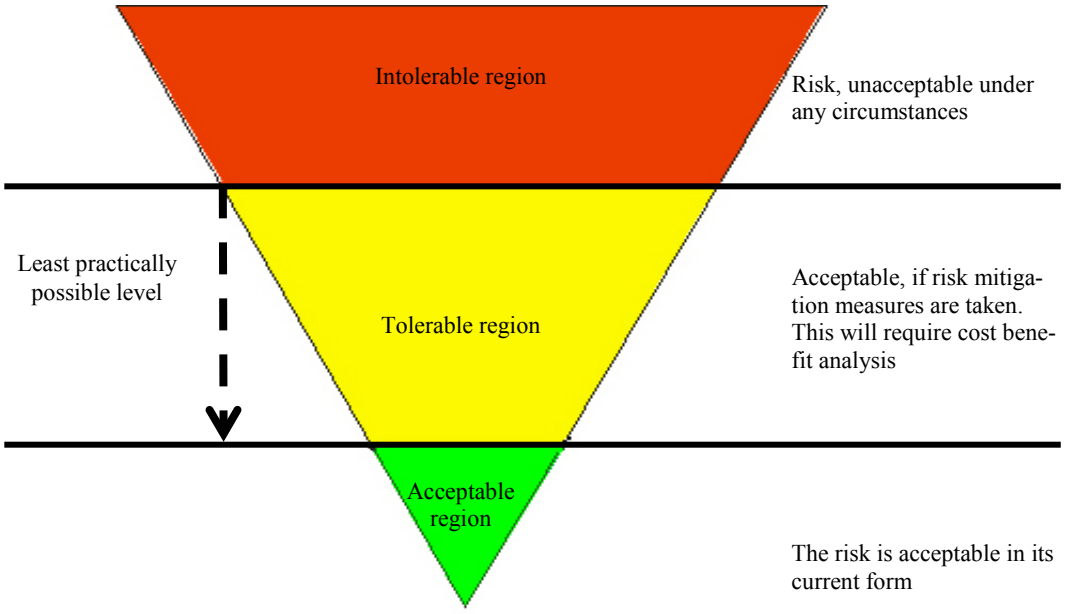
INTRODUCTION

Nowadays both positive and negative influence of human factor (HF) on flight safety (FS) is widely acknowledged (Reva, 2006). Aviation "first line" operators, namely pilots and ATCs play the main role in this process. That's why ICAO clearly recognized that ATCs, as air operators (AO) that execute specific proactive functions to control and overcome the threats, are "the last line of defense for minimizing the impact of threats on management of air traffic and support of the threshold levels of flight safety" (ICAO, 2008).

These FS threshold levels form ICAO's "Risk Triangle" (Fig.1) (ICAO, 2009), solving which has a great scientific and practical urgency and importance. The term "risk" in the context of our research (Petrovskiy and Roshevskiy, 1990) means the capacity of some decisions (and actions that result from these decisions) to worsen the situation and lead to potentially dangerous conditions that may turn into an aviation accident (AA) and result in unwanted consequences. Risk may also be defined as the probability (chance) of the current situation turning into AA. From this point of view, every decision taken by ATCs influences risk as probability by increasing or decreasing its value.

Table 1. ICAO's Threats Classification

Internal ATSP threats	External ATSP threats	Airborne threats	Environmental threats
1	2	3	4
Equipment	Airport Layout	Pilots	Weather
Workplace Factors	Navigational Aids	Aircraft Performance	Geographic conditions
Procedures R/T	Airspace Infrastructure/Design	Communication	-
Other Controllers	Adjacent Units	Air Traffic	-



Pic. 1. "ICAO's Risk Triangle"

There are almost no situations without actions, which can significantly increase the risk or immediately turn the situation, into AA. There are absolutely no situations without risk. Risk value may be vanishingly small and security may be based on several fault-tolerant systems but nevertheless risk always exists. It is also important to point out that there is a huge difference between situations where risk depends only on probability ("stochastic", chance based situations) and situations where risk depends only on the subject's capabilities ("habitual" situations). It was found that everything else being equal, AOs are more likely to risk in situations of "habitual" nature, but not in the chance-based ones. A possible explanation is that a person feels able to influence situations of "habitual" nature (ICAO, 2009; Kozeletskiy, 1975).

According to current ICAO paradigm of FS (ICAO, 2009), every AA is viewed not simply as an event caused by one specific reason but as an event whose causes could potentially be accumulated for a long time. Once a critical mass of reasons (regardless of their nature) is accumulated, the safety system reaches a bifurcation point, which leads to the destruction of the system and loss of its ability to respond to any kind of threat. In this situation, any action that should be parried by the safety system leads to an AA. This approach is also observed in Reason's model (Fig. 2) (Reason, 1990; Reason, 1998). It should be noted that such behavior is inherent to systems that have one or more bifurcation points while moving along phase space.

SOURCES ANALYSIS

The first mention of the "Risk Triangle" was made in the first edition of "Safety management manual (SMM)" (2006) without any recommendations about its solution. But in the same year, the solution was proposed by a Ukrainian scientist, professor Oleksiy Reva at the international ICAO, Airbus and MAK congress on the impact of human factor (HF) on FS. This solution was based on the specific points of utility estimating functions of AO's pro-

fessional activity (Fomenko, 2006; Reva, 2008).

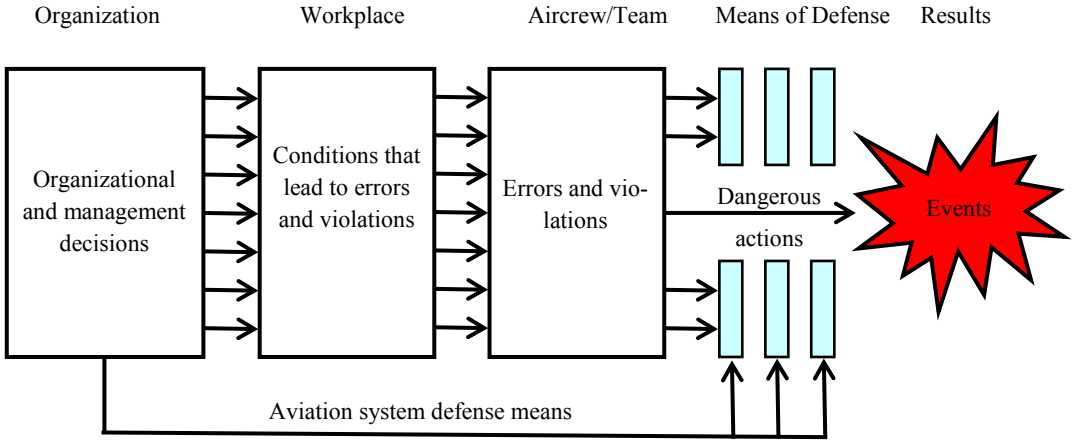


Fig. 2. Reasons model

All further research on solution of the "Risk Triangle" is connected with the next SMM edition (2009) and is summarized in the paper by Reva et al (2013). This includes ICAO's recommendations on qualitative solution of the "risk triangle" based on different combinations of qualitative linguistic indicators of the frequency of unwanted events with qualitative indexes of their danger (table 2). These recommendations were checked with the help of desirability coefficients of improved Harrington scale and respective Harrington functions. The paper by Reva et al (2013) found some important weaknesses of ICAO's proposals as it showed that if they were true, desirability coefficients of event frequency and danger must change in a linear way, which is nonsense from the point of view of FS threshold levels.

This paper also confirmed the possibility to solve the "risk triangle" with the help of key-points of ATC's professional activity utility estimation functions built with the help of special procedures for the closed (limited number of points) and open (theoretically unlimited number of points) decision making (DM) tasks (Reva, 1998; Reva, 2007; Reva and others, 2013).

The paper by Chinchenko (2011) states that qualitative characteristics of probability (chances) of unwanted event and its consequences significance may be given with the help of term-set (TS) of corresponding linguistic variables (LV), that have their own membership functions (MF) built for fuzzy argument "danger level" that has no strict quantitative definition. Moreover it is not shown how to get an aggregated safety estimate with the help of fuzzy operations based on the partial estimates, how to defuzzyficate it and normalize it to the ICAO risk categories transition.

The paper by Reva and Borsuk (2013), shows the research results on constructing the LV MF "danger level" for the distance between aircrafts of 10 km. This distance corresponds to the ICAO flight level norms for cases when air traffic control occurs for aircrafts on the same flight level (same path line or crossing path lines) using DME and/or GNSS. In this case, the forward aircraft must hold velocity, which exceeds that of the following aircraft by 27 km/h (20 knots) or more (ICAO, 2009). The size of correspondent TS is equal to 7 taking into account human psychophysiological capabilities on following and tracing limited number of objects, which is Miller's number (Kozeletskiy, 1975; Miller, 1956; Gerasimov and Kamishin, 2009). On the base of fuzzy concentration and dilation operations the initial size of the "risk of potentially conflict situation (PCS) appearance" LV TS was reduced to the ICAO requirements. The paper also provided defuzzyfication of the results and determined the distance norms that correspond to the ICAO's "risk triangle" threshold values.

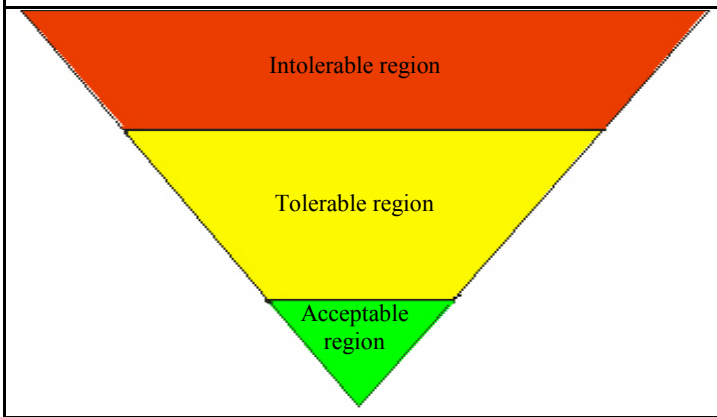
PROBLEM DEFINITION

From the above analysis it is clear that ICAO's "risk triangle" solution is possible with the help of building MF of corresponding LV as a fuzzy model of ATC's attitude to the violation of flight level norms. It is important that the

investigated value of 10 km. is not the only norm of air traffic flight levels. And it is obvious that violation of all these norms are unwanted states (potentially dangerous states), or such working conditions under which unpredicted air situation decreases FS threshold levels. Unwanted states that appeared as a result of ineffective control over threat factors or mistakes, can lead to the situations that threaten the whole flight or decrease FS threshold level of air traffic control.

Hence, proactive formation of ATC's knowledge about the possibility of unwanted states occurrence and identifying their personal attitude to their threat should become an effective way of preventing AA at air traffic control. This is the goal of this paper.

Table 2: ICAO's recommendations on "risk triangle" solution

Tolerability description	Index of risk assessment	Suggested criteria
	5A, 5B, 5C, 4A, 4B, 3A	Unacceptable under the existing circumstances
	5D, 5E, 4C, 4D, 4E, 3B, 3C, 3D, 2A, 2B, 2C	Acceptable if risk mitigation measures are implemented. It may require management decision.
	3E, 2D, 2E, 1A, 1B ,1C, 1D, 1E	Acceptable

Note:
Risk probability: 5 – «frequent», 4 – «occasional», 3 – «remote», 2 – «improbable», 1 – «extremely improbable»;
Risk severity: A – Catastrophic, B – hazardous, C – major, D – minor, E – negligible.

SOME PECULIARITIES OF ATC'S PROFESSIONAL ACTIVITIES

Peculiarities, stated above, include ensuring continuous process of ATC's decision making concerning control of the aircraft that ATC is responsible for. Taking into account the specificity of such control, ATC's decisions could be divided into strategic, tactical, operational and extreme ones (Blom and Bakket, 2011; Kantas and Maciejowski, 2011).

Strategic decisions influence general en-route flight process and, as a rule, they are taken in advance, at least one hour before the DM point to the point of its highest influence on the flight process and safety. They don't increase the AO's stress level or the risk of emergency situations in the ATC's responsibility area.

Tactical decisions are taken by ATCs in the air traffic control process with time reserve from an hour to 15 minutes, and have a significant influence on the flight process only for a small time interval. As a rule, they do not lead to a significant increase in the ATC's or pilot's stress level or to the risk of emergency situations in the ATC's responsibility area.

Operational decisions are taken by ATCs in air traffic control process with the time reserve from 15 to 5 minutes both during ordinary situations as well as emergencies or situations that can lead to emergencies in a short timeframe. These decisions are associated with a significant chance of potential conflict and require ATCs to precisely estimate potential safety, current safety and alternative management control influences.

Extreme decisions are taken by ATCs within the time interval of 5 minutes or less. These decisions are associated with huge chances of not only potential conflict but also potential disaster and require ATCs to quickly estimate potential safety and alternative management control influence actions.

During DM, ATCs continuously estimate not only the chances of air emergencies but also any changes in the current

situation as a result of the chosen action. This risk estimate is one of the most important components of the ATC's work and its results directly influence general flight process safety. ATCs as any other human beings tend to estimate risks in a fuzzy way ("big risk", "small risk", etc.). Such estimation is best described in terms of fuzzy sets and LV, which appeared in the middle of the previous century, and is continuously used in humanistic systems sciences (according to Zadeh (Zadeh, 1976)), ergonomics, cybernetic, and others where human beings takes part. Moreover, we believe that the methods of the possibility theory, which is based on fuzzy mathematics, should receive more attention for estimation of ATC's risk. The possibility theory is based on the foundations of the probability theory, the theory of fuzzy sets, as well as on the dependency of probability of the event's occurrence on the subject's knowledge about that event (Zadeh, 1976; Dubois and Prade, 1980; Dubois, Prade and Sandri, 1993; Dubois and Prade, 1995; Kaufman, 1982; Yager, 1986; Dubois and Prade, 1990; Pytiev, 2000].

Thus with the help of fuzzyfication/defuzzyfication procedures (Kamishin and others, 2012) of fuzzy notions it is easy to translate the notions easily understood by people into numbers and perform the necessary analysis. The results of this analysis can provide the foundation for construction of the fuzzy model of ATC's risk attitude.

RESEARCH RESULTS

Research was carried out in the National Aviation University (Kiev, Ukraine) and Kirovograd Flight Academy (Kirovograd. Ukraine). It involved 131 students of 4-th and 5-th years of study. Their attitude to risk was checked by special questionnaire(Cooper-Harper rating scale (Cooper 1957; Dobrolenskiy and others, 1975) modified according to research requirements) where they had to estimate risk value of PCS appearance during flight level rules violation using the following linguistic scale (Fig. 3):

$$T^{M}(risk\ level) = \overset{VH}{very\ high} + \overset{H}{high} + \overset{AA}{above\ average} + \overset{U}{usual\ (average)} + \overset{BA}{below\ average} + \overset{L}{low} + \overset{VL}{very\ low}$$

Where « + » stands for terms logical aggregation.

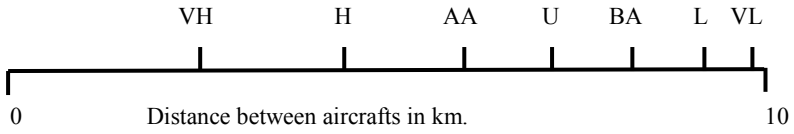


Fig. 3. Application of Cooper-Harper's scale for expert information collection about flight safety during flight level rules violation

As it was already mentioned, we used the following premises choosing scale dimension. It is well known that people have significant psycho-physiological limits as to the number of objects or ideas that they can pay attention to simultaneously. For the majority of human beings, this number ranges from 5±9 objects or ideas, which form the space of the so-called "Miller's magic number". Keeping them in fast memory, human-operator (H-O) is able to manage them, analyze, combine, and unite into clusters or groups. To determine the relationship between each measured value and all possible terms of LV's "risk level", the respondent has to compare his/her subjective opinion about the measured values (distance between aircrafts) with the subjective opinion about this value's belonging to a certain term, which provides a qualitative estimate about its risk level. At first, it seems that any number of LV terms may be taken for correct MF plotting but there is a reason to doubt this approach. Under the arbitrary number of LV terms greater than 9, ATCs as H-Os will face the problem of single distance correct comparison with all acceptable terms at the same time. In this case all comparing should be done rather in pairs or along with terms number subdivision into groups with further necessity to compare value with group terms. It is obvious that appearance of second polling phase will increase additional complexity. For example, some comparing will be executed as if separate group is whole TS diapason, leading to the mistakes. Using 7 terms (taking 7 as average number of objects efficiently kept by human in mind at the same time) person will be able to make one comparison of given value with all LV terms at the same time and define the best fit.

During the poll respondents were proposed segment of 100 mm. as an equivalent of 10-km. distance between aircrafts in the space, where respondents were asked to mark distances (Fig. 3) that corresponds to risk levels according to the scale.

Given flight level norm of 10 km. is defined by ICAO for the aircraft flights that follow same route on same level in

AC (Approach Control), TMA (Terminal Control Area) area using automated air traffic control system (ICAO, 2007). Respondent were asked to define risk levels for the distances. Given poll conditions and rules revealed great opportunities of firstly taking expert information with "point on the scale" scheme (Reva and others, 2010; Utkina and Kruchkov, 1988) and secondly use supplementary matrix (Reva and others, 2010; Borisov, Krumberg and Fyodorov, 1990). Results, received in this way are presented in the initial matrix (table 3).

Table 3: Initial poll results matrix

Risk level	Distance between aircrafts in km.									
	1	2	3	4	5	6	7	8	9	10
Very high	122	113	79	29	15	9	5	2	1	0
High	9	14	43	76	51	23	14	5	2	1
Above average	0	4	5	18	46	55	35	16	8	3
Usual	0	0	4	8	15	36	39	42	27	18
Below average	0	0	0	0	3	5	28	32	26	22
Low	0	0	0	0	0	2	8	29	44	23
Very low	0	0	0	0	1	1	2	5	23	64

For normalization supplementary matrix S was found in following way. Let a_{ij} be matrix element received as a result of polling $i=\overline{1,l}$, $j=\overline{1,m}$. Then we find $a_i^* = \sum_{j=1}^{m} a_{ij}$ and we get $s_i = \frac{\max(a^*)}{a_i^*}, \forall i$ where s_i is element of S.

$$S = (s_1, s_2, ..., s_i, ..., s_n)^T$$

Using supplementary matrix we find matrix B, where each element

$$b_{ij} = a_{ij} s_i = a_{ij} \frac{\max(a_i^*)}{a_i^*}$$

$$B = (1; 0.6347; 0.5067; 0.3093; 0.2827; 0.256)^T$$

Second normalization step is to find maximal number in every row and to multiply all row values into that maximal. Values of calculated "risk level" LV MF are given in table 4.

Table 4: "Risk level" LV MF values

Risk level	Distance									
	1	2	3	4	5	6	7	8	9	10
Very high	1	0,9262	0,6475	0,2377	0,123	0,0738	0,041	0,0164	0,0082	0
High	0,1184	0,1842	0,5658	1	0,6711	0,3026	0,1842	0,0658	0,0263	0,0132
Above average	0	0,0727	0,0909	0,3273	0,8364	1	0,6364	0,2909	0,1455	0,0545
Usual	0	0	0,0947	0,1894	0,3551	0,8523	1	0,9375	0,6392	0,4261
Below average	0	0	0	0	0,0938	0,1563	0,875	1	0,8125	0,6875
Low	0	0	0	0	0	0,0455	0,1818	0,6591	1	0,5227
Very low	0	0	0	0	0,0156	0,0156	0,0313	0,0781	0,3594	1

Figure 4 illustrate "risk level" LV MF as fuzzy model of ATC students attitude to the flight level rules violation. Lets analyze it starting from transition point proposed by L. Zadeh (Zadeh, 1976): $\mu_{\widetilde{T}_i}(s) = 0{,}5$. Now we consider that if any distance s between aircrafts in observed interval of $0 \div 10$ km. will have value of MF greater than transition point $\mu_{\widetilde{T}_i}(s) > 0{,}5$, then it's more likely to belong to corresponding term $\widetilde{T}_i$ and it's more likely not belong to corresponding term, if $\mu_{\widetilde{T}_i}(s) < 0{,}5$.

Hence, according to Fig. 4, we notice that all intersection points of neighboring terms has greater values than transition point.

$$\mu_A(s = 3{,}1\,km.) = 0{,}6;\ \mu_B(s = 4{,}8\,km.) = 0{,}75;\ \mu_C(s = 6{,}3\,km.) = 0{,}93;$$
$$\mu_D(s = 7{,}7\,km.) = 0{,}98;\ \mu_E(s = 8{,}6\,km.) = 0{,}90;\ \mu_E(s = 9{,}7\,km.) = 0{,}74.$$

This means that polled students as future professionals makes clear difference in degreases of observed flight level rules violation danger and operates with qualitative-quantitative estimates comparing. It is easy to find numerical values for intervals which correspond to qualitative linguistic estimates (table 5).

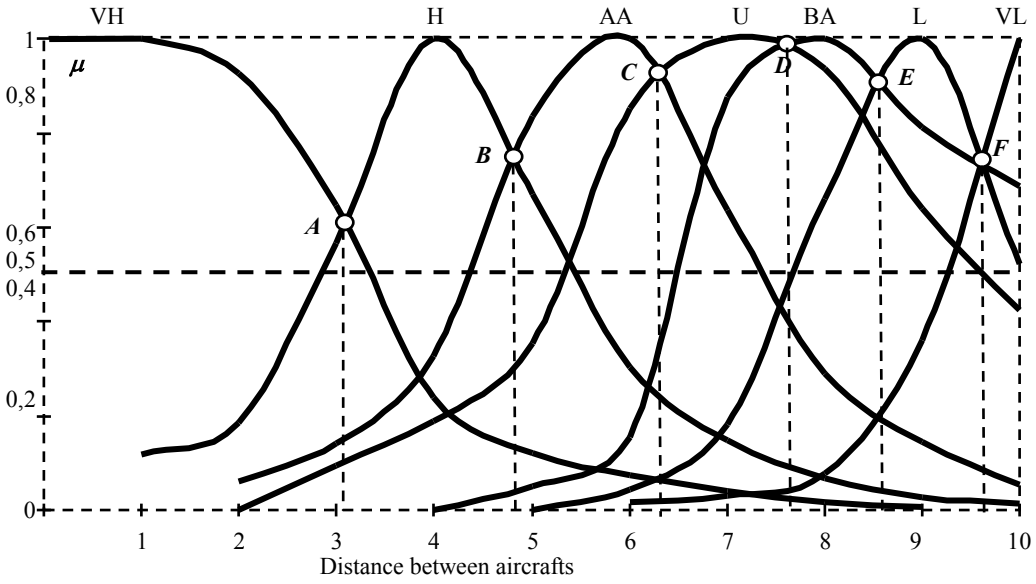


Fig. 4. "Risk level" LV MF during flight level rules violation

Table 5: Quantitative-qualitative differentiation of PCS occurrence risk

PCS occurrence risk	Distance between aircrafts, km.
Very high	$0 < S \le 3.1$
High	$3{,}1 < S \le 4{,}8$
Above average	$4{,}8 < S \le 6{,}3$
Usual	$6{,}3 < S \le 7{,}7$
Below average	$7{,}7 < S \le 8{,}6$
Low	$8{,}6 < S \le 9{,}7$
Very low	$9{,}7 < S$

Since ICAO recommends to use five qualitative levels of unwanted event danger and happening frequency (table 2) it is quite easy to present them as corresponding TS for danger and frequency:

$$T^M(risk\ probability) = frequent + occasional + remote + improbable + extremel\,improbable$$

$$T^M(risk\ severity) = catastrophic + hazardous + major + \min o + negligible$$

In context of current paper researches let's find out how to reduce scale used in polling to the ICAO scale size. Bearing on the fuzzy sets and LV theory (Zadeh, 1976; Kaufman, 1982; Yager, 1986; Reva and others, 2010; Utkina and Kruchkov, 1988; Borisov, Krumberg and Fyodorov, 1990) we will keep in mind the following.

Firstly during creation of "risk level" LV TS the "very" modifier was used, that may connect neighboring terms with help of fuzzy concentration and dilation operations. Thus, by inverting operations we may reduce initial polling scale by two terms.

$$\begin{cases} \mu_{(very\ high)}(S) = \mu^{2}_{high}(s) \\ \mu_{(very\ low)}(S) = \mu^{0.5}_{low}(s) \end{cases}$$

Secondly there exists such aggregation operation that allows to unite terms in following way:

$$\forall s \in S : \mu_{\widetilde{T}_k}(s) = \mu_{\widetilde{T}_i \cup \widetilde{T}_j} = MAX\left[\mu_{\widetilde{T}_i}(s), \mu_{\widetilde{T}_j}(s)\right]$$

Where $\widetilde{T}_i, \widetilde{T}_j$ are neighboring scale terms that are united, $\widetilde{T}_k$ is new scale term received as a result of $\widetilde{T}_i$ and $\widetilde{T}_j$ aggregation operation. Taking into account all mentioned above initial pole TS indeed may be reduced into ICAO TM from figure 5. Analysis of that figure shows that intersection points in this case has also values greater than L. Zadeh transition point, that allows us to switch to the ICAO "tolerability matrix" solution (Fig. 6). Apparently from figures 5 and 6 according to ATC students opinions catastrophic risk appears if flight level rules violation reaches 40% of norm distance $S = 10\,km.$ between aircrafts and all unacceptable risk zone reaches 63% of that distance. At the same time acceptable zone takes 37%. It would be interesting to compare results with other flight leveling norms and with other opinions namely from real ATC.

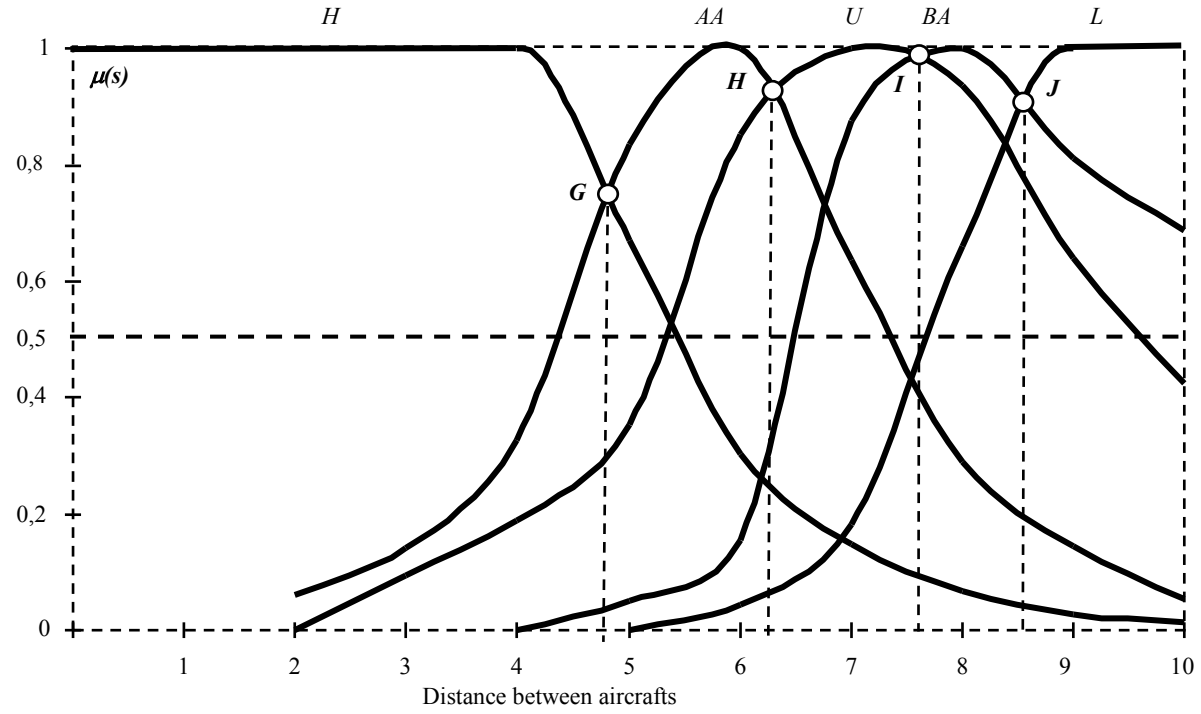


Fig. 5. Terms/levels risk distances after LV MF terms aggregation

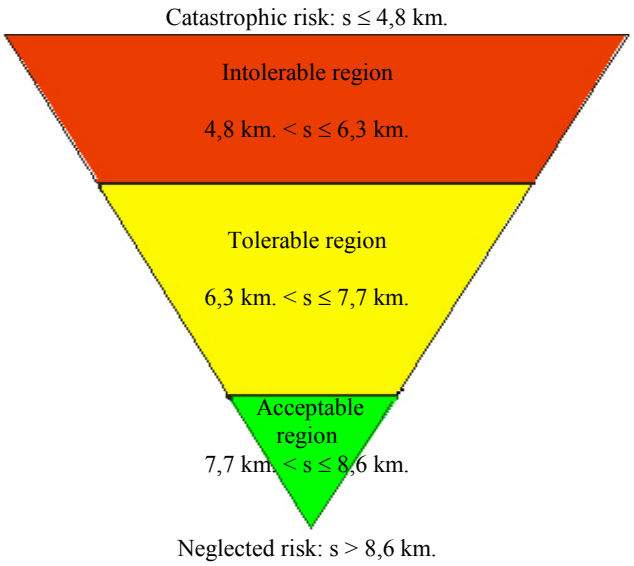


Fig. 6. "Tolerability matrix" solution for S=10 km. flight level rules violation

CONCLUSIONS

Taking into account scientific results achieved and presented in this paper we may state the fact of ICAO "tolerability matrix" solution with help of ATC students attitude to the aircraft flight level norm violation fuzzy models, that opens prospects of researching the whole family of such functions along total range of flight level norms. Special feature of given results that differs them from ICAO proposition is scientific base of quantitative relation between distance norm deviation and risk categories. Other results includes following.

Number of ranks in the qualitative linguistic scale used to estimate risk during flight level norm violation was grounded basing on psycho-physiological capabilities of H-O memorizing and recognition of limited number of observed objects as well as on fuzzy mathematics methodology. Plotted "risk level" LV MF during flight level norm violation of 10 km. It was discovered that all MF intersection points of neighboring terms has values greater than L. Zadeh transition point value that proves about good respondents understanding of correspondence between qualitative risk levels and quantitative distance values. PCS quantitative-qualitative risk norm is found. Taking into account ICAO requirements developed experimental scale was reduced with help of fuzzy operations like concentration, dilation and aggregation. MF received in this way allowed us to solve ICAO "tolerability matrix" in quantitative way.

Further researches should be held in following directions:

- receive MF family for all effective aircraft flight level norms to define regularity in ATC students opinions;
- receive MF family to define "normal" frequency of unwanted events occurrence during air traffic control and aggregate danger and frequency estimates;
- receive MF family from professional ATC and compare them with student-wise;
- carry out simulation in order to form ATC skills in conditions of flight level norms violations.

REFERENCES

Blom H. A. Safety of advanced airborne self separation under very high en-route traffic demand / H. A. Blom, G. J. Bakker // First SESAR Innovation Days, 29th November - 1st December 2011. - Toulouse, France, 2011. - P. 1-9.

Borisov A. N. Decision taking basing on fuzzy models: usage examples/ A. N. Borisov, O. A. Krumberg, I. P. Fyodorov. – Riga : Zinatne, 1990. - 184 p.

Chinchenko Y. V. Riaks and threats estimation on the air traffic controllers workplace with help of fuzzy sets / Y. V. Chinchenko // National aviation university bulletin: scientific magazine. / Ministry of science and education of Ukraine, National avia-

tion university – K.: NAU-print, 2011. - № 2. – P.44-49.
Cooper G. E. Understanding and interpreting pilot opinion / G. E. Cooper // Aeronautical Engineering Review, 1957. - N 3.- P.P. 47-51.
Dobrolenskiy Y. P. Methods of engineering-psychological researches in aviation / Y. P. Dobrolenskiy, N. D. Zavalova, V. A. Ponomarenko, V. A. Tuvaev; editorship. Y. P. Dobrolenskiy. – M. : Mechanical engineering, 1975.– 280 p.
Dubois D. Fuzzy sets and systems - Theory and applications. / D. Dubois, H. Prade; Academic press, New York, 1980. – 393 p.
Dubois D. On possibility/probability transformations / Didier Dubois, Henri Prade, Sandra Sandri // Proceedings of Fourth IFSA Conference, Kluwer Academic Publ. 1993. – p. 103-112.
Dubois D. Possibility theory as a basis for qualitative decision theory. / D. Dubois, H. Prade. // In Proceedings of the 14th International Joint Conference on Artificial Intelligence, Morgan Kaufmann, San Francisco, 1995 – p. 1924-1930.
Dubois D. Opportunities theory. Appliance to knowledge presentation in informatics: translated from French. / D. Dubois, A. Prade. – M. : Radio and communications, 1990. – 288 p.
Fomenko Y. M. Risk triangle in system analysis of ait traffic controllers professional activity / Y. M. Fomenko // Problems of informatization and control: scientific digest. / Ministry of science and education of Ukraine, National aviation university. – K.: NAU, 2006. - № 3. – P. 147-151.
Gerasimov B. M. Organizational ergonomics : Methods and algorithms of research and design: monograph/ B. M. Gerassimov. V. V. Kamishin. – K. : Informational systems, 2009. – 212 p.
ICAO: A unified framework for collision risk modelling in support of the manual on airspace planning methodology for the determination of separation minima: Cir. ICAO 319 – AN / 181 (Doc. 9689). – Montreal, Canada, 2009.
ICAO: Threat and Error Management (TEM) in Air Traffic Control : Cir. ICAO 314 – AN / 178. – Montreal, Canada, 2008
ICAO: Safety Management Manual (SMM): Doc. ICAO 9859 – AN / 474. – Montreal, Canada, 2009
ICAO: Air traffic management: Doc. ICAO 4444 – ATM / 501. – Montreal, Canada, fifteenth edition, 2007
Kamishin V.V. Estimation scales points fuzzification/defuzzification procedure / V. V. Kamishin, O.M. Reva, L.M. Makarenko, O. M. Medevedenko // Electronics and control systems: scientific magazine. / Ministry of science and education of Ukraine, National aviation university.– K.: NAU-print, 2012. – № 3 (33). – P. 53-62
Kantas N. Analysis of conflict resolution needs of the A3 operational concept. [iFly Deliverable D5.2.] / N. Kantas, J. M. Maciejowski, A. Lecchini- Visintini etc. - 2011. - 4l p.
Kaufman A. Introduction into fuzzy set theory: translated for French. V. B. Kuzmina/ A. Kaufman; editorship. S. I. Travkina. – M.: Radio and communications, 1982. – 432 p.
Kozeletskiy Y. Psychological solutions theory : translated from Polish. G. E. Mintz, V. N. Porus / Y. Kozeletskiy; editorship B. V. Biryukov. – M. : Progress, 1979. – 504 p.
Miller G. The magical number seven, plus or minus two : some limits on or capacity for processing information / G. Miller // Psychological Review, 1956. – N 63. – P. 81-97
Petrovskiy A. V. Psychology: Dictionary /general editorship. A. V. Petrovskogo, M.G. Roshevskogo. – P. : Politizdat, 1990. – 494 p.
Pytiev Y. P. Opportunity. Theoretical elements and appliance/ Y. P. Pytiev. - M. : Editorial YRSS, 2000. - 192 p.
Reason J. Human Error / J. Reason // Cambridge University Press. - Cambridge, 1990. – 41 p.
Reason J. Managing the Risks of Organizational Accidents / J. Reason; Ashgate Aldershot, 1998. – 266 p.
Reva O. M. Fuzzy models of flight precision ergonomical qualimetry : monograph / O. M. Reva, V. V. Kamishin, V. A. Shulgin, S. V. Nedbay; editorship. O. M. Reva. – Rivne: Ovid, 2010. - 106 p.
Reva A. N. Human factor and flight safety: (Research of proactive influence) : monograph / A. N. Reva, K. M. Tymishev, A. A. Bekmuhmambetov; scientific editorship A. N. Reva, K. M. Tymishev. – Almati, 2006.– 242 p.
Reva O. M. Proactive risk management on basis of human factor in civil aviation / O. M. Reva, S. I. Osadchiy, O. M. Medvedenko, Y. M, Fomenko // Railroad transport of Ukraine: theoretical and practical magazine, 2008. – № 6. – P. 54-59.
Reva A. N. Urgent area of ICAO “risk triangle” solution proactive models development / A. N. Reva, V. I. Vdovichenko, S. P. Borsuk [and others] // Modern transport power-plants, technologies and maintenance equipment: 4-th Ukrainian theoretical and practical conference. – Kherson, 9-11 october 2013, - Kherson: HDMA, 2013. – P. 334-338.
Reva O. M. Methods of aviation operator a priori attitude to risk as human’s that takes decisions: lectures conspect on course «Grounding of decision taking theory» / O.M. Reva. – Kirovograd: DLAU, 1998. – 45 p.
Reva O. M. Proactive estimate of aviation personnel attitude to risk and safe activities / O. M. Reva // National aviation university bulletin: scientific magazine. / Ministry of science and education of Ukraine, National aviation university. – K. : NAU, 2007. – № 2. – P. 36-42.
Reva O. M. Air traffic controllers pretension level with workload parameters / O. M. Reva, B. M. Mirzoev, P. S. Mukhtarov, S. S. Nasirov// Air-space technique and technology: theoretical practical magazine. / Ministry of science and education of Ukraine, National airspace university named after M. E. Zjukovskiy «HAI». – H., 2013. - № 8 (105). – P. 273-281.
Reva O. M. Fuzzy model of air traffic controller attitude to risk of potentially conflict situation happening / O. M. Reva, S. P. Borsuk // Ministry of science and education of Ukraine, National airspace university named after M. E. Zjukovskiy «HAI». – H., 2013. - № 10 (107). – P.214-221.
Utkina V.F. Reliability and efficiency in technique: reference book in 10 volumes. - V. 3 : Technical systems efficiency / general editorship V. F. Utkina, Y. V. Kruchkova. – M. : Mechanical engineering, 1988. – 328 p.
Yager R.R. Fuzzy sets and opportunity theory : last achievements: translated from English. V. B. Kuzmina; editorship. S. I. Travkina / editorship. R. R. Yager. – M. : Soviet radio, 1986. – 408 p.
Zadeh L. The concept of a linguistic variable and its application to approximate reasoning: translate from English. N. I. Ringo / K. Zadeh; editorship. N. N. Moiseeva, S. A. Orlovskogo. - M. : Mir, 1976. - 165 p.

Section 21

Aviation:

Human Factors Issues in Air Traffic Management III

Air Traffic Controller Workload Impact by Integration of Datalink System in an Advanced Airport Controller Working Position

Mohamed ELLEJMI, Stephane Dubuisson, Marc Bonnier and Roger Lane

EUROCONTROL Research Centre
Centre du bois des Bordes, CS41005
Brétigny Sur Orge, 91222, France

ABSTRACT

Integrating a new system in the air traffic control system is always a big challenge, in particular when the new system impacts the controller human machine interface (HMI). The Controller Pilot DataLink Communication system for airport controller is a new concept that aims to reduce the voice communication and increase safety at airport. To prove and evaluate the concept, an experiment took place at airbus in Toulouse last December 2013 with the participation of air traffic controller from Germany, Italy and France. In this paper we explain the concept, we detail the experiment, we discuss its impact on the controller mental workload, we determine the side effects and we provide some recommendation for future implementation.

Keywords: Controller Pilot DataLink Communication, Advanced Controller Working Position, safety, workload

INTRODUCTION

During the last decade, different tools were developed to ease the work of Air Traffic controller (ATCOs) to handle the increase on the traffic. Most of the time, when a new tool is introduced, new machines are added to the Tower Facility and the ATCOs have to handle more systems, more screens and more materials.

The project Advanced Controller Working Position, from the SESAR program[1], tackles this issue by proposing the integration of the current and future system in a concept called Advanced Controller Working Position (A-CWP). The project studies also the impact of the integration of the different systems on the ATCOs situational awareness, workload and efficiency.

Among the system integrated into the A-CWP, the Airport datalink function called D-TAXI. D-TAXI is the exchange of standardised messages between a mobile (either a vehicle or an aircraft) and the airport ATCOs as defined by the datalink standard under development jointly by the jointly by the RTCA SC-214 with EUROCAE WG-78.

The purpose of the D-TAXI is to reduce the voice communication between the Tower and mobile at the airport by exchanging non-time critical messages which should solve two major issues: ATCOs workload and Driver or Pilot clearance message understanding. Some messages depend on the existing of other A-SMGCS function such as the routing function.

[1] SESAR is a joint undertaking founded by the European commission and EUROCONTROL, the SESAR programme is building the future European air traffic management system

BACKGROUND

The first studies on D-TAXI focused on the technical feasibility and on-board guidance were conducted during the EMMA2 project[2] and provide a very important contribution to the understanding of the concept and some recommendation for the future studies.

SESAR Project 6.7.3 (A-SMGCS Guidance Function) works on D-TAXI based on the work started on EMMA2 and the one performed jointly by the RTCA SC-214 with EUROCAE WG-78 to develop standards to define the safety, performance and interoperability requirements for Air Traffic Services supported by data communications.

SESAR Project 6.9.2 (Advanced Controller Working Position) focuses in developing a homogenous Tower Controller Working position based on the different concept developed in the other airport SESAR projects such as the guidance function, the main purpose of the project is to identify and describes in detail each role and responsibility which will interface with the future HMI taking into account, human, safety and other performance aspects.

Different validation exercise took place in the context of SESAR, to prove the concept and to investigate the impact of the integration into the advanced working position. A questionnaire and interviews were conducted to have ATCOs feedbacks. In the following paragraphs we start by explaining the concept, than we will describe the experiment and discuss its results.

OPERATIONAL CONCEPT AND OPERATING METHOD

The operational concept and detailed operating Method are described in the Operational Service and Environment Definition document (OSED) from the A-SMGCS Guidance function.

The D-TAXI and Data Link for Vehicles services aim to reduce voice communication (R/T) by exchanging non-time critical messages between ATC and mobiles (Start Up, Push back, Taxi and Revised Taxi) by data link. R/T should be used on first contact with the ATCO and has to remain available at any time in case the Flight Crew, Vehicle Driver or ATCO need to revert to voice communication.

A prerequisite for D-TAXI to be used at an airport is that the following services are in place,

- Data Link Initiation Capabilities (DLIC) enabling data link communication between ATC Ground and mobile systems. DLIC is initiated when the mobile wants to use Aeronautical Telecommunication Network (ATN) for communication purposes.

The DLIC service provides the log-on procedure to the ATN and exchanges the required application information. The DLIC process supports addressing requirements for Air Traffic Service Communication applications such as Controller Pilot3 Data Link Communications (CPDLC)

The DLIC supports the update of application information;

The DLIC service propagates application information, implementing the contact procedure a given time before the centre exit;

- ATC Communications Management (ACM) service provides automated assistance to the Flight Crew or Vehicle Driver, current and next controllers for conducting the transfer of ATC communications. Note: Only required if transfer of communication will be performed by data link.

[2] EMMA2 "European Airport Movement Management by A-SMGCS, part 2" (EMMA2), an Integrated Project launched by the European Commission in its sixth framework programme

[3] In this case Pilot indicates both Pilot and Vehicle drivers.

The D-TAXI service for aircraft consists of the following seven sub-services:

- Departure Taxi Route Information –information provided prior to departure on the departure taxi route as well as other departure information, it is foreseen that the Departure Taxi Route Information will be part of the DCL in the future
- Start-Up – Approval for aircraft engine start-up as well as departure information.
- Push-back – Approval for aircraft push-back from a stand as well as departure information.
- Taxi-Out – Taxi route instructions for aircraft to taxi from a point on the airport to another point on the airport surface.
- Arrival Taxi Route Information – The expected arrival taxi route as well as other arrival information provided to Flight Crews while still in flight.
- Taxi-In – Taxi route instruction for aircraft to taxi from landing runway exit to another point on the airport, provided after clearing the landing runway.
- Taxi Revision – Change to any previously delivered taxi route.

The Data Link for Vehicles service for vehicle drivers consists of the following two sub-services:

- Proceed / Tow –instructions for vehicle drivers to proceed/tow from a point on the airport to another point on the airport surface.
- Proceed / Tow Revision – Change to any previously delivered route.

The following figure provides an example of the operating method of D-TAXI for a pushback request.

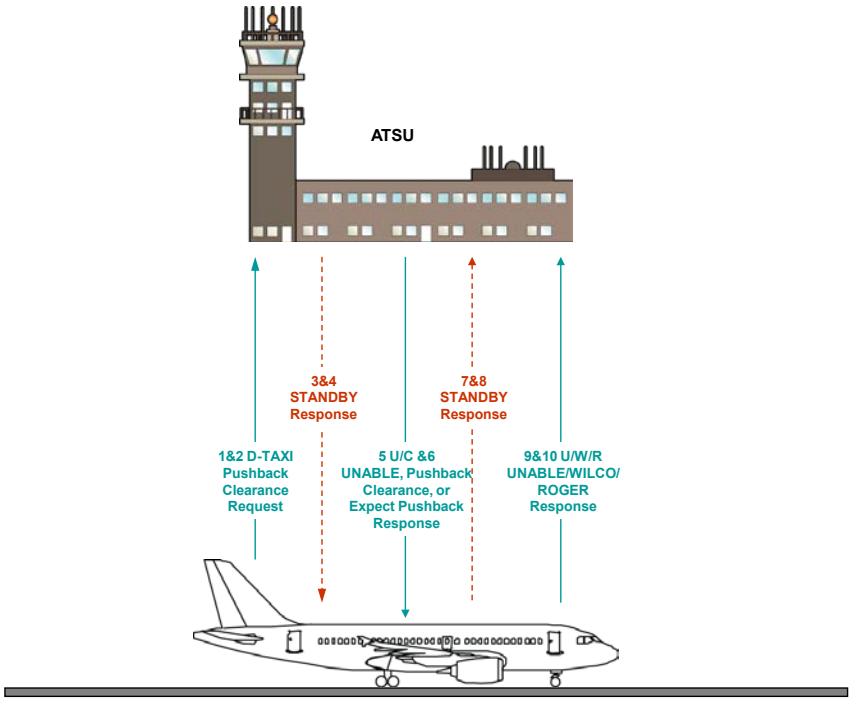


Figure 1 D-TAXI Push-Back Request Operating Method Diagram

D-TAXI TIMERS MANAGEMENT

A different timers were defines to deal with the different use cases such as delay in a response to a request message.

The list of timers as defined in EUROCAE WG-78/RTCA SC-214 is given in the following table:

Timer	Description
tr	Technical response timer. This timer is used by a sending system in LACK regions to detect the absence of an expected LACK response in an acceptable period of time.
ttr	Termination timer (Receiver). This is used by a receiving system to detect the absence of a response (system or human) to a received message in an acceptable period of time.
tts	Termination timer (Sender). This timer is used by a sending system to detect the absence of an operational response from the remote system in an acceptable period of time.
t-CPDLC-end	End timer. This timer is used by the ATSU System after generation of a CPDLC-end request to detect the absence of an End Response from the aircraft in an acceptable period of time.

Table 1: List of Timers

Figure 2 illustrates graphically the use of D-TAXI timers for a request issued by the flight crew on the airborne side. The timers tr, tts and ttr can be implemented on the ground side and/or on the airborne side.

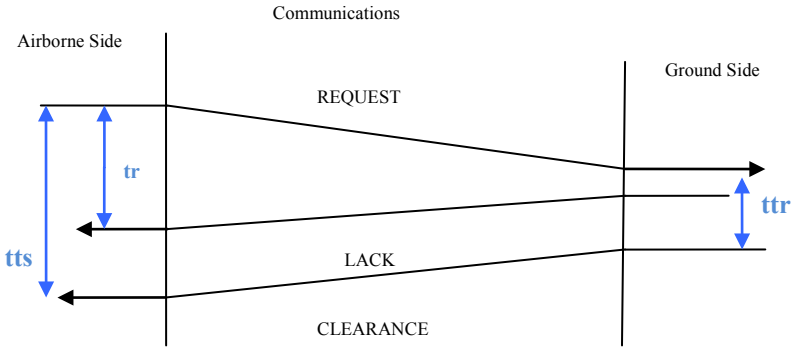


Figure 2: Overview of D-TAXI timers

Operational values to be used for the ground timer ttr and tts are defined in Table 2:

D-TAXI transaction context	Timer value
Start-up, Push-back	45s
Update and simple taxi-out instruction	45s
Simple taxi-in instruction	45s
Simple arrival expected taxi route	45 s
Holding position instruction	45 s
Complex taxi in instruction (for large TMA)	100 s
Complex arrival expected taxi route (for large TMA)	100 s
Departure expected taxi route	180s
Complex taxi-out instruction	180s

Table 2: Timer values with respect to D-TAXI transaction

D-TAXI VALIDATION

To prove the concept, a validation process was followed based on the E-OCVM European Operational Concept Validation Methodology (E-OCVM version 3.0) by EUROCONTROL.

The following diagram provides the workflow followed to validate the concept.

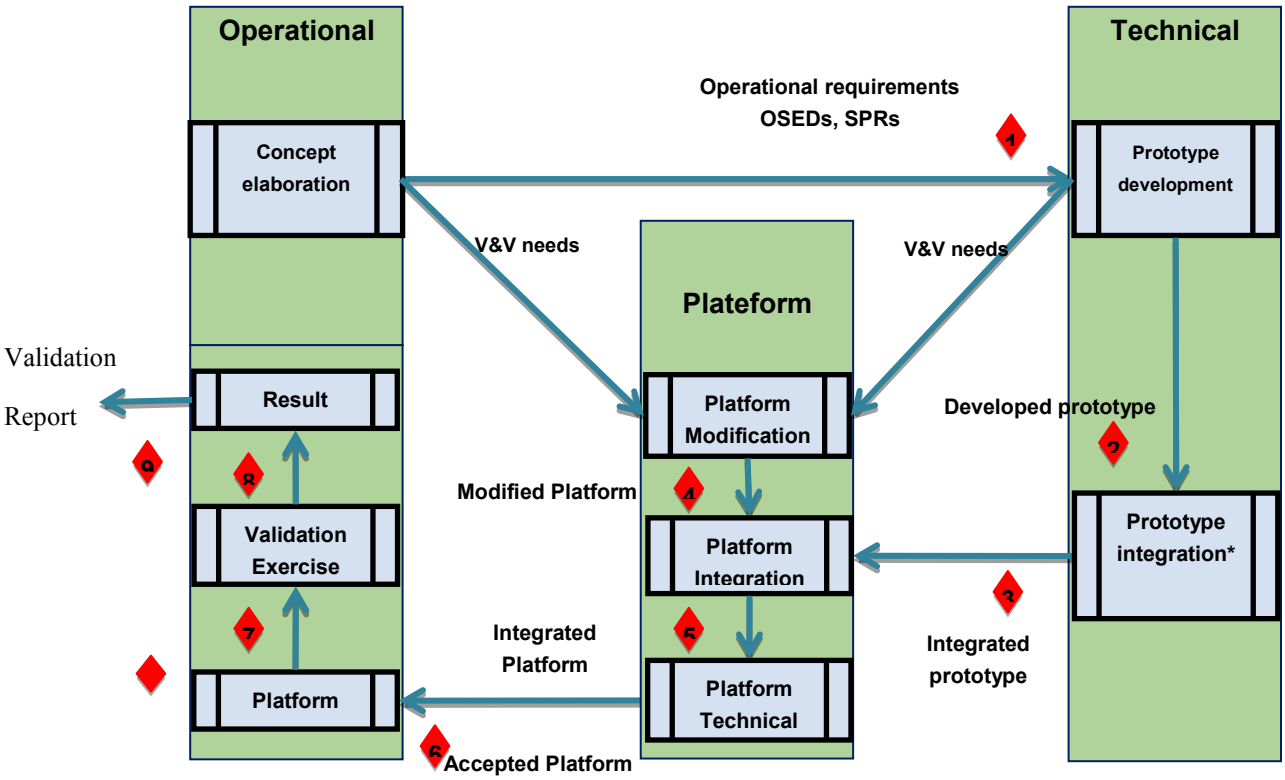


Figure 3: V&V life cycle

PLATFORM DESCRIPTION AND LIMITATIONS

The exercise used a validation infrastructure composed of integrated components:

- EUROCONTROL Integrated Tower Working Position (ITWP) simulator for Tower Runway and Tower Ground Controller Working Positions,
- EUROCONTROL ITWP providing flight plans, airspace environment, tracks, pseudo pilot positions and time server.
- Parkair System providing routing server system
- AIRBUS cockpit simulator MOSART providing the AIRBUS 340 cockpit environment used to simulate 2 flights per exercise one on departure and one on arrival

Validation infrastructure was installed at AIRBUS Toulouse premises. Interconnection of simulated platforms was performed using an agreed interfaces description.

The D-TAXI hardware infrastructure is not in the scope of the work, technical feasibility studies for the hardware was conducted on EMMA2 and continues in others SESAR projects (project 15.2.7 “Airport Surface Datalink” and project 09.16 “New Communication Technology at Airport” are addressing this technology. These 2 projects are contributing to EUROCAE WG82/RTCA SC223, the international WiMAX forum and the ICAO Aeronautical Communications Panel is planning to start developing Standards and Recommended Practices (SARPs))

DATALINK MESSAGES

The following list details the Data link message set used to perform the validations. The messages come from the Working Group WG78 SPR version J. (the Version M was released after the validation plan was started)

Downlink Messages: WILCO,UNABLE,STANDBY, ROGER, AFFIRM, NEGATIVE,DUE TO [due to reason],Free text (From FC Only),Free text (for BTV EXIT),LOGICAL ACKNOWLEDGMENT,DE-ICING COMPLETE,REQUEST DEPARTURE CLEARANCE [departure clearance request],DE-ICING STARTED,FOR DE-ICING,ABLE INTERSECTION [intersection] RUNWAY [runway],REQUEST PUSHBACK,[pushbackpositionO], REQUEST START-UP,REQUEST TAXI [taxi requestO],REQUEST EXPECTED TAXI ROUTING [taxi request].

Uplink Messages: UNABLE, STANDBY, [departureClearanceR], CONTACT [unit id] [frequencyO], MONITOR [unit id] [frequencyO], LOGICAL ACKNOWLEDGMENT, CURRENT ATC UNIT [unit id], START-UP APPROVED [assigned timeO], PUSHBACK APPROVED [pushback positionO] [assigned timeO], EXPECT [clearance typeR] [assigned time], EXPECT TAXI [taxi route] [taxi durationO], RESUME TAXI [taxiresumeconditionO], RUNWAY [runwayO] TAXI [taxi route], DE-ICING APPROVED, HOLD POSITION, FOR DE-ICING, CAN YOU ACCEPT INTERSECTION [intersection] FOR DEPARTURE RUNWAY [runway], REVISED [revision reasonO], [distance ground] AVAILABLE, HOLD SHORT [ground locationR]

VALIDATION PROCEDURES, SCENARIOS AND ASSUMPTIONS

According to the E-OCVM guideline, a validation plan was produced in which the procedures were detailed; the following is an extract from the procedures that were applied during the validation exercise:

- During simulation runs with D-TAXI the participants should try to use data link at all times unless instructed not to by the simulation supervisor or in cases where the participant needs to use R/T to clarify a situation or where it is considered that the data link messages available are not sufficient to cover the operation need (e.g. a complex push back instruction)
- On first contact with a new sector (e.g. Tower Ground Controller) the Flight Crew (FC) shall make a radio check before commencing to use data link. This R/T check is not required if the previous sector has instructed the FC by R/T or data link to Monitor the next frequency.
- The mixing of R/T and data link instructions is permitted, e.g. if the FC commence to use data link and then use R/T it is ok to revert to using data link.
- When a Data link message sent by the FC times out the FC can either send the message again or contact ATC by R/T

- When a Data link message sent by the ATCO times out the ATCO can either send the message again or contact the FC by R/T

The operational environment used for the validation was based on the Paris-Charles de Gaulle (LFPG) airport environment. The traffic samples were based on real flights using Paris CDG, however the parking stands have been modified to suit the northern part of the airport environment. It is not possible to control the arrival flights during the approach phase, so they were pre-programmed for normal spacing on Runway 27R with some landing also on Runway 27L. Departure flights had a programmed EOBT (Estimated Off Block Time) and it was up to the ATCOs to decide whether the flights could commence at these times, subject to the operational situation at the time during the exercise run.

The validation objectives includes more subjects than D-TAXI and its integration into the HMI, but hereafter we focus only on the objectives related to the D-TAXI.

In order to access the D-TAXI objectives, the samples had 85% of the aircraft data link equipped. Each traffic sample contained several flights where the aircraft type was an Airbus A340, and the flights were either arrivals or departures and had different assigned parking stands. Any one of these flights could be selected prior to the start of an exercise run to be the flight performed by the AIRBUS Cockpit simulator. The flight was selected according to the objective that was tested during the exercise run. The Ground South position managed flights arriving to and departing from Terminal 2 B and D Aprons but only planned to use the northerly runways.

4 traffic samples were defined.

Traffic sample 1 = 85% Data link Equipped – 60 minutes with 36 ARR, 39 DEP and 3 Vehicles

Traffic sample 2 = 85% Data link Equipped – 60 minutes with 36 ARR, 36 DEP and 3 Vehicles

Traffic sample 3 = Low Visibility – 60 minutes – Based on R67B_100H sample with 30% less Arrivals and Departures (28 ARR, 28 DEP) and 3 Vehicles

Traffic sample 4 = De-Icing planned for 40% of departures – Based on R67B_200H sample (36 ARR, 36 DEP (of which 15 require remote DE-ICE) and 3 Vehicles)

A total of 12 runs were conducted, each run last one hour. Each run corresponds to a pre-defined scenario using one of the traffic samples.

4 ATCO from France, Italy and Germany and 8 pilots participated to the validation for one week.

In order to prepare the ATCOs for the Validation, a dedicated training session was organised at EUROCONTROL EEC Bretigny using a 3D Tower simulator and the ITWP platform.

D-TAXI VALIDATION ANALYSIS METHOD

The main source of data comes from using subjective methods of analysis: observations, questionnaires and debriefings.

Observations

Observation was used to gain an indication of:

- Controllers situation awareness, for example failure of controllers to identify certain events / potential hazards
- The usability and utility of the HMI, for example controllers' difficulty in the HMI, and possible errors they may perform when interacting with the Platform; suggestions for improvement and other general observations, comments and remarks.

Questionnaires

A number of different questionnaires were used throughout the evaluation to gain the necessary feedback from the participants regarding the D-TAXI functions and its integration into HMI.

Post exercise questionnaire

The post exercise questionnaire contains open questions for controllers to note down any additional comments relevant to the exercise just completed.

Post simulation questionnaire

A bespoke post simulation questionnaire was used to gather all the relevant information regarding the Data link D-TAXI functions and its integration into the HMI.

The final questionnaire was distributed to participants at the beginning of the evaluation, to enable them to understand the feedback required from the evaluation in terms of utility and usability. However, the participants were requested to fill it out on the last day of the evaluation.

The Human Performance KPA has been addressed by assessing the impact of the new concept on situation awareness and mental workload areas. Questionnaires have been created for both domain based on models developed in the EUROCONTROL SHAPE (Solutions for Human Automation Partnerships in European ATM) project which objective was to assess the impact of new automation on the human operator. The SHAPE project has yielded a set of questionnaires which assess the impact of automation on workload AIM-s (Assessing the Impact of Automation on Mental Workload), situational awareness SASHA (Situation Awareness for SHAPE) but also teamwork and trust.

Debriefings

Post exercise technical debriefing

After the controllers have completed the questionnaires a short post-exercise debriefing led by the operational experts took place. The debriefing gave controllers and pilots the opportunity to discuss any issues which have arisen during the previous exercise.

Post-evaluation debriefing

A post evaluation debriefing was held at the very end of the evaluation session after the final questionnaire have been completed and collected.

The success criterion for the evaluation of every objective was depending on the controllers feedback collected during the simulation session. Using the rating scale, controllers rated the degree to which they agreed/disagreed with the statements developed by the validation project team to assess the usability and utility of every function under test. Each item has a six point scale ranging from from 1 “strongly disagree” to 6 “strongly agree”.

Analysis of the collected data quantified the degree of agreement/disagreement with regards to the addressed validation objectives.

The Objectives set are related to the D-TAXI messages content, the utility of the message and the usability of the HMI, for example:

Utility & Usability of Revised Taxi Function
TU01. The Revised D-TAXI is useful **Results – ATCO=6 APRON=4 Mean=5**
TU02. The Revised D-TAXI Taxi Update is easy to use **Results – ATCO=3.5 APRON=3 Mean=3.25**

RESULTS ANALYSIS

During the experiment, we had two groups of people: ATCOs managing the Ground and the runway movements and Apron Managers.

Situation Awareness

The overall resulting means for the exercises where the D-TAXI operations were used are:

- For the Tower Ground North Controller **4.02.** This corresponds to a "Good" level of situational awareness.
- For the Tower Runway Controller **4.91.** This corresponds to a value close to a "Very Good" level of situational awareness.

These results indicate that the Data Link function, with its associated HMI elements, maintains the controller's situational awareness to a suitable level to perform his/her tasks; the figure below details the ATCOs' answers to the SASHA questionnaire:

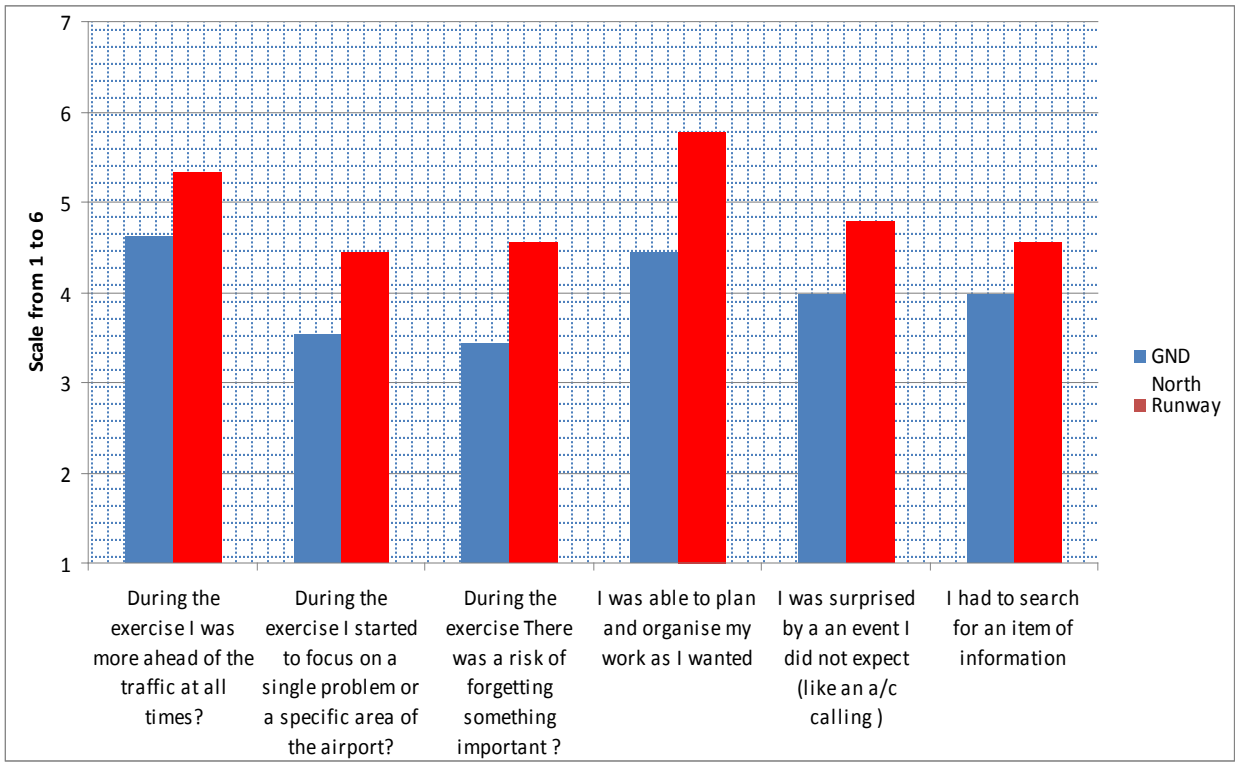


Figure 4: Situation Awareness results for Ground North and Runway controllers

Questionnaires were also filled in for 1 baseline exercise without the Data Link function activated. The purpose was to assess the difference, if any, of the controller workload between the two scenarios.

- For the Tower Ground North Controller **5.33.** This corresponds to a "Very Good" level of situational awareness. Compared to the result in Data Link environment, **4.02**, there is a decrease of situation awareness on the Tower Ground position in Data Link environment. This can be explain that the initial exercises were still sort of training exercises for the controllers and results were hence affected. Nevertheless the situation awareness in D-TAXI environment remains good and acceptable for the Tower Ground controller to perform his/her tasks.
- For the Tower Runway Controller **5.** This corresponds to a value close to a "Very Good" level of situational awareness. Compared to the result in D-TAXI environment, **4.91**, there is a nearly no change in controller situation awareness which can be explained by the fact that Data Link do not concerned too much the Tower Runway position.

Safety

The safety aspects linked with the introduction of D-TAXI and Routing functions were addressed in the questionnaires where three questions were asked to ATCOs and Apron managers.

The table below provides the results for both groups of people.

Statement	Mean ATCOs	Mean APRON
Safety		
The use of data link is reducing the probability of misunderstanding between ATCOs and flight crews	**5,00**	**2,50**
The use of data link for D-TAXI operation is reducing (negative impact on) safety of airport operations on the movement area	**1,50**	**4,00**
The use of data link for D-TAXI operation is increasing (positive impact on) safety of airport operations on the movement area	**4,50**	**3,50**

Table 3 Safety results for ATCOS and Apron managers

We can observe that the ATCOs are positive concerning safety aspects linked to the introduction of the new concept as no critical safety element was identified.

The Apron Managers are more sceptical and not convinced of the safety impact either negative or positive of the new concept on apron operations. This can be explained by the fact that it was the first time that data link and routing functions were simulated on the Apron area and tested by Apron managers. This clearly shows that apron managers are lacking experience of the concept and need to perform additional evaluations. It is to note that these results are also in line with the situation awareness results with the same explanation

Mental Workload

ATCOs were asked to fill in an AIM-s (Assessing the Impact of Automation on Mental Workload) questionnaire after every exercise and were asked to rate the degree to which they agreed/disagreed with eight statements in order to measure their level of mental workload during the just completed exercise. The final result provides a mean on a scale from 1 to 6 where the level of mental workload can be considered as 1: none 2: very little, 3: little, 4: much 5: Very much 6: Extreme.

The overall resulting means for the exercises where the D-TAXI operations were used are:

- For the Tower Ground North Controller was **1.84** which indicates a value between **"None" and "Very little"** workload**,** which is a satisfactory.
- For the Tower Runway Controller was **1.43** which also indicates a value between **"None" and "Very little"** workload, which is again satisfactory.

The result for the baseline exercise without D-TAXI gives a result of **1.4** for the Tower Ground North controller and **1.00** for the Tower Runway controller. This indicates that without D-TAXI, the mental workload was even lower which is normal since it corresponds to today operations.

Detailed results for controller workload are represented in the figure below. We have mentioned the questions on these two graphics, nevertheless the beginning of each question has been deleted "During the exercise how much effort did it take…".

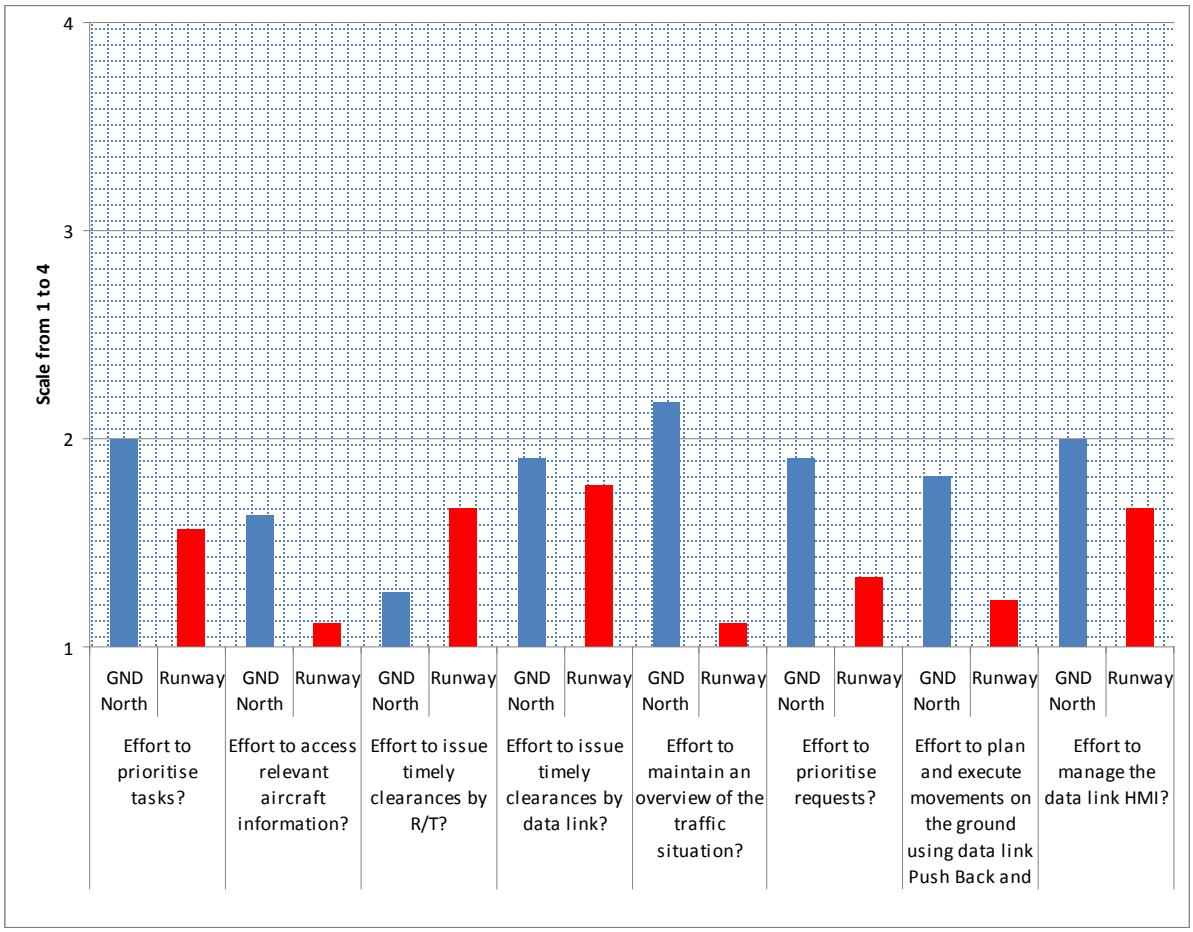


Figure 5: Mental Workload results for Ground North and Runway controllers

CONCLUSIONS

This paper describes some of the result of the validation exercise conducted in the context of the SESAR program to validate the integration of the D-TAXI concept into an Advanced Controller working position.

The results of the experiment and the very positive feedbacks collected from the Air Traffic controllers' show that not only implementing datalink function has a very little impact on their mental workload but it can improve the safety at airport. All ATCOs agree that they were not confused when applying procedures using both datalink and R/T and that they can manage the datalink communication appropriately.

Despite the lack of training on the new system, the controllers were able to learn quickly and use efficiently the advanced controller working position, they found most of the D-TAXI messages useful but depending on the controller expertise and local implementation procedure some of the messages are less important for implementation.

The controllers highlighted a Human factor issue as in a voice communication they can have an additional feeling about the pilot readiness to execute a clearance; this feeling is removed by using datalik.

Another known issue is the situation awareness of the other users during a "silent communication" as in today's operation; the other users listening to the frequency can have more information about what is happening around them. A possible enhancement can be to offer to the user the possibility to subscribe to all datalink exchange that is for interest for him, this feature should be further investigated.

REFERENCES

ED-78A, Guidelines for Approval of the provision and use of Air Traffic Services supported by Data Communications
ICAO Doc 9694, Manual Of Air Traffic Services Data Link Applications, 1st Edition 1999
ICAO Doc 9476, Manual of Surface Movement Guidance and Control Systems (SMGCS), 1st Edition 1986, 2nd Amdt 1987
ICAO Doc 9830, A-SMGCS Manual, 1st Edition 2004
ICAO Annex 14, 5th Edition 2009, Vol. I 10th Amdt.
Jörn Jakobi et al (2009), EMMA2 Recommendations report.
RTCA SC-214/EUROCAE WG-78, Air Traffic Services Safety and Interoperability Requirements, Version H, 03/02/2010
SESAR Joint Undertaking, http://www.sesarju.eu.
SESAR, DEL06.07.03-D10-Final OSED V00.01.00, 2012
SESAR, DEL06.07.02-D07-Data link routing information - development (V2) V00.01.00, 2012
SESAR P06.07.02 D80 Validation Plan (VALP) D80 V1.0, 10 December 2013
SESAR P06.07.03 D81 Preliminary VARL Advanced routing V0.0, 29/01/2013
Thomas Ludwig et Al, Implementation and validation of TAXI-CPDLC functionality within the EU Project EMMA2, ICAS2010

Air Traffic Controller Support by Speech Recognition[1]

Oliver Ohneiser, Hartmut Helmke, Heiko Ehr, Hejar Gürlük, Michael Hössl[2], Thorsten Mühlhausen

Institute of Flight Guidance
German Aerospace Center (DLR)
Lilienthalplatz 7, 38108 Braunschweig, Germany

Youssef Oualil, Marc Schulder, Anna Schmidt, Arif Khan, Dietrich Klakow

Department of Computational Linguistics and Phonetics
Saarland University (UdS)
Campus, 66123 Saarbrücken, Germany

ABSTRACT

Air traffic controllers (ATCO) are a core element of the flight guidance process. Decision support systems with accurate output data could increase the controllers' performance or reduce their workload. Nowadays radar data based identification of controllers' intent causes delays of up to 30 seconds. The intents could be predicted better and earlier if their spoken commands would be taken into account. Therefore, the project AcListant® combines an arrival manager (AMAN) with an automatic speech recognizer (ASR). Spoken commands are automatically recognized and forwarded to the AMAN. The AMAN updates its plan (e.g. sequence and aircraft trajectories). The ATCO also receives a direct feedback of ASR recognition performance via an (optional) visual interface. The ATCO gets better support from the assistance system without additional personal effort. This has an impact on the controller's work and behavior like adhering closely to the radio telephony procedures and articulating more clearly. Together with the suggestion of most probable commands depending on the air traffic situation, this would lead to higher speech recognition rates and result in better AMAN assistance. Current results of a pre-study performed by the Düsseldorf Approach Area foreshadow significant positive effects of ATCO support by ASR.

Keywords: Air Traffic Controller, Automatic Speech Recognition, Situation Data Display, Speech Recognition Log, Arrival Manager, Context, AcListant®

INTRODUCTION

Today controllers communicate with pilots via radio telephony (R/T) for flight guidance of aircraft. They are responsible for the safe, punctual, cost efficient, and eco-friendly processing of air traffic movements within their

[1] The work was conducted in the AcListant® project, which is supported by DLR Technology Marketing and Helmholtz Validation Fund. The DLR colleagues were responsible for arrival management system and validation purposes, UdS colleagues provided the automatic speech recognition.

[2] Michael Hössl now works as a first officer on a Boeing 737 aircraft.

assigned airspace. Air traffic controllers obtain information about the status of the air traffic from a radar screen, which displays each aircraft with a configurable label containing information like speed or altitude. Controller commands and clearances via radio telephony for each aircraft are recorded on flight strips, which can either be paper or digital. According to the phase of a flight, there are several controller working positions in order to assure maximum safety. Approach controllers for instance are in charge of the final approach phase of flight. Their flight guidance can be supported by an arrival manager (AMAN), which plans efficient sequences based on radar data input and generates advisories for efficient approaches. The AMAN creates the optimal sequence and trajectories depending on flight status and position of each aircraft. Plan deviations can occur when the controller generates another sequence or "vectors" an aircraft in order to accommodate air traffic management or airline needs. The AMAN then detects these deviations from the radar data resulting from the flown aircraft trajectory. This leads to a new planning cycle. As the clearances given by the controller are unknown to the AMAN (especially the controllers' intent), the plan possibly has to be updated several times until the flight status (e.g. flight level, speed etc.) of the corresponding flight remains stable. This results in a delayed recognition time of the controller intended plan by 30 seconds or even more. For the controller this reduces the benefit of the AMAN.

It is expected that automatic speech recognition (ASR) in the ATC (Air Traffic Control) environment will improve the usage of an AMAN (Helmke et al., 2013), which in turn provides the controller with timely and stable support (see Figure 1). Although the usage of data link in ATC is discussed at least since the 90s, voice communication will definitely remain a pillar of air traffic control. The Strategic Research & Innovation Agenda (SRIA) of ACARE (ACARE, 2012) or Flightpath 2050 (European Commission, 2011) do not expect a fully automated ATM (Air Traffic Management) system in the next decades.

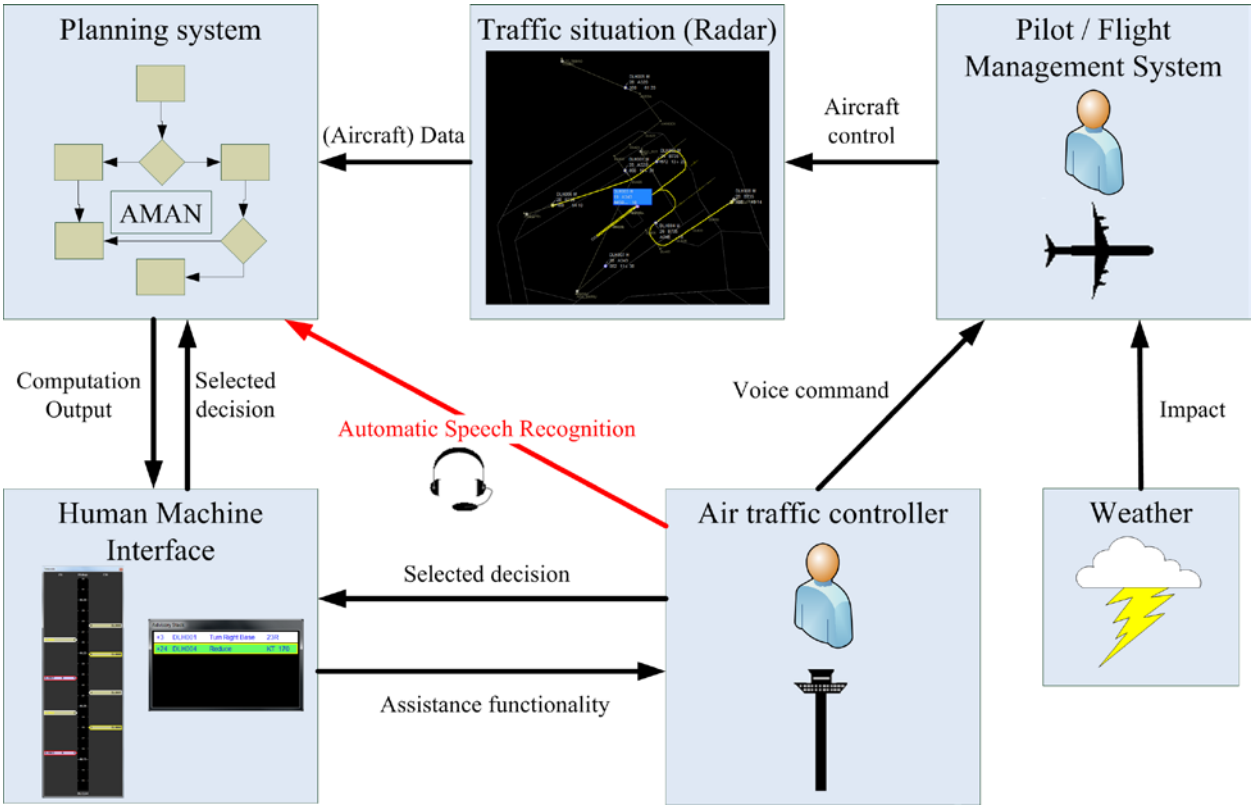


Figure 1. Today's operational procedure with standard AMAN: Interaction of air traffic controller with pilot and machine environment (modified from Elzer, 2001); red arrow: new, with an additional sensor for automatic speech recognition

The following chapters begin with the operational concept of using an ASR-supported AMAN. In addition, the recording of ATCO voice data, the necessary transcription and the extraction of concepts are described. To achieve small error rates, the generation of air traffic context information as well as the inclusion of this context in automatic speech recognition is required. This is explained in two further chapters. Afterwards the visualization of ASR's output in an HMI is detailed. The pre-study evaluating the context-dependent AMAN supported ASR is outlined whereas the last chapter summarizes the work and gives an outlook on future steps of the project.

OPERATIONAL CONCEPT FOR USING AMAN WITH ASR

As described above, the delayed update of the AMAN can be solved by the use of ASR. An assistant system which actively follows the human communication solves this task by analyzing controller-pilot communications and incorporating it as additional sensor information. The project AcListant® follows this new approach, by integrating ASR into the assistant system of an AMAN. The AMAN-ASR-system transforms speech into text and extracts the semantic intent (e.g. descending to a flight level). Listening to controller commands in ATM requires a highly reliable speech recognizer. Therefore, the approach makes explicit use of dynamic context information being created by the AMAN and used by the ASR system. This context information is derived by the AMAN from its other input sources (e.g. radar data, flight plans). It informs the ASR of expected and possible controller advisories, i.e. it continuously creates context information. Consequently, these hypotheses help the speech recognizer to detect spoken commands with improved reliability. Based on the extracted information from the controller-pilot communications the assistant system itself can more quickly adapt its own model, i.e. its knowledge concerning possible future system states.

The described approach solves the problem of delayed system reaction whereby optimal planning is ensured at all times. The better the system is informed about the airspace situation as well as about operator intents, the better the quality of the planning as a combined effort of the AMAN and the human air traffic controller becomes. The improved planning results directly in improved support functionality of the assistant system.

Besides that, the new ASR-based assistance system, AcListant®, operates in a seamless mode, i.e. controllers are not disturbed in their natural way of work. The speech recognizer operates in the background, which means the submission of the detected commands to the AMAN will not be noticed by the controllers. However, it is possible to display the detected commands to the controllers to indicate the performance of the speech recognition and hence to assist the controllers in their understanding of the system. Figure 2 schematically depicts the layout of the AcListant® controller's working position without this additional display. Beneath the situation data display, the flight strip bays are located, depicted as blue rectangles. The blue solid line beneath the flight strip bays merely symbolizes the seamless design of AcListant®, i.e. the method of operation does not change for the air traffic controller. The controllers are actually not aware of the ASR and its coupling to the AMAN, meaning the air traffic controllers operate in their daily routine using R/T and flight strips in order to guide the air traffic.

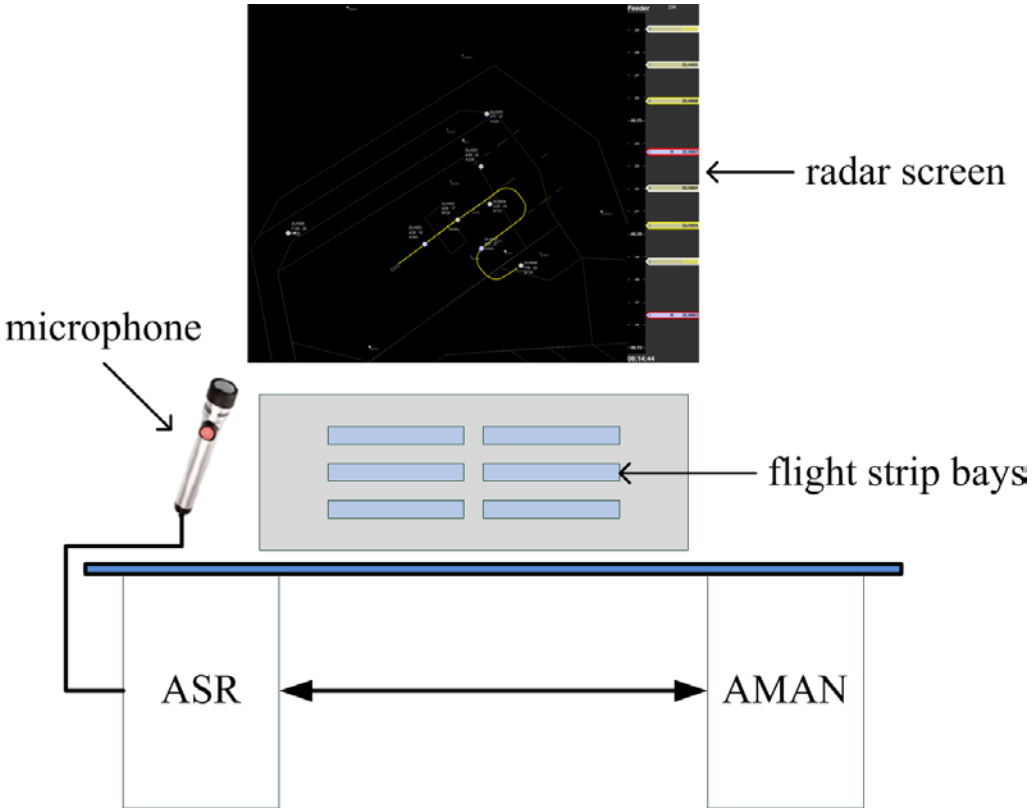


Figure 2. Schematic layout of AcListant® controller working position

The AMAN receives recognized commands from the ASR. After a first plausibility check, missing information which is not available in the speech recognizer process is added by the AMAN, e.g. for the command "SPEED 180"

the AMAN compares the speed value with the aircraft's actual speed and assigns the command type REDUCE or INCREASE. The same procedure happens to the command "FLIGHT LEVEL 80". The AMAN determines by the actual altitude value whether a DESCEND or CLIMB command has been issued. "TURN HEADING 270" can be transformed to TURN_RIGHT_HEADING or TURN_LEFT_HEADING if the present heading of the aircraft is known. If the command is stated as correct, it immediately influences the actual planning cycle, e.g. sequence and trajectory are updated. Derived from the trajectory, the distance-to-go and the distance-to-top-of-descent are computed.

As mentioned above, a plausibility check is performed on the recognized commands. Wrong ASR commands may be detected if the aircraft's radar data do not fit the recognized command. If the assumed value of a command type is not reached within a certain time, the command is ignored again by the AMAN. This is the case when the actual time is greater than the command time plus an estimated operation time, e.g. an assigned heading change of 30 degrees should be achieved after at most 40 seconds. All valid commands could be considered for the next computation cycle as far as their values do not exceed certain limits (see chapter "VISUALIZATION OF RECOGNITION OUTPUT IN AMAN HMI").

VOICE DATA RECORDING, TRANSCRIPTION AND CONCEPT EXTRACTION

Achieving good command recognition rates is a basis for delivering correct feedback for an AMAN and, therefore, increasing user acceptance (Hah and Ahlstrom, 2005). The automatic recognition of spoken natural language is difficult for a variety of reasons. Machines process sound very differently than humans and successful recognition of words depends on a variety of factors. Apart from volume, speed and clarity of articulation, the accuracy of machine recognition is also influenced by the speaker's gender, physical and psychic constitution (e.g. stress or fatigue), whether they are native speakers of the language and which dialect they speak. Hence, every single word can be produced in ways, although they all mean the same. In fact, even a single speaker will never say a word exactly the same twice. The automatic speech recognizer thus has to learn from a great amount of training data to successfully extract patterns of pronunciation from individual renditions of sound. The Wall Street Journal WSJ0 corpus (Paul and Baker, 1992) is a commonly used source for training speech recognition systems for American English. It comprises eight hours of recordings by 82 speakers reading articles from the stock exchange journal.

Although this corpus, together with some similar corpora (Hofbauer et al., 2008), is a good basis for speech recognition training, the phraseology and speaking habits of air traffic controllers are different. The scope of content during communication with pilots is limited and regulated. The ICAO (International Civil Aviation Organization) air traffic communication regulation defines a standard (ICAO, 2007). Every radio contact between controller and pilot must follow the defined rules and syntax. However, in practice there are many deviations (Karlsson, 1990). Also, with their individual experience every controller has different flight guiding preferences. The form and semantics of the speech input therefore even deviates at the same controller working position.

Contrary to the specified limited language area, so called out-of-grammar phrases often occur in an ATCO command. Voice communication usually starts with a greeting like "hello" or "bonjour" and ends with "bye" or "tschüs" in different languages. In addition sounds of clearing one's throat and hesitation sounds like "um" or "ähm" may disturb the understandability of commands. As those elements are both completely out-of-grammar (i.e. the word is not part of any allowed sentence) and not part of command-critical phrases, they can be filtered out quite reliably. It is more difficult to interpret deviations (Zokić et al., 2012) in command-critical phrases, e.g. saying "reduce your speed one eighty" instead of the official version "reduce speed one eight zero knots" both introduces illegal words ("eighty") and omits required ones ("knots"). Another challenge is the use of acceptable words in incorrect contexts, e.g. "100" may be spoken as "one hundred" instead of "one zero zero" in flight levels, but not when part of a callsign.

To collect realistic and extensive data (De Cordoba et al., 2006) concerning the concrete test field Düsseldorf approach, we work together with the DFS (Deutsche Flugsicherung GmbH). We recorded several hours of voice communication and corresponding radar data during a simulation of Düsseldorf approach at their place of business in Langen. The speech data was recorded with a headset and a sample rate of 16,000 Hz regarding the "wav"-files. The quality of the recordings varied with the correct and stable "connection" from mouth output to headset input. The complete voice stream was then manually cut into short audio files of individual commands.

During simulation runs in the DLR's ATMOS (Air Traffic Management and Operations Simulator) shown in Figure 3 an input trigger was used for activating ASR recording, capturing start and end times of commands and splitting different commands. The trigger could be set via a foot switch or microphone button, preferably in a way that does not add to the cognitive load of the controller.

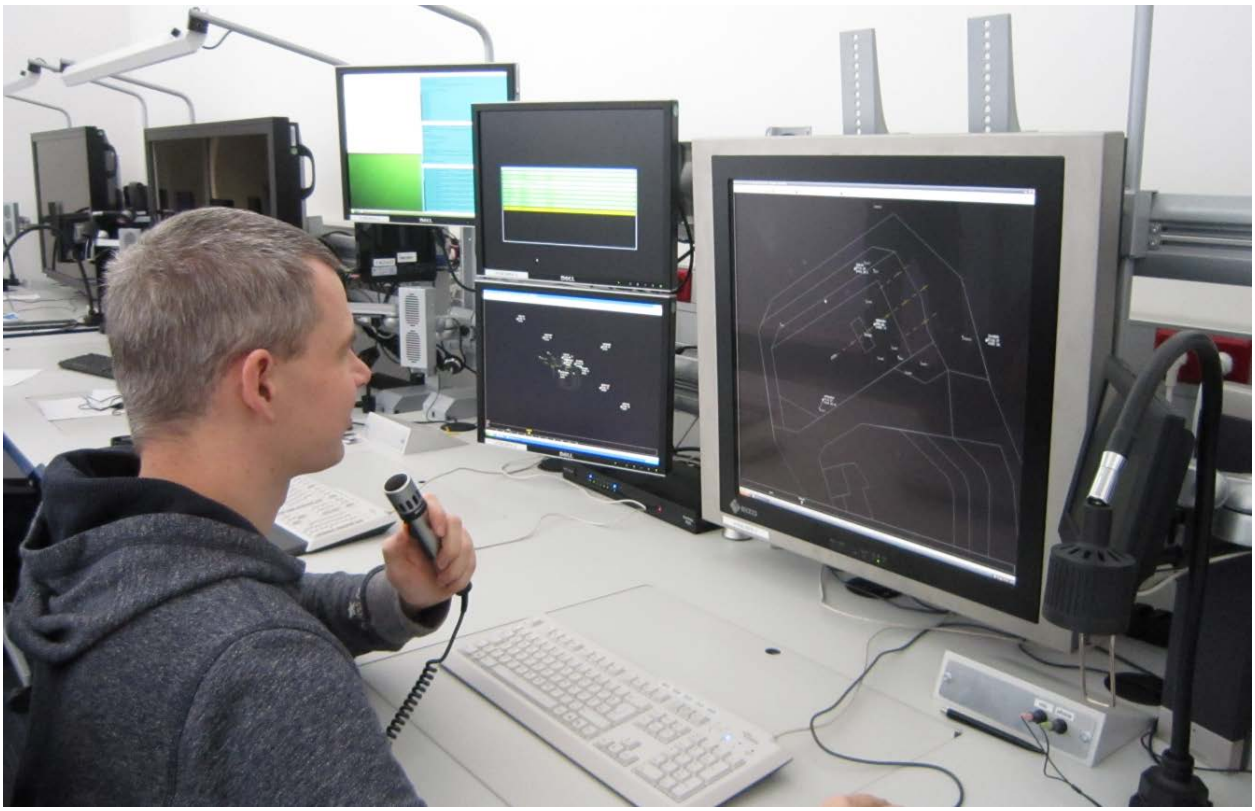


Figure 3. Simulation of Düsseldorf approach with recording of voice and radar data

A large amount of voice data has to be prepared for further electronic processing in different steps. The recorded utterances need to be transcribed as text and be annotated with additional information. The tool "Annotator" (see Figure 4) assists the human auditor with annotating commands.

The upper left window Name shows a list of all command files found in the specified directory. The color of the file names symbolize if the annotation of the utterance was already done (black) or not (blue). Next to it the amplitude and duration of the currently selected recording are shown. The lower half of the window contains the ASR system's prediction on the left and corrected version of the human auditor on the right. Each side consists of fields representing the utterance in plain text, XML (Extensible Markup Language) syntax and as AMAN commands. The XML annotation represents semantic markers that allow easy machine extraction of information from free speech, e.g. which command is meant by a certain utterance or whether a series of digits is part of a flight number, a speed value, a height value, etc. Lastly, the commentary field in the upper right allows auditors to leave remarks about individual recordings, such as unresolved questions or special observations.

The ASR prediction offers a starting point for the auditor, who compares it to the audio recording and can copy it over to the transcription fields if correct. If the prediction contains mistakes, the auditor corrects these in the plain text field (Text transcription). To receive an automatic suggestion on how to annotate the corrected sentence with XML, the auditor presses the GetXML button. The result can again be revised if necessary. In a similar fashion, commands are generated, based on the XML text, by means of the GetCommand button. In cases where a single utterance contains more than one command, they are listed sequentially and each command is headed by the callsign given in the utterance. The resulting auditor-approved utterance representations (text, XML and command) are used either for training the ASR system or to evaluate its word error rate (WER) and command error rate (CmdER).

The annotation effort concentrates on the correct extraction of concepts. Callsigns, command types and the command values constitute concepts and the combination of those three concepts forms a command. Examples for callsigns are DLH437 "Lufthansa four three seven", BAW86V "Speedbird eight seven victor", UAE55 "Emirates five five" or LOT927 "Pollot nine two seven". The command type followed by a possible value is e.g. "Reduce 180 KT ('reduce one eight zero knots')", "Descend FL 80 ('descend flight level eight zero')", "Rate of Descent 1000 ft ('rate of descent one thousand feet')", "Direct DL455 ("direct to delta lima four five five")', "Turn Left heading 260 ('turn left heading two six zero')".

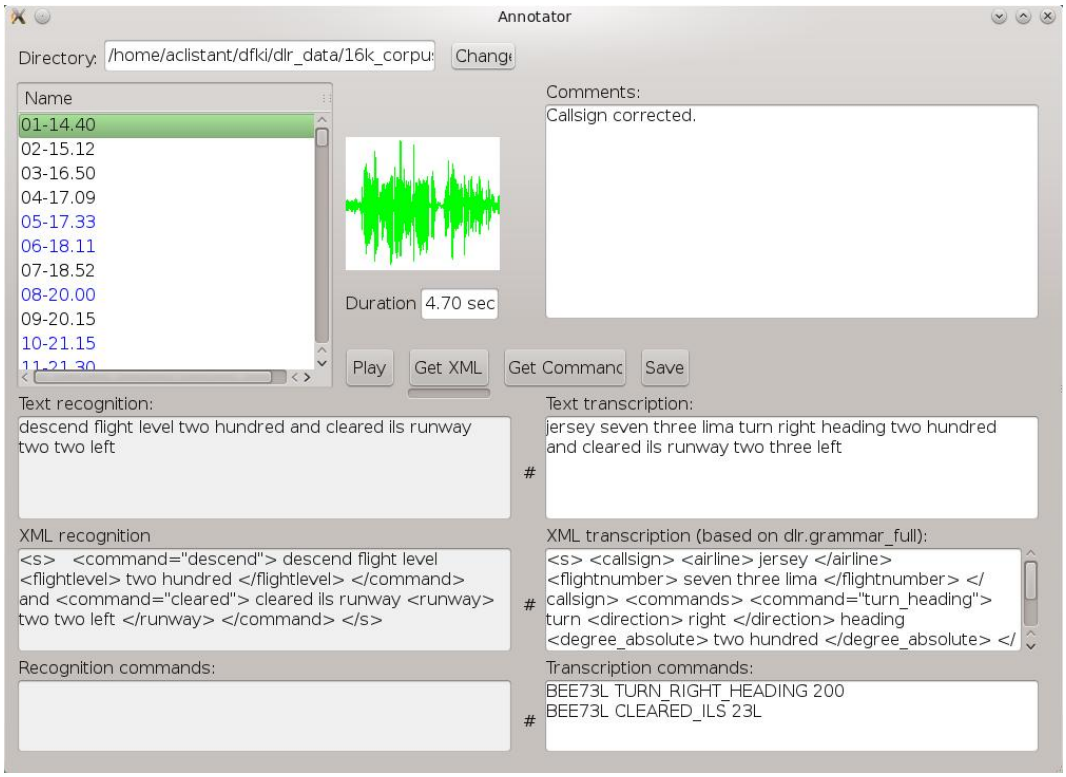


Figure 4. Annotation tool for extraction of concepts out of controller voice commands

GENERATION OF AIR TRAFFIC CONTEXT INFORMATION

Speech recognition for aviation applications benefits from the high level of standardization and procedural structure that is used for all commercial airline operations (ICAO, 2007). Especially the standardized grammar limits the number of possible word combinations. In an effort to further reduce the word error rate in ATC speech recognition, a reduction in the number of possible predefined commands for the speech recognizer to choose from could be beneficial. A smaller number of combinations imply a better chance of recognizing the correct terms. This limitation can be achieved by anticipating air traffic controller commands, i.e. by excluding commands which are not possible due to operational constraints from the context.

Radar flight track analysis and air traffic controller interviews represent the basis of controller command prediction, as real life air traffic control operations include a great number of variables that can cause the air traffic controller to deviate from published procedures. Reasons are e.g. different airplane types and their aerodynamic characteristics, payload, airline standard operating procedures, environmental considerations, weather and above all, the omnipresent necessity to keep approach flow high. Even though these deviations may seem random at first, it was shown (Hössl, 2013) that they do follow certain rules, ultimately allowing a prediction of possible controller commands. Typical deviations from published approach procedures are shown in Figure 5.

Düsseldorf Int. Airport, located in the busy airspace of the Rhine-Ruhr metropolitan area in Western Germany has been selected as example airport for this project. The basic principle of our considerations, however, is universally applicable. This is achieved by strict separation of generic and airport specific data.

Based on the aircraft's parameters, such as position, altitude, airspeed and heading, the amount of possible commands can be significantly reduced. The implementation of this concept is called Mapped Exclusion Method (MEM). The MEM is based on the idea that future ATC commands are primarily influenced by the aircraft's position and configuration. The MEM-function starts with the assumption that all commands are possible all over the approach sector, i.e. the set of possible commands for an aircraft contains all thinkable commands. Subsections of the airspace can be modeled as polygonal areas on a map. Areas have attributes such as preconditions (e.g. "Alt>50" means that only aircraft above 5000 feet are considered to be within the polygon) and restrictions (e.g. "Hdg(20,40)" means that all heading commands have to be within the range of 20° to 40°).

If an aircraft is located within an area and all preconditions are met, the set of possible controller commands is reduced by the restriction-attributes of that area, i.e. the previous set of possible commands is intersected with all restrictions of the area. If it should be required to add commands again in certain sub-areas, expansions can be added as an opposing function to restrictions. Finally the aircraft is checked against all areas and the set of possible controller commands is reduced by the restrictions of respective areas. As a result, only operationally likely commands are added to the list of possible instructions for each aircraft.

To facilitate the MEM, the mapping application JOSM of the OpenStreetMap (OSM) project was used (Ramm et al., 2010). Originally designed to create open source world vector maps, the JOSM interface serves the project's purposes for the allocation of the approach airspace into polygon areas required for the MEM (see Figure 6). These polygons can include properties which are used for preconditions, restrictions and expansions. As the main algorithm remains generic, only airport specific data needs to be created. Using JOSM there is no need of any coding skills of the user.

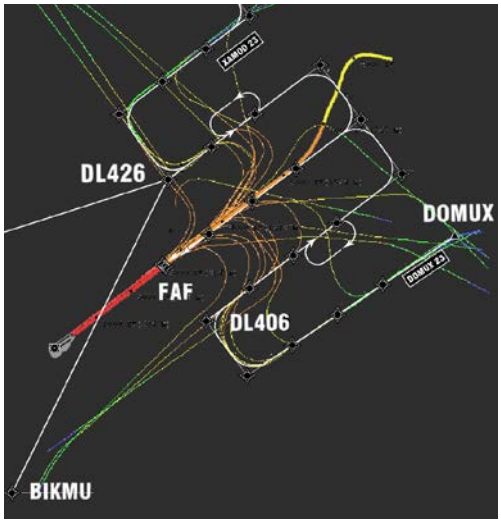


Figure 5. Aircraft flight tracks (colored lines) deviating from published procedures (white lines) at the approach into Düsseldorf Int.

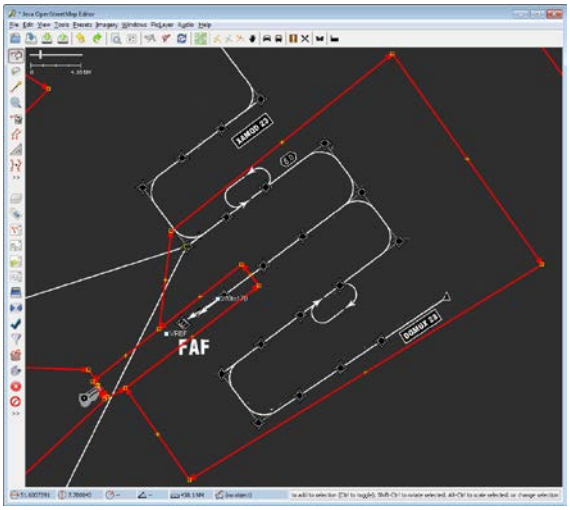


Figure 6. Polygon area (red lines) creation in JOSM

A conducted test series of the MEM algorithm showed an average context error rate of 0.9%, meaning 99.1% of commands issued by the controller in the test cases were predicted (982 out of 991). Additionally, the amount of possible commands could be reduced by an average of 77.3%, meaning only 22.7% of the total context remains within the algorithm's output (204,308 out of 899,577 commands).

AUTOMATIC SPEECH RECOGNITION BASED ON CONTEXT

The proposed Automatic Speech Recognition (ASR) system is based on Kaldi software (Povey et al., 2011), which is an emerging speech recognition toolkit that is freely available under the Apache license. This framework is very flexible and easy to extend to include new ideas and methods. In order to increase the performance of our ASR system, an acoustic model has been trained on five hours of in-domain data, collected in Air Traffic Control (ATC) simulations and annotated using the annotation tool described above.

This model is then combined with an ATC grammar (a language model) that defines the language rules and structure of the ATC commands. The grammar is designed in a scenario-agnostic way that allows easy adaptation for different airports, e.g. callsigns are defined as one airline (out of a list of several dozen), followed by any combination of up to five numbers or letters and digits, irrespective of which airlines and callsigns are actually encountered at the airport in question.

The combination of the acoustic and language model leads to a recognition network (a search space) defining all expected utterances and the different ways of pronouncing them. The recognized sequence of words is returned as the most likely path in this network. The concept extractor is then applied to this sequence to extract the spoken ATC command.

The scheme described above is rather naïve as it does not include any ATC knowledge, e.g. how frequently commands are spoken, how likely commands are, radar information, etc. The presence of such information can

strongly improve the recognition performance (Shore et al., 2012). This idea is integrated into our system using the context information (described in chapter "GENERATION OF AIR TRAFFIC CONTEXT INFORMATION") that is updated every 40-60 seconds. The received context is used to build a lattice (see example in Figure 7) that replaces the recognition network, strongly reducing both the number of possible callsigns and commands.

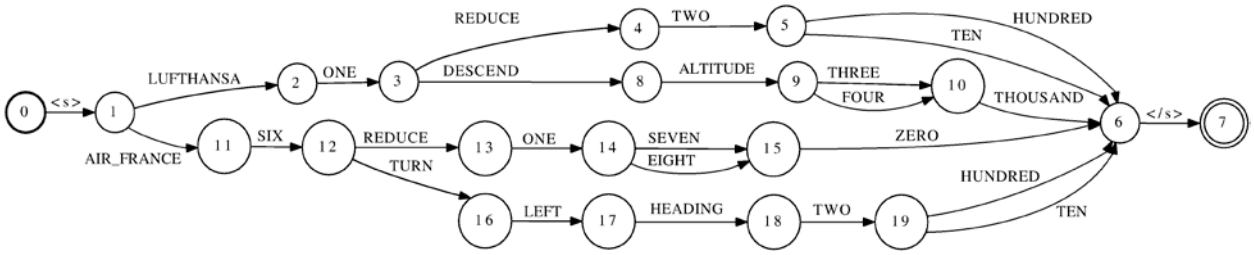


Figure 7. A simplified lattice resulting from the application of context to the recognition network. Each path represents a sentence that can be recognized by the ASR. Note that for the Lufthansa aircraft a reduce or descend command is allowed, while the Air France can receive reduce or turn. In the original network, any callsign could have received any command.

The search space for callsigns (airline and flight number) is limited to the aircrafts that appear on the radar, reducing it from millions of possible combinations to between ten and twenty. Similarly, the number of possible command-value combinations can be reduced to 30 to 40 per aircraft by discarding those that are not anticipated in the context information. Furthermore, the possible probability values that could be received with the context information can be used to distinguish between more likely callsigns and commands in the reduced search space itself, e.g. aircrafts that are close to the terminal are more likely to be addressed than ones that are still far away.

The proposed approach reduces the Word Error Rate (WER) by 47% relative, dropping it from 12.0% to 6.3%. More importantly, it reduces the Command Error Rate (CmdER), i.e. the number of commands with at least one mistake in callsign, command or value, by 75% relative, decreasing it from 61% to 15%. The gap between the improvement of the CmdER compared to the WER is due to the fact that the recognizer mostly commits errors while recognizing keywords such as digits, which define flight numbers and command values. As a result, spoken sentences that were correctly recognized except for a single digit will still be counted as completely miss-recognized in the CmdER, leading to this discrepancy in the two metrics.

VISUALIZATION OF RECOGNITION OUTPUT IN AMAN HMI

The human machine interface RadarVision visualizes output data of the database used by the arrival management system. RadarVision mainly serves as a prototypic situation data display for air traffic controllers. The largest part of the screen is filled with a conventional radar screen. The radar display includes all aircraft positions shown in a two-dimensional way completed by alphanumeric details on callsign, aircraft type, altitude and speed.

In addition to the current runways in use, significant waypoints, route segments and borders of important airspace areas can be shown. Beyond simply displaying the current situation, the calculated support functionalities of the AMAN can be visualized to enhance the radar screen as shown in Figure 8. This includes for instance the planned touchdown sequence, the computed trajectory for every aircraft or a distance-to-go prediction. The timescale on the right consists of aircraft labels assigned to a certain landing time on a specific runway. The whole scale is moving on downwards to the actual touchdown time as time goes by.

Regarding the automatic speech recognition process running in the background, the controllers have no additional work. They neither have to click or type anything nor need to change habits. Nevertheless, the ATCOs can choose to receive direct feedback about their voice input. The recognized command is therefore prepared for a so called speech recognition log (see Figure 9). This display stack visualizes the system's understanding of the controller's utterance. The speech recognition log also distinguishes between different categories of recognized commands. Every command is checked for plausibility and compared to the given context.

A command is only plausible if its type-dependent value lies within predefined limits within the air traffic approach environment. Speed commands span a valid range between 150 and 300 knots. Altitude commands must lie within the interval from 2000 feet to 400 flight levels. Rates of descent are valid in steps of 500 feet from 1000 up to 3500 feet per minute.

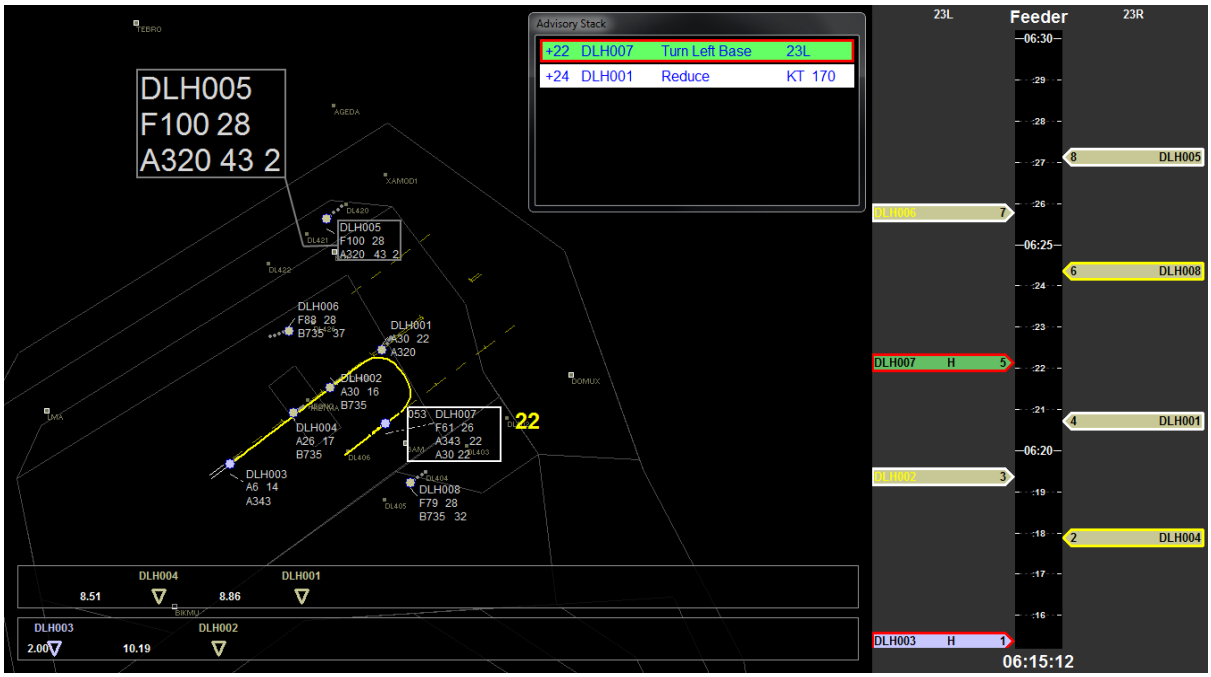


Figure 8. *RadarVision* shows a radar screen with AMAN data in the form of an advisory stack, a timeline with sequences and planned touchdown times, values for distance-to top-of-descent and threshold and a centerline separation range

A heading command consists of valid values between 0 and 360 degrees. In case of a *transition*, a *direct to*, a *handover* or a *cleared ILS* command, the respective route, waypoint, controller position or runway has to exist in the database of the modelled approach control area. During the further process, the combination of type, value and actual aircraft state is checked. An aircraft already flying a speed of 180 knots cannot reduce its speed to 220 knots. This speed value would only be possible when accelerating with an increase command. The same way a command with a descent to FL 90 is negated if the aircraft is already at flight level 80. Such commands will be assigned as invalid.

Speech Recognition Log

DLH2NC	Descend	Alt 3000
AUA151Y	Reduce	KT 150
UAE55	Descend	Alt 2000
BER630	Descend	FL 100
AUA151Y	Reduce	KT 150
DLH2NC	Reduce	KT 210
SXS1CF	Reduce	KT 320
AUA151Y	Reduce	KT 170
BER947V	Descend	Alt 2000
AFR1906	Descend	Alt 2000
BER2356	Turn Left Base	78L
DLH437	Descend	FL 270
BAW938	Descend	FL 110
DLH61A	Reduce	KT 150
UAE55	Descend	FL 70
AUA151Y	Descend	Alt 2000

Figure 9. Speech Recognition Log shows valid, out-of-context and invalid commands as a direct controller feedback

In the second phase of investigating plausibility, the AMAN's context is used. The command is compared with the context information regarding the current timeslot. If the command is not included in this context, the command is marked as "not in context". This could only happen if ASR is used without context information for demonstration and comparing purposes. All other commands that do not fit the range of invalidity or out-of-context are supposed to be valid.

All spoken and recognized commands of the last five minutes then are displayed in a vertical list on the speech recognition log (see Figure 9). Valid commands have a green, out-of-context commands a turquoise and invalid commands a yellow background. The color red is not used here, because a missed recognition should not provoke an alarming effect. During the first five seconds the font of the command text is bold and the background color brighter to highlight the last commands. By using the mouse-over functionality, the background of the aircraft is highlighted in purple (see example "AFR1906 Descend Alt 2000" in Figure 9). All corresponding information of this aircraft is also highlighted on the radar screen, the timescale and in other supporting windows. The recognition log allows the controllers to get direct feedback on how well their commands are understood by the system. Nevertheless, there is no need to repetitively check the evaluation of the recognition output. Therefore, current work attitudes are not intruded into. The visualization of the voice recognition outcome, however, may lead to a cognitive response on good and bad recognition. The personal awareness of a controller may be that strictly following the ICAO aircraft radio regulations causes good results in being automatically understood. However, the recognition rate should still be acceptable when making intentional deviations from the allowed syntax. Perhaps this presents an incentive to achieve better recognition rates by articulating more clearly and closer to the standard.

Another effect on the controller could be a better acceptance of the decision support system AMAN due to more accurate and up-to-date information particularly in high-workload (Cordero et al., 2012) air traffic situations. Low word error rates of the automatic speech recognizer also mean good command recognition quality. The AMAN analyses these mostly correct commands and includes them in the current calculation cycle. Discovering the controller's intent through this data should then lead to a better controller support. Relevant values would be more up-to-date regarding the true plan of the controller. The shown values, as for instance the distance-to-threshold for each aircraft, would become more reliable to the controller. The distance-to-threshold is often useful for controllers and pilots in situations of sequence rescheduling due to emergencies or unexpected behaviors. So controllers could experience some kind of benefits with the described better support by an assistance system using automatic speech recognition. Following this, their acceptance and active encouragement of support systems might rise. As a last consequence even their behavior could adapt to profit from an air traffic controller support system with speech recognition.

PRE-STUDY FOR VALIDATION OF CONTEXT-DEPENDENT SPEECH RECOGNITION SUPPORTED AMAN

Depending on the operational concept (described in chapter "OPERATIONAL CONCEPT FOR USING AMAN WITH ASR") and the correlated validation goals, six scenarios of 45 minutes each, which are clustered in two groups, are envisaged. These two groups differ in traffic and the exceptional event occurring during the run. However, both groups contain a base scenario with no controller assistance and an ASR-scenario with AMAN support. In addition, the first scenario cluster also contains an AMAN scenario with no additional input to validate the benefit of ASR. The second scenario cluster includes an AMAN scenario, where the AMAN can get additional input via mouse and keyboard. By including this scenario, the workload difference between manual input and the automatic input via speech recognition can be validated. Table 1 gives an overview of the six scenarios and their main characteristics.

Table 1: Scenario overview

Base scenario without AMAN	Traffic with 23 arrivals with one priority flight
Standard AMAN scenario	
ASR AMAN scenario	
Base scenario without AMAN	Traffic with 20 arrivals and a short runway closing
Manual input AMAN scenario	
ASR AMAN scenario	

To access the benefits described in the operational concept, several performance parameters describing the air traffic control and the speech recognition processes are analyzed (as listed in Table 2). These parameters include air traffic information, characteristics of spoken commands or different error rates which describe the quality of the speech recognition.

Table 2: Overview of validation goals and accessed parameter

Validation goals	Accessed parameter
Improved prediction quality of ETA (Estimated Time of Arrival)	Comparison of ATA (Actual Time of Arrival) and ETA for each aircraft
Improved vertical flight path	Comparison of flight time and trajectory length for each aircraft
Improved flight duration	
Reduced number of clearances	Number of clearances given
Reduced frequency utilization time	Accumulation of controller radio time
Earlier detection of sequence change	Time, when landing sequence is established in AMAN
Reduced error rate (at least 5%) already improves AMAN performance	
Higher degree of conformance of planned and actual flown trajectory	Deviation between radar position and AMAN position of each aircraft
ASR reduces workload due the absence of manual data entry	Number of mouse and/or keyboard inputs
Number of active controller corrections is reduced	Number of corrections
Time of active controller corrections is reduced	Accumulated time of corrections
Context information keeps response time of speech recognition process in acceptable boundaries	Computation time of speech recognition process with and without context information
Use of context information for robust Digits Error Rate (DER)	Calculation of DER
Context information improves Word Error Rate (WER) as well as Concept and Command Error Rate (ConER/CmdER)	Calculation of Error Rates
Even imperfect context information improves WER and ConER/CmdER	
Acoustic models for non-native speakers improve WER and ConER/CmdER	
Different acoustic models for male and female speakers improve WER and ConER/CmdER	
Increasing robustness against out-of-grammar utterances improves WER and ConER/CmdER	

During the pre-study for validation, which was conducted with a controller from the Düsseldorf approach sector, several aspects of improvement for the validation cycle were detected and implemented. These changes include scenario-related topics like the adaptation of the flight plan, the aircraft behavior or the look and feel of the controller working position, but also validation related topics like the implementation of a training scenario or the alignment of the validation storyboard. With these results the next validation cycle will be approached. The intention is to have three more validation cycles in total, with a steady increase in test persons. The final validation trials will be conducted in early spring 2015 with at least ten controllers.

CONCLUSIONS AND OUTLOOK

Air traffic controllers are responsible for the flight guidance process. They could be supported by decision support systems. The assistance quality of these systems depends on the accuracy of input data, e.g. the update rate of the radar information. Taking into account the controller's intent would improve the support quality. This intent could be derived by analyzing the controller-pilot voice channel via speech recognition. Therefore, the project AcListant® combines an arrival manager (AMAN) with an automatic speech recognizer (ASR). For training of the acoustic model of ASR, speech data is recorded and transcribed to extract flight guidance concepts. The ASR uses AMAN generated air traffic context information to achieve high command recognition rates. The recognition performance is visualized in an HMI as an optional direct feedback for the controller. Due to high speech recognition quality, the assistance system delivers better support. As voice communication is the usual task of an ATCO, no additional workload is caused, which highly increased the acceptance probability. The performed pre-study using the Düsseldorf Approach Area foreshadows that the controller's behavior could change positively with regard to articulation and radio telephony procedures. Higher speech recognition performance improves the AMAN assistance and vice versa.

As a next step it is planned to take the contextual likelihood of each possible air traffic command into account. This would further enhance the ASR performance. Final validation trials at the end of the project should reveal the effect of the AMAN-ASR combination on ATCOs.

REFERENCES

ACARE (2012), "*Realizing Europe's vision for aviation - Strategic Research & Innovation Agenda (SRIA)*", Vol. 2, Sep. 2012.

Cordero, J.M., Dorado, M., de Pablo, J.M. (2012), "*Automated Speech Recognition in ATC Environment*", International Conference on Application and Theory of Automation in Command and Control Systems (ATACCS'2012), ISBN 978-2-917490-20-4.

De Cordoba, R., Ferreiros, J., San-Segundo, R., Macias-Guarasa, J., Montero, J.M., Fernandez, F., D'Haro, L.F., Pardo, J.M. (2006), "*Air Traffic Control Speech Recognition System Cross-Task & Speaker Adaptation*", in: IEEE A&E Systems Magazine, Universidad Politécnica de Madrid, IEEE.

Elzer, P. (2001), "*Recent Developments in the Field of Human-Machine-Interfaces*", in: at - Automatisierungstechnik 49, No. 1/2001, pp. 15–29. – ISSN 0178–2312. – Oldenbourg Verlag.

European Commission (2011), "*Flightpath 2050, Europe's Vision for Aviation Maintaining Global Leadership & Serving Society's Needs -- Report of the High Level Group on Aviation Research*".

Hah, S., Ahlstrom, V. (2005), "*Comparison of Speech with Keyboard and Mouse as the Text Entry Method*", in: Proceedings of the Human Factors and Ergonomics Society, 49th Annual Meeting.

Helmke, H., Ehr, H., Kleinert, M., Faubel, F., Klakow, D. (2013), "*Increased Acceptance of Controller Assistance by Automatic Speech Recognition*", in: Tenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2013).

Hofbauer, K., Petrik, S., Hering, H. (2008), "*The ATCOSIM Corpus of Non-Prompted Clean Air Traffic Control Speech*", in: Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), Eds.: Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Tapias, D., European Language Resources Association (ELRA).

Hössl, M. (2013), "*ATCpredict: Predicting Air Traffic Control Anticipation of Approach Controller Intentions at the Example of Düsseldorf Airport*", DLR-IB 112-2013/40, Bachelor Thesis, University of Applied Sciences Bremen.

ICAO (International Civil Aviation Organization) (2007), "*Doc 4444, ATM/501: Procedures for Air Navigation Services, Air Traffic Management, Fifteenth Edition*", Chapter 12: Phraseologies.

Karlsson, J. (1990), "*The Integration of Automatic Speech Recognition into the Air Traffic Control System (FTL Report R90-1)*", Flight Transportation Laboratory, Department of Aeronautics and Astronautics, M.I.T. Cambridge, Massachusetts.

Paul, D.B., Baker, J.M. (1992), "*The design for the wall street journal-based CSR corpus*", in: Proceedings of the workshop on Speech and Natural Language, ACL HLT '91.

Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hannemann, M., Motlicek, P., Qian, Y., Schwarz, P., Silovsky, J., Stemmer, G., Vesely, K. (2011), "*The Kaldi Speech Recognition Toolkit*", in: IEEE 2011 Workshop on Automatic Speech Recognition and Understanding.

Ramm, F., Topf, J., Chilton, S. (2010), "*OpenStreetMap: Using and Enhancing the Free Map of the World*", UIT Cambridge, ISBN 978-1906860110.

Shore, T., Faubel, F., Helmke, H., Klakow, D. (2012), "*Knowledge-Based Word Lattice Rescoring in a Dynamic Context"*, Interspeech 2012, ISCA.

Zokić, M., Boras, D., Lazić, N. (2012), "*Say Again*", in: International Journal of Education and Information Technologies, Issue 4, Volume 6, 2012, International Association of Computer Science and Information Technology Press (IACSIT Press).

What About the Next Generation? Assessing Experts' Judgments of Human Abilities Required for Working in a Future ATC Environment

Dirk Schulze Kissing and Hinnerk Eißfeldt

Department for Aviation and Space Psychology
German Aerospace Center
Hamburg, Germany

ABSTRACT

With the implementation of the European ATM Master Plan (SESAR, 2012) job requirements for air-traffic controllers (ATCOs) will alter significantly. Especially the introduction of time-based operations is considered to impose large changes in task operations. A controller-in-the-loop simulation study was designed to assess experts' judgments on the main human abilities required for working in time-based arrival scenarios with different levels of assistance by an arrival-management system. Five male controllers (mean age: 32.5 years; mean job experience: 10.8 years) licensed for the simulated sector performed on one baseline scenario and four future scenarios. Gaze-data were tracked and questionnaires on situation awareness and workload, as well as a modified Fleishman Job Analysis Survey (F-JAS) were applied. The experts judged the ability to identify problems produced by automation, as well as the ability to stay vigilant as being increasingly required in a highly automated time-based environment. The high number of gaze transitions between aircraft-targets adds objective indication: Under the simulated future conditions ATCOs were hardly able to build up attention guiding expectations which are necessary to keep up with the system- and traffic-parameters. This was also reflected in low situation awareness and high workload ratings.

Keywords: Strategic Job Analysis, Eye-Tracking, Air-Traffic Control, Controller-In-The-Loop Simulation, Supervisory Control, AMAN

INTRODUCTION

With the changes that are discussed for future air traffic management (ATM) the roles of air-traffic controllers (ATCOs) will alter significantly. For example, it is assumed that an exchange of accurate data between ATM actors will enable to build an optimum traffic sequence on a pre-tactical level, so significantly less tactical intervention by ATCOs will be needed for traffic control, in both En Route sectors and terminal maneuvering areas (TMA). Within the Single European Sky ATM Research (SESAR) program it is assumed that the ATCOs tasks may generally become more strategic in nature.

Different evolutions of the future ATM systems with different automation design philosophies are considered in literature, implying different ATCOs roles, from scenarios where the human is the assisted by advanced automation (e.g. in SESAR) up to far looking concepts where the computer takes the controller role. From the human-factors design point of view it has to be guaranteed that, whatever the ATCOs future role will be it has to be in accordance with human capabilities. From a personnel selection point of view it is then essential to proactively assess how such evolutions of the human role may change the required abilities to perform in the job, so selection profiles for ATCOs will be modified accordingly in time.

This paper focuses on analyzing the job requirements and their corresponding abilities in a hypothesized future approach control setting. Several scenarios are constructed in which controllers are supported by automation, with different levels of assistance up to the point where the human has to supervise a self-executing automation. The roles under scrutiny all referred to the tactical level. The ATCOs were focused on traffic situations in TMA as displayed at the controller working position (CWP). A change in task allocation between the human operators was also introduced.

As monitoring is supposed to be a main occupation not only in the current but also in a future control tasks (Broach, 2013), ATCOs' gaze-data were tracked to gain insight into the controllers' visual sampling strategies. The impact of these working conditions on situation awareness and workload are measured for control.

THEORETICAL BACKGROUND

Future Oriented Job Analysis

With the effort that is put into the deployment of advanced automation into ATC the prognosis of abilities for future ATCOs is co-emerging as a relating human factors topic. Alexander et al. (1989; Ammerman et al., 1987) conducted a job analysis in which they assessed how future controller work would change after the introduction of advanced automation into ATC. Nickels (Nickels, Bobko, Blair, Sands, & Tartak, 1995) developed a list of future ATC-work requirements which was the foundation of a selection system for the hiring of air-traffic controllers. Manning and Broach (1992) conducted a strategic job analysis to assess the changes of the ATCO-job as it is predicted to exist in the future as a result of increasing automation. After an analysis of the present job, they gathered information about the kinds of issues, like technological and organizational developments which may affect the job in the future. Eißfeldt (2009) used a technique for future oriented job analysis first introduced by Schneider and Konz (1989). They held several workshops with subject matter experts (SME) to discuss future issues of ATCO and pilots jobs. After being primed with the discussion outputs the SME filled out the Fleishman Job Analysis Survey (F-JAS, Fleishman & Reilly, 1995) once rating the requirements of the job in the present exists and once for the job of the future. Because none of them had ever performed the job using the future automation, it is unclear how valid these SME predictions of the future job requirements are. It may be advantageous to provide the SME with some experience on using new automation to gain higher validity. Eißfeldt, Deuchert and Bierwagen (1999) followed this approach. They assessed future job requirements of ATCOs under simulated future ATC conditions. The simulation features were data-link communication, a stripless system, planning-tools, and automatic conflict detection aids. In a one hour exercise four brief simulation scenarios were performed by the SMEs. Before and after a session they filled out the F-JAS. Eißfeldt et al. (1999) found a general decrease of cognitive requirements for future ATC, with the strongest effects occurring for oral expression and oral comprehension, written expression, and number facility. These decreases were attributed to the effect of data-link communication.

Broach (2013) recently conducted a strategic-job analysis for the tower controller working position. He used the overall profile of abilities derived by Nickels et al. (1995) to compare current aptitude to future aptitude requirements. He performed an extensive review of current aptitude profiles and of assessable documentation on technology changes which are expected to be introduced until the year 2018. Based on the current profile and the expected technology changes he infers a future aptitude profile. Broach resumes that scanning of visual sources will be one of the abilities which will become more important for future performance.

New Roles and Responsibilities for ATCOs

As an increased level of computerized support for ATCOs appears to be mandatory to maintain safety and efficiency of the future air traffic operations (Kirwan, 2001), it has to be considered what the future role of the human should

be. In the time-based environment proposed in the SESAR ATM target concept the task of tactical vectoring in the TMA in a way that all aircraft (a/c) exactly meet their preplanned times of arrival is supposedly beyond the ATCOs current performance envelope. So for a 4D-plan conformant tactical traffic handling one will have to take advantage of the strength of automation processing power to quickly calculate the optimal trajectories and assign the referring advisories.

In a human-centered approach for TMA design automation is assumed as a means to expand the ATCOs functional envelope by calculating a/c trajectories with respect to the landing time on ground and providing the ATCO with according tactical guidance advisories. Automation is not expected to take over the whole task, as trajectory management in the TMA also requires a level of adaptability, flexibility, and problem-solving skills which currently can only be provided by the human operator. To keep the human within the decision loop in such a system-support setting it has to be guaranteed that automation supports the ATCOs in developing a mental picture of the traffic situation (Mogford, 1997), instead of degrading it or even removing the picture from the human domain into the automation.

Another paradigm is to let the ATCOs chose what is to be done and delegate the task execution to other agents (humans or automation) with the possibility to intervene if required. Such a role change is envisaged in the SESAR concept of self merging and spacing a/c which are suited with an airborne separation assurance system (ASAS). Here the ATCOs delegate certain spacing and positioning tasks in relation to a specified target a/c to the pilot, who takes the responsibility for identifying the target a/c and establishes separation based on instructions from the ground. With the delegation of specific tasks to other agents it is assumed that ATCOs will be relieved so they could adjust to a more strategic work-style in trajectory management (SESAR, 2012).

In some more far-reaching concepts it is supposed that in the long term future automation will take over (Truman & de Graaff, 2007). Arnaldo et al. (2012) discuss the question whether humans will remain within the future trajectory management processes and will perform some strategic functions where their skills come into play, or if tasks which include tactical decisions in the future may imply autonomous and fully automated processes. As the ATCOs would only be able to assess if the automation advisories are reasonable on a safety level, they would rather be reduced on the safety monitor role where the automation controls and the human monitors only, and over-rides when necessary. In this case the ATCO will have the role of the supervisor of automation. Truman & de Graaff (2007) also envisage that the human intervention and control will be at an absolute minimum in such future ATM system which are self-monitoring and self-controlling.

ATCOs' Monitoring Behavior

Broach (2013) resumes that scanning of visual sources will be one of the abilities which will become more important for future ATC job performance. Thus, the strategic job analysis presented here particularly focusses on the ATCOs' visual behavior. Visual scanning can be equalized with situation awareness on level 1, which is by definition 'the perception of the elements in the environment within a volume of time and space" (Endsley, 2006). For ATCOs working with advanced automation the elements to be scanned not only comprise all moving elements within the sectors, but also the automation itself. Parasuraman, Sheridan and Wickens (2008) suppose that operators scan the raw information sources less frequently when tasks are delegated to an advanced automation. Rovira and Parasuraman (2010) accordingly assume that in this case ATCOs also exhibit fewer fixations to the radar display compared with manual control, because they invest their freed cognitive resources to focus on secondary tasks. Generally, good monitoring behavior in ATC requires both focused and distributed attention (Moore & Gugerty, 2010). Even while ATCOs are focusing on a specific area of the radar screen, they must continue to be aware for what is occurring around their sector. Consequently, controllers should not neglect to utilize both focused and distributed attention strategies.

METHOD

In the current experiment the approach of Eißfeldt et al. (1999) is followed to assess ATCOs judgments about future job requirements after they experienced some future ATCO roles. Additionally, their interaction with the simulation environment was assessed by tracking their eye movements to gain further objective indication for changes in work-style. As the most important information for an ATCO is related to the moving targets on a situational display, the analysis focusses on attentional switches (transitions) between dynamic area of interests (dAoI) which are defined by the moving zone each a/c blip with related data-block covers on the radar screen (cf. Gross, Friedrich, &

Möhlenbrink, 2010). The analysis of fixation transitions within experimental runs is a means to investigate the co-activation of a/c within the overall traffic-representation, the mental picture, which is assumed to underlie the ATCOs monitoring behavior. As the controllers picture is relational in nature (Niessen & Eyferth, 2001), it could be claimed that a number of transitions between a/c reflects the activation of meaningful relations within this overall traffic-representation. Additionally the transition number presumably also indicates the priority of that constellation with regard to expectations or intended actions.

Research Question 1 (explorative): What do experts assess to be the main changes in job abilities required for their working position after performing new roles in highly automated scenarios within their familiar TMA environment?

Research Question 2: Is there evidence for changes in work-style when the new roles imply that central task of TMA management are delegated to cockpit and/or to automation

Hypothesis: Under task delegation conditions, the ATCOs use freed resources to switch to a more strategic work style, i.e. to focus on secondary tasks or to attend to what is occurring around their sector. Consequentially, fewer fixations to the radar display compared with manual control and/or more transitions of fixations between distant a/c on the radar screen are expected. A more strategic work style should also be reflected in a comparably higher level of situation awareness. The freeing of resources should also be indicated in lower levels of experienced workload.

To allow a prognosis of what changes are up to come the key features in the European ATM Master Plan (SESAR, 2012) were taken into consideration. Time-based operations are the key concept to be introduced with SESAR Step 1 (interval-based metering). In a time-based environment the initial synchronization of the traffic arriving in high traffic density TMA will be generally ensured via the allocation of controlled times of arrival (CTAs) on initial approach fixes (IAFs). The introduction of a/c capable of meeting a CTA with appropriate accuracy improves the performance and reliability of an Arrival Manager (AMAN) system. AMANs normally are a means to avoid overcrowded TMA by controlling the flow en-route before reaching the initial approach fixes (IAF), where a/c are handed over from area control centers (ACC) to TMA. However, the DLR AMAN research prototype 4D-CARMA (Ehr & Uebbing-Rumke, 2013), which is classified as a strong executing system which is compliant to the SESAR requirements, also is calculating trajectories with respect to the landing time on ground and generates according tactical guidance advisories. These advisories are timely precise shown to the approach controllers at all approach phases on the radar screen. From the ATC-perspective the implementation of the concept of time-based operations may make a strong execution of automated systems necessary.

Experimental Design

Several kinds of future scenarios were constructed with the AMAN as the central element to create different roles for ATCOs. The AMAN can be configured to various levels of automation (Sheridan & Verplank, 1978; Endsley, 1987; Endsley & Kaber, 1999). On the five-level LoA hierarchy proposed by Endsley (1987) the AMAN can be configured to the LoAs ‘manual control’, ‘decision support’, and ‘monitored AI’. In the Endsley and Kaber (1999) taxonomy this would correspond to the LoAs ‘manual control’(MC), ‘decision support’(DS) and ‘supervisory control’ (SC) as the second highest level of automation. However, these levels of automation correspond to different ATCOs roles. They can be assigned by using a taxonomy proposed by in the RHEA (Role of the Human in the Evolution of ATM systems; Nijhuis, 2000) project. Accordingly, in four experimental conditions the ATCOs fulfilled the following roles:

- Experimental run 1 (baseline): ATCO handles traffic according to the first-come-first-serve principle
- Experimental run 2: ATCO delegates responsibility of certain tasks to suitably-equipped a/c and aircrew
- Experimental run 3: AMAN predicts what needs to be done for certain tasks and presents proposals to the ATCO
- Experimental run 4: AMAN controls – ATCO monitors only, and over-rides when necessary
- Experimental run 5: Automation controls - ATCO monitors only, and over-rides when necessary, with both working together in the control loop during an emergency event.

In the experimental runs only incoming traffic was simulated. All conditions were simulated within the structure of the sector Frankfurt (FRA) Approach (DFAPP) in its configuration valid until October 2011 (i.e., no third landing runway integrated). DFAPP designates the sector in the terminal maneuvering area (TMA) that gets all a/c with

destination FRA Airport (EDDF). These flights approach the area via four different arrival routes and Initial Approach Fixes (IAF). These are OSMAX (western sector inbound), ROLIS (northern sector inbound), GEDERN (eastern sector inbound), and PSA (southern sector inbound). The arriving a/c are lined-up on the downwind and then 'fed' onto the final approach. The sector controls all levels from surface up to flight-level 115. The arrival routes are designed in a way so as not to interfere with each other. In the experimental trials only a sector configuration for using runways 25 is applied.

Experimental run 1 (baseline) resembled current operations in which controllers had standard systems support. The AMAN was running during the baseline scenario because it is a standard tool in all Langen ACC sectors. Experimental run 2 (ATCO delegates Control to A/C) also resembled current operations from the point of system support, but the concept of a/c controlled merging and spacing was introduced for the piloted aircraft. From the pilots' perspective the new requirements during in-flight procedures were (1) to meet the target time of overflow at the IAF, calculated on ground by AMAN and displayed in their CDTI, and (2) to follow the Merging and Spacing instructions given by ATC, which comprises to keep up distance and time to their target a/c ahead. In experimental run 3 (ATCO Controls - AMAN Proposes) the AMAN supported the controller by recommending level, speed and turn-to-base advisories and their timing to meet the time-based criteria for the incoming flights. The AMAN advisories for ATCOs were displayed within the corresponding flight data blocks (= Labels) on the radar screen. From the cockpits perspective run 3 functioned as a repeated measurement with the same requirements as in experimental run 2. In experimental run 4 (AMAN Controls – ATCO Monitors) all of the controller's tasks, i.e. determining speeds, headings, and levels to the a/c and building an arrival sequence are delegated to the AMAN. The work of the controller here moves toward supervisory control. From the pilots perspective in comparison to the baseline requirements change in receiving ATC advisories from ground-based automation via data-link and reading it back via radiotelephony (r/t) to keep the supervising controller informed about their intent. The ATCO monitored the r/t and always was able to intervene via r/t. In experimental run 5 (AMAN and ATCO in Control Loop) again, the controller was in the supervisory role but had to intervene due to a simulated emergency. From the pilots perspective the requirements were analogue to experimental run 4.

In all experimental task-settings the participants were advised to control or monitor the incoming traffic for the whole TMA. Generally the participating ATCOs were advised to guide the down-winding traffic via the standard arrival routes (STARs). All controllers reported to be very familiar with this assignment, as this is a standard setting in real-life during periods of lower traffic load. The participants also worked in a one-man setting without the assistance of a coordinate controller. The experimental baseline scenario (experimental run 1) is oriented at the actual method of operation of approach controllers. The ATCOs tasks were to receive incoming traffic at the IAF, coordinate these a/c according to the traffic flow and provide sufficient lateral and vertical separation between a/c, as well as guiding them to the final approach. The AMAN was active, but just displaying a planned sequence on a time ladder. To the subsequent experimental runs the CTA and merge-&-space operation, a strong executing AMAN, and CPDLC-communication were selectively introduced. For the first future condition (experimental run 2) the AMAN was active, but not in a strong executing mode. The job includes all core-tasks of the baseline, added by the tasks related to the merge and space procedure, like instructing the pilots and monitoring the on-going operations. For the next, more advanced future condition (experimental run 3) the AMAN was set into strong executing mode. In this setting the ATCO is advised to generally instruct the pilots with the AMAN advisories, if no safety concerns would stand against it. Also the core tasks to determine headings, level, and speeds according to the intended approach sequence where suspended as AMAN made recommendations to follow. However, the list again was expanded by the new core-task to evaluate AMAN advisories before communicating them via r/t. For the most advanced future condition (experimental run 4) the AMAN was set into a self-executing mode. Almost all core-tasks are now performed by the AMAN via data-link communication with the pilots. In the role of supervisory controller, the new core-tasks are to monitor AMAN performance on the AMAN display, monitor a/c behavior on the radar screen, monitor r/t communication, and to intervene to take manual control when necessary should emerge. This most advanced future scenario was performed a second time with shuffled a/c signs and a provoked emergency event that forced the ATCO to perform tactical interventions (experimental run 5).

Simulator Preparation

The Human-in-the-Loop Simulations were performed at DLR's integrated Air-Ground Simulator AviaSim (Bruder et al., 2013). The ATCO is seated in front of four displays with the radar screen as the main window in front. On the left the electronic flight strip bay is located. On the right the 4D-CARMA generated timeline for the arriving a/c, a display for the advisories the automation currently is sending to a/c, as well as a scale showing distances of a/c on the final approach (the so called mileage-metering scale) are shown. The head-up display shows labels of a/c approaching the TMA on a timeline according to their controlled times of overfly (CTOs) at the referring IAFs. All

a/c are simulated by pseudo pilots except for three that are simulated by fully equipped simulation cockpits operated by airline pilots. Pseudo pilots operate the non-piloted a/c according to standard procedures until they are transferred to the TMA. At the beginning of the scenario a pseudo ACC-controller overtakes communication with piloted a/c and transfers it to the ATCO when the a/c reaches the TMA border.

Dependent Variables and Procedure

The experiment was conducted in 5 sessions over three days. In total five controllers from Deutsche Flugsicherung GmbH (DFS) Area Control Center (ACC) Langen with valid licenses for TMA Frankfurt and 15 pilots from various airlines participated. Only controller data will be reported here. The controller sample was all male with a mean age of 32,5 years (SD: 6,42), and a job experience of 10,8 years in the mean (SD: 6,85). A session comprises four runs of a future scenario, one baseline scenario, and five training runs for the controllers or three training runs for pilots, respectively. A single simulation run involved one controller, three pilots and two pseudo pilots. The controllers were working the whole Frankfurt Terminal Maneuvering Area (TMA) in a one-man operation modus. Pilots were starting airborne within the lower airspace and entered TMA via three different initial approach fixes (IAF). The controllers, pilots and pseudo-pilots communicated using r/t or CPDLC/Data-Link transmissions.

For the controllers the first exercise day started with an introduction of questionnaires, calibration on the eye-tracking system and an about 90-minutes in-depth walk-through the F-JAS dimensions. The second part of the first day included two initial trainings runs of about one hour to get acquainted with the simulator and the Arrival Manager (AMAN) interface (first trainings run) and to familiarize with pseudo-pilots interactions (second training run). The second day started with a recapping trainings session in which the merging and spacing procedures were practiced with pseudo-piloted a/c. This was followed by a checkup of the controllers' dynamic vision abilities. Afterwards the first joint trainings scenario was executed to familiarize controller, pilots, and the pseudo-pilots cooperation. This was followed by a second joint training in which merging and spacing procedures were practiced with the piloted aircraft. The measured traffic sample contained 24(+x) flights with destination Frankfurt (EDDF), with at least six a/c entering the TMA on each of the four IAF during the scenario run. The scenario contained a mix of a/c of heavy and medium weight class planned to land in staggered mode on the two parallel runways 25R and 25L, which have to be operated dependently. Next, the controllers executed the five experimental runs, one at the end of day two and four on day three. For comparison reasons, the same traffic sample was used for all runs. However, to avoid familiarization with the traffic sample, all a/c signs were shuffled or substituted between each run. Furthermore, pilots switched working positions after each run to let them experience different traffic situations. Experimental run four and five were always performed consecutively (en bloc) and the baseline run was always performed first, i.e. at the end of day one. The runs 2, 3 and the block 4+5 were counterbalanced to avoid practice effects. Each run was performed only after a break.

Quantitative and qualitative data were gathered during and after each simulation run. As a subjective measure of workload and situation awareness the instantaneous self-assessment (ISA; Kirwan & Flynn, 2001) was used. Controllers were prompted four times during a scenario (appointed according to the flight phases of the piloted aircraft, i.e. RTA IAF, Merge, Follow, and Turn) to give their ratings on current situation awareness and workload on two three-point scales. During the simulation runs eye-movements were tracked to gain indication for participants' visual sampling behavior, like the change from focused to distributed sampling. The transitions (switches) of fixations between a/c (dAoIs) within trials are used as an indicator for the actual relational structure of the ATCOs mental picture. The fixation duration on objects is used as an objective measure of situation awareness related to these objects. Immediately after each simulator run, the participants filled out (a) a 14-Component Version of the Situation Awareness Rating Technique (SART; Taylor, 1990) , (b) the NASA Task-Load Index (NASA tlx; Hart & Staveland, 1988), a questionnaire to assess the impact on mental workload (AIM; Dehn, 2013), and the SHAPE Automation Trust Index (SATI; Dehn, 2013). At the end of the last experimental run of exercise day two and exercise day three the participants filled out a modified version of the Fleishman Job Analysis Survey (F-JAS; Fleishman & Reilly, 1995) and cognitive interviews were held directly afterwards to gain in-depth information about experienced job requirements. Additionally, a questionnaire on user satisfaction and acceptance was completed at the end of day three.

RESULTS

The controllers agreed on several usability scales that the system was easy to learn, and that the time needed to learn

using the system was adequate. The controllers also agreed on the statement that the scenarios they had worked on were interesting and inspiring. They regarded the traffic-load of the scenarios to be comparable to the current real-life traffic.

Questionnaire on Job Ability Requirements

The ATCOs judgments on the abilities required for future job performance were assessed once for the presented future scenarios (overall) and compared to the ability-profile the experts initially identified for the baseline scenario. These ability-requirement profiles are put together from a of 37 ability-requirement scales the ATCOs filled-in, with 32 of them selected from the Fleishman Job Analysis Survey (F-JAS), supplemented by five scales covering additional requirements developed by DLR in a format similar to F-JAS. A paired-sample t-test on the single scales

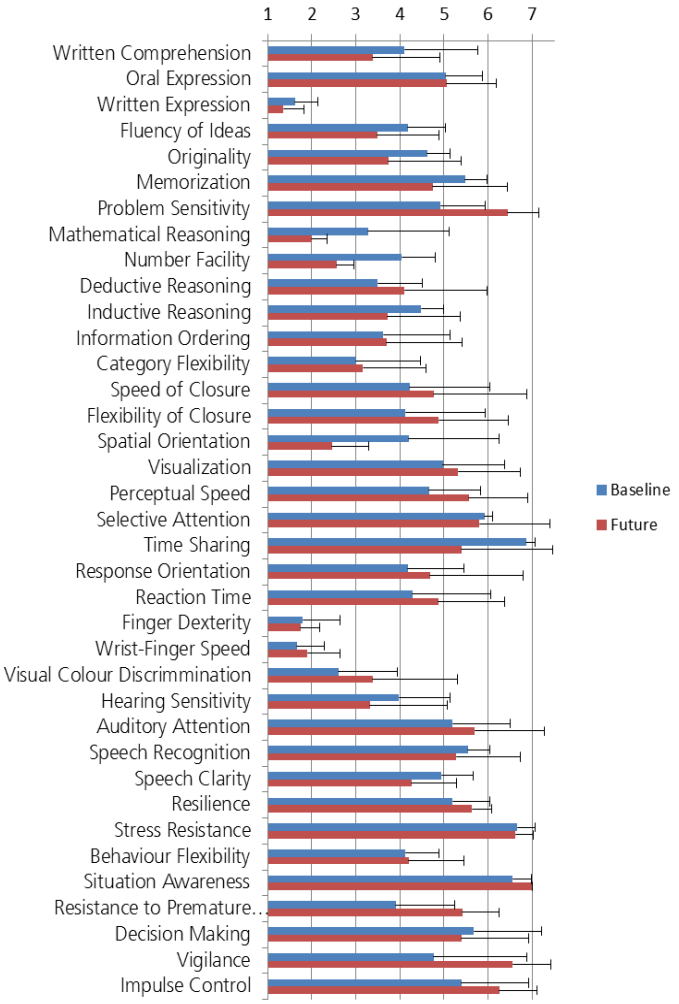


Figure 1: Mean and standard errors of SMEs (n=5 ATCOs) F-JAS type scale ratings on the 37 ability requirement dimensions plotted for the (a) the simulated baseline scenario and (b) the simulated future scenarios

revealed a significant effect of future changes (versus baseline) on required Problem Sensitivity (cognitive scale; future increase) and Number Facility (cognitive scale; future decrease) abilities (see Table 1). With respect to practical relevance the marginal differences in required Spatial Orientation (cognitive scale; future decrease), Perceptual Speed (cognitive scale; future increase), Situation Awareness (interactive/social scale; future increase), Resistance to Premature Judgment (interactive/social scale; future increase), and Vigilance (future increase) abilities may also be considered (see Table 1). A Wilcoxon signed-rank test on the rank of means for the 38 ability requirement scales showed no fundamental change in the ability requirement profiles from current (baseline) to

future scenarios. Figure 1 presents the means and standard errors of SMEs' F-JAS-type scale ratings related to the baseline and the future scenarios on the 37 ability requirement scales. For the baseline scenario time sharing received the highest mean rating and is thus considered as the most important ability required for the current job, which is a typical finding.

Table 1: Mean ratings of F-JAS type scale ratings for the baseline and future scenarios. Only those scales are listed for which a paired-sample t-test yielded a significant or marginal within-group effect (N=5)

F-JAS Scales	ATC				
	Baseline	Future	T	df	P (2 way)
Problem Sensitivity	4,92	6,44	- 3,919	4	.017
Number Facility	4,04	2,56	.2,929	4	.043
Spatial Orientation	4,20	2,46	2,400	4	.074
Perceptual Speed	4,66	5,56	- 2,535	4	.064
Situation Awareness	6,56	7,00	- 2,269	4	.086
Resistance to Premature Judgment	3,90	5,42	- 2,513	4	.077
Vigilance	4,76	6,56	- 2,126	4	.066

In the resulting future ability-requirement profile, Situation Awareness[1] - which received the highest mean ratings - followed by Stress Resistance[2], Vigilance[3], Problem Sensitivity[4], Emotional Control[5], Selective Attention[6], Auditory Attention[7], Resilience[8] and Perceptual Speed[9] received higher values than the ability of Time Sharing[10] (cf. Figure 1). The increasing importance of Situation Awareness, Vigilance, Problem Sensitivity, and Perceptual Speed should be considered as important developments in task requirements, as they are ranked in the top ten required future abilities when compared by their mean ratings (cf. Figure 1). Significant future decreases (see Table 1) were only determined for scales of minor importance (namely Number Facility and Spatial Orientation).

Gaze Analysis

Transitions. In Figure 2 for each experimental condition the state-transition matrices containing the total number of gaze transitions between a/c are displayed. The referring a/c are labeled according to their position in the landing sequence. For each experimental condition the degree an a/c was co-activated (and thus related) with the other a/c can be read out by first tracing the respective line and then the respective column of the a/c. For example, to determine for the baseline condition the degree the a/c in the fifth position of the landing sequence is co-activated with other a/c, the counted transitions for the preceding a/c can be read out in the line labeled '5th' (counts for a/c ordered in their spatial distance to the a/c under consideration are 26, 14, 34 and 11), the transition numbers for the succeeding a/c can be traced in the column correspondingly labeled '5th' (counts for a/c ordered in their spatial

1 Ability to perceive the current automation-traffic interaction, understand the reasons behind, and project the future development.

2 Capability of dealing with stress situations in such way that control is maintained and the objective achieved.

3 Ability to maintain attention and alertness over prolonged periods of time

4 Ability to tell when something is wrong or likely to go wrong.

5 Ability to stay in self-command in emotionally challenging situations, when irritating, unexpected, difficult or stressful factors occur.

6 Ability to concentrate on a task without getting distracted.

7 Ability to focus on a single source of auditory (hearing) information in the presence of other irrelevant and distracting sounds.

8 Ability to rapidly recover normal energy and enthusiasm following a discouraging situation.

9 Involves the degree to which one can compare letters, numbers, objects, or patterns, both quickly and accurately. This ability also includes comparing a presented object with a remembered object.

10 Ability to shift back and forth between two or more sources of information.

distance to the a/c under consideration are 37, 24, 12, 14, 5, 13, 0, 2, 0, 0, 0). To simplify the interpretation the colors indicate the suspected co-activation of these relations in the ATCOs mental picture, from red (indicating high relevance) to white (medium relevance) to blue (minor or no relevance). Obviously, for all experimental conditions the ATCOs allocated much attention to relations of a/c neighbors on the final approach. Maximum transition numbers between neighboring a/c are counted for (a) the experimental condition in which the ATCO delegated responsibility of certain tasks to suitably-equipped a/c and (b) the condition where the AMAN controlled and the ATCO monitored only, and over-rode when necessary. However, for both of these conditions, and especially for the supervisory control condition, even relations between distant a/c (preceding and succeeding) received more attention than in the baseline condition.

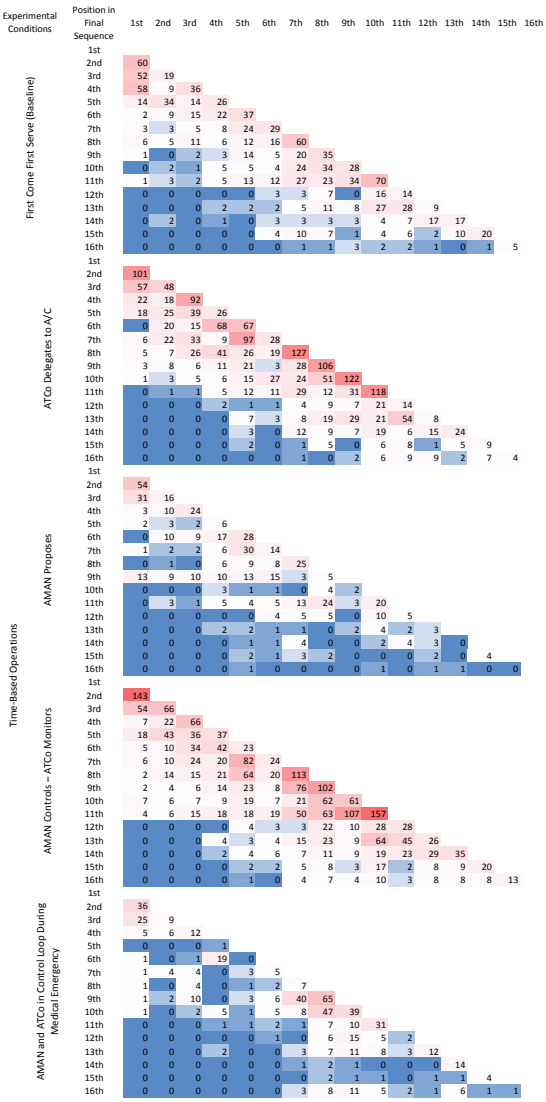


Figure 2. State-transition matrices containing the total number (summed up for N= 5) of gaze transitions between dynamic areas of interest (blip and data block) representing the single a/c on the radar monitor during an experimental run

Fixations per screen. The numbers of fixations counted for the radar display for all experimental conditions are all higher compared to the baseline (cf. Figure 3). This is contrary to the directed hypothesis predicting that the number of fixations counted for the radar display should decrease when trajectory management tasks are delegated to other agents. As the data was not normally distributed the most appropriate way to statistically test the differences from the baseline was Wilcoxon signed-rank test. There was a significant increase from baseline to the 3rd condition (ATCO Controls - AMAN Proposes) in the number of fixations counted for the radar screen, Z = -2,023, < .05, and also to the 5th condition (AMAN and ATCO in Control Loop: Z= - 2,023, p< .05). The 4th condition (AMAN

Controls – ATCO Monitors) only marginally differed from the baseline (Z= -1,753; p= .080). Thus, there were changes in visual sampling strategies (and work style) during scenarios, but not as expected. As flight strips information has relevance for the ATCOs planning of future incoming traffic the numbers of fixations counted for the strip-bays display should increase in case the ATCOs have used free attentional resources to focus on some more strategic considerations. Figure 3shows that under all experimental conditions the ATCOs only rarely attended to this planning information. However, there was a significant increase from baseline to the 5th condition (Z= - 2,032, p= .042) indicating some planning activity was involved during the emergency handling. As also can be seen in Figure 3 the AMAN (4D-CARMA) screen was more attended to under the 4th (Z= -1,753; p= .080) and 5th condition (Z= - 2,023, p= .043). This change in visual sampling is mainly attributable to the ATCOs' close monitoring of the HMI the automation was displaying its instructions given to the cockpits.

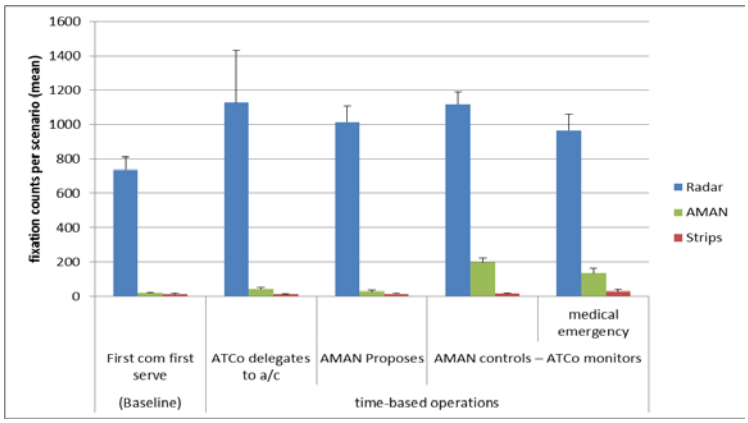


Figure 3. Number of fixations on each of the three monitors of the controller working position for each scenario run

Control Variables

Repeated Measures Analyses of Variance (rANOVA) were used to statistically compare the questionnaire outcomes for the various runs. Partial eta-squared (η_p^2) is given as a measure of effect size. For each analysis $p < .05$ was used.

Workload. For each run workload measures with NASA TLX were obtained (SME). An ANOVA on the NASA TLX Overall Scores for the five runs showed only a marginal effect [$F(4,16) = 2{,}694$, $p < .10$, $\eta_p^2 = .402$]. The within subject single contrasts showed that experimental run 5 imposed significantly higher workload compared to the baseline condition [$F(1,4) = 9{,}476$, $p < .05$, $\eta_p^2 = .703$]. No effect for the five runs on the AIM total score was measured. Single contrasts indicate significantly higher demands on monitoring for experimental run 2 (ATCO Delegates to A/C) compared to experimental run 1 (baseline), $F(1,4) = 16{,}000$, $p < .05$, $\eta_p^2 = .800$. A 4 (Moment of Measurement) x 5 (ATCO Role)-factorial rANOVA for the ISA scores showed a significant effect for the factors 'Moment of Measurement", $F(3,6) = 14{,}600$, $p < .05$, $\eta_p^2 = .880$. These findings contradict the hypothesis trajectory management task delegation would release the ATCO. There is a distinctive trend for a workload increase with the scenario progressing. The experimental factor (ATCO Role) had no significant effect on workload. Single contrasts revealed a significant higher ISA workload ratings only for the experimental run 2 (ATCO Delegates to A/C) compared to run 1 (baseline), $F(1,4) = 64{,}000$, $p < .05$, $\eta_p^2 = .970$.

Situation Awareness. An ANOVA for the 5 runs showed a significant effect for the 10-D SART scale values, $F(4,16) = 4{,}063$, $p < .05$, $\eta_p^2 = .504$. Single comparisons with the baseline run revealed significantly lower situation awareness under 3rd experimental run (ATCO Controls - AMAN Proposes: $F(1,4) = 9{,}732$, $p < .05$, $\eta_p^2 = .709$), as well as under the fourth experimental run (AMAN Controls – ATCO Monitors: $F(1,4) = 24{,}500$, $p < .05$, $\eta_p^2 = .806$). This is contrary to the directed hypothesis which predicts the allocation of attentional resource released by tasks delegation to enhance situation awareness. A 4 (Moment of Measurement) x 5 (ATCO Role)-factorial ANOVA for the ISA scores on 'situation awareness' showed a significant effect for the factors 'Moment of Measurement', $F(1{,}931, 3{,}861) = 8{,}826$, $p < .05$, $\eta_p^2 = .815$. There is a distinctive trend for situation awareness to decrease with the scenario progressing. There was no significant effect for the experimental factor (ATCO Role).

CONCLUSIONS

The ability analysis revealed the specific job requirements which occur for ATCOs when moving from the current to a future TMA-task profile. The SME experienced higher sensitivity with regard to problems produced by automation will be required (problem sensitivity). In their opinion a system-supervising ATCO has to continuously assess if the automation instructions are reasonable or not. This requires higher abilities to stay alert over a long period of time (vigilance). In the passive supervisor roles ATCOs experienced they were unable to from expectations on traffic behavior on the basis of their own intentions, so they were more heavily occupied with tracing and integrating automation behavior into their mental picture to understand what is happening and project what the system is going to do. According to the SMEs' appraisal these new job requirement more heavily goes back to the abilities to obtain situation awareness.

The gaze transitions between dynamic areas of interest representing the single a/c on the radar monitor are interpreted to indicate the relational structure of the ATCOs current mental picture. Higher transition numbers show a/c relations which are more closely monitored, relate to a/c which are more frequently updated, and thus refer to a/c representations which are more highly activated in the ATCOs mental picture. The transition counts show most highly activated a/c relations for neighboring a/c on the final approach. They are most often monitored in succession. The highest counts for transitions between those focal a/c relations were observed for scenarios (a) where the ATCO delegated separation tasks to a/c and (b) when the ATCOs monitored the automation controlling the traffic. These findings may indicate how monitoring behavior is changing when it is no longer guided by the ATCOs intentions, but by their expectations and recurrent checks of other agents (pilots or automation) actions. This increase in focal switches between relevant a/c constellations reveals more frequent updates of focal elements in the ATCOs mental picture. The increase in number of transitions between distant a/c also indicates higher distributed-attention requirements for both scenarios. The ATCO more frequently had to update what else was going on in the sector when other actors also were in the decision loop.

As flight strips information has relevance for the ATCOs planning of future incoming traffic, it was expected that the numbers of fixations counted for the strip-bays display should increase in case the ATCOs delegates tasks and has free attentional resources to focus on some more strategic considerations. Under all experimental conditions the ATCOs only rarely attended to this planning information. So the ATCOs generally only took the tactical level into account. Only in the experimental condition the ATCO had to handle an emergency a/c gaze data suggest some planning activity was involved. However, questionnaire data did not proof a workload-reducing effect of task delegation. In fact, ATCOs experienced higher workload which might be attributable to higher monitoring efforts. Even for those experimental future conditions where the ATCOs supposedly updated their mental picture more frequently questionnaire data indicate that situation awareness was deteriorating. Thus, changes in visual sampling strategies and work style can be attributed to the introduction of the future experimental conditions, but they were not as they were expected in the hypothesis.

It may be questioned if receiving tactical advises from automation to comply with 4D-plan requirements is a means to expand the ATCOs functional envelope. When ATCOs in our simulation were advised by automation when and how to act, they were hardly able to assess if the automation advisories were reasonably 4D-plan compliant. As the ATCOs actually had to decide on situational matters which they could not assess, they were only acting as a transducer of system advisories to the pilots. This might have caused their experienced decrease in situation awareness. One could argue the system could equally well directly interact with the receiver in this case, namely the a/c. However, the monitoring of such a system would require keeping up with a large quantity of rapidly changing system parameters. Based on the current data this would imply a shift of job requirements for TMA ATCOs to a higher demand for vigilance, paired with high demands for perceptual speed, problem sensitivity and situation awareness. If such a shift already can be expected with the introduction of 4D-trajectory management will depend on system design decisions. Whatever these decisions may be it is recommended here that the human performance assessment in the system validation activities should also encompass the use of job analysis tools. This is to ensure that ATCO selection will sooner than too late look out for employees who are capable to fill in the roles the new ATM system assigns to the human operator.

REFERENCES

Alexander, J. R., Alley, V. L., Ammerman, H. L., Fairhurst, W. S., Hostetler, C. M., Jones, G. W. et al. (1989). FAA air traffic control operations concepts Volume VII: ATCT tower controllers Washington, DC: Federal Aviation Administration Advanced Automation Program (AAP-110).

Ammerman, H. L., Becker, E. S., Bergen, L. J., Claussen, C. A., Davies, D. K., Inman, E. E. et al. (1987). FAA air traffic control operations concepts Volume V: ATCT/TCCC tower controllers Washington, DC: Federal Aviation Administration Advanced Automation Program (AAP-110).

Arnaldo, R., Sáez, F. J., & García, E. (2012). HALA!: Towards Higher Levels of Automation in ATM. Position Paper - State of the Art and Research Agenda. In.

Broach, D. (2013). Selection of the Next Generation of Air Traffic Control Specialists: Aptitude Requirements for the Air Traffic Control Tower Cab in 2018 (Rep. No. DOT/FAA/AM-13/5). Oklahoma City, OK: Federal Aviation Administration, Civil Aerospace Medical Institute.

Bruder, C., Eissfeldt, H., Grasshoff, D., Gürlük, H., Friedrich, M., Hasse, C. et al. (2013). Aviator II: Simulator-based Research on Operational Monitoring and Decision Making for Human Operators in Future Aviation (Rep. No. Forschungsbericht 2013-04). Köln: Deutsches Zentrum für Luft- und Raumfahrt (DLR).

Dehn, D. (2013). The SHAPE Questionnaires. http://www.eurocontrol.int/sites/default/files/content/documents/nm/safety/safety-shape-leaflet.pdf.

Ehr, H. & Uebbing-Rumke, M. (2013). Interface AviaSim 4D-CARMA (Rep. No. IB 112-2010/43). Braunschweig: DLR.

Eissfeldt, H. (2009). Aviator 2030 - Determining ability requirements in future ATM systems. In 15th International Symposium on Aviation Psychology Dayton, OH.

Eissfeldt, H., Deuchert, I., & Bierwagen, T. (1999). Ability Requirements for Future ATM Systems Comprising Datalink: A Simulation Study Using an EATCHIP III based Platform (Rep. No. Forschungsbericht 1999-15). Köln: Deutsches Zentrum für Luft- und Raumfahrt (DLR).

Endsley, M. R. (1987). The application of human factors to the development of expert systems for advanced cockpits. In Human Factors and Ergonomics Society (Ed.), (pp. 1388-1392). Santa Monica, CA.

Endsley, M. R. (2006). Situation awareness. Salvendy, G. Handbook of human factors and ergonomics. 3, 528-542. New York, Wiley.

Endsley, M. R. & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, 42, 492.

Fleishman, E. A. & Reilly, M. (1995). Administrator's guide to the Fleishman Job Analysis Survey (F-JAS). Potomac, MD: Management Research Institute.

Gross, H., Friedrich, M., & Möhlenbrink, C. (2010). Development of a process for the evaluation of eye-fixation data using dynamic areas of interest. Braunschweig: DLR.

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (task load index). Results of theoretical and empirical research. In P.A.Hancock & N. Meshkati (Eds.), Human Mental Workload (pp. 139-181). North Holland: Elsevier.

Kirwan, B. (2001). The role of the controller in the accelerating industry of air traffic management. Safety Science, 37, 151-185.

Kirwan, B. & Flynn, M. (2001). Controller conflict resolution strategies for CORA2. In 4th USA/Europe ATM R&D Seminar Santa Fe.

Manning, C. A. & Broach, D. (1992). Identifying ability requirements for operators of future automated air traffic control systems (Rep. No. DOT/FAA/AM-87/26). Washington, DC: Federal Aviation Administration.

Mogford, R. H. (1997). Mental models and situation awareness in air traffic control. International Journal of Aviation Psychology, 7, 331-341.

Moore, K. & Gugerty, L. (2010). Development of a novel measure of siutation awareness: The case for eye movement analysis. In HUMAN FACTORS and ERGONOMICS SOCIETY 54th ANNUAL MEETING (pp. 1650-1654). Santa Monica, CA: Human Factors & Ergonomics Society.

Nickels, B. J., Bobko, P., Blair, M. D., Sands, W. A., & Tartak, E. L. (1995). Separation and Control Hiring Assessment (SACHA) final job analysis report Washington, DC: Federal Aviation Aviation Administration Office of the Assistant Administrator for Human Resources Management.

Niessen, C. & Eyferth, K. (2001). A model of the air traffic controllers' picture. Safety Science, 73, 187-202.

Nijhuis, H. B. (2000). Role of the Human in the Evolution of ATM (RHEA): Final Report (Rep. No. RHEA/NL/SPR/01). Amsterdam: NLR.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. Journal of Cognitive Engineering and Decision Making, 2, 141-161.

Rovira, E. & Parasuraman, R. (2010). Transitioning to Future Air Traffic Management: Effects of Imperfect Automation on Controller Attention and Performance. Human Factors, 52, 411-425.

Schneider, B. & Konz, A. M. (1989). Strategic Job Analysis. Human Resource Management, 28, 51-63.

SESAR (2012). The European ATM Master Plan - The Roadmap for Sustainable Air Traffic Management. https://www.atmmasterplan.eu/

Sheridan, T. B. & Verplank, W. L. (1978). Human and computer control of undersea teleoperators. Cambridge, MA: MIT Man-Machine Systems Laboratory.

Taylor, R. M. (1990). Situation awareness rating technique (SART): The development of a tool for aircrew systems design (Rep. No. AGARD-CP-478). Neuilly Sur Seine: NATO-AGARD.

Truman, T. & de Graaff, A. (2007). Out of the Box - Ideas About the Future of Air Transport, Part 2 Advisory Council for

Aeronautic Research in Europe (ACARE).

Evaluation of an Arrival Coordinator Position in a Terminal Metering Environment

Michael Kupfer, Joey Mercer, Todd J. Callantine, Vimmy Gujral and Ashley Gomez

Human-Systems Integration Division
San Jose State University/NASA Ames Research Center
Moffett Field, CA 94035, USA

ABSTRACT

A human-in-the-loop (HITL) simulation investigated the performance of an integrated arrival-management solution developed by NASA, called the Interval Management Terminal Area Precision Scheduling System (IM-TAPSS). The focus of this paper is on the operator who, during the simulation, served as both the terminal radar approach control (TRACON) traffic management controller and Arrival Radar Coordinator (ARC), and played an integral part in coordinating, adjusting, and instantiating the arrival schedule computed by the IM-TAPSS system. Analyses of the simulation data highlight the role of the ARC within the terminal metering environment, describe his planning strategies, tool interactions, and coordinations with controllers, and lend insights to the impact of the ARC's actions on the arrival problem. High levels of comfort and confidence were reported when working with the IM-TAPSS system. Challenges included the sequencing of unscheduled satellite arrival flights during periods of peak demand, in response to which, the ARC had to identify sequence slots while minimizing disruptions to the overall schedule.

Keywords: Terminal-area arrival management, terminal metering, traffic management strategy, collaborative work

INTRODUCTION

During peak traffic periods, the current air transportation system around terminal areas is impaired by flight inefficiencies, such as frequent vectoring maneuvers and extended level-off flight segments; often at low altitudes. This translates into high noise contamination for effected communities, increased fuel consumption and consequently increased emission output. In consideration of forecasted increases in air traffic demand, Europe's Single European Sky ATM Research Program (SESAR) and the United States' Next Generation Air Transportation System (NextGen) initiative focus on developing new Air Traffic Management (ATM) technologies, systems and procedures, such as optimized profile descents, to mitigate these inefficiencies (SESAR Joint Undertaking, 2013), (Federal Aviation Administration, 2013).

NASA, the FAA, and industry partners are currently working jointly on a multi-year effort, called the ATM Technology Demonstration-1 (ATD-1) (Prevot et al., 2012). The Airspace Operations Laboratory (AOL) and two other research groups at NASA Ames and NASA Langley are supporting the ATD-1 efforts by conducting HITLs of which the goal is to first integrate the different system components into the laboratory setting, to refine the system components, tools, procedures and phraseology, and then to study the overall system performance and compare it with current day operations (Callantine, Kupfer, Martin and Prevot, 2013; Murdoch, Wilson, Hubbs and Smail,

2013; Thipphavong et al., 2013). This work focuses on the most recent simulation conducted in the AOL, with particular emphasis on the strategies employed by the operator working a position that combined the roles and responsibilities of a TRACON Traffic Management Controller (TMC) and an Arrival Radar Coordinator (ARC).

BACKGROUND

IM-TAPSS includes three components: a Traffic Management Advisor for Terminal Metering (TMA-TM) scheduling element, Controller-Managed Spacing (CMS) tools, and Flight-Deck Interval Management (FIM) system. The TMA-TM is an extension to the currently fielded TMA, and provides a timeline graphical user interface (TGUI) to display schedule information to controllers and TMCs. The CMS tools, comprised of slot markers, timelines, early/late indicators and speed advisories, are TRACON decision-support tools that assist controllers in managing delays and delivering aircraft in accordance with the schedule (Kupfer, Callantine, Martin, Mercer and Palmer, (2011). The Airborne Spacing for Terminal Arrival Routes algorithm is used onboard FIM-equipped aircraft to support flight crews during airborne spacing tasks. The simulation discussed in this paper was limited to the first two components. The FIM component will be added in a subsequent simulation. The following description of the operational concept excludes FIM.

The ATD-1 range of operation extends from the TRACON into Center airspace. While still in cruise TMA-TM assigns runways and computes estimated times of arrival (ETAs) at various scheduling points, such as meter-fixes, terminal-area merge points, and runways. Using minimum spacing information, a scheduled time of arrival (STA) is computed for each aircraft at every schedule point. At a particular distance the STAs are frozen, providing controllers with a stable control target. Center controllers then begin to reduce schedule delays using speed control and path assignments. After transitioning into the terminal area, controllers use the CMS tools to issue speeds to unequipped aircraft to ensure correct inter-arrival spacing.

EXPERIMENT

A HITL simulation was conducted in September 2013 – referred to as CA5-2 – to investigate terminal metering for fuel-efficient arrival operations during periods of high traffic demand. This section describes the elements of the simulation in detail.

Previous Simulation Studies

Preceding the CA5 simulation series the AOL conducted four other HITLs under into the ATD-1 project. The 'CMS ATD-1' simulations in January, April, and June 2012 – referred to as 'CA-1,' 'CA-2,' and 'CA-3,' respectively – served to: ensure technical and procedural integration of the IM-TAPSS components, allow researchers to identify additional requirements, develop and iteratively refine procedures and new functionalities, and conduct initial investigations on how the controller tools perform in a FIM-equipped/non-FIM-equipped mixed environment (Callantine, Cabrall, Kupfer, Omar, Prevot, 2012). The objectives of the CA-4 simulation, conducted in March 2013, were to continue to explore FIM operations in a mixed-equipage environment, quantify the effect of preconditioning the arrival flows, study CMS tool modifications (Callantine, Kupfer, Martin and Prevot, 2013).

The CA5 simulation series is comprised of three phases serving to compare ATD-1 operations with current-day PHX operations. Traffic scenarios were designed to closely resemble actual PHX traffic and winds data. The CA5-1 simulation was conducted in July 2013 and served as baseline study. Only the TMA-TM component was included. Arrival problems were simulated with the TMA-TM configured to operate like the deployed TMA version. The CA5-2 simulation, conducted in September 2013, examined limited IM-TAPSS operations without FIM. This paper will focus on the CA5-2 simulation and, since data analysis is ongoing, a selection of its results. The third phase, CA5-3, which will include FIM, is expected to occur soon.

Airspace, Routes and Scenarios

Figure 1 shows the test airspace including the test sectors and the routes, showing the waypoints with altitude and speed restrictions. The simulation's test airspace was comprised of three high-altitude sectors (sectors 37, 50, and 93) and five low-altitude sectors (43, 39, 38, 46, and 42) from Albuquerque Center (ZAB), four high-altitude sectors (40, 60, 35, and 36) from Los Angeles Center (ZLA), and four Phoenix TRACON (P50) sectors; two feeder sectors (206-Apache, 203-Quartz) and two final sectors (205-Freeway, 204-Verde). Sectors 37 and 43, 40 and 60, and 35 and 36 were combined. The simulation was supported by several confederate ghost controllers: a departure ghost controller handled the departing traffic out of KPHX and the surrounding satellite airports, a satellite ghost controller handled the arriving flights into neighboring airports, a tower ghost position controlled KPHX arrivals during the last few miles of the final approach until landing, and two en-route ghost controllers (ghost high, ghost low) worked the substantial amount of remaining traffic in the surrounding airspaces. Real-world sector boundaries and altitude strata were used. Additionally, positions for a ZAB TMC and a P50 ARC, who also doubled as a P50 TMC, were staffed.

The CA5 studies simulated both west- and east-flow airport configurations. Published RNAV OPD routes and non-RNAV routes (see Figure 1), were adapted based on existing Standard Terminal Arrival Routes and Instrument Approach Procedures. The adapted RNAV OPD routes were: MAIER5, EAGUL5, KOOLY4, and GEELA6. The adapted non-RNAV routes were: COYOT2, JESSE1, SUNSS8, and ARLIN4. During west-flow operations the RNAV routes from the east connected to the runway 25L and 26 approach procedures. During east-flow operations the RNAV routes from the west connected to the runway 07R and 08 approach procedures. RNAV routes from the other approach directions included downwind segments that were unconnected. The non-RNAV routes were unconnected as well and required vectoring to the final approach course.

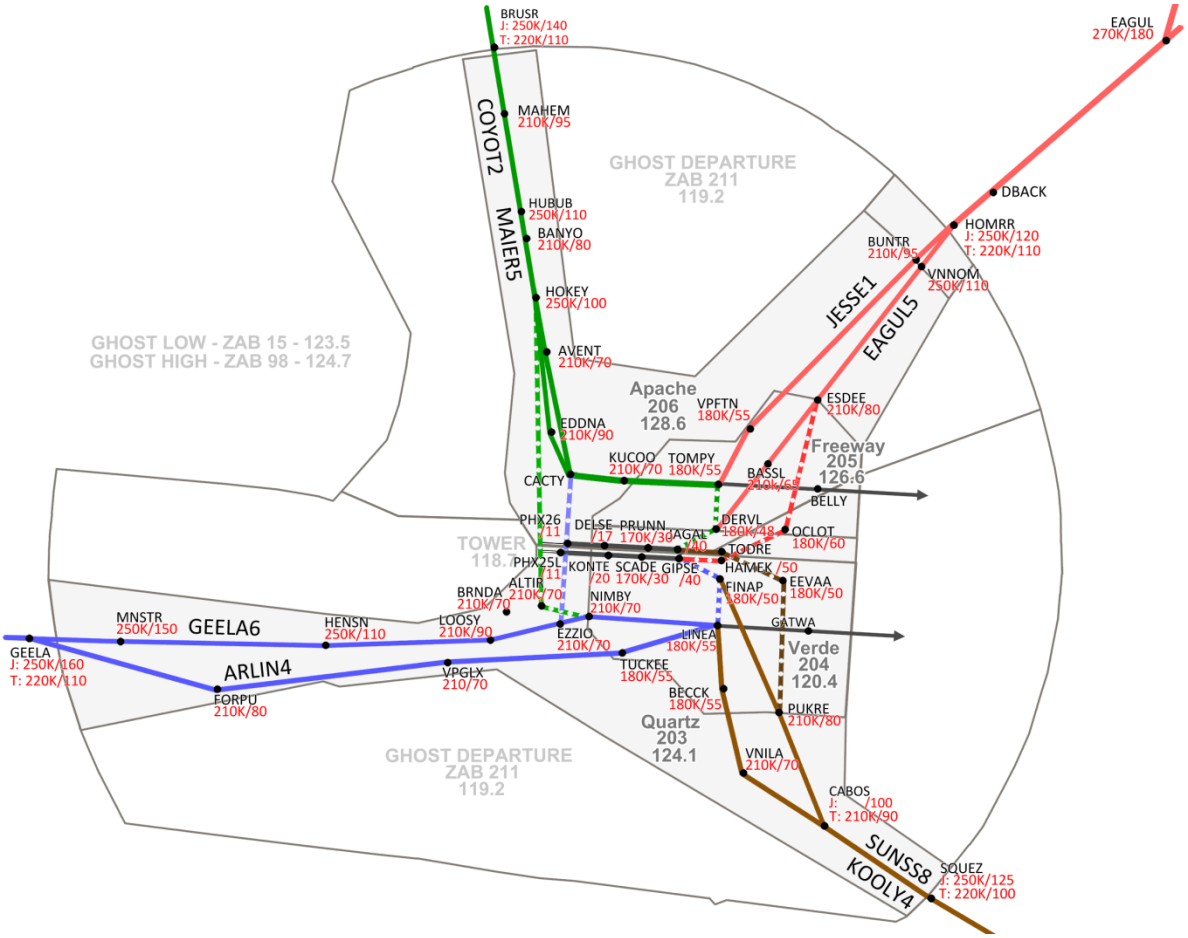


Figure 1. Simulation Airspace showing the test sectors and adapted routes for west-flow configuration.

For each airport flow direction, one base scenario was designed. The scenarios mirrored real-world high-demand arrival flows under realistic wind conditions, with few differences from the live traffic recordings. These scenario files served as the starting point for the final traffic scenarios. Each traffic scenario is comprised of flights arriving at KPHX (subsequently referred to as arrivals), departures leaving KPHX and over-flight traffic departing and landing from airports other than KPHX. The arrivals consisted of a range of aircraft performance groups (i.e., jets, high- and low-performing turbo-prop aircraft and piston aircraft). Only the west-flow scenario contained low-performing turbo-prop and piston aircraft, which did not follow any of the RNAV and non-RNAV routes used by the jets and higher performing turbo-prop aircraft. These flights approached KPHX under the control of the satellite ghost controller's airspace. Their routes were not adapted in the TMA-TM system. Copies of the two traffic scenarios were created that included additional KPHX arrival flights (six in east flow and five in west flow) in order to create a period of sustained increased traffic demand.

Participants

The Center and TRACON sectors were staffed with recently retired ZAB and P50 controllers, with an average of 26.25 years of experience. CA5-2 also included a ZAB TMC as well as a position combining both theP50 TMC and the P50 ARC (referred to throughout this document as the ARC). The ARC had 29 years of experience as tower/terminal controller, nine of those years as the TMC/ARC at P50 TRACON. He was retired since December 2011. The two ZLA controllers, the tower controller, the departure controllers, the two en-route "ghost" controllers responsible for the areas surrounding the test sectors, and the satellite ghost controller responsible for traffic arriving at local airports, were staffed with retired confederate controllers. Prior to the CA5-1 simulation, none of the participant controllers had any previous experience with the IM-TAPSS concept and associated tools or the MACS simulation software. Boeing glass-cockpit type-rated pilots, many of whom had participated in previous CA simulations, worked single-pilot mid-fidelity desktop flight simulators, while local commercial and student pilots who were familiar with MACS, worked pseudo-pilot positions.

During the simulations, operator workstations were arranged such that the ZLA, ghost high, ghost low, departure and tower controllers were located in one room. Located in a separate room, were the satellite ghost TRACON controllers, and the ARC. The TRACON controller workstations' arrangement was such that it replicated the arrangement in use in P50. A third room housed the high-altitude ZAB sector 93, and low-altitude ZAB sectors 42 and 46, while located in a fourth room were the low altitude ZAB sectors 50, 39, and 43. The single-aircraft pilots were grouped together in one area, while the pseudo pilots were distributed over two separate areas in the AOL.

ARC Position and Procedures

In today's operations, P50's ARC position is normally combined with the TRACON TMC. All P50 personnel are expected to fully comply and cooperate with the direction issued by the ARC. The ARC is required to participate in operational briefings, and to provide a plan for: integrating the KPHX arrivals not sequenced by Apache and Quartz feeder sectors (i.e., satellite arrivals), accommodating aircraft approaching from one side of the airport that need to land on the opposite side (i.e., 'cross-overs'), and changing the landing runway of an aircraft when needed (Federal Aviation Administration, 2011).

The ARC coordinates with the TRACON controllers to identify and create slots in the arrival sequence in which to accommodate the satellite arrivals. Because current-day operations are distance-based, rather than schedule-based, the ARC does not interact with the TMA TGUI to modify STAs or sequences. P50 controllers simply maneuver aircraft as needed to maintain safe separation, while the ARC helps coordinate workable sequences. In contrast, the schedule-based operations of the IM-TAPSS concept require the ARC to find naturally occurring slots in the arrival sequence, or if necessary, create a slot by adjusting the STAs of adjacent flights while minimizing the overall impact to the schedule. In this context, the ARC's 'big-picture' perspective enables him to develop a plan for the unscheduled arrivals using the various TMA-TM TGUI functions for manipulating STAs: move, swap, change runway, reset (re-computation of ETA and STA), and several types of reschedule (re-computation of STAs). He then coordinates his plan with the Center TMU and the TRACON controllers.

The ARC created a custom timeline layout to help him perform his tasks (see Figure 2). Four timelines were configured to show a one-hour time period, with the outer left timeline displaying meter-fix STAs (i.e., TRACON entry times) for aircraft coming over the northern meter-fixes, and the outer right timeline displaying meter-fix STAs for aircraft coming over the southern meter-fixes. More specifically, the outer left timeline's left-side displayed the northwest meter-fix STAs, while the right-side displayed the northeast meter-fix STAs. Similarly, the outer right

timeline's left- and right-sides displayed the southwest and southeast meter-fix STAs, respectively. The call-signs displayed in these two timelines were time-shared with the aircraft's assigned runway (left inset in Figure 2); allowing the ARC to quickly identify any cross-over flights. Delay values were displayed next to the callsigns in the meter-fix and runway STA timelines (right inset in Figure 2).

The TGUI's two middle timelines provided schedule information for the runways. The center-left timeline displayed runway ETAs (in green), arranged to show north runway ETAs on the left-side, and south runway ETAs on the right-side. The center-right timeline displayed runway STAs (in blue if frozen, yellow if un-frozen), also with north runway STAs on the left-side, and south runway STAs on the right-side. Because the runway timelines accepted runways PHX08 and PHX26 as north runways, and runways PHX07R and PHX25L as south runways, this custom TGUI configuration worked for both east- and west-flow airport configurations. The separate arrangement of runway ETAs and STAs is based on the TGUI configuration in current use at P50. The ARC explained during one of the runs: "VFR airplanes coming from the Center [airspace], they are never scheduled, but they are always estimated". The two middle timelines then, help the ARC to easily assess if slots are available on one of the runways. Although VFR aircraft were not included in the CA5 series of simulations, this example demonstrates the TGUI's importance to the ARC's tasks. Similarly, orange-colored leader-lines on the TGUI gave the ARC additional information to incorporate into his decisions. The orange leader-lines indicated aircraft that either had not yet departed, or were currently flying in a Center airspace that does not feed data to P50.

The simulation environment posed some limitations on the realism of the operations. In current-day operations, satellite arrivals are held on the ground until a 'call for release' is approved by the ARC. However, in the CA5-2 simulation, those flights began flying at pre-specified times, without the option to delay their departure until a slot in the KPHX arrival schedule was available. The slots identified for those arrivals by the ARC sometimes required that large amounts of delay be absorbed. In an effort to meet those slots, the satellite ghost controller often had to vector the aircraft. Additionally, the routes of the satellite arrivals were not adapted in the version of TMA-TM used during the simulation, resulting in unreliable STA computations. The ARC consequently suspended the STAs of those flights (removing them from the STA timeline), developed an arrival plan, and coordinated with the TRACON controller team.

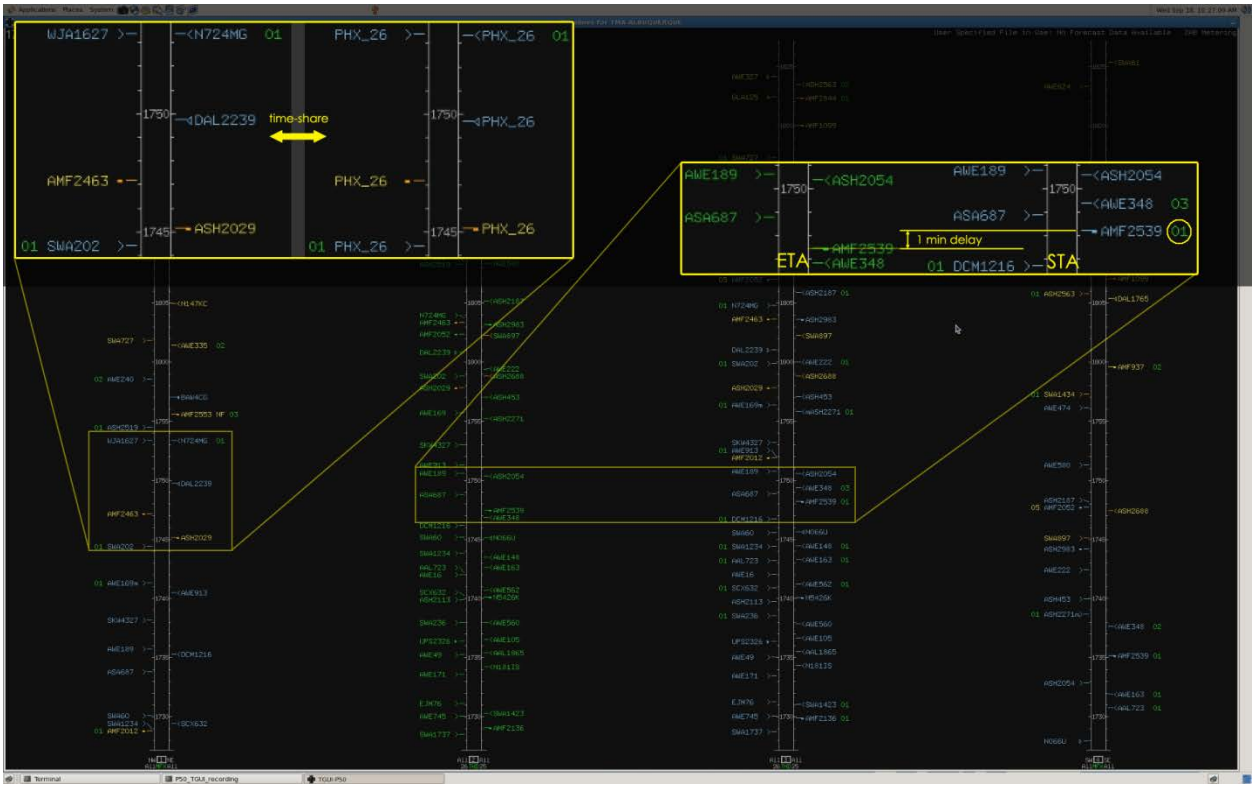


Figure 2. TGUI setup used by the TRACON TMC / ARC position. Timelines for meterfix STAs are on the left (northern routes) and on the right (southern routes). Runway ETAs for both runways are shown on the center-left and runway STAs for both runways are shown on the center-right timeline.

Training and Data-Collection Schedule

The simulation began with a general introductory briefing, followed by four days of training before the data-collection period. Center and TRACON controllers were first briefed on the airspace, routes, procedures, systems and technologies, then received training in separate, parallel simulations. Pilots flying the mid-fidelity single-pilot simulators received a separate initial preparation as well. On the third training day, controllers and pilots participated together in joint training simulations with lighter traffic densities that, by the end of the day, increased to the traffic densities that were simulated in the data collection runs. The training was concluded at the end of the week with a debrief discussion. The data collection period took place over four days during the following week. Experimental trials were conducted in each of the airport flow configurations under four realistic forecast-wind error conditions with current-day and increased-load traffic scenarios, for a total of 16 one-hour data-collection runs. A post-run questionnaire and a 20-minute break followed each trial. The study was concluded with a post-simulation questionnaire and an in-depth final debrief discussion.

Recorded simulation data logged various metrics such as trajectory and flight state information, pilot and controller entries, schedule data, etc., collected from every controller, pseudo-pilot and ASTOR workstation. Voice communications between controllers and pilots were recorded using an emulation of the FAA's Voice Switching and Communication System (VSCS). An individual recording was obtained from the ARC using a collar audio reorder. Additionally, the displays of all workstations, along with the ambient audio, were captured via screen-recording software. Other logs captured TMA-TM's schedule calculations. Using the Air Traffic Workload Input Technique (ATWIT) (Stein, 1985), controllers rated their current workload every three minutes, on a scale from 1 (low) - 6 (high). The next section highlights data regarding the role and tasks of the P50 TRACON TMC/ARC.

RESULTS

The CA5-2 simulation provided a wealth of information and insights on how the IM-TAPSS concept would look like if introduced into actual operations at KPHX. The following section provides salient results that highlight how the ARC adapted to new requirements of the IM-TAPSS system. First, objective data lends insights into which TGUI features the ARC used to accomplish his task. Subsequently, subjective data provides details on tool acceptance and traffic management techniques. To emphasize the complexity of merging flights into an already dense arrival stream and to highlight ARC-controller coordination two explicit examples of arrival coordination problems are presented.

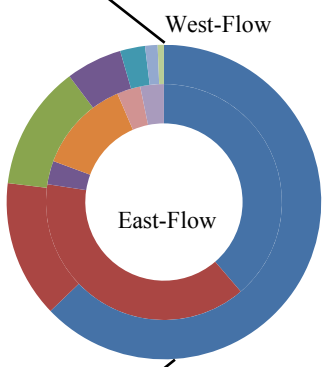
TGUI Interactions

The analysis of TGUI log files, cross-referenced with screen recordings, identified the frequency of the ARC's interactions with the TGUI. Due to software-stability issues identified in the first half of the simulation, this analysis only includes data from the last eight runs of the simulation, all of which were increased-load traffic scenarios. Table 1 shows the frequencies of the ARC's various interactions with the TGUI. For comparison, in each of the four east- and four west-flow runs an average of 64 flights entered the TRACON.

There was a clear difference between the east- and west-flow conditions, as the unscheduled satellite arrivals were only included in the west-flow scenarios. Of those actions that affected aircraft STAs, the 'Move' function, which allows changing an aircraft's STA by a simple drag-and-drop mechanism, was the most frequently used feature, followed by the 'Re-schedule [aircraft] only' function, which automatically finds another slot for the respective aircraft. The 'Reschedule [aircraft] and after' function was used less frequently, since it also changed the STAs of all following aircraft. The ARC's decision to twice use the 'Reschedule [aircraft] and after' function impacted the STAs of 15 other flights. The Aircraft Data Window in the TGUI offers detailed flight plan and schedule data. The ARC frequently took advantage of this information to aid his planning decisions. The 'Suspend flight' function excludes the respective flight from being considered by the scheduler, which the ARC often used on the satellite arrivals.

Table 1: Total frequencies of ARC TGUI interaction events (for the eight increased-load trial runs).

Event	East-Flow	West-Flow	
Change of Frozen STA	12	92	
Move (drag and drop to change single aircraft's schedule)	8	50	
Reschedule [aircraft] and after (re-computes STAs for this and all subsequent aircraft)	0	2	
Reschedule [aircraft] only (re-computes the STA for the selected aircraft)	4	25	
With reschedule and after (shows STA changes indirectly caused by 'reschedule and after')	0	15	
Open A/C Data Window	12	22	
Suspend Flight	0	20	
Swap Assigned Runway	1	9	
Reset STA of Flight	0	4	
Open Traffic Count Window	4	0	
Swap two STAs	0	2	
Open Separation Matrix Window	1	0	
Resume STA Computation	0	1	
Open Status &. Schedule Window	1	0	



Questionnaire Responses

After the conclusion of each run, the ARC answered questionnaires through which he provided feedback on procedures and tool usefulness, and also described his work strategies. In general, the ARC was very confident in using the TMA-TM [7 on Likert scale: 1 Not at all confident – 7 Very Confident]. The ARC rated the initial sequence and schedule provided by the TMA as a very workable starting point [7 on Likert scale: 1 not at all workable – 7 very workable]. Although the ARC felt he had sufficient information to manage his traffic effectively, he identified some clutter on his display [4 on Likert scale: 1 No clutter – 7 Considerable clutter]. Despite this, the ARC gave ratings of 'not at all frustrated' by the metering tools or the metering procedures [1 on Likert scale: 1 Not at all frustrated – 7 Very frustrated]. Along with these positive ratings, he provided some suggestions on how to improve, alter, or develop his tools to make them better. "A TMA 'what-if scenario' feature would be helpful for probing aircraft-pair runway swaps or STA schedule tweaks at the meter-fixes."

Sequence Coordination

The west-flow traffic scenarios contained flights which were departing from airports close to KPHX (typically, small turbo-prop and piston aircraft). They were not following the routes used by large turbo-prop and jet aircraft, but instead were following routes through airspace controlled by the satellite ghost controller. When asked which barriers the ARC saw to being able to use the IM-TAPSS metering concept in real operations, he elaborated on the problem of scheduling satellite arrivals: "In heavier traffic situations, local traffic would have to be scheduled into TMA. [For aircraft without] fixed routings (such as for VFR traffic or traffic from satellite airports), it would be difficult to determine these times. [But] it would be possible to hold slots [for those aircraft] as long as controllers can hit them." For the traffic densities simulated in the study, the ARC rated incorporating the satellite arrivals into the arrival stream as somewhat easy [5 on Likert scale: 1 very difficult – 7 very easy].

The ARC indicated that the new metering environment and the availability of the metering tools significantly changed the way he coordinated with his controller team [1 on Likert scale: 1 Coordination decreased a lot – 7 Coordination increased a lot]. He also reported that the operations demanded a lot of adjustment of his traffic management strategies to accommodate the schedule [6 on Likert scale: 1 No changes – 7 A lot of changes]. The ARC developed a general strategy for accommodating the satellite arrivals. First, he obtained a detailed picture of the arrival situation by coordinating with the TRACON controller team and by studying the TGUI. The ARC explains: "The slot markers dictated [the] scheduled aircraft sequence". Then, he suspended the STAs of the satellite arrivals. Without TMA-TM tracking the satellite arrivals, no CMS tools (slot markers, timeline ETA/STA information, and

speed advisories) were available to guide the controllers. Without the slot markers for example, controllers had to rely on other means to ensure safe spacing. Next, he identified any pre-existing excess spacing (i.e., 'gaps') between consecutive flights in the arrival stream to accommodate these suspended (i.e., now unscheduled) arrivals. If the gap was not large enough he tried to extend the gap by modifying the neighboring STAs of other aircraft. Because the goal was to minimize the impact on the overall schedule, the ARC preferred schedule manipulations resulting in only one flight being affected, leading to his often use of the 'Move STA' function. To enlarge pre-existing gaps, the ARC tried making room on the schedule for the satellite arrivals by slightly increasing the delay of the adjacent flights. Other tools at his disposal were runway or aircraft swaps. For example, the ARC checked if at the required time period any space in the arrival stream to the other runway was available. If this was the case, assigning one of the neighboring flights in the sequence to the other runway was a valid option to increase a schedule gap. Finally, he then coordinated his schedule/sequence solution with the satellite ghost controller and the TRACON controller team. This work flow was repeated throughout the simulation run.

The following two examples were chosen to illustrate these interactions. They are representative of the typical ARC work flow and strategy. Both stem from a west-flow run that includes satellite arrivals. Figures 3 and 4 visually support the communication excerpts used to frame the course of the coordination events. They provide an overview of the track plots of the flights involved in the example problems. The tracks were color-coded by relative simulation time, from between 600s after simulation start to the end of the run at 3600s. The information in the two examples is based on data from TGUI and controller screen recordings, as well as audio transcriptions.

Example 1 (Figure 3) emphasizes how the ARC coordinated an arrival sequence, focusing on flights from the south side of the TRACON, where two satellite arrivals need to be fit into the arrival stream. During this period, the TMA-TM assigned cross-over runways to three aircraft.

At 832s the ARC, coordinating the upcoming sequence with the satellite ghost controller, discusses two south-side arrivals: "I don't think [CFS806] can get ahead of the AMF2136. I have one spot behind the guy on the west side [SWA1423]; the only one. And then [AMF2837 will follow]. And after that, I am full for quite some time." At 935s he briefed the south feeder controller about a cross-over flight on the KOOLY4 route, coming from the south: "You should have a [runway] 26 [flight] (AWE745) coming up there." At 1146s he briefed the north feeder controller about two other cross-over flights on the EAGUL5 route, coming from the north: "Two [runway] 25's are coming: one [N181IS, (pointing at the screen)] and two [AWE105, (pointing at the timeline)]."

A few minutes later, at 1210s, the ARC coordinates the arrival sequence with the south final controller: "So, the Am-flight [AMF2837] is supposed to go behind Southwest [SWA1423], but there is a Falcon [N181IS] coming from the north-east." The south final controller then asks for clarification of the sequence: "Where is the Falcon [N181IS] going? Back here?" The ARC responds: "[With] the Falcon [N181IS] just go [behind the] Am-flight [AMF2837]." The south final controller points out: "Well, if [the satellite ghost] doesn't keep him tight it's not gonna work." The ARC then verifies with the satellite ghost controller: "You'll work him right behind Southwest [SWA1423], right?" The ARC and the south final controller agree that AMF2837 will barely fit into the slot. At 1270s, the ARC comments: "Just eyeballing while I watched him appear, I thought this [SWA] 1423 and [AMF] 2136 is gonna be close [...]."

In this first example, the ARC first discusses the arrival sequence with the satellite ghost and later with the final controller. He also points out cross over-flights to the controllers. The south final controller and ARC agree that the selected slot will be tight fit for the satellite arrival.

Example 1b. Later the ARC is elaborating on the cross-over problem. At 1520 s: "She's [north feeder controller] got no idea what's going [on] on the other side, unless I tell her […], and he's [south final controller] got no idea he [N181IS] is coming to his runway. So unless we widen him [N181IS] out ..." The 205 north final verifies: "Widen him out?" The ARC confirms: "Yea." The north final controller gives the clearance: "November-181-Indie-Sierra Phoenix Approach, roger, fly heading 170, vector for spacing." At 1629 s the north final controller points out the cross-over flight to the south final controller: "He's [N181IS] on the 170 heading right now." The south final controller comments: "Seems like he fits better on your runway, but I'll take him." The ARC confirms his plan to the south final controller and briefs him about AWE105 crossing over from the north to the south runway: "And you don't know this yet, he [AAL1865] goes behind him [N181IS]. There's another [cross-over] from the north [AWE105] who follows America [AAL1865]." The south final controller clarifies with the ARC: "Who goes behind who?" The ARC is re-iterating the sequence: "Indie-Sierra [N181IS] is going behind Am-flight [AMF2837 coming from the south], American [AAL1865] is going behind Sierra [N181IS]. There is another guy coming from

the north [AWE105] that goes behind America [AAL1865]." The north final controller: "Oh, he's going to … uh ... Oh! AWE105 reduce speed to 190. ... Alrighty then. ... I see that now." The embedded TGUI screen snippet in Figure 3 shows the final arrival sequence: for runway PHX26: SWA1737, AWE745, EJM76, AWE171, AWE49; for runway PHX25L: AMF2136, SWA1423, AMF2837 (unscheduled satellite arrival), N181IS, AAL1865, AWE105.

This section highlighted the difficulties the controllers had to recognize the intended arrival sequence on their own. They assessed the situation differently than the ARC, without his broader situational awareness. The excerpt even indicates elements of surprise on behalf of the TRACON controllers regarding the arrival sequence. Controllers verified the order with the ARC, and he then had to clearly state which flight followed which.

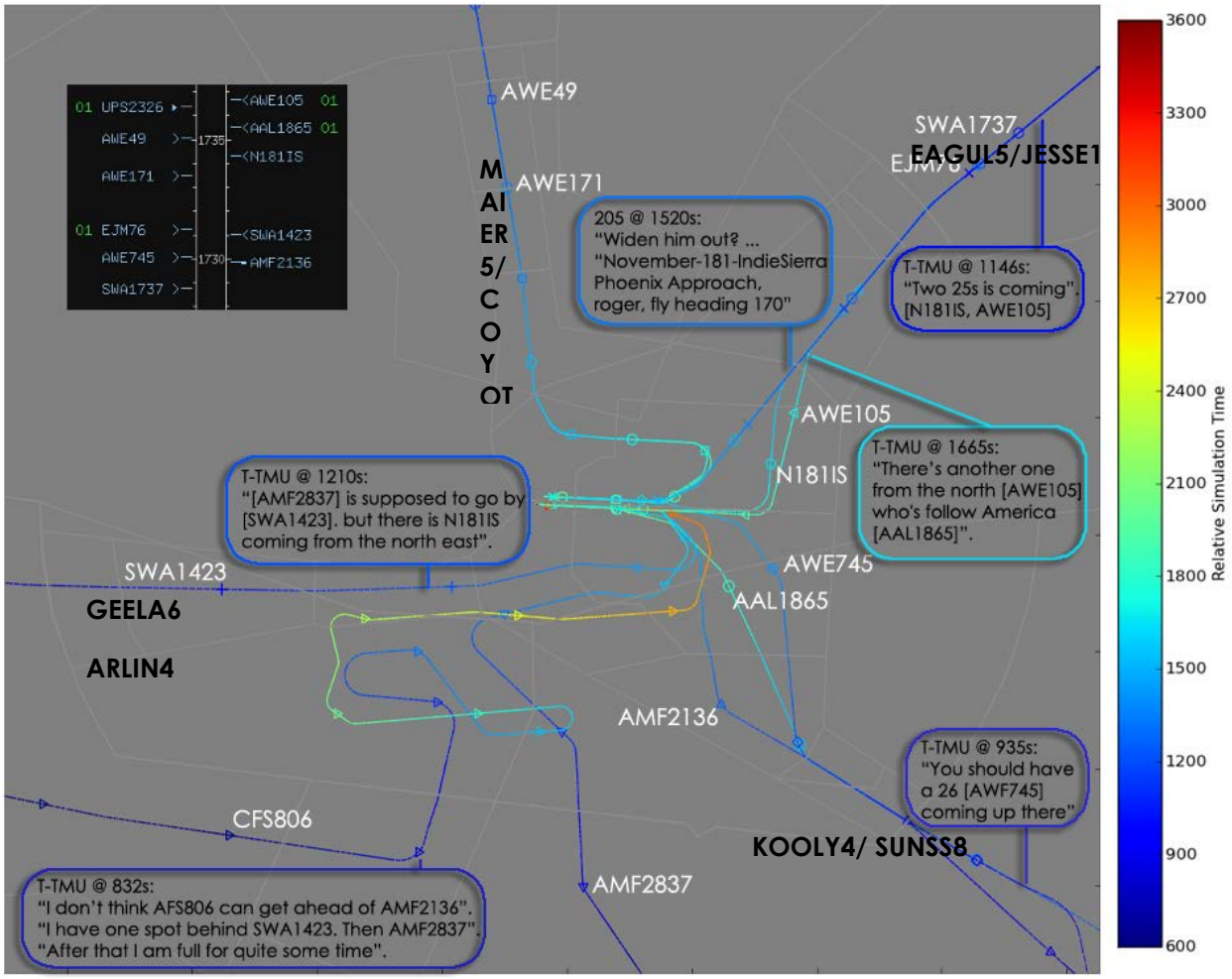


Figure 3. Overview of the track plots of the flights involved in example problem 1.

Problem 2 (Figure 4, Figure 5) emphasizes specific TGUI interactions by the ARC. The ARC is working with the TGUI trying to open slots to fit in satellite arrivals. In the meantime, the satellite ghost controller has to issue delay vectors to several flights until any arrival slots are available.

At 1849 s the ARC modifies the arrival schedule via the TGUI interface (Figure 5). He moves the frozen STA of AWE169 from 17:56:20 back to 17:56:50, and suspends the STA of AMF2012 to open up a gap after SKW4327. At 1890 s the ARC discusses the sequencing problem with the satellite ghost controller: "What do you got? Three [flights in the] south and one or two [flights in the] north? Three north?" The satellite ghost responds: "I've got three [flights in the] north, I've got two [flights in the] south, but I've got two flashing [in handoff status]." The ARC replies: "Well, I'm gonna have two slots for you in a little bit." The satellite ghost answers: "Alright then, I'll bring that guy [N17DM, landing at Phoenix Deer Valley Airport] to the south, and that's probably gonna do it for me for a while. Not a lot of room here."

At 2235 s the ARC informs the satellite ghost about an open slot: "I got one slot on [runway] 26 behind him [SKW4327]." The satellite ghost controller replies: "Just one, ok. Here's my man." He clears AMF2627 (from the north) for the descent: "AMF2627 descend and maintain 6000. I've got a slot for you here." At 2578 s the ARC moves the frozen STA of ASH2983 from 18:03:40 back to 18:04:30 to open up a gap after SWA897. At 2786 s the ARC briefs the satellite ghost about two more gaps available for the satellite arrivals: "He's [CFS864 from the south] gonna go between him [SWA897] and him [ASH2983]. […] and the other one [AMF2012] goes behind him [DAL2239]."

This section illustrated the strategies of the ARC when interacting with the TGUI. He identified and enlarged pre-existing gaps in the schedule by utilizing the allowable STA range and/or adding small amounts of delay to flights. The satellite ghost had more flights to control than there are sequence slots available. The result of this shortage is shown by the drastic delay vectors of the satellite arrivals shown in Figure 4. The workload histogram for the satellite ghost position embedded in Figure 4 shows a workload increase at 1800 s. At this time he had six flights, most of them on vectors, in his control (the histogram uses the same color coding as Figure 3 and Figure 4).

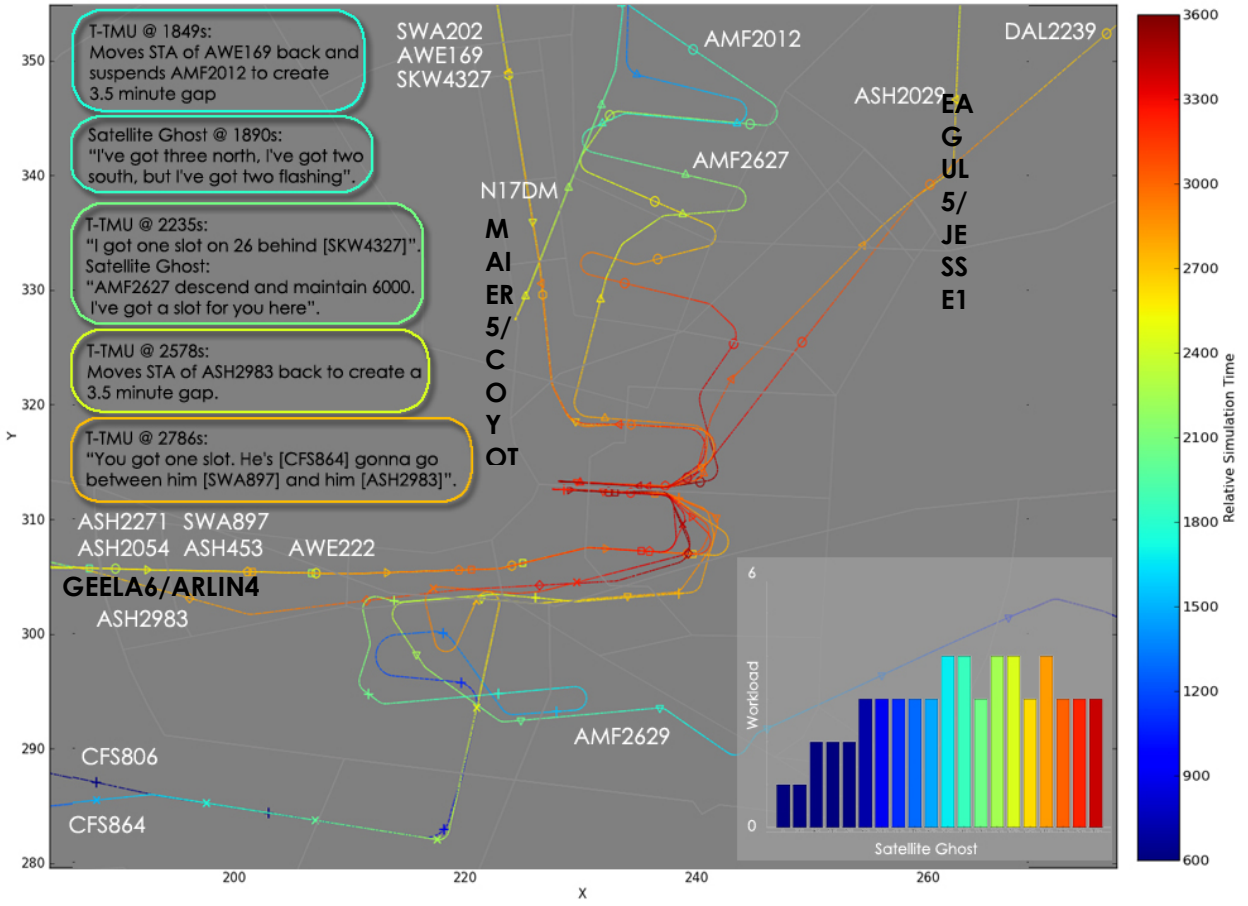


Figure 4. Overview of the track plots of the flights involved in the example problem 2.

Figure 5 shows the schedule at 17:30:25 UTC / 1825 s on the left, and at 17:43:00 UTC / 2580 s on the right. On the left, shown in yellow, are pre-departures that are not yet frozen (i.e., ASH2688 and AMF2634), and one satellite arrival (AMF2012). The pre-departures were scheduled to times past 18:05:00 UTC (not shown in Figure 5) and the AMF2012 was suspended. The STAs of AWE169 and ASH2983 were moved back to slightly later times (indicated by the 'm' near the leader line), to create room for the unscheduled satellite arrivals. As indicated by the TGUI screen snippets in Figure 5, the final arrival sequence was as follows: for runway PHX26: SKW4327, AMF2627 (satellite arrival), AWE169, ASH2029, SWA202, DAL2239, AMF2012 (satellite arrival); for runway PHX25L: ASH2054, CFS806 followed by AMF2629 (both satellite arrivals), ASH2271, ASH453, AWE222, SWA897, CFS864 (satellite arrival), and ASH2983.

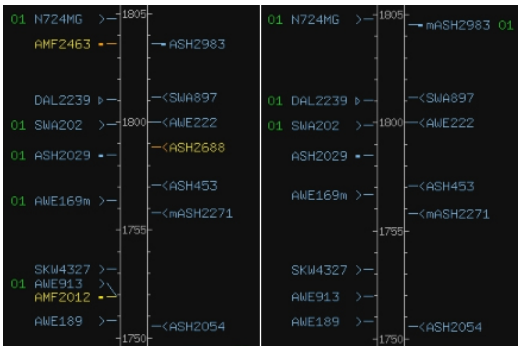


Figure 5. TGUI runway STA timelines at 17:30:25 UTC (left) and at 17:43:00 UTC (right).

CONCLUSIONS

Future schedule-based arrival operations, such as the IM-TAPSS concept, rely on the integrity of an arrival schedule to provide system benefits (e.g., sustained, high throughput and fuel savings by enabling continuous descent operations). However, even if most of the arrival traffic is accounted for by future scheduling systems, unforeseen events likely will still occur. Unscheduled (VFR) traffic or off-nominal flights (e.g., go-arounds), are examples of such unexpected events. Accommodating those flights then, consequently will impact the arrival schedule.

The simulation described in this paper investigated the role, responsibilities and strategies of an operator who, during the simulation, served as both the P50 TRACON TMC and ARC. He had, relative to the individual TRACON sector controllers, a more complete picture of the overall arrival operations. Therefore, he was well-suited to develop efficient plans and solutions. The simulation illustrated his function as schedule manager and team strategist, and showed that he played an integral part in coordinating, adjusting, and instantiating the arrival schedule computed by the IM-TAPSS system.

The results underline the importance of such an arrival coordinator position for future schedule-based arrival management solutions. To create room in the arrival sequence, flights in current-day operations are maneuvered as necessary to maintain safe separation from other aircraft, with no concern for meeting specific arrival times. In a schedule-based environment however, the arrival schedule imposes additional constraints on the possible control actions. The ARC is able to identify natural occurring gaps in the sequence to help accommodate any flight not tracked by the scheduler. If necessary, he has to adjust other aircraft STAs to increase the gaps, while limiting disruptions to the overall schedule. If traffic densities are high, and available gaps are sparse, it is especially difficult to accommodate any unscheduled flights. Vectors have to be issued to delay the flights until a slot is available. This translates to increased fuel consumption and emissions, and increased controller workload.

Given the complexity of the arrival problems, clear coordination between the ARC and the controller team about the intended plan is crucial to avoid any misperceptions. In the examples provided in this paper, controllers seemed to be sometimes confused about the intended arrival sequence (e.g., “Where is the Falcon going?”), or even surprised (e.g., “Oh! I see that now.”). Besides the unscheduled arrivals, cross-over flights to other runways also necessitated coordination between the ARC and the controllers: controllers had to check back with the ARC to verify the desired sequence.

Similar to findings from (Martin, Mercer, Callantine, Kupfer, Cabrall, 2012), a ‘what-if-feature’ built into the TGUI would help the ARC to independently play-out several strategies before implementing them into the active schedule. In the current software however, each modification of aircraft STAs by the ARC immediately impacted the CMS tools, leading to artifacts such as slot marker jumps.

The CA5-2 study provided valuable insights on the role of an ARC position within a schedule-based terminal metering environment. Upcoming research will investigate the ARC’s role when FIM operations complement the controller managed spacing tools.

REFERENCES

Callantine, T., Cabrall, C., Kupfer, M., Omar, F., & Prevot, T. (2012), "*Initial investigations of controller tools and procedures for schedule-based arrival operations with mixed flight-deck interval management equipage*", Proceedings of the 12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, Indianapolis, IN, USA.

Callantine, T., Kupfer, M., Martin, L., & Prevot, T. (2013), "*Simulations of Continuous Descent Operations with Arrival-Management Automation and mixed Flight-Deck Interval Management Equipage*", Proceedings of the Tenth USA/Europe Air Traffic Management Research and Development Seminar, Chicago, IL, USA.

Federal Aviation Administration (2011), "*Order P50 TRACON 7110.1A*", Washington, D.C., USA.

Federal Aviation Administration (2013), "*FAA NextGen implementation plan*", Washington, D.C., USA.

Kupfer, M., Callantine, T., Martin, L., Mercer, J., & Palmer, E. (2011), "*Controller Support Tools for Schedule-Based Terminal Operations*", Proceedings of the Ninth USA/Europe Air Traffic Management Research and Development Seminar, Berlin, Germany.

Martin, L., Mercer, J., Callantine, T., Kupfer, M., & Cabrall, C. (2012), "*Air Traffic Controllers' Control Strategies in the Terminal Area under Off- Nominal Conditions*", Proceedings of the 4th International Conference on Applied Human Factors and Ergonomics, San Francisco, CA., USA.

Murdoch J., Wilson, S., Hubbs, C., & Smail, J. (2013), "*Acceptability of Flight Deck-Based Interval Management Crew Procedures*", Proceedings of the AIAA Modeling and Simulation Technologies (MST) Conference, Boston, MA, USA.

Prevot, T., Baxley, B., Callantine, T., Johnson, W., Quon, L., Robinson, J., & Swenson, H. (2012), "*NASA's ATM Technology Demonstration-1: Transitioning fuel efficient, high throughput arrival operations from simulation to reality*", Proceedings of the International Conference on Human-Computer Interaction in Aerospace (HCI-Aero), Brussels, Belgium.

SESAR Joint Undertaking (2013), "*SESAR release 2013*", Brussels, Belgium.

Stein, E. (1985), "*Air Traffic Controller Workload: An Examination of Workload Probe*", (DOT/FAA/CT-TN84/24), Federal Aviation Administration Technical Center, Atlantic City International Airport, NJ, USA.

Thipphavong, J., Jung, J., Swenson, H., Witzberger, K., Lin, M., Nguyen, J., ... Smith, T. (2013), "*Evaluation of the Controller-Managed Spacing Tools, Flight-deck Interval Management and Terminal Area Metering Capabilities for the ATM Technology Demonstration #1*", Proceedings of the Tenth USA/Europe Air Traffic Management Research and Development Seminar, Chicago, IL, USA.

Identifying Markers of Performance Decline in Air Traffic Controllers

Tamsyn Edwards, [a] Sarah Sharples [a], Barry Kirwan [b], John Wilson [a] and Nora Balfe [c]

[a] Human Factors Research Group
Faculty of Engineering
University of Nottingham
University Park, Nottingham
NG7 2RD, UK

[b] Eurocontrol Experimental Centre
Bretigny/Orge, France

[c] Trinity College, Dublin
College Green, Dublin 2, Ireland

ABSTRACT

Air traffic control (ATC) is a safety critical environment in which air traffic controllers (ATCOs) must maintain a consistently high standard of human performance in order to maintain safety. It is therefore essential to understand the impact, of human factors such as workload on human performance. Traditionally, a wide range of methods have been utilized to investigate the association between performance-influencing factors and human performance, from experimental to research in the wild (Sharples et al., 2012). However, little research has focused on the identification of 'signs and symptoms', or indicators, that a factor may be negatively influencing human performance. The aim of this study was to identify indicators that were commonly associated with specific factors (workload, stress and situation awareness) and the associated influences on human performance. Findings identified internal and external, observable indicators that were common across all controllers and signified when ATCO performance was threatened. Findings also suggest that controllers use internal markers to adapt control strategies in order to maintain performance. Findings have relevance for existing aviation training, online measurement of performance during live operations, and measurements during simulations. Further research should investigate the application of indicators in measurement of live operations.

Keywords: Air Traffic Control, Observation, Behavioural Markers, Human Performance, Workload, Stress, SA

INTRODUCTION

Air traffic control (ATC) is a safety critical environment (Chang & Yeh, 2010). Air Traffic Controllers (ATCOs) are responsible for the safety and efficiency of all air traffic. Unlike other safety critical industries, there are no physical barriers or defences that protect aircraft in flight. It is therefore essential that controllers maintain a consistently high standard of human performance in order to maintain flight safety.

With a large potential for incidents, Kirwan (2011) suggests that Air Traffic Management (ATM) is remarkably reliable. However, when aircraft have breached standard regulated minima, termed an 'incident', human error has been attributed as a primary or secondary cause in 75-90% of cases (Mackieh & Cilingir, 1998). Human factors (such as workload, fatigue, inadequate communications) are "major determiners of a human error" (Park & Jung, 1996, p330) and have been repeatedly shown to negatively affect human performance (Chang & Yeh, 2010; Park & Jung, 1996).

Traditionally within the discipline of human factors, the investigation of human performance, and the association with influencing factors such as workload and stress, has focused on a wide range of methods, from experimental to research in the wild (Sharples et al., 2012). However, little research has focused on the identification and application of 'signs and symptoms', or indicators, of the negative association between human factors and performance, or indicators that may provide a signal that a controller is reaching their personal 'edge' of performance.

An investigation of the indicators of potential performance decline in ATCOs may enhance knowledge of the use of indicators within an operational environment, and examine if indicators could be utilized to support performance prior to a performance decline, potentially contributing to preventing performance – related incidents. In addition, within the air traffic domain, identifying indicators of potentially declining ATCO performance may have implications both for research and live operations. A challenge of conducting research in the safety critical domain of air traffic control is the restriction of measurements during live operations. Consequently, air traffic control research is predominantly conducted within simulated, or other controlled, environments. Knowledge of consistent indicators would provide a basis for observation metrics, which may be employed in operations rooms, supporting research in live operations. In addition, knowledge of consistent indicators may support training for operations personnel. Results may provide a standardized foundation of knowledge for identifying when controllers maybe reaching the 'edge of performance'. Supervisors and unit competency examiners may be provided training to ensure a standardized foundation of knowledge. Performance could then be supported, potentially preventing performance decline and subsequently, performance-related incidents.

METHOD

Design

A total of 22, one hour, face to face semi-structured interviews were conducted with en-route ATCOs. Participants were selected to represent several sub-groups of controllers, based on age, sex and experience, role and nationality. Interviews took place in the Maastricht Upper Area control centre (MUAC). An interview schedule was developed to guide the semi-structured interview; participants were asked pre-designed lead questions which were then followed by probes. A protocol was used to standardize the interview method procedure. Participants were asked to list indicators of potential performance decline that were perceived to be associated with a specific factor (e.g. "What indicators of potential performance decline are associated with the influence of high workload on performance?"). The factors included in the interview study were: workload, stress and SA. Interviews were tape recorded and were then orthographically transcribed. Thematic analysis was applied as the analysis strategy.

Participants

In total 22 en-route controllers were interviewed. The majority of participants were male (17, or 77.27%), compared to female participants (5, or 22.73%). All participants worked as en-route controllers in the Maastricht Upper Airspace control (MUAC) centre. Participants' ages ranged from 21-60. Years of experience as an ATCO (excluding training) ranged from 1.5-31 (M=14.55, SD=8.68).

RESULTS AND DISCUSSION

A total of 22 one-hour interviews were orthographically transcribed and analysed using thematic analysis. Recurrent and relevant themes are presented below with illustrative quotes from participants.

Use of indicators of potential performance decline in air traffic control

All participants (22/22) were familiar with the concept of indicators and agreed that they occurred in the ATC operations room. All controllers stated indicators are commonly used in ATC to provide information regarding when a controller may not be controlling optimally, or a factor (such as fatigue, stress) is negatively influencing performance. All controllers monitor their own personal indicators as well as colleagues' indicators*: "…We work close together, we monitor each other you would be very aware of the person sitting beside you whether they're on the ball or whether they're slightly less, whether they're tired, whether they're distracted by whatever or something, it's part of the job and you do, you make allowances"* (Participant 2).

Controllers used indicators for the primary purpose of gaining information, which subsequently led to supporting performance. After identification of a marker, controllers frequently reported applying a supportive mechanism to maintain performance, such as a change in control strategy to mitigate the issue that was negatively influencing the controller: *"...it's that point [of recognising something is wrong] where you have to, well in my opinion you have to change the way that you're controlling the traffic"* (Participant 10). This process occurs both for controllers to support their own performance, and controllers who apply adaptive strategies to support their colleagues' performance. The application of the compensation strategy can result in the maintenance of performance even when associated with negative factor influences. The experienced negative influence may therefore not be reflected in the controllers' performance even though the influence is present. Therefore, some indicators will not strictly associate with ATCO performance, but rather the ATCO's subjective state.

Categorisation of indicators used by air traffic controllers

Figure 1 provides an initial broad categorisation of indicators based on qualitative data from controllers. It should be noted that the categories are not mutually exclusive. An indicator can be both experienced internally and also overtly reflected in performance, such as a feeling of being 'slow' and falling behind with traffic. Participants distinguished between internal, subjectively experienced indicators and external, observable indicators.

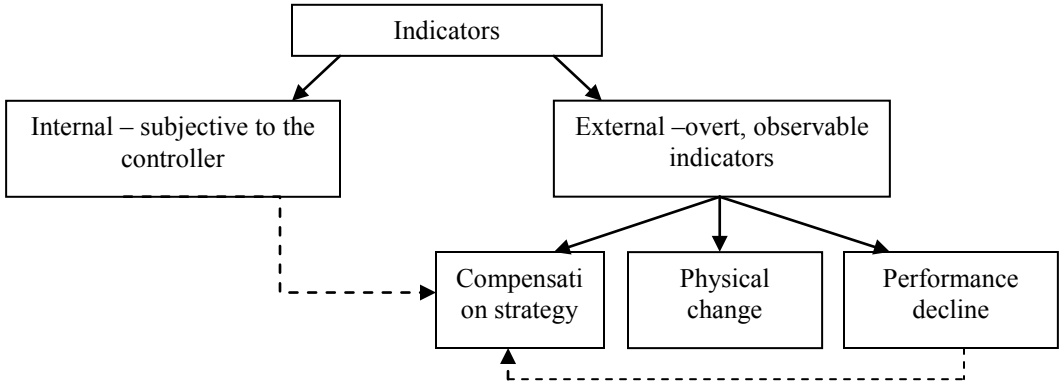


Figure 1. Categorisation of indicators produced from qualitative data

Internal indicators

Internal indicators may alert the controller to specific state or negative influence on performance*: "I know that when I start thinking, 'Oh it's going fine' I've learned that I force myself to tighten the bolts and to really pay extra attention"* (Participant 1).

External indicators

By contrast, external indicators are overt and observable. Three broad sub-categories of external indicators emerged from participant responses (Figure 1). Changes to a controller's performance, such as a performance decline, serve as an external marker to the controller his/herself and their colleagues that the controller may not be comfortable. Another sub-category of external indicators is behavioural and physical changes in a controller. Examples include a face becoming red, or fidgeting, which provide information to other controllers: *"You see it coming, you see them getting nervous, you see them talking faster"* (Participant 2). Adaptive changes to the control strategy to mitigate negative influences on performance are a sub-category of external indicators. Compensation strategies as indicators suggest that a controller is feeling uncomfortable, but is aware of the potentially negative influence, and is attempting to protect and maintain performance. As one controller summarised *"When somebody is just extra careful, I suppose that it's because they feel that they have to be extra careful"* (Participant 1).

If a controller is aware of internal indicators, a compensation strategy may be selected and applied. This is represented in Figure 1 by an arrow with a dashed line connecting the internal marker to the compensation strategy. This is then observable in the method of controlling. Alternatively, a decline in performance may alert the controller that a compensation strategy is required to protect performance. This is represented by an arrow with a dashed line

connecting performance decline and compensation strategy. In addition, colleagues may observe the controller's performance decline and apply their own compensation strategy to support the controller's performance.

Sub-categories of external indicators may provide distinct information

Performance declines or errors may provide a more serious indication that a controller is experiencing difficulties with the task: *"If I see that someone is correcting themselves very often then I would pay a lot of attention to what he's actually doing … I really follow every single clearance. I try to support as well like giving hints"* (Participant 11). A physical change (e.g. red face, yawning, laid back posture) may indicate a change in controllers' cognitive state (Sharples et al., 2012), although may not be related to feelings of discomfort. Finally, adaptive changes in control strategy inform colleagues that the executive controller (EC) is experiencing discomfort with the control task, although the EC is aware of this and is attempting to maintain performance with the application of a compensation strategy.

Indicators of specific factor influences

Participants were presented with a list of factors (workload, stress, situation awareness), accompanied by standardised definitions, developed from previous research (see Edwards, Sharples, Wilson & Kirwan, 2012). Participants were asked to identify internal or external indicators that they believed to be associated with a specific influencing factor. Although not contained in the study questions, all participants naturally reported adaptive compensation strategies that were applied in response to the detection of potential performance decline. These strategies can also serve as observable indicators. Therefore compensation strategies that controllers use in response to factor influences are also reported.

Workload

Controllers differentiated between high workload and low workload. Each form of workload was reported to be associated with different indicators.

High workload

Participants reported internal (Table 1) and external (Table 2) indicators of potential performance decline that were associated with high workload. Findings were grouped into categories. Changes to subjective feelings and performance changes were reported as important indicators that a controller may be reaching the edge of performance.

Table 1: Indicators of high workload internal to the controller

Category	Indicators
Cognitive changes	Don't know the next steps Increase focus Calls are a surprise
Changes to control	More reactive No back-up plan Future plan reduces in minutes ahead

Table 2: Observed indicators of high workload

Category	Indicators
Perception changes	Can't talk to executive/ executive doesn't hear you
Performance changes	Miss actions Can't see simple solutions Overlook aircraft
Visible cues	Fidgeting Moving closer to the screen

Controllers reported using specific compensation strategies in high taskload periods if they were aware of potential performance decline. These were primarily control strategies such as reducing efficiency to ensure safety, or going 'back to basics' to ensure all aircraft are safe. Preparation for a high taskload was reported to be the most effective strategy.

Low workload

Table 3 and Table 4 list the internal and external and compensation strategy indicators of reaching the edge of performance during low workload. In comparison to high workload, the indicators reflect a potential influence on performance through potential boredom or relaxation, leading to distraction: *"In low workload, there's nothing to do so you start doing other things, boredom becomes an issue and then you start talking or having a chat or doing whatever and it's, yeah, you can miss things"* (Participant 10). One particularly interesting marker during low workload is that controllers report leaving a problem develop for longer or creating complex situations to reduce boredom.

Table 3: Indicators of low workload internal to the controller

Category	Indicators
Cognitive changes	Pay less attention Easily distracted Reduced awareness
Changes to control	Leave situations develop for longer Trying to create more complex situations Less safety buffer
Subjective feeling	Boredom Relaxed

Table 4: Observed indicators of low workload

Category	Indicators
Visible cues	Sit back in chair Talking to colleagues
Performance changes	Overlooking aircraft Forgetting aircraft Fall behind traffic due to distraction

Several distinct types of performance decline are associated with workload, such as overlooking aircraft (vigilance issues). Workload may therefore negatively influence other factors such as vigilance and situation awareness which are then observed to be causal factors of performance decline. This is important for understanding underlying causes of performance decline that may manifest as a different factor.

Stress

Respondents differentiated between stress resulting from personal situations and task-related stress, which were both reported to negatively influence performance. Respondents suggested that the indicators and effects of stress were the same regardless of cause. Respondents also differentiated between positive, or 'excited', stress and stress which results in negative feelings and potentially performance change.

Only indicators of stress that influenced controllers negatively were discussed. Respondents emphasised changes in subjective feeling, such as feeling tense, uncomfortable and anxious, as unambiguous indicators of stress (Table 5). This suggests that stress may affect subjective experience and associated cognitive changes rather than performance directly. Several observable indicators (

Table 6) were manifestations of emotional responses, such as frustration and demonstrations of anger, and associated physiological changes such as vocal changes, shaking and fidgeting. Compensation strategies (**Error! Reference source not found.**) were designed to counteract the influences of stress on the controller, such as emotion regulation and practical strategies such as reducing rate of speech. Support was sought from the CC to further protect performance.

Table 5: Indicators of stress internal to the controller

Category	Indicators
Cognitive changes	Start to think slower
Physiological changes	Heartbeat increase Sweat
Subjective feeling	Not coping Feeling doing badly/ uncomfortable (negative) Anxious (negative) Nervous Tense

Table 6: Observed indicators of stress

Category	Indicators
Visible cues	Fidgeting Red cheeks/neck, flushed Flushed Sit closer to the screen Shaking
Performance changes	Falling behind Incorrect instructions
Changes to voice	Speaks faster (negative) Speaks higher (negative) Speaks louder Speaks quieter

Table 7: Compensation strategies which are also indicators of stress

Category	Indicators
Verbal changes	Speak slower More authoritative in instruction
Emotion regulation	Reduce stress Sit back, reduce anxiety Relax

Situation Awareness

Controllers referred to a decline or loss of situation awareness (SA) as 'losing the picture'. The loss of SA was reported to be progressive and occur in stages which were associated with different indicators: *"It starts off by just falling behind a bit. So you might just be a few steps behind what you're supposed to be doing and if that builds up too much then you will get to the point where you start to lose the picture"* (Participant 20). Therefore, the internal and observable indicators for SA were categorised into indicators which indicate a controller is progressing to losing the picture, or that a controller has lost the picture, and therefore SA (Table 8, Table 9). The context of high or low taskload was reported to influence SA, and the associated indicators of SA, differentially. The progressive decline of SA was only reported under conditions of high taskload. In low traffic the loss of awareness was more instantaneous: *"We sort of relaxed, 'Oh, it's done now', eating a sandwich and both of us had forgotten about it [the aircraft]"* (Participant 4). Controllers felt that recovery of SA was relatively easy in periods of low traffic as compared to high traffic.

Compensation strategies protect performance when a controller is losing the picture. The compensation strategies from the EC attempt to make the situation safe when awareness is degraded. Conversely, compensation strategies by the CC are tactical and appear to facilitate the EC in rebuilding the picture. The more degraded awareness is, the

more reliant the EC may be on the CC to protect performance and rebuild the picture.

Table 8: Indicators of reduction and loss of the picture internal to the controller

Category	Indicators internal to the controller of losing the picture	Indicators internal to the controller having lost the picture
Cognitive changes	Difficulty selecting priorities Thinking whilst giving the clearance Tunnel vision/hearing	Lose awareness Everything a surprise No plan Cannot see a solution
Changes to control	Reduction of future plan	Reactive control
Subjective feeling	Under confident	Panic

Table 9: Observed indicators of reduction and loss of the picture

Category	Observable indicators of losing the picture	Observable indicators of having lost the picture
Visible cues	Slow at task	Zig-zagging head movement of where to look 'Blacked out'/ silent
Performance changes	Running behind Time of planning ahead degrades Missing calls	Unsafe clearance Unexpected decisions Jumping from one aircraft to another Don't know who's calling

The importance of awareness of indicators, and effects of individual differences

All participants reported that indicators provide information that performance is, or may soon be, negatively influenced by factors such as workload and stress. Indicators, and controller awareness of these indicators, therefore appear to have a critical role in the maintenance of an extremely high level of human performance, even when potentially negatively influencing factors are present. However, a common theme throughout the results was that employing a compensation strategy to support performance was dependent on the controller's awareness of the available internal and external indicators. One participant summarised: *"I'd say 300%, if you know that you're not being top performing today then that's fine, just adapt your working style and you'll get through the day…if you don't know it and you're still trying to do the same then it might end in tears"* (Participant 12).

When an influencing factor is present (e.g. extremes of workload, stress, reduction in SA), performance may be protected by several 'barriers' (created from awareness and compensation strategies) before becoming vulnerable to factor influences. Based on participant responses, if an influencing factor occurs (e.g. workload extremes) internal indicators such as feelings of discomfort may alert the EC and trigger the application of a compensation strategy. Performance may then be maintained. If an internal marker did not occur or was not detected, another opportunity to detect the issue may occur through observable indicators. For example: *"It's getting busy… you start speaking fast*

and then somebody says "Say again" and then that's it, you have a hint. 'Okay good, I have to slow down'" (Participant 1). However, if the EC is not aware of indicators, protection of performance is dependent on a colleague's (i.e. the CC's) awareness: *"You're not aware that you're working to the edge of your performance then you need to rely on other people to tell you"* (Participant 15). If neither controller notices an issue, participants suggested that performance is more likely to decline than if a compensation strategy was applied. This concept may not be applicable in all situations, although captures many elements of controllers' answers.

Controllers reported that indicators of potential performance decline are not formally taught but are learned through experience. Therefore, differences in experience may limit a controllers' awareness of internal and external indicators. Controllers need to build awareness of the respective indicators prior to being able to identify the specific indicators. This is particularly relevant to inexperienced controllers.

Individual differences in awareness may alter the extent that a controller is aware of these indicators. Participants appeared to differ in the extent of conscious awareness of personal indicators. A minority of controllers (3/22) suggested that they personally 'sense' or 'just know' when they are reaching a performance limit but could not identify how they knew. In contrast, most controllers could identify personal indicators.

OVERALL DISCUSSION

Expert opinion regarding indicators of potential performance decline was generated by conducting one hour face-to-face interviews with 22 en-route controllers from MUAC. Interviews were orthographically transcribed and analysed using thematic analysis. Results revealed that indicators were used in an ATC setting by all respondents, as indication of when a controller was reaching the edge of performance, or a factor was negatively influencing performance. It was considered a natural process that all controllers used. Participants differentiated between internal indicators, representing a subjective experience, and external indicators, which were observable. Three sub-categories of external indicators were identified: changes in performance, physical signs and application of a compensation strategy.

Participants confirmed that specific factor influences on performance were associated with specific internal and external indicators. Indicators were identified for the factors of workload, stress and SA. Participants also reported adaptive changes in control strategy that were applied to maintain performance when the influencing factor was present, which resulted from the identification of indicators. Compensation strategies are therefore an integral aspect of the application of indicators in ATC. In addition, compensation strategies were also identified to be used as external indicators themselves of controller discomfort.

Specific factors influenced performance differentially, which in turn influenced the associated indicators. The influences of workload and stress may not be visible in performance but manifest as other factor influences, such as overlooking aircraft (a vigilance issue). It is therefore important for aviation professionals to acknowledge the underlying issues of performance declines to gain a valid and comprehensive understanding about factor influences and to understand how to best protect performance. Factors such as SA appear to influence performance directly.

Awareness emerged as an integral element in the use of indicators; controllers needed to be aware of their own or colleagues' indicators in order to apply a compensation strategy. It was suggested that there were individual differences in overall levels of awareness. In addition, controllers suggested that it was harder to be self-aware than aware of colleagues' indicators. This was especially true for inexperienced controllers who were perceived to not have the experience in order to identify indicators and apply adaptive strategies. Awareness is critical in the identification of indicators and subsequent application of adaptive strategies to protect performance. Supervisors and controllers may benefit from a more standardised list of generic indicators, and in addition, possibly learn different indicators to monitor from each other. Workshops that provide standardised indicators for which to monitor, and support development of awareness of indicators for self and colleagues, may support controllers in protecting and maintaining performance in the presence of negative influences. This may also support trainees in protecting performance whilst developing the required experience to identify specific, personal indicators.

CONCLUSIONS

The findings from this qualitative interview study with 22 air traffic controllers suggests that indicators of potential performance decline are currently utilized in air traffic control operations rooms. Indicators are reported to have a functional purpose which enables controllers to adapt to, and mitigate, potential negative influences on performance,

therefore facilitating controllers' maintenance of an exceptionally high level of performance. In order for indictors to be utilized to maintain performance, operational personnel's awareness of indicators emerged as a central theme. This finding suggests operational personnel such as supervisors may benefit from a formal training to raise awareness of these indicators, or utilization of a list of generic indicators. Findings have implications for both live operations, and research in the air traffic control domain. Further research must now investigate the validity and reliability of utilizing these indicators as a basis for observation of potential performance declines. In addition, further research may investigate the application of standardized indicators to highlight to supervisors when a controller's performance may benefit from supportive strategies. In this way, performance declines, and ultimately, performance-related incidents, may be prevented.

REFERENCES

Chang, Y. & Yeh, C. (2010), Human performance interfaces in air traffic control. *Applied Ergonomics, 41*, 123-129.

Edwards, T., Sharples, S., Wilson, J. R., and Kirwan, B. (2012). Factor interaction influences on human performance in air traffic control: The need for a multifactorial model. *Work: A Journal of Prevention, Assessment and Rehabilitation, 41(1)*, 159-166.

Kirwan, B. (2011). Incident reduction and risk migration. *Safety Science, 49(1)*, 11-20.

Mackieh, A. & Cilingir, C. (1998). Effects of performance shaping factors on human error. *International Journal of Industrial Ergonomics, 22*, 285-292.

Park, K. S. & Jung, K. T. (1996). Considering performance shaping factors in situation-specific human error probabilities. *International Journal of Industrial Ergonomics, 18*, 325-331.