Advances in Intelligent Systems and Computing 1212

Neville Stanton *Editor*

Advances in Human Aspects of Transportation

Proceedings of the AHFE 2020 Virtual Conference on Human Aspects of Transportation, July 16–20, 2020, USA

Springer

Advances in Intelligent Systems and Computing

Volume 1212

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within "Advances in Intelligent Systems and Computing" are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink ****

More information about this series at http://www.springer.com/series/11156

Neville Stanton
Editor

Advances in Human Aspects of Transportation

Proceedings of the AHFE 2020 Virtual Conference on Human Aspects of Transportation, July 16–20, 2020, USA

Springer

Editor
Neville Stanton
Boldrewood Innovation Campus
University of Southampton
Southampton, UK

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-030-50942-2 ISBN 978-3-030-50943-9 (eBook)
https://doi.org/10.1007/978-3-030-50943-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Advances in Human Factors and Ergonomics 2020

AHFE 2020 Series Editors

Tareq Z. Ahram, Florida, USA
Waldemar Karwowski, Florida, USA

11th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences

Proceedings of the AHFE 2020 Virtual Conference on Human Factors in Transportation, July 16–20, 2020, USA

Advances in Neuroergonomics and Cognitive Engineering	Hasan Ayaz and Umer Asgher
Advances in Industrial Design	Giuseppe Di Bucchianico, Cliff Sungsoo Shin, Scott Shim, Shuichi Fukuda, Gianni Montagna and Cristina Carvalho
Advances in Ergonomics in Design	Francisco Rebelo and Marcelo Soares
Advances in Safety Management and Human Performance	Pedro M. Arezes and Ronald L. Boring
Advances in Human Factors and Ergonomics in Healthcare and Medical Devices	Jay Kalra and Nancy J. Lightner
Advances in Simulation and Digital Human Modeling	Daniel N Cassenti, Sofia Scataglini, Sudhakar L. Rajulu and Julia L. Wright
Advances in Human Factors and Systems Interaction	Isabel L. Nunes
Advances in the Human Side of Service Engineering	Jim Spohrer and Christine Leitner
Advances in Human Factors, Business Management and Leadership	Jussi Ilari Kantola, Salman Nazir and Vesa Salminen
Advances in Human Factors in Robots, Drones and Unmanned Systems	Matteo Zallio
Advances in Human Factors in Cybersecurity	Isabella Corradini, Enrico Nardelli and Tareq Ahram

(continued)

(continued)

Advances in Human Factors in Training, Education, and Learning Sciences	Salman Nazir, Tareq Ahram and Waldemar Karwowski
Advances in Human Aspects of Transportation	Neville Stanton
Advances in Artificial Intelligence, Software and Systems Engineering	Tareq Ahram
Advances in Human Factors in Architecture, Sustainable Urban Planning and Infrastructure	Jerzy Charytonowicz
Advances in Physical, Social & Occupational Ergonomics	Waldemar Karwowski, Ravindra S. Goonetilleke, Shuping Xiong, Richard H.M. Goossens and Atsuo Murata
Advances in Manufacturing, Production Management and Process Control	Beata Mrugalska, Stefan Trzcielinski, Waldemar Karwowski, Massimo Di Nicolantonio and Emilio Rossi
Advances in Usability, User Experience, Wearable and Assistive Technology	Tareq Ahram and Christianne Falcão
Advances in Creativity, Innovation, Entrepreneurship and Communication of Design	Evangelos Markopoulos, Ravindra S. Goonetilleke, Amic G. Ho and Yan Luximon

Preface

Human Factors and Ergonomics have made a considerable contribution to the research, design, development, operation and analysis of transportation systems. This includes road, rail, aviation, aerospace and maritime vehicles as well as their complementary infrastructure. This book presents recent advances in the Human Factors aspects of Transportation. These advances include accident analysis, automation of vehicles, comfort, distraction of drivers (including 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 interactions of humans with systems in the environment. Moreover, many of the chapter topics cross domain and discipline boundaries. This is in keeping with the systemic nature of the problems facing human factors experts in rail and road, aviation and aerospace, and maritime research– it is becoming increasingly important to view problems not as isolated factors 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 ten sections:

Human Factors in Transportation: Road & Rail

Section 1 Traffic Behavior and Driver Performance
Section 2 Driving Automation
Section 3 Accidents
Section 4 Comfort and Posture

Section 5 Vulnerable Road Users
Section 6 Transport Planning and Infrastructure Design
Section 7 Route Choice, Navigation and Wayfinding

Human Factors in Transportation: Maritime

Section 8 Transportation: Maritime

Human Factors in Transportation: Aviation and Space

Section 9 Exploiting Contemporary Technology in Flight Deck Design to Improve Flight Safety
Section 10 Aviation and Space

This book will be of interest and use to transportation professionals who work in the road and rail, aviation and aerospace, 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 experts and academic leaders in their respective fields, and their help is very much appreciated, especially as they gave their time to the project. Special thanks to Giorgio Musso and Nancy Currie-Gregg for their contribution to the Space program.

Road and Rail

C. Allison, UK
G. Balbinotti, Brazil
K. Bengler, Germany
S. Birrell, UK
G. Burnett, UK
P. Chapman, UK
F. Chen, Sweden
D. Coelho, Portugal
B. Colucci Rios, Puerto Rico
G. Craveur, France
L. Dickson-Bull, USA
L. Dorn, UK
I. Glendon, Australia
I. Grabarek, Poland
R. Grice, USA
R. Happee, The Netherlands
S. Jamson, UK
D. Kaber, USA

J. Krems, Germany
M. Lenné, Australia
E. Macioszek, Poland
F. Mars, France
D. McAvoy, USA
A. Mills, UK
R. Philipsen, Germany
K. Revell, UK
R. Risser, Austria
P. Salmon, Australia
G. Sierpiński, Poland
S. ur Rehman, Sweden
D. Valdes Diaz, Puerto Rico
G. Walker, Scotland
K. Young, Australia

Aviation

V. Banks, UK
M. Biella, Germany
C. Borst, The Netherlands
T. Edwards, USA
M. Feary, USA
A. Haslbeck, Germany
B. Hooey, USA
J. Huddlestone, UK
D. Kaber, USA
A. Majumdar, UK
L. Martin, USA
J. Mercer, USA
M. Mulder, The Netherlands
K. Plant, UK
J. Syversen, Norway
S. Verma, USA
D. Yacht, USA
K. Vu, USA

Space

D. Bedini, Italy
J. Beierle, USA
T. Bernard, USA
R. Capra, Italy
M. Ferrino, Italy
E. Gaia, Italy
S. Hauplick, Austria

K. Latorella, USA
K. Nieuwenhuis, The Netherlands
R. Ricci, Italy
A. Salatino, Italy
I. Schlacht, Italy
D. Tedone, Italy

Maritime

A. Alkan, Turkey
D. Andrews, UK
J. Dahlman, Sweden
G. Di Bucchianico, Italy
P. Ferreira, Portugal
D. Gray, USA
M. Grootjen, The Netherlands
M. Hänninen, Finland
T. Koester, Denmark
S. MacKinnon, Canada
S. Mallam, Norway
M. Musio Sale, Italy
S. Nazir, Norway
S. Parisi, Greece
G. Praetorius, Sweden
A. Ratti, Italy
A. Roberts, UK
J. Smith, Canada
A. Vallicelli, Italy
F. van de Merwe, Norway
H. van den Broek, The Netherlands
N. Ventikos, Greece

July 2020 Neville A. Stanton

Contents

Traffic Behavior and Driver Performance

Determining Infrastructure- and Traffic Factors that Increase the Perceived Complexity of Driving Situations

Anika Boelhouwer[1(✉)], Arie Paul van den Beukel[1], Mascha C. van der Voort[1], and Marieke H. Martens[2,3]

[1] University of Twente, 7522NB Enschede, The Netherlands
{a.boelhouwer, a.p.vandenbeukel, m.c.vandervoort}@utwente.nl
[2] Eindhoven University of Technology, 5612AE Eindhoven, The Netherlands
[3] TNO Traffic and Transport, 2496RZ The Hague, The Netherlands
marieke.martens@tno.nl

Abstract. When designing experimental studies in the driving domain, an important decision is which driving scenarios to include. It is proposed that HMI need to be adaptive to the complexity of the driving situation, in order to avoid overloading the driver. To further study adaptive HMI a comprehensive list of factors that determine the perceived complexity of a driving situation is required, yet absent. In this, infrastructure- and traffic characteristics that may influence the perceived complexity of a driving situation were collected from literature. Next, four sets of driving scenarios of varying complexities were created and validated in an online survey. The results of this study include: 1) a list of infrastructure- and traffic characteristics that influence the overall complexity of a driving situation, and 2) validated scenarios of varying complexities. These outcomes help researchers and designers in setting up future driving studies.

Keywords: Automated driving · Human-machine interaction · Driving situations · Complexity

1 Introduction

It is currently studied how the Human Machine Interfaces (HMI) in cars may support drivers in their new supervisory role in partially automated cars. It is suggested that the HMI should be adaptive in order for the driver to fully understand the automation's capabilities and limitations, and to use the automation safely and efficiently [1–3]. The HMI may be adaptive to any of the three core factors in a driving situation: 1) the driver, 2) the car state, and 3) (the complexity of the) driving situation. That is, any changes in these core factors may require the driver to get different information about the automation. In this study we focus on (the complexity of the) driving situation, as the required information to safely operate a partially automated car may depend on this for several reasons. First, the driver needs situated information about the environment in order to be able to understand and anticipate their own role. Second, drivers need

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 3–10, 2020.
https://doi.org/10.1007/978-3-030-50943-9_1

information that is adaptive to the environment as their processing capability is dependent on the complexity of the driving situation [3, 4]. Studies have already shown that an increase in several environmental elements, such as traffic density or traffic signs, can increase drivers' workload and take-over times [5–8]. Information interfaces in the car may need to be adaptive in these situations to avoid overloading the driver. For example, drivers may require more condensed or simplified information from the HMI in highly complex situations.

While the complexity of a driving situation may largely influence the information needs of a driver, no base scenarios of varying complexities exist to this day to further investigate this. Furthermore, while other studies [5, 6] have investigated the effect of one specific road- or traffic element on drivers' mental workload, there is no condensed overview of elements which may contribute to the overall complexity when designing scenarios. Also, the listed studies mainly focused on highway scenarios. It is unclear how the results translate to urban and rural scenarios. The work by Fastenmeier [9] presents a great base for defining the complexity of road sections and will be used as the foundation of this research. Still, this research only focusses on the infrastructural aspects but not the variation in traffic.

The goal of this study was to identify core environment characteristics that contribute to the overall complexity while creating and validating a set of reusable scenarios of varying complexities. This was achieved by applying various infrastructure- and traffic characteristics found in literature to five basic road types (straight, curve, intersection, roundabout, highway). These were then rated on their overall-, infrastructure- and traffic complexity in an online survey.

2 Methods

The road analysis schema described by Fastenmeier [7] was used as a base to create scenarios of different complexities for both urban and highway settings. This schema is based on both the physical and mental processing load of a driver in specific infrastructure sections. Based on other studies like those by Radlmayr and Zeitlin [5, 9] we added several traffic- and infrastructure characteristics to the scenarios. An overview of all scenarios can be found in Fig. 1. Then, an online survey was conducted to assess the complexity of driving scenarios, and the elements that may increase the complexity. In this survey, participants rated each scenarios on its overall-, traffic- and infrastructure complexity.

2.1 Participants

A total of 143 respondents completed our online survey. As the survey was distributed on the American online crowdsourcing platform MTurk, all respondents were from the US. The youngest respondent was 21 years old, while the oldest was 73 years old (M = 34.06, SD = 8.83). 51% of the respondents was male and 49% was female. The education level was relatively high with 58.1% having at least a bachelor's degree. Respondents were required to have a driver's license to be able to participate. The majority (76.2%) had their driver's license for more than 10 years, 18.2% had it between five and 10 years and 5.6% had it for less than 5 years.

2.2 Scenarios

Infrastructure. Five frequently occurring base scenarios were chosen: straight road, curve, intersection, roundabout and highway (Fig. 1 column a). The road analysis schema by Fastenmeier [9] was subsequently used in order to vary the infrastructure complexity of these five base scenarios. This schema clusters road elements on their physical and mental processing load. The work of Fastenmeier [9] is an integration of own research and prior work by (amongst others) Benda and Hackman [10, 11].

In the schema by Fastenmeier [9], certain road elements either increase or decrease the complexity for the driver. Road elements that increase the complexity according to the schema are: sharp turns, slopes, obstacles, road narrowing, absence of road lines, and having to give priority regulated by signs. Elements that should relieve the load of drivers are unobstructed views, signalized intersections (instead of priority to the right or signed priority) and separated lanes by for example a center divider. Other elements that may increase the complexity are unique or unfamiliar road constructions. Similarly, the study by Matthews et al. [12] showed increased stress when encountering unfamiliar road situations. Therefore, we included a Dutch style roundabout with integrated bicycle paths. This was expectedly an unfamiliar sight for the respondents that were from the US. To create scenarios of higher infrastructure complexity the workload inducting elements were applied to the five base scenarios (Fig. 1 column b).

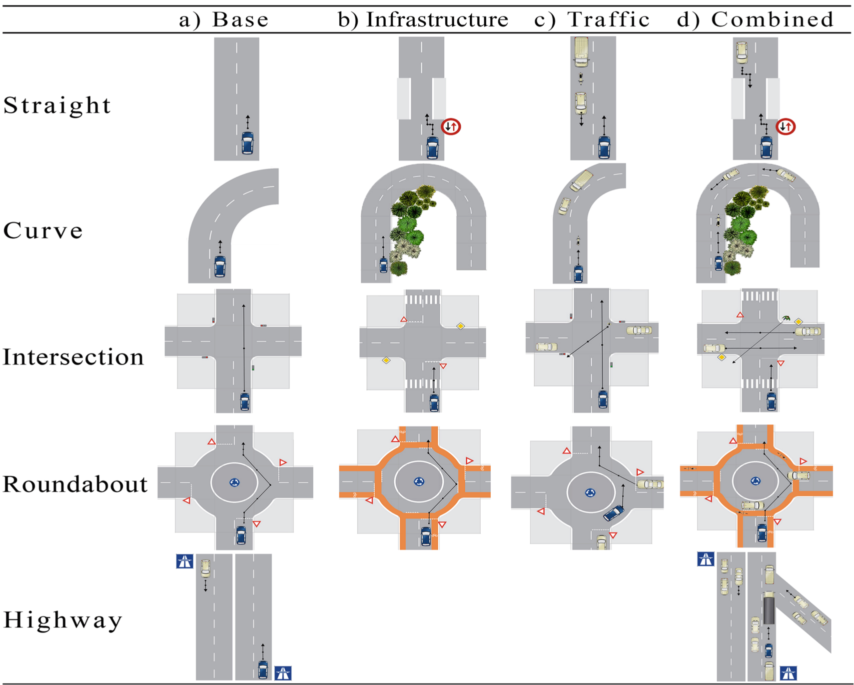


Fig. 1. The constructed scenarios: a) base scenarios, b) base and added infrastructure complexity, c) base and added traffic d) traffic- and infrastructure changes combined.

Traffic. Next, traffic was added to the base scenarios (Fig. 1 column c). Several traffic factors may influence the complexity of a driving situation. First, prior studies like that of Radlmayr et al. and de Waard [5, 13] have already shown the effect of traffic density on drivers' mental workload. Still, there are other traffic related factors that can make driving situations more or less complex for the driver. A second important overall factor is the predictability of other road users. The predictability of other road users showed to have a large impact on the perceived difficulty of a driving situation [14]. Unpredictable traffic behaviour may include road users that break traffic rules. Furthermore, the variety in road user types and differences in travel direction can decrease predictability and increase the situation complexity.

To create situations of higher traffic complexity, the factors of traffic density, traffic predictability and giving priority without support were applied. For the traffic density, the number of road users was increased. For the predictability, a variety of road user types, travel directions and road users that break the rules were included. To include priority without support, road users were included that should receive priority from the ego vehicle. Lastly, both the infrastructure- and traffic adaptations were combined to create the last scenarios (Fig. 1 column d).

2.3 Materials and Procedure

This experiment consisted of a survey distributed through an online crowdsourcing platform by Amazon, MTurk. This platform allowed us to gather a large amount of respondents relatively fast for a small compensation [15]. The average completion time of the survey was 24 min. Each participant was compensated $2 for their time. The start of the survey contained four demographics questions regarding the respondents age, gender, driver's license and education level. The questionnaire contained 18 different scenarios, each was displayed at random and twice in the experiment. Each scenario was rated on the overall-, infrastructure- and traffic complexity. The answer scales ranged from 1 (very simple) to 9 (very complex).

3 Results

3.1 Overall Complexity

First, the overall complexity of each of the five basic road types (straight, curve, intersection, roundabout, highway) were inspected without any infrastructure- or traffic additions. The base straight, curve and highway scenarios scored relatively low with respective overall scores of $M_{Straight} = 1.4$ ($SD = 1.32$), $M_{Curve} = 1.7$ ($SD = 1.39$) and $M_{Highway} = 1.8$ ($SD = 1.39$). The intersection and roundabout scenarios appeared to have higher scores with $M_{Intersection} = 2.3$ (SD = 1.52) and $M_{Roundabout} = 4.1$ ($SD = 1.39$). A One-way ANOVA confirmed that the road types scored differently on their complexity ($F(4,710) = 70.22$, $p < .001$). Following t-tests showed that all base road types were significantly different from each other (all $p < .05$), just the highway and curve scenarios were not scored differently from each other ($t(284) = -.43$, $p = .67$).

Following, it was analyzed whether the traffic and infrastructural changes that were applied to increase the overall complexity actually increase the overall scores. The base scenarios scored an average of $M_{Base} = 2.3$ ($SD = 1.22$). The scenarios that had their infrastructure changed to increase complexity were scored an average of $M_{Infrastructure} = 3.6$ ($SD = 1.35$) while the scenarios with changed traffic scored an average of $M_{Traffic} = 3.7$ ($SD = 1.22$). The scenarios that combined the infrastructure and traffic changes scored an average of $M_{Combined} = 5.2$ ($SD = 1.20$). Figure 2 shows an overview of the complexity scores for all scenarios.

A One-way ANOVA confirmed that the base-, infrastructure-, traffic- and combined scenarios were scored differently ($F(3,568) = 128.67$, $p < .001$). Paired t-tests showed that the base scenarios were scored significantly lower than the infrastructure ($t(142) = -17.80$, $p < .001$), traffic ($t(142) = -19.02$, $p < .001$) and combined scenarios ($t(142) = -27.12$, $p < .001$). The infrastructure and traffic altered scenarios were not found to be different ($t(142) = -1.34$, $p = .18$). Still, both the infrastructure ($t(142) = -16.58$, $p < .001$) and traffic ($t(142) = -20.17$, $p < .001$) altered scenarios were scored lower than the scenarios in which the infrastructure and traffic changes were combined.

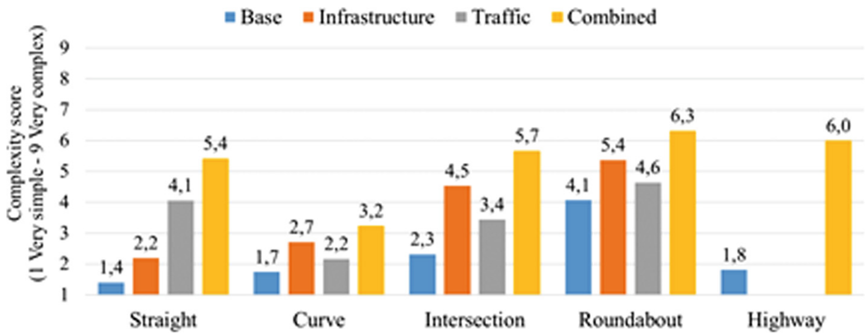


Fig. 2. Overview of the complexity scores for all scenarios.

3.2 Influence of Infrastructure- and Traffic Complexity on Overall Complexity

It was unclear to what extend the infrastructure and traffic contributed to the overall complexity of the scenarios. As each scenario was scored on the infrastructural-, traffic- and overall complexity, this could be analyzed through a multiple regression model. Scatterplots confirmed linear relationships between the infrastructure complexity scores and overall complexity scores. The same was found for the traffic complexity scores and overall complexity scores. Multicollinearity was not of concern (TrafficComplexity, Tolerance = .267, VIF = 3.74; InfrastructureComplexity, Tolerance = .267, VIF = 3.74). First a regression model with only traffic complexity scores as a predictor was created. This model showed that the traffic complexity scores explain the overall complexity scores well with an R^2 of .77 (F (1,141) = 464.032, $p < .001$). Following, the infrastructure complexity scores were added as a predictor. In this combined regression model, the infrastructure- and traffic scores explain the overall complexity very well with an R^2 of .86 (F(2,140) = 431.153, $p < .001$). In this model the overall complexity score is equal to .261 + .357 (InfrastructureScore) + .594 (TrafficScore). Increased traffic complexity scores appear to contribute to overall complexity scores more than infrastructure.

4 Discussion

During this study, we identified several infrastructure- and traffic elements in literature that were proposed to increase the perceived complexity of a driving situation. The infrastructural elements included: sharp turns, slopes, obstacles, road narrowing, absence of road lines, unfamiliar constructions and having to give priority regulated by signs. Traffic elements that were used to increase complexity were based on several studies and included: traffic density, varying road user types, multiple travel directions of other road users, road users that break the traffic rules and having to give priority without infrastructural support. These elements were integrated into multiple scenarios of varying complexity and validated through an online survey.

The results of this study show that both the infrastructural- and traffic changes indeed increased the overall complexity of scenarios. By combining these infrastructure- and traffic elements the overall complexity increased even further. While both contribute to the overall complexity, it is noteworthy that the traffic had a larger influence compared to the infrastructure. When constructing scenarios that need to be of high complexity, it therefore seems advisable to focus on (but not solely rely on) introducing traffic elements.

Interestingly, not only the infrastructural- and traffic changes contributed to the overall complexity of the scenarios. The overall complexity changes also based on the core road type (straight, curve, intersection, roundabout and highway). Curved roads were perceived as the least complex followed by the straight roads. Intersections and highways were considered slightly more complex. The roundabout scenarios were perceived to be more complex than the other scenarios. As described earlier, unfamiliarity may increase the complexity of a scenario. This may have contributed to the increased complexity as the respondents were American. Roundabouts, especially with dedicated bicycle lanes, are still relatively new in the US compared to in European countries. There are some elements that may further increase (or decrease) the complexity of driving scenarios that were not included in this study such as non-road environment, weather and communication by other road users. A city environment with highly dense and variable building is more complex to process compared to a rural road surrounded by similar meadows. Furthermore, verbal and behavioral communication by other road users can increase complexity.

Concluding, we identified infrastructure- and traffic elements that contributed to the overall complexity of driving situations which can be used to create new scenarios of varying complexities for driving studies. Furthermore, the study presents a validated selection of driving situations of varying complexities in Figs. 1 and 2. The results of this study are particularly useful to further study and develop adaptive driver feedback and support systems in (partially) automated cars.

Funding Statement. This research is supported by the Dutch Domain Applied and Engineering Sciences, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs [project number 14896].

References

1. Boelhouwer, A., van Dijk, J., Martens, M.H.: Turmoil behind the automated wheel. In: Krömker, H. (ed.) HCI in Mobility, Transport, and Automotive Systems, pp. 3–25. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22666-4_1
2. Lilis, Y., Zidianakis, E., Partarakis, N., Antona, M., Stephanidis, C.: Personalizing HMI elements in ADAS using ontology meta-models and rule based reasoning. In: International Conference on Universal Access in Human-Computer Interaction, pp. 383–401. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58706-6_31
3. Mueller, M.: Deficiency drive. Vision Zero International, pp. 44–47 (2014)
4. Kroon, E.C.M., Martens, M.H., Brookhuis, K., Hagenzieker, M., Alferdinck, J.W.A.M., Harms, I., Hof, T.: Human factor guidelines for the design of safe in-car traffic information services (2016)
5. Radlmayr, J., Gold, C., Lorenz, L., Farid, M., Bengler, K.J.: How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 2063–2067 (2014). https://doi.org/10.1177/1541931214581434
6. Lyu, N., Xie, L., Wu, C., Fu, Q., Deng, C.: Driver's cognitive workload and driving performance under traffic sign information exposure in complex environments: a case study of the highways in China. Int. J. Environ. Res. Public Health **14**, 1–25 (2017). https://doi.org/10.3390/ijerph14020203
7. Gold, C., Körber, M., Lechner, D., Bengler, K.J.: Taking over control from highly automated vehicles in complex traffic situations. Hum. Fact. J. Hum. Fact. Ergon. Soc. **58**, 642–652 (2016). https://doi.org/10.1177/0018720816634226
8. Jamson, A.H., Merat, N., Carsten, O., Lai, F.C.H.: Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. Transp. Res. Part C Emerg. Technol. **30**, 116–125 (2013). https://doi.org/10.1016/j.trc.2013.02.008
9. Fastenmeier, W.: Die Verkehrssituation als Analyseeinheit im Verkehrssystem. In: Fastenmeier, W. (ed.) Autofahrer und Verkehrssituation - Neue Wege zur Bewertung von Sicherheit und Zuverlässigkeit moderner Straßenverkehrssysteme, pp. 27–82. Verlag TÜV Rheinland, Köln (1995)
10. Benda, H.V.: Scaling of the dangerousness of road traffic situations. Part 1: A classification system for road traffic situations from the drivers' perspective (1977)
11. Hackman, J.R.: Tasks and task performance in research on stress. In: McGrath, J.E. (ed.) Social and Psychological Factors in Stress, pp. 202–237. Rinehart & Winston, New York (1970)
12. Matthews, G., Tsuda, A., Xin, G., Ozeki, Y.: Individual differences in driver stress vulnerability in a Japanese sample. Ergonomics **42**, 401–415 (1999). https://doi.org/10.1080/001401399185559
13. De Waard, D.: Driving behaviour on a high-accident-rate motorway in the Netherlands. In: Weikert, C., Brookhuis, K.A., Ovidius (eds.) Proceedings of the Europe Chapter of the Human Factors Society Annual Meeting (1991)

14. Zeitlin, L.R.: Estimates of driver mental workload: a long-term field trial of two subsidiary tasks. Hum. Fact. J. Hum. Fact. Ergon. Soc. **37**, 611–621 (2006). https://doi.org/10.1518/001872095779049327
15. Casler, K., Bickel, L., Hackett, E.: Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. Comput. Hum. Behav. **29**, 2156–2160 (2013). https://doi.org/10.1016/j.chb.2013.05.009

Development of Statistical Models for Predicting Automobile Seat Fit of Drivers

Baekhee Lee[1], Kihyo Jung[2], and Jangwoon Park[3](✉)

[1] Body Test Team 3, Hyundai Motor Company, Suwon, South Korea
x200won@hyundai.com
[2] School of Industrial Engineering, University of Ulsan, Ulsan, South Korea
[3] Department of Engineering, Texas A&M University-Corpus Christi, Corpus Christi, TX, USA
jangwoon.park@tamucc.edu

Abstract. The present study is intended to develop statistical models for predicting automobile seat fit based on the relationships between seat dimensions and subjective seat fit. The evaluations of the subjective seat fit for 43 different driver seats and the seat dimensions at six cross-sectional planes (three for the seatback and the other three for the cushion) were measured and evaluated by eight seat-engineers. The best subset logistic regression analyses were conducted to quantify the relationships between the measured seat dimensions and evaluated subjective seat fit at each of the cross-sectional planes. As a result, significant seat dimensions, such as insert width or bolster height, on the subjective seat fit were identified. The developed logistic models show 90% overall classification accuracy at each section with 80% accuracy with five-fold cross-validation. The developed models would be particularly useful to support seat engineers by providing recommended seat dimensions, which could increase seat fit. In addition, the model is useful to reduce development costs for an automobile seat and increase work efficiency in the digital evaluation process of an automobile seat.

Keywords: Seat fit · Seat dimension · Automobile seat · Logistic regression analysis

1 Introduction

Seat dimensions (e.g., height, depth, and width) are important measures affecting seating comfort, fit, and satisfaction in vehicles. Kolich [1] reported four factors affecting seating comfort, which were seat factor (e.g., geometry, dimensions, stiffness), package factors (e.g., seat height, knee room), social factors (e.g., brand, purchase price), and individual factors (e.g., anthropometry, posture). Jones et al. [2] developed statistical models for predicting a driver's seating pressure based on a driver's anthropometric features and seat dimensions. In addition, Reed and Flannagan [3] developed logistic regression models to predict the distribution of subjective fit for optimizing headroom height design.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 11–15, 2020.
https://doi.org/10.1007/978-3-030-50943-9_2

Although optimizing seat dimensions are important for better seating comfort, measuring seat dimensions are time-consuming work, and repetitiveness of the measurements are relatively low. Because seat engineers at automobile companies measure several seat dimensions using a computer-aided-design (CAD) software, and the measuring process (e.g., find landmark location on 3D scan or CAD file) is relying on manual tasks [4]. In order to improve the efficiency and repeatability of measuring seat dimensions, Park et al. [5] and Lee et al. [6] developed computerized systems that can measure and evaluate various seat dimensions automatically, which are defined as Society-of-Automotive-Engineers (SAE) standard J2732 [7] as shown in Fig. 1. The developed system can efficiently and accurately evaluate seat dimensions less than 2–3 min. However, the system does not provide seat fit estimation function because there was limited research on the relationships between seat dimensions and subjective seat fit.

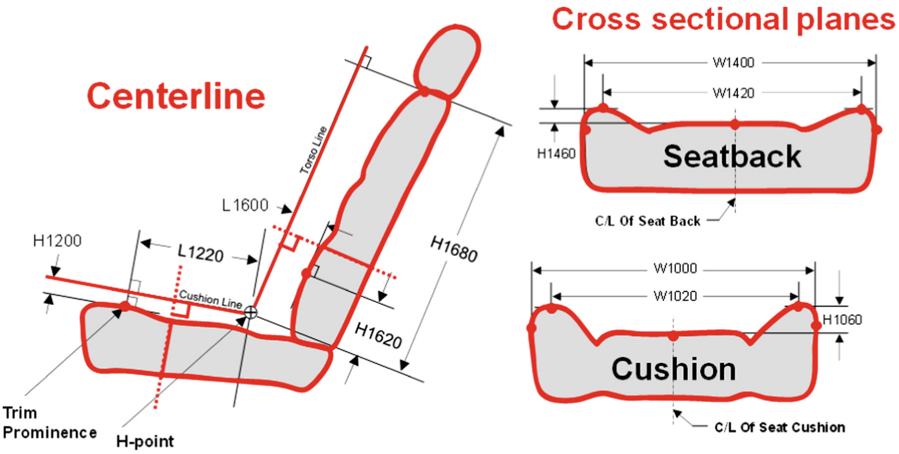


Fig. 1. Illustrations of SAE J2732 (2008) seat dimensions.

The present study is intended to develop seat fit prediction models by incorporating SAE seat dimensions. In this paper, we will show the data collection as well as model development and validation.

2 Data Collection

The subjective seat fit evaluations of 44 different driver-seats, including prototype user experience mock-up or prototype levels, were collected. Those subjective evaluations were recorded by eight-seat engineers at the automotive company R&D center. Those seat engineers are experts to evaluate subjective seat fit, and they are more sensitively and accurately evaluate seat fit compared to the general population. Across six cross-sectional planes on a driver seat, 20 SAE J2732 dimensions as well as H-Pt, Toro & Thigh angles, and cushion lengths were summarized in Excel worksheet. In addition, to classify the subjective fit evaluations, the subjective evaluations were coded into three classes (short, right, and wide) with numbers: −1 for short, 0 for right, or +1 for wide. The seat engineers used different terms when they evaluated subjective seat fit, such as narrow feeling, pressure, or support. We have unified those terms as subjective “fit” since all they are related to subjective seat fit. In addition, we classified the seat data based on the target market (domestic, international, or Europe), size of the vehicle (small, medium, or large), and type of the vehicle (sedan, SUV, or CUV).

3 Model Development and Validation

The model development and validation process consist of four steps: First, the data reliability of expert assessments was verified by using a contour plot. The contour plot is a 3D surface plot of expert evaluations on insert width (x-axis), bolster height (y-axis), and fitting evaluations (short, right, and bulky; z-axis) with colors (See Fig. 2). Second, important variables were selected based on their technical significance and statistical significance (see Fig. 3). Technically significant variables can be chosen that are directly related to the fitting results (e.g., insert width dimension is directly affecting seat fit evaluations of insert width). Statistically significant variables were chosen based on its statistical significance (e.g., $\alpha = 0.1$) as well as Spearman’s rank correlation coefficient with seat fit evaluations. Third, logistic regression models were developed that predict the probability, $p(\text{x})$ of seat fit (short, right, or bulky), based on the seat dimensions (x) (e.g., insert width, bolster height). If a predicted probability ($p(\text{x})$) is above the threshold (typically 0.5 is a default threshold), then it is determined to be of right fit; otherwise, it is not of the right fit, either short or bulky. In this step, we used best-subset logistic regression analysis to create models for predicting seat fit of three-seat dimensions (insert width, inboard bolster height, and outboard bolster height)

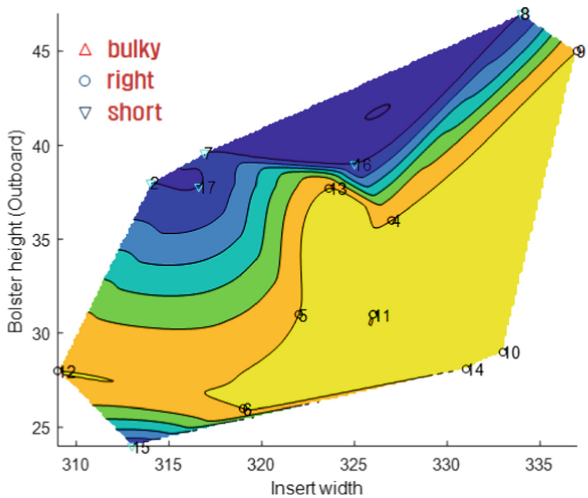
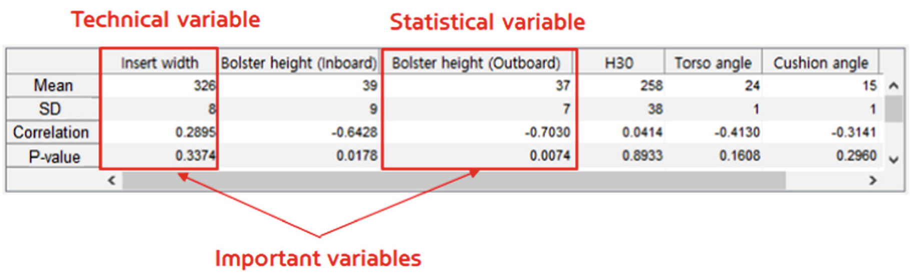


Fig. 2. An example of a contour plot of experts’ subjective seat fit evaluations.



	Insert width	Bolster height (Inboard)	Bolster height (Outboard)	H30	Torso angle	Cushion angle
Mean	326	39	37	258	24	15
SD	8	9	7	38	1	1
Correlation	0.2895	-0.6428	-0.7030	0.0414	-0.4130	-0.3141
P-value	0.3374	0.0178	0.0074	0.8933	0.1608	0.2960

Fig. 3. An example of variable selections to develop seat fit classification model.

at each of the six cross-sectional planes. The definition of the best subset is a subset (combination) of the most predictive variables among several important (independent) variables. Lastly, the developed models were validated with a 5-fold cross-validation process. As a result of the validation, we found that the developed models have around 90% classification accuracy and 80% cross-validated accuracy.

4 Discussion

The developed seat fit classification models are particularly useful to design optimal seat dimensions providing proper fit. Although seat engineers want to know the effects of seat dimensions on subjective seat fit, it is difficult to seamlessly check and evaluate the seat fit all the time. However, because the developed models can predict seat fit classes (short, right, or bulky) based on the seat dimension changes, the seat engineers can design the seat dimensions by continuously evaluating the estimated seat fit. This system can advance the existing seat-dimension design technology by providing great efficiency, and it allows the seat engineers to check the design issues in an early design stage.

Acknowledgments. The present study was supported by Hyundai Motor Company.

References

1. Kolich, M.: A conceptual framework proposed to formalize the scientific investigation of automobile seat comfort. Appl. Ergon. **39**(1), 15–27 (2008)
2. Jones, M., Park, J., Ebert, S., Kim, H., Reed, M.: Statistical models to predict seat pressure in automotive seats. In: SAE Congress (2017)
3. Reed, M.P., Flannagan, C.A.: Modeling population distributions of subjective ratings (No. 2001-01-2122). In: SAE Technical Paper (2001)
4. Park, J., Kim, H., Jones, M., Ebert, S., Park, B., Reed, M.: Development of an automatic seat-dimension extraction system. In: SAE Congress (technical paper no.: 2016-01-1429), Detroit, USA (2016)

5. Park, J., Park, B., Lee, B., Lee, B., Kwak, M., Kim, Y.: Development of an automated seat dimension evaluation system. In: SAE Congress (technical paper no.: 2019-01-0401), Detroit, USA (2019)
6. Lee, B., Jung, K.: Development of logistic regression models to classify seat fit (No. 2020-01-0869). In: SAE Technical Paper (2020)
7. SAE International Surface Vehicle Recommended Practice, "Motor Vehicle Seat Dimensions," SAE Standard J2732 (2008)

How Personal Identity Influences the Driving Behavior-Correlation Analysis with Naturalistic Driving Data

Caecilia von Lienen[1,2](✉), Jana-Sophie Effert[2], Fabian Schwarzenberger[1], Lars Hannawald[1], and Guenther Prokop[2]

[1] Automotive Technology and Safety, University of Applied Sciences Dresden, Friedrich-List-Platz 1, 01069 Dresden, Germany
{Caecilia.von-Lienen, Fabian.Schwarzenberger, Lars.Hannawald}@htw-dresden.de

[2] IAD, Technische Universität Dresden, George-Baehr-Strasse 1b, 01069 Dresden, Germany
janasophie.effert@googlemail.com, Guenther.Prokop@tu-dresden.de

Abstract. In order to guarantee user acceptance and functional safety of automated driving functions, a large number of real driving maneuvers are needed to represent real possible traffic scenarios. Evidence has indicated that people had different driving characteristics based on their demographic and socio-economic characteristics as well as their sensation-seeking score or their assessment of driving costs and benefits. A detailed analysis of driver personality operated in a naturalistic driving study investigates the significant influence of driver characteristics on driving behavior using statistical analysis methods to identify different driver types.

Keywords: Human factors · Naturalistic driving · NDS · Driver profiling · Driving behavior · Driver factors · Driving style · Driver performance

1 Introduction

The motivation is based on the fundamental question how normal driving behavior can be defined. When analyzing driving maneuvers, it must not be forgotten that driving is more than a purely mechanical process to get from a point A to a point B. Driving is an interaction between human and machine, not only because the majority of the control is still taken over by a human being. Our driving experience is largely based on the influence of the human factor, which is determined by the person controlling the vehicle. We all differ in our personalities, which has an influence on the occurrence of certain behavior patterns. For example, some people tend to be more prone to show risky behavior or more fearful behavior than others, which can be assessed according to personality characteristics [1]. If these behaviors are transferred to the driving context, they are reflected in different driving styles such as risky driving or restrained driving. These different driving styles can hence be associated with certain personality characteristics that can in turn be useful when working out driving type classifications.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 16–23, 2020.
https://doi.org/10.1007/978-3-030-50943-9_3

2 Related Work

In order to gain knowledge of how people drive in real traffic, it is necessary to develop a method for collecting naturalistic driving data. A cost-effective solution is the use of a smartphone application to detect the driving performance. A mobile application has the ability to monitor the driving characteristics and provides a sufficiently good quality of the driving data. Critical events could be detected using smartphones based on fixed threshold values for the acceleration and yaw rate sensors and are almost the same quality as the results of permanently installed inertial measuring systems [2]. Further studies were able to analyze driving behavior in its various forms using smartphones and their sensors. The characteristics of the different driver profiles could be categorized so that the characteristics such as safe, risky, aggressive or non-aggressive driving styles were made describable [3, 4].

So far, the influence of personality on driving behavior has mainly been determined by self-assessments or simulation studies, whereby in many cases an influence of personality traits on driving behavior has been demonstrated.

A widely used model to describe personality is the Five Factor Model according to Costa and McCrae [5]. In this model, personality is described along five dimensions namely: neuroticism, extraversion, openness, agreeableness and conscientiousness. A high degree of neuroticism characterizes people who are rather insecure and react sadly. People with a low degree of extraversion describe themselves as reserved, independent and optimistic. Openness stands for pronounced cultural or public interests. Agreeableness stands for understanding for others, need for harmony and compassion. The fifth dimension conscientiousness is characterized by ambition, order and discipline. Lajunen [6] demonstrated a positive correlation between extraversion and the frequency of car accidents. There are also results that indicate a positive relationship between neuroticism and aggressive driving behavior [7]. Conscientiousness, on the other hand, was associated with less aggressive driving behavior [1]. Based on these results, we expect that higher levels of extraversion and neuroticism will lead to a riskier, more aggressive driving style, while conscientiousness will lead to a more cautious driving style. A significant correlation between agreeableness and openness to experience and driving behavior is not expected.

Sensation Seeking is another factor in personality research, which has often been studied in connection with driving behavior. It describes the need for changing, new, complex and intensive impressions and experiences. Results show a higher reported risky driving style and higher acceleration values in simulator studies, for high degrees of Sensation Seeking [8]. Based on these results, we expect that higher levels of Sensation Seeking will lead to higher acceleration values, generally higher speeds and a more offensive driving style.

Locus of control describes the extent to which individuals causally attribute the events they experience to themselves or to other factors (situational). External locus of control stands for the attribution of responsibility to external sources. Internal locus of control describes the tendency to ascribe behavior to internal, stable characteristics. A high degree of internal locus of control is associated with increased belt use and cautious driving [9, 10]. For this reason, we also expect that higher levels of internal control conviction will lead to a more restrained driving style in our real driving study.

In addition to these personality traits, Taubman-Ben-Ari [11] examined the relationship between driving style and individually perceived costs and benefits of driving. Driving styles were divided into reckless, angry, anxious and careful. Both reckless and angry driving styles showed a positive correlation to all components of perceived benefit. In contrast, participants who reported anxious styles estimate all costs of driving to be higher. Based on these results, we expect that a higher estimation of the individual cost components is associated with a cautious driving style, while a high estimation of the individual benefit components is associated with a risky and offensive driving style.

3 Methods

3.1 Driver Performance

To categorize different driver types, hypotheses were first drawn up that show a distinction in driving behavior. In Fig. 1, five driver categories have been set up, which differ in their driving dynamics and performance. The common driver is defined as the average of all drivers and represents the basis from which all other driving types differ. According to the hypothesis, the offensive and the risky types are described by a sportier and more aggressive driving style that approaches or exceeds the limits of the driving dynamics circle. The defensive and the unsafe driver types are characterized by lower values in their driving dynamic parameters. The acceleration potential of these categories shows a large deviation in relation to the driving dynamics limit. Both extremes (unsteady, risky) are classified as critical driving styles.

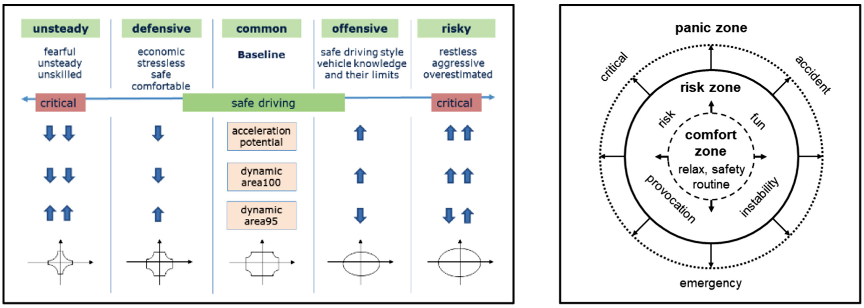


Fig. 1. Driver type hypothesis *(left)*, Three Zone Model based on [12, 13] *(right)*

In experiential education and psychology, the three-zone model describes the different areas of sensation in which a person could be. There are three distinctions: the comfort, risk and panic zone. The comfort zone describes the area of safety and routine action, while in the area of the panic zone personal skills are clearly exceeded. The risk zone defines a person's willingness to take risks, but also unsafe behavior. The zone model (Fig. 1, *(right)*) shows the personal traits that characterize people in their actions.

The ring-shaped structure of the zone model appears in a natural way when analyzing driver profiles. The representation of driving data is most commonly done via the driving dynamic circle (g-g diagram), which is a plot of the longitudinal and lateral acceleration potential. Here each person has its individual comfort zone (small radius, moderate acceleration), risk zone (larger radius, higher acceleration) and panic zone (outliers with large radius), see Fig. 2. The comfort zone is by definition the area of this circle where the driver spends most of its time, not reaching its individual driving limits. We defined this by calculating angle-wise the 95-quantile. Situations where the driver operates close to its own acceleration limits (above the 95-quantile) are in the risk zone. If these acceleration limits are exceeded, we speak of the so-called panic zone. This panic zone is also individually defined using the outlier test of Grubbs. We denote outer boundary of the risk zone (which does not contain the outliers) by the 100-quantile curve. The acceleration potential (lateral, longitudinal, curve acceleration) is an important factor in representing the driving performance. It corresponds to the potential of the respective driver in his personal limits and is usually below the physical limit.

3.2 Naturalistic Driving Application

The smartphone has high-quality sensors that make it possible to record driving behavior data. Using a specially developed app for the collection of naturalistic driving data, the hardware sensors (3-axis acceleration sensor, yaw rate sensor) and the position data (GPS) are recorded and stored automatically on a security server. The app is driver-related, so that profile settings such as age, gender and driving experience are stored and each driver's trips are assigned. This enables collectivized driving data with driver-related personal characteristics in a database.

3.3 Human Factors

Participants. 27 drivers (18 male and 9 female, Age 24–62 (M = 36.4, SD = 9.9)) participated in this study. The mean annual mileage was 18,889 km (SD = 10,689 km) and the mean driving experience was 17.4 years (SD = 8.3). The 27 evaluated participants are only a subset of the overall naturalistic driving study with 60 people.

Measures. *Demographic Questionnaire:* This questionnaire includes the questions on age, gender, driving experience and driving performance in km per year.

NEO-FFI [5]: The NEO-FFI is a validated scale measure for the assessment of the personality traits of the five factors personality model, which is considered the international standard model of personality research.

AISS-d (German Version of Arnett Inventory of Sensation Seeking [14]): The AISS-d is a validated scale measure for the assessment of Sensation Seeking. It consists of 20 items leading to the two subscales novelty and intensity.

IE-4 [9]: The IE-4 is a validated short scale measure for the assessment of locus of control. Two items each capture the internal and external locus of control.

The Driving Costs and Benefits Questionnaire [11]: This questionnaire was constructed to exploit subjective perceptions of the costs and benefits of driving.

Procedure. In October 2018 the described questionnaires were sent to the participants, who completed them and sent them back. The participants were also provided with an information sheet on how to install and use the app as well as with a free smartphone holder to ensure correct use of the app while driving. In the period from March to December 2019 naturalistic driving data were collected. The total number of kilometers driven is 14,309 km with a total time of 334 h. Collected kilometers ranged from 98.54–1600.79 km (M = 529.9, SD = 4105.6) per driver.

4 Results and Discussion

Descriptive methods impressively show that size of this circle (Area95 – area of comfort zone, Area100 – total area, exclusive outliers) pretty well differentiates between the individuals, see Fig. 2. This confirms the above mentioned hypothesis (Fig. 1). Thus, we used the feature Area100 as the main feature (target variable) to characterize the individual driving behavior. Other features like Area95, lateral and longitudinal potential showed extremely high correlation to Area100, such that an individual analysis of these features is omitted.

In order to categorize five groups of driver types, we performed k-Means cluster analysis (Fig. 2 (*left*)), which gives an overview of which persons are grouped and have a similar driving performance. Cluster 3 represents the group of a common driver. Cluster 4 and 5 are more aggressive and risky drivers, while the other persons in cluster 1 and 2 have a defensive driving style.

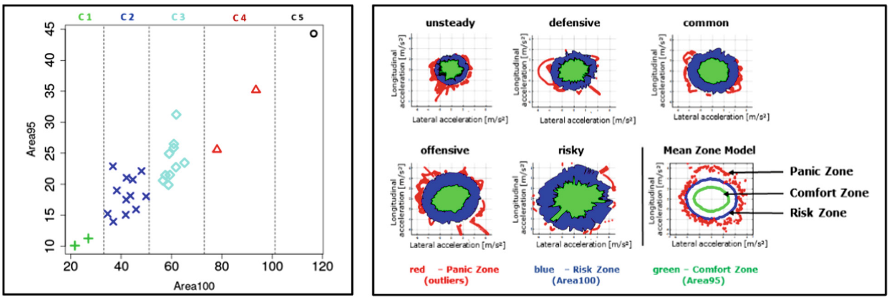


Fig. 2. Cluster analysis: k-Means clustering (*left*), Three zone modelling of different driver types *(one driver example per type)* (*right*)

The average of the zone boundaries for the comfort, risk and panic zone in the three zone modelling based on the g-g-diagram of each participant will be set as the baseline (Mean-Zone-Model) to describe the deviations of each driver type (Fig. 2, *(right)*). The common driver should be near the baseline. While the defensive type is round 10% and the unsteady driver is round 30% below the baseline. In contrast, the risky driver type has up to 50% larger zones. Starting from the unsteady to the risky driver, the surface areas increase significantly and larger acceleration values are achieved. Only the shapes in the 2D space does not confirm the hypothesis, especially the expected characteristics for an unsteady and defensive driver. There is no significant change in the resulting accelerations compared to the lateral and longitudinal acceleration potential.

Two of our hypotheses were partially confirmed. Sensation seeking correlated significantly with Area100 and shows a higher acceleration potential with a higher sensation seeking level ($r = 0.51$, $p = 0.007$). Similar results are available for the estimated benefits of driving. Again, there is a positive correlation ($r = 0.39$, $p = 0.04$) with the Area100. A further analysis with more participants could confirm this result. Unfortunately, none of the other hypotheses can be supported by any significant statistics. Future investigations will show, whether this really means that there are no dependencies between the variables or whether the sample of the current study was too small.

In order to further investigate the dependence of the driving type (i.e. Area100) on the psychological data, we performed multivariate regression analysis. While the full model containing Area100 as dependent and all the other variables of Fig. 3 (besides Pzone) as independent variables would be overdetermined, we decided to utilize model selection via Mellows Cp. Here we ended up with the covariates: age, external locus of control, neuroticism, openness, agreeableness, conscientiousness and driving experience. All of them have individual p-values below 0.05 and the F-statistic is significant with $p = 0.001$.

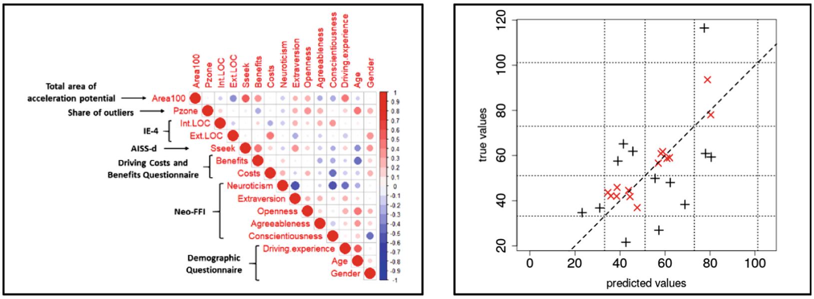


Fig. 3. Correlation between personal identity and driver performance *(left)* based on Sect. 3.2 – Measures, prediction via linear modeling *(right)*

When predicting the target variable with these features using the leave one out cross validation we obtain the values presented in Fig. 3 *(right)*. Though, it is clearly optimistic to train a model with only 26 samples, we already obtain quite satisfying prediction errors. Combining the predicted values with the classes we defined above, we predict the right class for 14 (out of 27) samples.

Several limitations restrict the statements of the results. With 27 persons a small number of participants are considered who previously provided a sufficient amount of driving data. The driving dynamic dataset is a combination of the vehicle and the driver. The influence of vehicle data is not considered separately. With the analysis in the 2D space, the results contain all speed levels of driving. The parameter space is limited and must be expanded by further factors (jerk, speed-up and braking). It has already been discussed that differentiation of the individuals by the size of circle works extremely well. However, due to our relatively small sample size it is not possible to reasonably differentiate individuals according to the shape of the circle. The proportion between longitudinal and lateral acceleration did not define a reasonable feature for our data. The analysis of the shape was hence omit-ted for this study.

5 Conclusion

This paper evaluated naturalistic driving data to define typical driver types and the influence of the personal identity of driving behavior. We found significant correlations between sensation seeking as well as the benefits of driving and a higher acceleration potential. In general, the results of this showed that increasing zone areas of the g-g-diagram predict offensive and risky driving behaviors. Although it is a small data set, we already have a number of significant results. The predictive power will continue to increase with the growing number of participants and other extracted features. Further investigations are urgently needed and very promising.

References

1. Paunonen, S.V.: Big five factors of personality and replicated predictions of behavior. J. Pers. Soc. Psychol., 411 (2003)
2. Paefgen, J., Kehr, F., Zhai, Y., Michahelles, F.: Driving behavior analysis with smartphones: insights from a controlled filed study. In: 11th International Conference on Mobile and Ubiquitous Multimedia, Ulm, Germany (2012)
3. Johnson, D., Trivedi, M.: Driving style recognition using a smartphone as a sensor platform. In: 14th International IEEE Conference on Intelligent Transportation Systems, Washington DC, USA (2011)
4. Eeren, H., Makinist, S., Akin, E., Yilmaz, A.: Estimating driving behavior by smartphone. In: Intelligent Vehicles Symposium, Alcala de Henares, Spain (2012)
5. Costa Jr., P.T., McCrae, R.R.: The five-factor model of personality and its relevance to personality disorders. J. Pers. Disord. **6**(4), 343–359 (1992)
6. Lajunen, T.: Personality and accident liability: are extraversion, neuroticism and psychoticism related to traffic and occupational fatalities? Pers. Individ. Differ. **31**(8), 1365–1373 (2001)
7. Kirkcaldy, B., Furnham, A.: Positive affectivity, psychological well-being, accident-and traffic-deaths, and suicide. Studia Psychol. **42**, 1–2 (2000)
8. McMillen, D.L., Smith, S.M., Wells-Parker, E.: The effects of alcohol, expectancy, and sensation seeking on driving risk taking. Addict. Behav. **14**(4), 477–483 (1989)
9. Kovalea, A., Beierlein, C., Kemper, C.J., Rammstedt, B.: Internale-Externale-Kontrollü berzeugung-4 (IE-4) (2014). https://doi.org/10.6102/zis184
10. Montag, I., Comrey, A.L.: Internality and externality as correlates of involvement in fatal driving accidents. J. Appl. Psychol. **72**(3), 339 (1987)

11. Taubman-Ben-Ari, O.: Motivational sources of driving and their associations with reckless driving cognitions and behavior. Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology **58**(1), 51–64 (2008)
12. Michl, W.: Erlebnispaedagogik, p. 40. Ernst Rheinhardt Verlag, Munich (2009)
13. Luckner, J., Nadler, R.: Processing the Experience: Strategies to Enhanced and Generalize Learning, p. 29. Kendall/Hunt Publishing Co., Dubuque (1997)
14. Roth, M., Mayerhofer, D.: Deutsche Version des Arnett Inventory of Sensation Seeking (AISS-d) (2014). https://doi.org/10.6102/zis73

Assessment of Driver Distraction Caused by Social Networking Activities Using the Smartphone: A Driving Simulator Study

Fabrizio D'Amico(✉), Alessandro Calvi, Chiara Ferrante, and Luca Bianchini Ciampoli

Department of Engineering, Roma Tre University, Via Vito Volterra 62, 00146 Rome, Italy
{fabrizio.damico, alessandro.calvi, chiara.ferrante, luca.bianchiniciampoli}@uniroma3.it

Abstract. The use of the mobile phone is one of the main contributors to driver distraction. Extensive scientific literature has addressed the issue of driver distraction, highlighting the negative effects of mobile phone use (i.e., calling or texting) while driving. Currently, only a few studies have investigated the effects of social networking activities while driving. The aim of this study is to assess the effect on driver performances and road safety of mobile phone use for social networking activities using a driving simulator experiment. The data collected are analyzed and compared between the scenarios in terms of driving performances and surrogate safety measures. In addition, further information are collected using pre-test and post-test questionnaires. Preliminary results demonstrates the negative effects of using the smartphone for social networking activities on driving performance and road safety, exposing drivers to high levels of risk.

Keywords: Human factors · Driver's behavior · Social network · Smartphone · Driving simulator

1 Background

The smartphone is one of the main sources of distraction for driving users. According to WHO report [1] more than one million of deaths per year are caused by driving distraction. In Italy, the national crash report [2] indicates that 40.8% of the accidents occurred in the extra-urban context has been caused by incorrect driving behavior, incorrect use of the rules, and driving distraction.

Although the use of the smartphone while driving is well known to increase the chances of accidents, in many countries it remains a growing issue that have forced the legislators to ban its use by the penalty of withdrawal of the driving license. Nowadays, most of the scientific studies on the effects of the use of the smartphone while driving focus on two device functions: talking and texting [3–9]. In fact, according to the European Commission, the main forms of distraction involving the use of the smartphone are the tactile and the visual distraction. Indeed, the presence of such a Secondary Task (ST) while driving shifts the driver's attention from the driving task, thereby compromising the road safety and increasing the number of accidents.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 24–31, 2020.
https://doi.org/10.1007/978-3-030-50943-9_4

Although the effect of the use of the smartphone for reading and texting while driving is quite an explored topic, much more effort has still to be oriented onto the distraction due to social networks.

In fact, only one published research [10] has focused on the use of social applications while driving (e.g. Facebook, Instagram, SnapChat). Among the results, the authors claim to have observed values of Brake Reaction Time and Time Headway significantly greater for the social-condition compared to the baseline. Furthermore, a topic-related questionnaire realized by Basch et al. in 2018 [11] for a sample of university students reported that about 43% of the 324 interviewed students admitted scrolling the home page of the social networks, reading posts and articles and even posting thoughts and photos, while driving.

A lack of knowledge in quantifying the effect of social activities on driving, highlights the phenomenon to require an in-depth investigation. Therefore, it becomes crucial to investigate such effects by means of a tool that allows to measure the driving performance decay in terms of kinematic and dynamic parameters. In these regards, the present paper aims to investigate how much the social activities compromise the driving performances.

2 Methodology

2.1 Equipment

The present study was conducted by using the fixed-base driving simulator located in the simulation laboratory of the Department of Engineering at Roma Tre University. The driving simulator has been previously validated [12] and its reliability for investigating driver's behavior is demonstrated, especially in terms of driving performances [13–15].

The driving simulator consists in a real vehicle equipped with standard hardware interfaces (pedals, wheels, etc.). The virtual scenarios are projected on a curved screen that covers a 180° field of view. During the simulation, the system records various dynamic parameters that allow for describing the driver's behavior (Fig. 1).

Fig. 1. Driving simulation using mobile phone.

2.2 Sample of Drivers

A sample of drivers was recruited and subjected to the driving tests. The selection criteria were: an age ranging from 20 and 35 years, driving license, smartphone and social networks familiarity. This age range has been chosen as it was reported, also by the questionnaire on students, that young drivers are the most involved in social activities. As a result of this selection, fifty-one drivers have been recruited, 57% males and 43% females with an average age of 26 years ranging from 21 and 32 years. During the experiment, five participants have prematurely finished the test due to simulation sickness. The other part of the sample has been statistically validated by two statistical criteria, the Chauvenet criterion and the Box-Plot technique, in order to identify the outliers. The statistical validation has been applied on the speed value at specific points of the scenario, thereby avoiding the affection of external conditions. At the end of the statistical procedure, a final sample of forty-four drivers has been analyzed.

2.3 Driving Scenario and Events

Three simulated scenarios are designed by combining random sequences of different road environment tracks, namely, urban, extra-urban and industrial, in order to avoid the memorization process of the road scenario. A simulated two-lane road with each lane of 3.50 m was reproduced according to the Italian rules and regulation. By means of the reproduction in the virtual environment of several 3D objects, a series of boundary elements such as markings and vertical sign, vegetation, buildings and other vehicles, have been included within the scenario in order to improve the level of realism of the simulation.

Several driving events have been simulated in order to test the driver's attention to possible hazards in the road environment. It is worthwhile to underline that these events have been selected for analyzing the drivers' behavior in highly risky conditions. Also, these events were located in the scenarios at different points to avoid the memorization of their sequence. However, they were designed with the same features, thereby assuring a reproducible situation for all the drivers. More in detail, three events are identified: a car-following (CF) maneuver, a legal pedestrian crossing (LPC) on the proper pedestrian crossing sign, an illegal pedestrian crossing (IPC) out of the proper sign.

The CF event is implemented in the scenario in order to analyze the driver's behavior when reaching a slower vehicle ahead (i.e. the leading vehicle) having a speed of 50 km/h and being positioned on the driver lane. The leading vehicle drives in this position for a distance of 1000 m. Under such CF condition, the leading vehicle suddenly brakes with a deceleration value of 5 m/s^2 and it stops in the center of lane, thereby requiring the driver to brake suddenly.

The LPC event is designed according to [16, 17]. The pedestrian starts to cross at the right edge of the road on the proper sign with a speed of 1.4 m/s when the driver is at 50 m from the crossing and, subsequently, has a TTZ (Time to Zebra) of 4 s (considering a driver approaching speed of 45 km/h).

The IPC event is designed based on the same studies with the same initial characteristics.

2.4 Secondary Task

The design of the ST is defined based on the wide range of scientific contribution concerning the driving distractions due to smartphone. Two social networks referred to two levels of task complexity are set: low task complexity involving the activities on Facebook (FB) and high task complexity including the activities on Instagram (IG).

In more details, two ad hoc social network profiles were created on the two social platforms. The required activities on FB were scrolling on the home page, liking and sharing posts. Instead, the activities on IG required a higher task complexity as they consisted in taking a selfie while driving and sharing it as an IG story.

2.5 Procedure

The driving simulator experiment was subjected to a strict protocol. As mentioned before, each participant was requested to drive three simulated scenarios and, parallelly, to perform the required ST described in the previous paragraph.

Two test session days with an interval of one week were defined: the first with one scenarios and the second with the remaining two. The order of the scenarios among the participants was randomized. Before the session each driver has to drive a training scenario for about 10 min to familiarize with the tool.

Moreover, a before-test questionnaire and an after-test questionnaire were submitted to the drivers by means of Google Form application, to collect general information about the drivers and their impressions and feelings on the test.

2.6 Data Collection

Several driving performances (speeds, distances, etc.) and surrogate safety measures such as reaction time and TTZ were collected, analyzed and compared between the two conditions (distracted driving and baseline) in order to investigate the influence of social networks on driving performances. Furthermore, the number of accidents was recorded for each event and each condition.

The variables to describe the driving performance in each event are listed below.

Car-following (CF):

- RT the time interval between the first braking of the leading vehicle's and the time when the driver starts braking;
- D_{min} is the minimum distance between vehicles at the end of the deceleration of the driver;
- D_{av} is the average distance between vehicles during CF;
- LP_{av} is the average lateral position during CF.

Legal Pedestrian Crossing (LPC):

- S_i is the initial speed where the driver started braking due to the pedestrian crossing;
- D_i is the initial distance between driver and pedestrian crossing when S_i is recorded;
- d_{max} is the maximum deceleration recorded during breaking maneuver;

- TTZ is the Time-to-Zebra, namely, the time interval from the collision between vehicle and pedestrian under the hypothesis of unvaried trajectory and speed [18]. It was recorded when the driver pressed the brake pedal.

Illegal Pedestrian Crossing (IPC):

- D_i is the initial distance (see LPC);
- D_{min} is the minimum distance between driver and pedestrian crossing when minimum speed during breaking maneuver is recorded;
- d_{max} is the maximum deceleration (see LPC);
- TTZ is the Time-to-Zebra (see LPC).

3 Results

3.1 Crash Occurrence

A preliminary descriptive analysis reported that a higher number of accidents in case of distracted driving is recorded respect to the baseline condition. In more details, in the CF event with braking of the leading vehicle, the drivers without any distractions never performed an accident, while in 16% and 9% of the cases accidents were observed for FB and IG scenarios, respectively. Similarly, in 7% of the LPC tests accidents have occurred in both FB and IG scenarios. The percentage of accidents was found to increase for IPC: 5% in baseline condition and 55% and 34% in FB and IG scenarios, respectively. Among these, some drivers have crashed without any reaction prior to the collision, with consequent high impact speed (up to 58 km/h). In particular, 50% in FB scenario and 67% in IG scenario, according to the higher workload.

3.2 Driving Performance

The variables referred to each maneuver (CF, LPC and IPC) were validated by means of statistical analyses. Table 1 summarizes the results of the statistical analyses conducted through Median, Kruskal-Wallis, Brown–Forsythe and ANOVA tests. The average values of each variable selected, along with the standard deviations (in parenthesis) and the results of the ANOVA or other analysis are given in the table.

Car-Following. RTs were observed to be greater in FB and IG conditions (up to 3.92 s) than in baseline conditions (1.32 s). D_{min} between vehicles was found to be lower than that recorded in the baseline condition (15.92 m) for FB scenario (13.21 m). Instead, IG scenario returned higher values than the baseline conditions, equal to 22.55 m.

Furthermore, D_{av} during CF was observed to be affected by the drivers' distractions; indeed, higher distances were recorded when the drivers were involved in social network activities (76.05 m in IG scenario and 45.74 m in FB scenario). LP_{av} values were significantly shifted on the right with respect to the ideal trajectory, that is fixed at 1.75 m for a 3.50 m two-lane road.

Furthermore, D_{av} during CF was observed to be affected by the drivers' distractions; indeed, higher distances were recorded when the drivers were involved in social network activities (76.05 m in IG scenario and 45.74 m in FB scenario). LP_{av} values were significantly shifted on the right with respect to the ideal trajectory, that is fixed at 1.75 m for a 3.50 m two-lane road.

Legal Pedestrian Crossing. Values of S_i, recorded where the driver started to decelerate, were observed to be lower and higher than the baseline condition, for IG and FB scenario, respectively. Values are 38.04 km/h for IG scenario and 49.16 km/h for FB scenario, compared to 46.51 km/h for the baseline condition.

D_i returned a greater value for IG scenario than both FB scenario and baseline condition. d_{max} showed the highest value of deceleration in FB scenario, up to −7.62 m/s^2. Finally, different results were observed for TTZ, with the highest value in IG scenario (1.19 s) and lowest value in FB scenario (0.73 s); in baseline condition it was 0.87 s.

Illegal Pedestrian Crossing. Also in this case, different values in the three conditions were recorded. In particular, D_i between vehicles at the moment the driver started the deceleration was greater in the baseline condition (26.30 m) than in the distracted scenarios (18.92 m in FB scenario and 18.56 m in IG scenario).

Similarly, D_{min} values were observed to be higher in the baseline condition than in the distracted conditions, while the results of d_{max} showed a greater value in FB scenario (9.31 m/s^2).

On the contrary, TTZ was found to be lower in the distracted drives with respect to the baseline condition.

Table 1. Average of the investigated parameters and their standard deviation (in parenthesis) and statistical analysis results for each event.

Phase	Parameter	Baseline condition	Facebook	Instagram	Statistical Results	
					F	P
CF	RT [s]	1.32 (0.29)	2.67 (2.96)	3.92 (2.72)	Median	<0.001
	D_{min} [m]	15.95 (7.99)	13.21 (8.60)	22.55 (12.97)	Median	<0.001
	D_{av} [m]	36.24 (15.98)	45.74 (29.26)	76.05 (44.39)	Median	<0.001
	LP_{av} [m]	1.76 (0.29)	1.90 (0.27)	2.06 (0.28)	12.78	<0.001
LPC	S_i [km/h]	46.51 (10.58)	49.16 (7.64)	38.04 (9.28)	15.67	<0.001
	D_i [m]	38.92 (15.03)	34.83 (6.05)	43.50 (17.08)	Median	0.005
	d_{max} [m/s^2]	−6.47 (2.36)	−7.62 (2.13)	−4.78 (2.55)	K. Wallis	<0.001
	TTZ [m]	0.87 (0.36)	0.73 (0.20)	1.19 (0.44)	Mediana	<0.001
IPC	D_i [m]	26.30 (7.49)	18.92 (3.77)	18.56 (5.66)	K. Wallis	<0.001
	D_{max} [m]	12.37 (3.56)	9.36 (4.74)	10.94 (4.98)	3.25	0.044
	d_{max} [m/s^2]	−7.81 (2.36)	−9.31 (1.42)	−7.46 (3.12)	Median	0.015
	TTZ [m]	0.57 (0.38)	0.42 (0.13)	0.53 (0.21)	B. Forshyte	0.002

4 Conclusions

In this study, the use of smartphone for social-network activities while driving was firstly studied through a general descriptive analysis and then by focusing on three events: CF, LPC and IPC. The results of the descriptive analysis have showed that the distracted driving caused a high number of accidents. Results about CF showed a significant impact of the use of the smartphone on driving in terms of longer reaction times. Furthermore, rightmost values were recorded for the lateral position. In terms of LPC and IPC, S_i, d_{max} and TTZ have shown the most significant decay of driving performances in case of distracted driving. The findings of this application are promising as they provide an insight of how the phenomenon of distracted driving due to social-network activities is serious and worrying for the drivers and road safety.

References

1. WHO: Mobile phone use: a growing problem of driver distraction, report (2011)
2. ACI-ISTAT: Incidenti stradali (2017). (in Italian)
3. Stayrinos, D., Jones, J.L., Garner, A.A., Griffin, R., Franklin, C.A., Ball, D., Welburn, S.C., Ball, K.K., Sisiopiku, V.P., Fine, P.R.: Impact of distracted driving on safety and traffic flow. Accid. Anal. Prev. **61**, 63–70 (2013)
4. Strayer, D.L., Drews, F.A., Crouch, D.J.: A comparison of the cell phone driver and the drunk driver. Hum. Fact. **48**(2), 381–391 (2006)
5. Tsimhoni, O., Smith, D., Green, P., Arbor, A.: Address entry while driving: speech recognition versus a touch-screen keyboard. Hum. Factors **46**(4), 600–610 (2004)
6. Yager, C.E., Cooper, J.M., Chrysler, S.T.: The effects of reading and writing text-based messages while driving. Hum. Fact. Ergon. Soc., 2196–2200 (2012)
7. Yannis, G., Laiou, A., Papantoniou, P., Christoforou, C.: Impact of texting on young drivers' behavior and safety on urban and rural roads through a simulation experiment. J. Saf. Res. **49**, 25–31 (2014)
8. Yannis, G., Laiou, A., Papantoniou, P., Gkartzonikas, C.: Simulation of texting impact on young drivers' behavior and safety on motorways. Transp. Res. Part F **41**, 10–18 (2016)
9. Young, K.L., Rudin-Brown, C.M., Patten, C., Ceci, R., Lennè, M.G.: Effect of phone type on driving and eye glance behavior while text messaging. Saf. Sci. **68**, 47–54 (2014)
10. McNabb, J., Gray, R.: Staying connected on the road: a comparison of different types of smart phone use in a driving simulator. PLoS One **11**(2) (2016)
11. Basch, C.H., MacLean, S.A., Romero, R.A., Ethan, D.: Health information seeking behavior among college students. J. Commun. Health **43**, 1094 (2018)
12. Bella, F.: Driving simulator for speed research on two-lane rural roads. AA&P **40**(3), 1078–1087(2008)
13. Calvi, A., Bella, F., D'Amico, F.: Diverging driver performance along deceleration lanes: driving simulator study. Transp. Res. Rec. J. Transp. Res. Board **2518**, 95–103 (2015)
14. Calvi, A., Bella, F., D'Amico, F.: Evaluating the effects of the number of exit lanes on the diverging driver performance. J. Transp. Saf. Secur. **10**(1–2), 105–123 (2018)
15. Calvi, A., Benedetto, A., D'Amico, F.: Investigating driver reaction time and speed during mobile phone conversations with a lead vehicle in front: a driving simulator comprehensive study. J. Transp. Saf. Secur. **10**(1–2), 5–24 (2018)

16. Bella, F., Silvestri, M.: Effects of safety measures on driver's speed behavior at pedestrian crossings. Accid. Anal. Prev. **83**, 111–124 (2015)
17. Bella, F.: Effects on driver's behavior of illegal pedestrian crossing. In: Advances in Human Aspects of Transportation, vol. 786, pp. 802—812 (2019)
18. Várhelyi, A.: Drivers speed behaviour at a zebra crossing: a case study. Accid. Anal. Prev. **30**(6), 731–743 (1998)

Personalized Driver State Profiles: A Naturalistic Data-Driven Study

Arash Tavakoli(✉), Mehdi Boukhechba, and Arsalan Heydarian

Department of Engineering Systems and Environment, Charlottesville, VA, USA
{at5cf,mob3f,ah6rx}@virginia.edu

Abstract. Previous studies suggest that variation in driver's states, such as being under stress, can degrade drivers' performance. Moreover, different drivers may have varying behaviors and reactions in different road conditions and environments (contexts). Thus, personalized driver models given different contextual settings can assist in better predicting the drivers' states (behavioral and psychological); this can then allow vehicles to adjust the driving experience around the driver and passengers' preferences and comfort levels. This paper aims at developing personalized hierarchical driver's state models by considering driver's heart rate variability (HRV) in relation to the changes in various contextual settings of road, weather, and presence of a passenger. Results from 12 participants over 150 h of driving data suggest that drivers are on average less stressed in highways compared to cities, when being with a passenger compared to alone, and when driving in non-rainy conditions compared to rainy weather.

Keywords: Naturalistic driving studies · Driver's heart rate · Environmental factors

1 Introduction

Studies suggest that emotional changes can affect driving behaviors. For instance, negative emotions may increase the risk perception among drivers compared to positive emotions [1]. Similarly, studies conducted through driving simulators suggest strong emotions, such as anger, can degrade the driving performance [2]. Further, environmental conditions have also been indicated to impact a driver's emotional states and impact his/her driving behaviors. For instance, high traffic density might cause frustration. Frustration can potentially result in a higher number of hard brakes in driving.

Although there are different studies that evaluate the impact of environmental conditions on driver's emotions and driving behaviors, many of these studies are conducted in driving-simulators, where the driver is monitored for a short term and/or under controlled settings such as [3]. While driving simulators allow us to better investigate the relationships among specific variables of interest, naturalistic studies provide a more comprehensive insight into the interplay of multiple environmental factors that may happen simultaneously in a realistic environment. Naturalistic environments are accompanied by real-world noise that is oftentimes eliminated from an experimental simulation study. However, to have a proper understanding of the

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 32–39, 2020.
https://doi.org/10.1007/978-3-030-50943-9_5

influence of different changes in the environment, naturalistic studies should be conducted longitudinally and collect longer duration of data. Collecting longitudinal data gains more importance when considering the dynamic nature of human psychophysiological states, how it varies momentarily under environmental conditions, and how it can be affected by many factors simultaneously.

This study aims to develop personalized hierarchical driver's state models by analyzing the drivers' physiological activation under variations of the driving context. In this paper, we specifically model drivers' heart rate variability (HRV) and how it fluctuates based on context (e.g. weather and road condition), by incorporating the naturalistic data from 12 participants. We analyze the HRV under the specific contextual settings of various (1) weather conditions, (2) road type, and (3) the presence of a passenger. This study answers the following questions:

- How does the heart rate variability of drivers vary under varying conditions of the road, weather, and presence of passengers?

2 Related Work

With the current improvements in ubiquitous computing and wearable devices, it is now viable to assess driver's physiological measures in a less intrusive fashion, and for a long-term. Driver's physiological measures have been used to classify and predict the state of the driver in the past [4]. The advantage of using physiological measures over other modalities of data (e.g. video) is that it gives valuable information without interfering with user's privacy.

In this study, we focus on drivers' heart rate as the physiological measure. Considering the previous studies, the general procedure for correlating driver's heart rate to his/her states is as follows. First heart rate is retrieved from the photoplethysmogram (PPG) sensor. PPG sensor is a technology that can evaluate the heart rate from examining the blood volume in the veins in a non-intrusive fashion. Then the heart rate signal will go through a de-noising process which smooths the heart rate signals to help better recognize each heartbeat. In this process filters such as adaptive filters [5] are applied to heart rate signals. These filters will help identify real peaks in the heart rate signal that shows the pulses. Then, by using the automatic multiscale-based peak detection (AMPD) algorithm [6] the beat to beat intervals are retrieved. Finally, multiple features of heart rate variability (HRV) will be computed using the computed beat to beat intervals. HRV is referred to as the statistical difference between beat-to-beat of heart rate. HRV features can be as simple as the mean value of beat to beat intervals. Previous studies have correlated the square root of the mean of the sum of the squares of differences between adjacent NN intervals (RMSSD) to the human states such as stress level and emotions (e.g. happiness, sadness, etc.). Studies suggest that an increase in stress level decreases RMSSD [7], an increase in amusement increases RMSSD [8], and an increase in anger decreases RMSSD [8]. Thus RMSSD can potentially be used as a measure for understanding the current situation of the driver and how he/she feels.

Additionally, studies in the past have evaluated how environmental factors such as different weather conditions, road design, and presence of passengers may impact the driver's state and respective driving behaviors. Considering the weather as an environmental factor, results from a survey suggest that adverse weather condition is among the top stress-related factors for drivers [7]. Weather conditions can predominantly affect driving behavior by decreasing the speed of drivers in rainy or snowy conditions [9]. Road design can affect a driver's state as well. For instance, the scenery has shown to have a significant effect on a driver's state of mind, including emotional states and stress levels [10]. In their study, authors by applying logistic regression analysis on data collected from a male driver for 21 days, concluded that city driving is more stressful for drivers. Other studies have noted that the presence of a passenger can have a calming effect on the driver [11]. Most of the previous studies have been conducted in driving simulators. Real-world driving includes many factors happening simultaneously, which makes it more difficult to understand the variations in driver's states under different conditions. Additionally, even the long-term naturalistic studies in the past have not considered monitoring driver's physiological measures such as heart rate for a long time. To the best of our knowledge, this is the first study that collects longitudinal physiological sensors measurements in a naturalistic driving environment.

3 Data Collection and Processing

This study has implemented a naturalistic driving study platform that collects data from both the outside and in-cabin environments. The platform includes: (1) a BlackVue DR-750S-2CH dual dash camera recording both the cabin and the frontal road, (2) an Android smartwatch that uses an in-house built android app for collecting heart rate and movement data. The app is programmed to collect 1 Hz heart rate, 100 Hz accelerometer and gyroscope, 1/60 Hz GPS, and 100 Hz photoplethysmogram (PPG) data used to measure heart rate variability (HRV) [12]. The smartwatch is connected to the participant's phone, thus it synchronizes the time regularly with the phone. The data was stored on the smartwatch and was unloaded every two weeks. A summary of the collected data can be viewed on Table 1.

Table 1. Details of the collected data to date

Road type (hours)		Weather type (hours)		Presence of a passenger (hours)	
City	Hwy	Non rain	Rainy	W/O passenger	With passenger
55.4	106	135.3	26.1	77.6	83.9

In total, 18 participants have been recruited (7 males, 11 females). In this paper, we present the data collected from 12 of the participants, as the other 6 did not have enough physiological data collected as of now. The recruitment and data collection is an ongoing process. The participants were from 21–33 years old. Participants' occupation was either a student or a faculty member. The devices stayed in participants' cars for up to two months, unless the participant decided to leave the study. Inspecting the GPS data revealed that the data was mostly generated in the central, east and northeast regions of the United States.

Using road videos, each driving scenario was categorized based on weather conditions and road types. Road types were manually annotated to be one of the city streets or highway. The weather condition was also manually annotated to be rainy or not. Using the in-cabin videos, the presence of a passenger for each video was retrieved (Fig. 1). Using the time frame of each video clip, the same epochs of heart rate data were produced by searching through the bulk heart rate data from the smartwatch. Thus, for every driving scenario, one video of the face, one video of road, and one epoch of data from heart rate were produced with tags of weather, road, and presence of a passenger.

Fig. 1. Different road, weather, and presence of passenger conditions

Using the heart rate data from the smartwatch, different HRV features were calculated. Given the PPG signal is very sensitive to motion artifacts, which may lead to poor HRV estimation if false peaks are detected, the signal was first denoised using adaptive filters [5]. Then we used the automatic multiscale-based peak detection (AMPD) algorithm [6] to estimate beat-to-beat intervals. Using a 30-s sliding window, several HRV features were computed such as Root mean square of successive RR interval differences (RMSSD) and Standard deviation of the beat-to-beat (NN) intervals (SDNN). As the results for different HRV features were similar, we only present only RMSSD results in this paper.

4 Data Analysis

To model the RMSSD variation under varying conditions of road, weather, and presence of a passenger, a hierarchical modeling approach was implemented. The fixed factors in this study are the type of road, weather conditions, and the presence of a passenger. The dependent variable is the RMSSD of the heart rate of each driver. The random factor is the participant. Hierarchical modeling which is also known as mixed effect modeling was chosen for its ability to address dependent data and to differentiate between different participants' data. This is important as each person's heart rate might have a specific intercept and it has its own internal dependency. A hierarchical model is a generalization of linear models in which the coefficients of the model are based on another model themselves. In a hierarchical modeling approach, the model finds the best predictors for the dependent variable by accounting for different intercepts and slopes among the random factors [13]. When modeling data under a hierarchical approach, the model can find a different intercept, slope, or both for each participant. This means that the model can find that every participant's RMSSD can have a separate intercept and can change with a unique rate when being under the fixed factors. A hierarchical model with a random intercept only was chosen for this data.

5 Results

Using a random intercept for the participant, a mixed effect model was fit to study the effect of context on drivers' HRV. The result of the model suggests that all three fixed factors are significant predictors of driver's RMSSD. In this section, we discuss each factor separately. The summary of the model can be viewed in Table 2. Additionally, the RMSSD in each condition is depicted in Fig. 2. The road factor estimate of the model is 3.36 meaning that on average Driver's heart rate variability increases in highway driving compared to city driving (Fig. 2). This implies less stress in highway driving. This is aligned with previous research suggesting that driving in open landscapes was accompanied with less stress [14]. Also, as the city environment includes many triggers of stress such as pedestrians, stop signs, bikes, etc. [4], this can be the cause of higher stress in city streets. Participants number 9, 10, 13, and 14 are outliers and have lower RMSSD in highways, which implies that they are more stressed out in highways comparing to city streets. Although this inconsistency can be due to a lack of data, it can also be a sign of having two different behaviors among individuals depending on their situation, character, or environment. Although this finding is promising, we require more additional to asses this. It is important to note that in our study we did not consider the traffic density of the highways. Higher traffic density can lead to higher stress. As part of our future work, we will add the traffic density factor by using computer vision techniques such as semantic segmentation to assess this issue.

Table 2. Summary of the hierarchical model

Random effects			Fixed Factors					
Groups	Variance	Std.Dev.	Factor	Estimate	Std. Error	df	t-value	p-value
Participant	2.884	1.698	(Intercept)	9.4e+00	4.5e-01	1.3e+01	20.6	1.9e-11
Residual	23.001	4.796	Passenger	5.0e-01	6.1e-02	6.7e+04	8.2	2.3e-16
			Road	3.3e-01	4.4e-02	7.0e+04	7.5	3.7e-14
			Weather	-2.2e-01	6.4e-02	7.0e+04	-3.5	0.00046

The weather factor had an estimate of −2.24 meaning that on average driving in rainy conditions was accompanied by lower RMSSD and higher stress level (Fig. 2). This is aligned with previous research suggesting the adverse weather as a cause of higher stress levels [7].

The passenger factor had an estimate of 5.03 in the model meaning that on average participants were less stressed when being accompanied by a passenger. Previous research suggests that the passenger might have a calming effect on a driver's state of mind [11]. It should be noted that although on average the participant had a calming effect, this effect might vary among individuals and how is the relation between driver and passenger. For instance, participant number 12 has a drastic decrease in RMSSD while being accompanied by a passenger implying a higher stress level. As the participants poll increases we will assess the relationship and engagement of the driver and passenger.

Based on the random intercept hierarchical model, some of the participants did not follow the average trend. One interpretation can be to consider that not everyone's heart rate RMSSD will follow the same pattern under stress and stress might have different levels as well. For instance, studies have shown that the effect of sadness and disgust on the human's heart rate is depended on the type of sadness and disgust. A study has shown that there are two types of sadness (namely activated and deactivated) that affects human's heart rate differently meaning that in activated sadness heart rate decreases and in deactivated sadness heart rate increases [15]. Another study suggested that the variation in heart rate under the emotion of disgust is dependent on type of disgust meaning that in core disgust heart rate increases and in body-boundary violation it decreases [15]. Thus one explanation could be that stress level and how that affects heart rate can have different levels. This will be explored in the future work of this study.

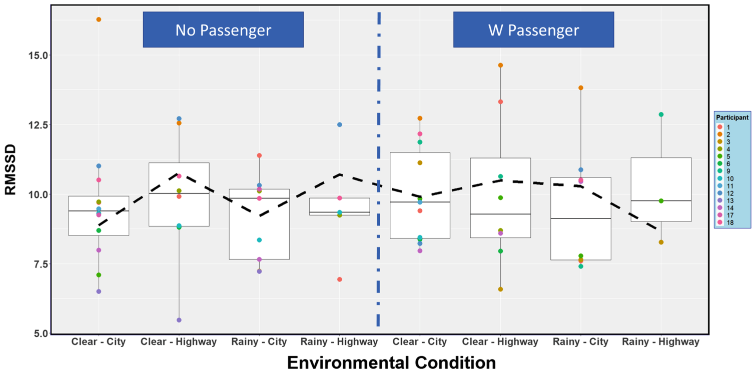


Fig. 2. Variation of drivers' RMSSD under different environmental conditions

6 Conclusion and Future Work

This study analyzes the variation of the heart rate variability of drivers under various environmental conditions. In contrast to previous studies, this paper takes a personalized approach in modeling the driver's state. On the average presence of a passenger, non-rainy weather and highway driving were accompanied with higher RMSSD.

Lower RMSSD in rainy weather, alone driving, and city streets can potentially imply more stress in these driving scenarios for those drivers. This study potentially suggests that physiological data for in vehicle application can be used as a long-term measure which is less-intrusive with lower computation burden compared to videos. In future work, we aim at increasing the participant poll, location, and time of this study to include a greater diversity among our data. As most of our data has been collected in the east coast and through summer and fall, the dataset currently lacks a seasonal effect (e.g. lack of having snowy condition in our dataset), and a geographic effect (e.g. lack of having data from west coast states). Moreover, we will include other factors such as traffic density, time of driving, and purpose of driving. In addition, we will use the computer vision applications to detect more specific events outside of the car and in an automatic fashion.

References

1. Hu, T.Y., Xie, X., Li, J.: Negative or positive? The effect of emotion and mood on risky driving. Transp. Res. Part F Traffic Psychol. Behav. **16**, 29–40 (2013)
2. Jeon, M., Walker, B.N., Yim, J.B.: Effects of specific emotions on subjective judgment, driving performance, and perceived workload. Transp. Res. Part F Traffic Psychol. Behav. **24**, 197–209 (2014)
3. Milleville-Pennel, I., Charron, C.: Driving for real or on a fixed-base simulator: is it so different? An explorative study. Presence Teleoperators Virtual Environ. **24**(1), 74–91 (2015)
4. Healey, J., Picard, R.W., et al.: Detecting stress during real-world driving tasks using physiological sensors. IEEE Trans. Intell. Transp. Syst. **6**(2), 156–166 (2005)
5. Eyal, S., Akselrod, S.: Heart rate variability (HRV): nonlinear HRV. In: Wiley Encyclopedia of Biomedical Engineering (2006)
6. Scholkmann, F., Boss, J., Wolf, M.: An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals. Algorithms **5**(4), 588–603 (2012). https://doi.org/10.3390/a5040588. http://www.mdpi.com/1999-4893/5/4/588
7. Hill, J.D., Boyle, L.N.: Driver stress as influenced by driving maneuvers and road- way conditions. Transp. Res. Part F Traffic Psychol. Behav. **10**(3), 177–186 (2007)
8. Punita, P., Saranya, K., Kumar, S.: Gender difference in heart rate variability in medical students and association with the level of stress. Natl. J. Physiol. Pharm. Pharmacol. **6**(5), 431–437 (2016)
9. Kilpeläinen, M., Summala, H.: Effects of weather and weather forecasts on driver behaviour. Transp. Res. Part F Traffic Psychol. Behav. **10**(4), 288–299 (2007)
10. Wilde, G.J.: Roadside aesthetic appeal, driver behaviour and safety. Can. J. Transp. **3**(1) (2009)
11. Shinar, D., Compton, R.: Aggressive driving: an observational study of driver, vehicle, and situational variables. Accid. Anal. Prev. **36**(3), 429–437 (2004)
12. Boukhechba, M., Barnes, L.E.: SWear: Sensing using WEARables. Generalized Human Crowdsensing on Smartwatches. In: 2019 IEEE 11th International Conference on Applied Human Factors and Ergonomics. IEEE (2020)
13. Osborne, J.W.: Advantages of hierarchical linear modeling. Pract. Assess. Res. Eval. **7**(1), 1 (2000)

14. Antonson, H., Mårdh, S., Wiklund, M., Blomqvist, G.: Effect of surrounding land-scape on driving behaviour: a driving simulator study. J. Environ. Psychol. **29**(4), 493–502 (2009)
15. Armony, J., Vuilleumier, P.: The Cambridge Handbook of Human Affective Neuroscience. Cambridge University Press, Cambridge (2013)

Human Car-Following Behavior: Parametric, Machine-Learning, and Deep-Learning Perspectives

Saeed Vasebi[1(✉)], Yeganeh M. Hayeri[1], and Jing Jin[2]

[1] School of Systems and Enterprises, Stevens Institute of Technology, 1 Castle Point, Hoboken, USA
Svasebi@stevens.edu
[2] Department of Civil and Environmental Engineering, Rutgers University, 500 Bartholomew Road, Piscataway, USA

Abstract. Emerging automated vehicles and mixed traffic flow have been substantially increased demand for modeling human driving behavior in both academia and industry. As a result, many car-following (CF) models have been proposed using parametric and data-driven approaches. Considering the large number of CF models, the critical question is which CF model or category of models (e.g. machine-learning) could accurately regenerate human CF behavior. This study conducts a cross-category comparison between one parametric model (intelligent driver model (IDM)), two new machine-learning CF models based on feedforward neural network (FNN) and recurrent neural network (RNN), and one novel deep-learning CF model (Deep-RNN) with long short-term memory (LSTM). The models are developed in TensorFlow and compared at local and global levels. At the local level, Deep-RNN significantly outperformed the others, followed by RNN and FNN. At the global level, IDM demonstrated the best performance, followed closely by Deep-RNN. The result illustrates there is no one-size fit model and the model should be selected given projects' characteristics. The result suggests a hybrid approach, which integrates parametric and deep-learning models, could precisely regenerate human car-following behavior.

Keywords: Human driving behavior · Car-Following · Deep-learning · Artificial neural network · Machine learning

1 Introduction

Automated vehicles are becoming a reality for the transportation sector. Several auto manufacturers have either launched or planning to launch their partially or highly automated vehicles (HAVs) in the next few years. Expansion of automated vehicles will create mixed traffic flow, in which human drivers will interact with machine drivers. In order to evaluate the impact of mixed traffic flow on transportation system, understanding human and HAVs driving behaviors is necessary. HAVs behavior follow certain algorithms which are known to their developers (e.g. image processing).

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 40–48, 2020.
https://doi.org/10.1007/978-3-030-50943-9_6

However, human driving behavior is more complex and hardly understood (e.g. phantom jam).

Human driving behavior is categorized into lane-changing behavior and car-following behavior. Car-following (CF) behavior has been widely studied and several models have been proposed to explain and predict CF behavior. The proposed CF models range from parametric engineering-perspective models (e.g. full velocity difference), to machine-learning models, and more recently, deep-learning models.

Regarding the large number of proposed CF models, the key question is which model or category of models result in higher performance. A robust comparison, across different modeling perspectives, could provide a valuable resource for academic and industry applications, like mixed traffic simulation and advanced driver-assistance systems. This study develops four CF models and conducts a fair comparison between them. The goal is not to suggest to eliminate one model but to evaluate which model or combination of models could result in higher accuracy given a particular situation.

2 Method

This section explains developing IDM and the artificial neural network CF models (i.e. FNN, RNN, and deep-RNN) and also the training and evaluation procedure.

2.1 Modeling

IDM is a popular CF model which explains drivers' CF behavior based on a desired velocity and a desired headway. IDM assumes drivers maintain a desired velocity in free-flow condition, and they switch to desired headway in synchronized or congested condition, to keep safe distance from the leading vehicle [1]. IDM's variables are gap (s), velocity (v), acceleration ($\dot{v}$), and velocity difference (Δv). IDM's calibrated parameters are maximum acceleration rate (a), comfortable deceleration rate (b), desire velocity (v_0), minimum gap (s_0), and desired safety time headway (T). This study uses Genetic algorithm (GA) [2] to calibrate IDM's parameters.

GA is an optimization technique for solving non-linear problems. GA generates an initial random population of N gens which are a set of calibration parameters for IDM. At each round, the best gens are selected to produce a child (i.e. crossover). Mutation is also used for exploring the solution area to avoid local optimums [2]. GA algorithm is coded in Python 3.6 using DEAP library.

Feedforward neural network (FNN) is a function approximator. A collection of neurons uses linear functions to translate inputs to outputs. Each neuron consists of weights and a bias. A sigmoid activation function is used for non-linear approximation. The weights and biases are updated at each epoch of training, using gradient-descent to minimize error. The neurons' operations are summarized as:

$$h_{it} = \sum_{j=1}^{n} w_{ij} x_j + b_i \tag{1}$$

$$Y_t = \sigma_y\left(\sum_{i=1}^{m} w_i h_{it} + b_y\right) \quad (2)$$

Where h_i is output for neuron i at time t, w_{ij} and w_i are weights, x_j is the model's input j, b_i and b_y are biases, Y_t is FNN's outputs, σ_y is sigmoid activation function, i indicates cells' number, and j indicates inputs' number [3].

FNN solely relies on information from current time step. Its application is limited for time series, where outcome relies on current status as well as previous statuses (e.g. stock market prediction and car-following behavior). RNN has an internal memory to find patterns in a sequence of data. Hence, RNN is able to model reaction time in car-following behavior. Figure 1 highlights the main difference between FNN and RNN cells. RNN cells use inputs at current time step (X_t) and hidden state at previous time step (h_{t-1}) to predict output (Y_t) [4]. Operations of RNN cells are:

$$h_{it} = \sum_{j=1}^{n} w_{ij} x_j + w_{ih} h_{it-1} + b_i \quad (3)$$

$$Y_t = \sigma_y\left(\sum_{i=1}^{m} w_i h_{it} + b_y\right)(4) \quad (4)$$

This study uses a deep recurrent neural network with LSTM for the deep-learning CF model (Deep-RNN). Deep-RNN consists of several RNN layers with LSTM cells. In Deep-RNN, the output of each RNN layer is used as input for the next layer, while each layer attempts to further improve accuracy. This cascading training process significantly increases Deep-RNN's capability to learn complex patterns [5].

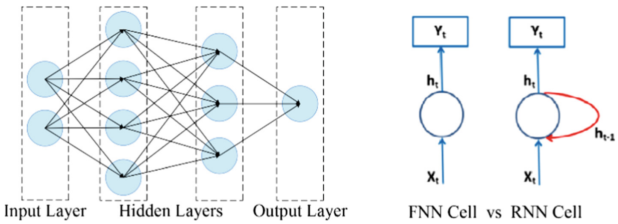


Fig. 1. Structure of FNN (left) and FNN and RNN cells' mechanisms (right).

LSTM is a neural network inside Deep-RNN cells learning important patterns in a sequence of data. LSTM has cell state (C_t) and hidden state (h_t). LSTM carefully alters the cell state by deciding to totally update, gradually update, or do not update. This decision is made through input, forget, and output gates [6].

Figure 2 shows the structure of Deep-RNN which is used in this study, with LSTM cells. All the neural networks are coded by the authors, using TensorFlow1.15. Further detailed information of each model is available from the authors upon request.

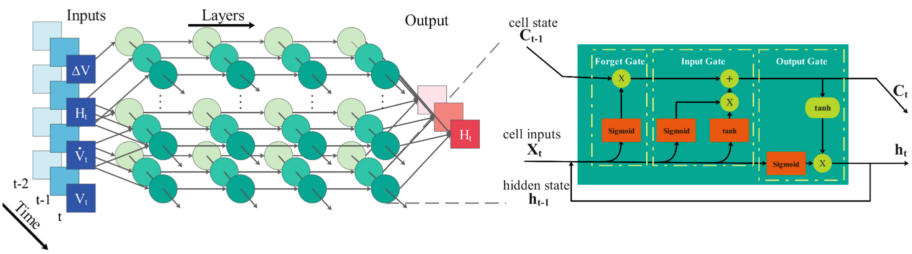


Fig. 2. Deep-RNN (left) and LSTM cell's (right) structures in this study.

2.2 Model Assessment

All the models receive the same inputs (i.e. velocity (v), acceleration ($\dot{v}$), headway (H) and velocity difference (Δv)) at time t and they predict headway at time $t + 1$.

This study uses Reconstructed Next Generation Simulation data (NGSIM) from Interstate 80 freeway (I-80) in Emeryville, California [7, 8]. This study uses 230,000 observations with 10 Hz resolution for training and calibration. NGSIM data has been normalized between 0 and 1. Next, the normalized data is divided into batches with 118 observations. The batch size is selected according to [9].

Performance of the models are compared at the local and global levels. At the local level, CF models receive the vehicle's inputs in time t and the model should accurately predict the vehicle's headway in time $t + 1$. This is a single-step assessment.

The global level is a multi-step assessment, in which the CF models regenerate vehicles' trajectories. At the global level, CF models receive the subject vehicle's inputs in time t and they predict the vehicle's headway in time $t + 1$. Then, the vehicle's attributes (e.g. acceleration) are updated for time $t + 1$, according to the predicted headway and physics-kinetic laws. Then, the subject vehicle's attributes at time $t + 1$ are used as inputs to predict its headway in time $t + 2$. This process is repeated to predict the headway and update the attributes at time $t + 3$, $t + 4$, …, up to $t + 100$.

This study uses percent absolute headway error (PHE) to compare the models' accuracy. The neural networks and genetic algorithm running many times with different number of populations, crossovers, layers, cells, and learning rates. The hyper-parameters which result in the highest accuracy are selected and used (Table 1).

Table 1. The CF models' hyper-parameters.

Hyper-parameter	GA	FNN	RNN	Deep-RNN
Population size	1000	–	–	–
Crossover size	500	–	–	–
Mutation rate	0.3	–	–	–
Number of layers	–	5	1	4
Number of cells	–	50, 100, 150, 80, 40	50	40
Learning rate	–	0.0001	0.001	0.001
Batch size	–	118	118	118

This study uses two different training-testing procedures, one for GA and another one for the neural networks (i.e. FNN, RNN, Deep-RNN) because GA is an optimization approach but neural network is a function approximation approach. The dataset for GA is divided into two parts: a training dataset (80%) and a testing dataset (20%). The neural networks have three sets of data: a training dataset (70%), a validation dataset (10%), and a testing dataset (20%). The datasets were divided by a set of random seeds which are identical across all the CF models. Hence, the models were trained and tested with identical datasets. Moreover, each model is run 5 times for a full cross-validation.

The training and testing procedures are shown on Fig. 3. GA is trained, using training dataset. After each epoch of training, the best performing gen's PHE is reported. If the PHE has not decreased for 20 epochs, the training process is stopped, and the best performing gen is selected to calculate PHE for testing trajectories.

The neural networks use gradient descent to find the best weights and biases, and reduce PHE. They are trained by the training dataset (Fig. 3). At the end of each training epoch, the models' PHEs are calculated for the validation trajectories. If the validation's PHE improves more than ε, training continues. Otherwise, the training process is stopped and PHE for the testing dataset is calculated. This stop-training strategy judges quality of learning based on an unseen dataset (i.e. validation data), preventing overfitting or underfitting. Unlike training based on "fix number of epochs", validation-based stop-condition evaluates training performance based on an unseen dataset. This adaptive stop-condition receives feedback from training process at the end of each epoch, to decide if the model needs more training or not. **ε** is chosen relatively small to the training's PHE. LSTM is initialized with randomly generated weights and biases. It is warmed up and trained through training process. Then, it is applied to the testing trajectories. Thus, it is not needed to warm up LSTM by testing dataset.

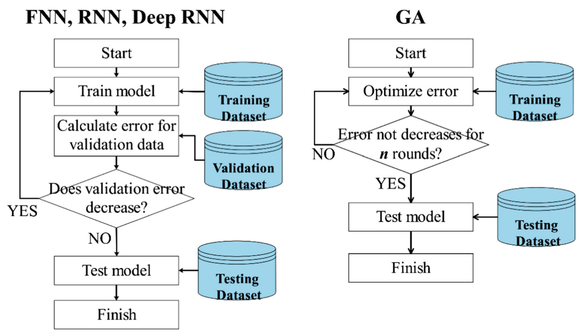


Fig. 3. GA and the neural networks' training and testing process.

3 Result

In this section, the car-following models are compared at the local and global levels. IDM parameters are calibrated and all the models' performance is shown in Fig. 4 at the local level. IDM has the highest error (0.06%). The error decreases from FNN (0.009%) to RNN (0.005%), and then to Deep-RNN (0.003%). In order to determine if the performance difference between the models is statistically significant, ANOVA test has been performed. The test found significant mean difference between the models (p-value = 1.2E-24). Then, t-test was conducted between each pair of the models. The result is reported on Table 2. T-test reveals significant performance differences between all pairs of models. At the local level, Deep-RNN significantly outperforms all the CF models and IDM significantly underperforms all the CF models.

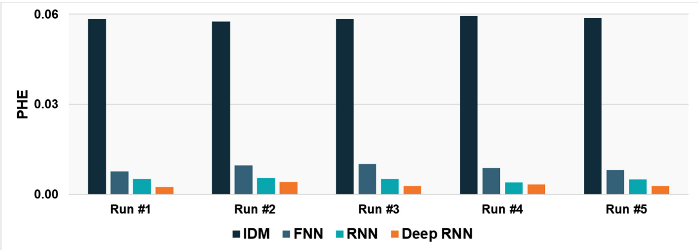


Fig. 4. Percent absolute headway error (PHE) for the CF models at the local level.

Table 2. Significance of performance difference at the local level.

Models	Run time(s)	PHE Ave.	PHE Var.	P-value		
				FNN	RNN	Deep-RNN
IDM	33,775	5.8E-02	4.8E-07	8.2E-08*	3.4E-08*	2.3E-08*
FNN	97	8.9E-03	1.1E-06		6.0E-04*	5.6E-05*
RNN	207	5.0E-03	3.2E-07			4.2E-03*
Deep-RNN	770	3.2E-07	4.6E-07			

At the global level, IDM and the neural networks were run 5 times, and their PHEs are reported on Fig. 5. Unlike the local level (Fig. 4), performance at the global level varies from one run to another. In order to achieve a reliable comparison, the models were run another 25 times (i.e. 30 runs in total). The models' average PHEs and variances are shown on Table 3. On average, FNN has the highest error (i.e. 25.8% error), with the widest performance variation. IDM demonstrates the lowest error (i.e. 12.4% error) and variation in performance. Deep-RNN has similar performance to IDM. RNN's performance is close to FNN.

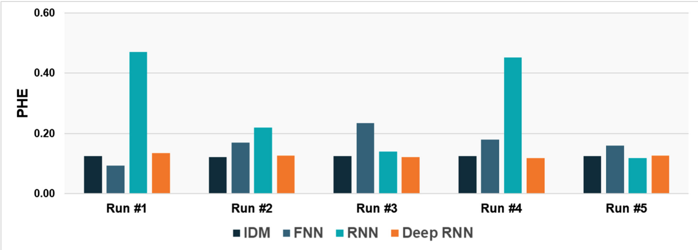


Fig. 5. Percent absolute headway error (PHE) for the CF models at the global level

Table 3. Significance of performance difference at the global level

Models	PHE Ave.	PHE Var.	P-value		
			FNN	RNN	Deep-RNN
IDM	12.4	3.0E-06	1.5E-02*	7.6E-06*	4.6E-03*
FNN	25.8	9.0E-02		8.3E-01*	2.2E-05*
RNN	24.5	2.1E-02			3.2E-05*
Deep-RNN	13.2	2.3E-04			

ANOVA test is applied to test the significance of mean difference between the CF models. ANOVA shows statistically significant performance difference among the models (p-value = 1.6E-3). Then, z-test is performed between each pair of the models. The p-values are reported on Table 3. Z-test demonstrates statistically significant mean difference between all CF models, except between FNN and RNN (p-value = 0.96), at $\alpha = 0.05$ level. This result reveals that at the global level, IDM performs significantly better than the other models, Deep-RNN significantly outperforms FNN and RNN, and there is no significant performance difference between FNN and RNN.

Figure 6 visually compares performance of the models at the global level, using "fan chart". A fan chart shows confidential intervals for each model, based on the models' performances in the 30 runs. FNN and RNN's accuracies significantly decrease after 10 steps. IDM and Deep-RNN could accurately predict drivers' headway up to 20 steps. Then, they only react to main stimulus (e.g. at step = 70 and 95). Although IDM and Deep-RNN have relatively similar mean performances, Deep-RNN shows more variation in performance that IDM. This variation in performance likely is due to the stochastic nature of function approximation in deep-learning versus optimization in IDM.

4 Discussion and Future Work

When it comes to CF models, traditional or artificial intelligent (AI), there is not one size fit all. It is important to use an appropriate model given a number of project specifics (e.g. types of available data and computational power). At the global level, all the CF models showed higher PHE than the local level. At the global level, error from each prediction step adds up and the error significantly increases at the final steps. IDM could maintain its accuracy for more steps due to its rule-based mechanism which always keeps safe headway or desired speed. But learning-based algorithms' (i.e. FNN, RNN, Deep-RNN) accuracies significantly decrease in multi-step prediction because they do not have any knowledge of human driving behavior, like the fact that headway cannot be zero or negative. Learning-based models learn all patterns and rules from data. When they are confronted with a high level of uncertainty at the global level, they may predict negative headway or velocity. In contrast, IDM is able to maintain stability by following its rules; for example, velocity cannot be negative.

For the data used in this study (NGSIM), we recommend developing a hybrid approach based on rule-based and learning-based algorithms. A hybrid approach could start with basic driving rules (e.g. maintaining safe headway) and learn further from pure data, using deep-learning algorithms. This approach would improve performance of car-following models at both local and global levels.

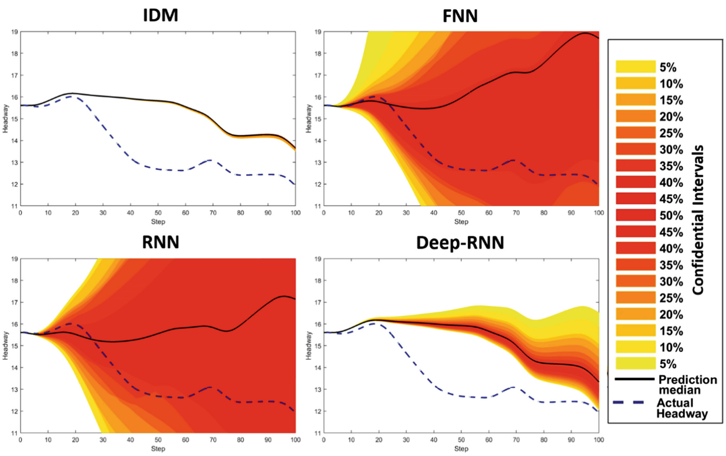


Fig. 6. The CF models' predicted headway vs real headway for 100 steps at the global level.

References

1. Treiber, M., Hennecke, A.: Congested traffic states in empirical observations and microscopic simulations. Phys. Rev. E **62**, 1805–1824 (2000)
2. Davis, L.: Handbook of Genetic Algorithms. Van Nostrand Reinhold (1991)
3. Nath, R., Rajagopalan, B.: Determining the saliency of input variables in neural network classifiers. Comput. Oper. Res. **24**, 767–773 (1997)
4. Elman, J.: Finding structure in time. Cogn. Sci. **14**, 179–211 (1990)
5. Vishnukumar, H. J., Butting, B.: Machine learning and deep neural network. In: Intelligent Systems Conference (2017)
6. Hochreiter, S.: Long short-term memory. Neural Comput. **9**, 1–32 (1997)

7. Federal Highway Administration, Next Generation Simulation (NGSIM) (2015). http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm
8. MULTITUDE, Multitude Dataset. http://www.multitude-project.eu
9. Zhou, M., Qu, X.: A recurrent neural network based microscopic car following model to predict traffic oscillation. Transp. Res. Part C **84**, 245–264 (2017)

Alert! Automated Vehicle (AV) System Failure – Drivers' Reactions to a Sudden, Total Automation Disengagement

Sarah El-Dabaja, Deborah McAvoy, and Bhaven Naik(✉)

Department of Civil Engineering, Russ College of Engineering and Technology, Ohio University, 28 W Green Dr., Athens, OH 45701, USA
{se954207,mcavoy,naik}@ohio.edu

Abstract. Despite the "driverless" term, drivers of automated vehicles still constitute an integral part of the human-machine team that will be driving the future. This study emphasized drivers by evaluating the attentiveness, stress levels, and reactions of 67 participants, ages 18–65+ years, during a sudden, total disengagement of automation in a driving simulator-based rural freeway setting. Attentiveness was characterized by a significant increase in gaze fixation and a significant decrease in fatigue, yet stress levels did not appear to significantly change. Regardless of age, gender, or level of non-driving related task involvement, participants reacted to the failure first by steering, requiring 12.30 s (50^{th} percentile) to 29.26 s (90^{th} percentile), followed by speed control after 18.26 s (50^{th} percentile) to 40.86 s (90^{th} percentile). These findings highlight the need for addressing the potentially dangerous implications of automation failure.

Keywords: Automated Vehicle (AV) failure · Driver reactions

1 Introduction

When automated vehicles (AVs) are mentioned, the first "image" that is often perceived is one of drivers cruising down a highway at a comfortable freeway speed, hands off the wheel and all attention diverted elsewhere, perhaps to a phone, book, other passengers, or even sleeping. After all, this is the image that is often included in presentations and promotions for AVs – the future of driving is portrayed as a task without hassle or worry, which may lead drivers to assume that they are now "free" to engage in other tasks as they please. However, this assumption could potentially become dangerous if the automated system suffers a sudden and complete shutdown. Drivers are now faced with a dire situation that they have never been trained for and certainly were not expecting, especially given the impression that the AV is completely responsible for driving. What happens next? As pessimistic as this scenario may seem, it is in fact a realistic problem that drivers of AVs may face given the possibility of computer glitches and cyberattacks. As the future of automated driving takes shape and the positive impacts of AVs are lauded, acknowledging the potential risks of AVs will likely ease many drivers' apprehensions about trusting an AV. This paper investigates a

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 49–55, 2020.
https://doi.org/10.1007/978-3-030-50943-9_7

scenario similar to the one described above in order to gain an understanding of the human-machine interactions that take place during an emergency automation failure.

2 Driver Transition Behaviors During AV Driving and Failures

Drivers' responses to AV failures have been studied primarily with longitudinal control systems and, more recently, in highly-automated vehicles, in the context of inappropriate functioning or driver-initiated disengagements [1–3]. A comparison of failures in the presence of longitudinal control only versus failures in highly-automated vehicles suggests that drivers exhibit inferior responses when the vehicle controls both aspects of driving. This was attributed to reduced situational awareness, since driver detachment from these basic tasks may induce secondary non-driving related tasks (NDRTs) or drowsiness. While the distraction and increased cognitive workload inherent to NDRTs could negatively affect the driver's ability to regain control of the vehicle, it is also conceivable that distractions might help mitigate the fatigue that would otherwise accompany automated driving [4–9]. Yet, investigations with respect to driver transition behaviors during normal AV driving have indicated that drivers' reactions to takeover requests (TORs) are not significantly impacted by NDRT involvement nor age, drowsiness, or duration of AV driving, but only appear to be influenced by drivers' familiarity with TORs [9–11]. During these transitions from highly-automated driving, which were primarily investigated in simulated environments wherein a non-critical controlled TOR was presented, drivers required 5 s (minimum) to take control of the vehicle, approximately 15 s to become fully cognizant, and up to 35–40 s to fully stabilize the vehicle [12–16]. Whereas simulator studies focused on driver-related factors, data obtained from California regarding actual on-road disengagements has indicated that drivers' reactions to the TORs varied by type of disengagement, roadway and weather conditions, and vehicle-miles travelled with AV activated [17]. On the basis of these studies, providing different TORs in both audio and visual form and educating drivers on the topic of TORs have been suggested as possible means for improving drivers' responses [18, 19].

3 Methodology

In this study, drivers' reactions (physical and physiological) were observed to determine: (i) drivers' stress levels during automation failure; (ii) perception reaction time (PRT) for responding to automation failure; and (iii) time to full control (TFC) for drivers to take reasonable control of all driving tasks following the automation failure. These reactions were obtained by developing a set of manual and automated driving scenarios on a 65 mph, rural, six-lane freeway using Ohio University's DriveSafety DS-600 research simulator [20]. The scenarios were designed such that drivers would be exposed first to 10 miles of manual driving followed by 10 miles of AV driving replicating Level 5 automation, with the ego AV maintaining a platoon position with 0.6 s headways [21]. The AV was actively driving until failure occurred, which was identified via an audible and visual TOR in a heads-up display fashion. At this point, the Level 5 automation abruptly ceased to function, which resulted in the vehicle slowing and drifting. Henceforth, control was relinquished back to the driver for the remaining five miles of the scenario. Throughout both scenarios, drivers experienced incidents in order to familiarize them with the feel of the simulator, provide a baseline behavior, and showcase the capabilities of the AV.

A total of 67 participants were recruited in accordance with IRB protocol [22] and were grouped by age according to previous studies of similar physical reaction capabilities [7, 23–25]. The following describes the age and gender (M/F) distribution of participants: 18–25 yrs (26 M, 8 F); 26–40 yrs (7 M, 6 F); 41–55 yrs (4 M, 7 F); 56–65 yrs (5 M, 2 F); and 65+ yrs (1 M, 1 F). Prior to driving, participants completed a pre-test questionnaire and were briefed about the study, which included a subtle reminder of their responsibility to resume control of the vehicle in the event that an issue were to arise. Participants' behaviors were quantified through both direct observations and the following recorded data: the driving simulator collected velocity (m/s), acceleration (scale of 0–1), braking (scale of 0–1), steering (radians), and lane position (m) data at 60 Hz; the FaceLab eyetracker collected gaze vectors and percent closure (PerClos) at 60 Hz; and the BioPatch logged heart rate (HR) in beats per minute (bpm) and heart rate variability (HRV) in milliseconds (ms), both at 1 Hz.

4 Results and Discussion

When presented with the failure TOR, drivers reverted their gaze back to the roadway and became more alert, as evidenced by a statistically significant ($p = 0.004$) increase in the percentage of time for which gaze was focused on the roadway from 66.4% (AV) to 73.6% (failure to PRT), and a statistically significant decrease ($p = 0.001$) in the percentage of time for which fatigue was observed from 2.01% (AV) to 0.05% (failure to PRT). At this point, drivers were trying to discern what was causing the problem and had not yet attempted to control the vehicle; therefore, this was not considered the "first reaction" for the purposes of defining PRT and TFC. Rather, PRT and TFC were based upon attempts to regain both lateral and longitudinal control, with PRT marked by the first attempt to control the vehicle and TFC defined by drivers beginning to control all aspects of driving. All 67 participants responded to the AV failure with steering first followed by acceleration. For some drivers, these reactions were immediate while others watched intently, skeptical about how to respond, as evidenced by the PRTs and TFCs in Table 1 and Fig. 1.

Table 1. PRT and TFC by Age, Gender, and NDRT involvement

	Avg PRT (s)	Min PRT (s)	Max PRT (s)	Std Dev
Age [N*]				
18–25 [32]	11.11	1.15	67.44	13.69
26–40 [12]	15.37	3.18	66.03	16.86
41–55 [10]	13.03	2.15	71.97	19.74
56–65 [7]	8.61	2.98	17.65	6.20
65+ [2]+	21.73	5.83	37.63	22.49
Gender [N*]				
Male [41]	10.51	1.15	71.97	13.02
Female [22]	15.53	3.18	67.44	17.66
NDRT [N*]				
None [33]	12.86	1.60	71.97	16.45
Task [30]	11.71	1.15	67.44	13.28
	Avg TFC (s)	Min TFC (s)	Max TFC (s)	Std Dev
Age [N*]				
18–25 [31]	12.36	4.02	42.63	9.24
26–40 [12]	19.65	4.68	41.02	11.71
41–55 [10]	13.00	6.02	74.42	5.96
56–65 [7]	13.19	3.57	23.55	8.11
65+ [2]+	27.35	14.17	40.53	18.64
Gender [N*]				
Male [41]	13.66	3.57	42.63	9.53
Female [21]	16.00	4.68	41.02	10.59
NDRT [N*]				
None [32]	13.35	3.57	40.53	9.13
Task [30]	15.63	4.83	42.63	10.66

*N denotes the number of participants in the associated category.
+Age group not included in statistical analyses due to small sample size.

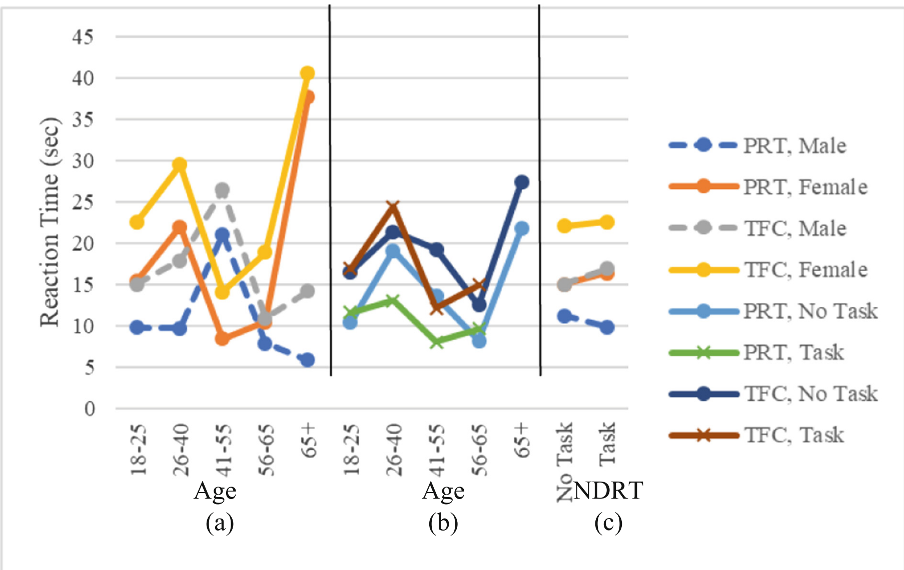


Fig. 1. Mean PRT and TFC for (a) Age by gender; (b) Age by NDRT; (c) NDRT by gender

A visual inspection of the data contained in Table 1 and Fig. 1 appears to suggest some age-based, gender-based, and NDRT involvement-based trends; therefore, a Factorial ANOVA was used to compare mean PRTs and TFCs across the variables of interest to determine if the differences in reaction times are statistically significant. To perform the Factorial ANOVA, the raw PRTs and TFCs were first log-transformed [26–29]. The results indicate that not enough evidence exists to suggest PRT is significantly affected by the main effects of the variables gender ($p = 0.072$), age ($p = 0.708$), and NDRT involvement ($p = 0.727$), or the interactions between gender*age ($p = 0.766$), gender*task ($p = 0.973$), age*task ($p = 0.997$), and gender*age*task ($p = 0.919$). Similarly, not enough evidence exists to suggest TFC is significantly affected by the main effects of the variables gender ($p = 0.113$), age ($p = 0.144$), and NDRT involvement ($p = 0.110$), or interactions between variables gender*age ($p = 0.269$), gender*task ($p = 0.667$), age*task ($p = 0.258$), or gender*age*task ($p = 0.328$). Therefore, it can be assumed that the critical PRTs and TFCs provided in Table 2 can be applied to all drivers, regardless of age, gender or NDRT involvement – a finding which is imperative for the design of an AV failure driver assistance system.

Table 2. Critical PRT and TFC Values

Percentile	PRT (sec)	TFC (sec)
50^{th} (Mean)	12.30	18.26
85^{th}	16.87	25.83
90^{th}	29.26	40.86
95^{th}	54.04	68.14

Post-simulation conversations highlighted that AV failure placed several participants in a situation of unknowns – at first, they were unsure of what was happening, wondering if the AV would take over again. Despite these uncertainties, participants' mean HR and HRV remained fairly constant. Repeated measures ANOVA revealed that HR during AV failure recovery was significantly, albeit slightly, lower than that observed during both manual (−4.5%, $p = 0.000$) and AV driving (−1.9%, $p = 0.003$), and HRV during AV failure recovery was not significantly different from that observed during either manual or AV driving. Post-simulation feedback also clarifies this observation – some drivers mentioned they were already stressed by the AV driving, which was evidenced by a drop in the minimum HRV, while other drivers noted that their response was diluted after being so relaxed and removed from driving.

5 Conclusions

Drivers' response to AV failure was characterized by statistically significant changes in gaze and fatigue, but not stress levels. While drivers refocused their attention and returned to a more aroused status, their initial reaction (driving-related) was to resume control of steering (PRT), followed by speed (TFC). PRT ranged from 1.15 s to 71.97 s (min – max) across all participants, while TFC ranged from 3.57 s to 74.42 s (min – max). Neither PRT nor TFC appeared to be significantly affected by age, gender, or NDRT involvement. Rather, reactions seemed to be dictated more by personality or risk-taking tendencies; however, this is merely a suggestion based on observations and requires further research. Thus, similar to PRTs for braking, it can be assumed that critical PRTs and TFCs can be adopted. Critical PRTs/TFCs were determined as: 12.30 s/18.26 s (50^{th} percentile); 16.87 s/25.83 s (85^{th} percentile); 29.26 s/40.86 s (90^{th} percentile); and 54.04 s/68.14 s (95^{th} percentile). With the 50^{th} percentile PRT nearly four times higher than the typical 2.5 s and 85% of drivers requiring 16.87 s PRT and 25.83 s TFC, developing a failure recovery system should be considered a top priority. At a speed of 65 mph, the 85^{th} percentile PRT and TFC translates to a distance of 742 ft. that the vehicle is traversing before the driver begins to resume lateral control and an additional 394 ft. (i.e., a total 1,136 ft.) for longitudinal control. Based on these results, it appears that the future of automated driving should be bolstered by a failure recovery system and exposure to AVs through drivers' education. While a driver assistance system for AV failure can be expected to introduce much needed redundancies, immersive drivers' education can also be expected to enhance driver performance during AV failures, especially since AV driving seems to lack a certain instinctiveness that would normally help drivers respond to incidents. Of course, these results should be viewed in light of the limitations presented by this study design, particularly the simulated environment and one-time exposure, to provide direction for future research.

References

1. Nilsson, J. et al.: Driver performance in the presence of adaptive cruise control related failures: implications for safety analysis and fault tolerance. In: 43rd Annual IEEE/IFIP Conference on Dependable Systems and Networks Workshop (DSN-W) (2013)
2. Strand, N., et al.: Semi-automated versus highly automated driving in critical situations caused by automation failures. Transp. Res. Part F **27**, 218–228 (2014)
3. Lv, C., et al.: Analysis of autopilot disengagements occurring during autonomous vehicle testing. IEEE/CAA J. Automatica Sinica **5**, 58–68 (2018)
4. Merat, N., et al.: Highly automated driving, secondary task performance, and driver state. Hum. Factors **54**(5), 762–771 (2012)
5. Kim, J., et al.: Take-over performance analysis depending on the drivers' non-driving secondary tasks in automated vehicles. In: 2018 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1364–1366 (2018)
6. Vogelpohl, T., et al.: Transitioning to manual driving requires additional time after automation deactivation. Transp. Res. Part F Traffic Psychol. Behav. **55**, 464–482 (2018)
7. Greenlee, E., DeLucia, P., Newton, D.: Driver vigilance in automated vehicles: hazard detection failures are a matter of time. Hum. Factors **60**(4), 465–476 (2018)
8. Jamson, A., et al.: Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. Transp. Res. Part C **30**, 116–125 (2013)
9. Naujoks, F., et al.: From partial and high automation to manual driving: relationship between non-driving related tasks, drowsiness and take-over performance. Accident Anal. Prev. **121**, 28–42 (2018)

10. Korber, M., et al.: The influence of age on the take-over of vehicle control in highly automated driving. Transp. Res. Part F Traffic Psychol. Behav. **39**, 19–32 (2016)
11. Feldhutter, A., et al.: How the duration of automated driving influences take-over performance and gaze behavior. In: Advances in Ergonomic Design of Systems, Products, and Processes (2017)
12. Favaro, F., Eurich, S., Nader, N.: Autonomous vehicles' disengagements: trends, triggers, and regulatory limitations. Accident Anal. Prev. **110**, 136–148 (2018)
13. Merat, N., et al.: Transition to manual: driver behavior when resuming control from a highly automated system. Transp. Res. Part F **27**, 274–282 (2014)
14. Mok, B., et al.: Emergency, automation off: unstructured transition timing for distracted drivers of automated vehicles. In: 2015 IEEE International Conference on Intelligent Transportation Systems, pp. 2458–2464 (2015)
15. Kuehn, M., Vogelpohl, T., Vollrath, M.: Takeover times in highly automated driving. In: 25th International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Transportation Safety Administration (2017)
16. Morgan, P., et al.: Manual takeover and handover of a simulated fully autonomous vehicle within urban and extra-urban settings. In: Advances in Human Aspects of Transportation: Advances in Intelligent Systems and Computing. AHFE (2017)
17. Dixit, V., Chand, S., Nair, D.: Autonomous vehicles: disengagements, actions, and reaction times. PLoS One **11**(12), e0168054 (2016)
18. Hergeth, S., Lorenz, L., Krems, J.: Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. Hum. Factors **59**(3), 457–470 (2017)
19. Funkhouser, K., Drews, F.: Reaction times with switching from autonomous to manual driving control: a pilot investigation. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 1854–1858 (2016)
20. DriveSafety. https://drivesafety.com/customers-research/
21. Nowakowski, C., et al.: Cooperative adaptive cruise control: driver acceptance of following gap settings less than one second. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 2033–2037 (2010)
22. El-Dabaja, S.: Drivers of "driverless" vehicles: a human factors study of connected and automated vehicles. PhD Dissertation, College of Engineering and Technology, Ohio University, Athens (2020)
23. Rumschlag, G., et al.: The effects of texting on driving performance in a driving simulator: the influence of driving age. Accident Anal. Prev. **74**, 145–149 (2015)
24. Woods, D., et al.: Age-related slowing of response selection and production in a visual choice reaction time task. Front. Hum. Neurosci. **9**, 193 (2015)
25. Salvia, E., et al.: Effects of age and task load on drivers' response accuracy and reaction time when responding to traffic lights. Front. Aging Neurosci. **8**, 169 (2016)
26. Takoa, G.: Brake reaction times of unalerted drivers. ITE J. **59**(3), 19–21 (1989)
27. Koppa, R.: Human factors. In: Gartner, N.. Messer, C., Rathi, A. (eds.) Traffic Flow Theory Transportation Research Board Monograph. National Research Council, Washington (2003)
28. Setti, J., Rakha, H., El-Shawarby, I.: Analysis of brake perception-reaction times on high-speed signalized intersection approaches. In: IEEE Conference on Intelligent Transportation Systems. IEEE (2006)
29. Shanker, R.: Lognormal distribution and its applications in biological and medical sciences. In: 4th International Conference and Exhibition on Biometrics and Biostatistics, OMICS International Conference, San Antonio (2015)

Traffic Behavior Analysis Using Mobile Base Station Data

Juyoung Kim[1(✉)] and Dongho Kim[2]

[1] Korea Transport Institute, Sejong City, South Korea
jykim@koti.re.kr

[2] Department of Transportation Engineering, The University of Seoul, Seoul, South Korea
kdh4756@koti.re.kr

Abstract. Most Koreans have mobile and their location information is collected based on location of the base station in one second increments. Mobile base stations are installed at intervals of 50 m in cities, and up to 2 km apart in rural areas. We developed an algorithm that builds an individual trip chain using mobile base station data and distinguishes home from work area by analyzing daily traffic patterns. The purpose of this study is to analyze traffic generation unit and traffic characteristics through seamless trip chain analysis of individual mobile base station data. A new method and experimental approach are established to estimate the passenger O/D based on mobile base station data. A new method has been analyzed to overcome many of the shortcomings of the existing O/D estimation methods that are based on household surveys, such as zero cells and inaccuracy due to low sampling rates.

Keywords: Mobile base station data · Traffic behavior · Trip chain analysis

1 Mobility Big Data

The movement of vehicles, public transportation, and people are critical to understanding the real situation in regards to mobility. Navigation data that gathers the locations of vehicle by the second is the information that enables us to understand vehicle movement accurately. The movement of public transportation users can be traced using the public transportation card data created when those users tag their cards each time they get on or off public transportation. The movement of people can be tracked using mobile data or GPS data collected every tenth of a second, relayed from the base station (Fig. 1).

Mobility big data provides the following benefits for the transportation sector. First, path-based data can be utilized as link information, such as travel time and speed without spatial discontinuation. Mobility big data ultimately enables us to collect data from throughout the country without spatial discontinuation and to create spatial and time series indicators. Second, high-quality and high-resolution data can be utilized that enables us to understand the temporal and spatial movement and behavioral characteristics of moving subjects. Third, temporally and spatially-customized transportation policies can be supported by considering these traffic behaviors and characteristics,

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 56–67, 2020.
https://doi.org/10.1007/978-3-030-50943-9_8

Fig. 1. Representation of mobility big data

because the root causes of traffic phenomena and problems can be understood and their implications can be readily identified.

2 Reference Review

Mobile data, among various types of big data, contains potential information which can estimate the location of individuals and the demand for passenger traffic. Related studies have been underway abroad, beginning in the early 2000s, to estimate passenger traffic demand using mobile data [1]. This study was focused on processing mobile data sent and received from base stations in the initial stages of the study. Studies working towards understanding the activity of individuals and their traffic characteristics, and estimating traffic volume at origin and destination points based on those understandings, were started in full five years ago. Many studies were conducted to estimate O/D using mobile CDR data. Representative studies include [2–4], and [5]. In those studies, the locations of activity were divided into home, work, and other, in consideration to travel time and visit frequency, whereas the purpose of trip was divided into from-home-to-work, from-home-to-other, and not-from-home trips. [6] studied a methodology for predicting the means of travel using the CDR data. The basic concept was to estimate the means of travel based on the number of people who departed from the same origin and arrived at the same destination, the departure time, and the travel time. [7] studied a methodology for estimating an O/D by using CDR data and Wi-Fi data. They proposed a method that estimates metro and commuter train O/D flow by matching movement traces, which are gathered using CDR and Wi-Fi data, together with information from the public transportation network, and estimates passenger car O/D flow among trips that travel above certain average speeds.

3 Existing O/D Estimation Method in Korea

The national transportation survey (NTS) conducted in 2016 cost more than 5.9 billion won to survey 1.25% of the nation's households, 613 facilities, 335 highway toll gates, 824 traffic survey points, and 1,824 chartered buses. Korea Transport Database (KTDB) conducts the O/D estimation process using the traditional four-step model for the national and six metropolitan areas. The sampling rate of the survey regarding the actual conditions of household trips fell from 2.54% in 2010 to 1.25% in 2016 due to the budget limitations of the survey and increases in the unit price of the survey, which seems to have had a direct impact on the reliability of O/D estimations. Zero cells in the O/D matrix can be pinpointed as a limitation on understanding of the actual condition of household trips due to the low sampling rate. When the survey data is simply converted into total number data, zero cells occupy 89% of the capital area, 85% of the Busan/Ulsan area, 84% of the Daegu metropolitan area, 79% of the Gwangju metropolitan area, and 92% of the Daejeon area. Even though zero cells are corrected for during the O/D estimation process when using other survey data (e.g., facilities surveys) and the performance data of public transportation, it sufficiently demonstrates that a low sampling rate imposes many restrictions on reliability improvement (Table 1).

Table 1. Status of zero cells in the result of O/D estimation (2016 KTDB)

Metropolitan area	O/D before zero cell correction			O/D after zero cell correction		
	Total number of cells	Number of zero cells	Ratio	Total number of cells	Number of zero cells	Ratio
Seoul/Gyonggi	1,288,225	1,152,628	89%	1,288,225	724,832	56%
Ulsan/Busan	178,929	152,564	85%	178,929	128,374	72%
Daegu	93,636	78,815	84%	93,636	70,912	76%
Gwangju	29,241	23,048	79%	29,241	17,669	60%
Daejeon	205,209	189,814	92%	205,209	130,465	64%

Source: Internal data on National traffic big data project team, The Korea Transport Institute (2018) [8]

Secondly, missing trip history can be identified as having arisen during the survey process. The representative example being that the number of trips reportedly taken during lunch and in the evening are very low (Table 2).

Table 2. Trips taken during lunch and in the evening (2016 KTDB)

Time zone	Area	Traffic frequency					
		0	1	2	3	More than 4 times	Total
From 11:30 to 13:30	Nationwide	91.56%	7.66%	0.78%	0.01%	0.00%	100.0%
	Seoul	95.30%	4.32%	0.37%	0.01%	0.00%	100.0%
From 18:00 to 21:59	Nationwide	28.96%	65.04%	4.77%	1.22%	0.02%	100.0%
	Seoul	21.63%	73.16%	4.11%	1.08%	0.02%	100.0%

Source: Internal data from the national traffic big data project team, The Korea Transport Institute (2018) [8]

Table 3 shows the results of traffic demand estimation using KTDB. The error rate between the observed and estimated traffic volume in the country in 2016 was 122% when the traffic volume was less than 5,000 vehicles, and it was about 25% when the traffic volume was over 60,000 vehicles. The error rate in the metropolitan area, where the zone size is divided into more detail, was 104% in the capital area when the traffic volume was less than 5,000 vehicles and 19.5% when the traffic volume was over 100,000 vehicles.

Table 3. Error ratio (%RMSE) by traffic volume level (2016 KTDB)

Item	Freeway		General highway		Local road		Provincial road		Total	
	Number of spots	Error rate	Number of spots	Error rate	Number of spots	Error rate	Number of spots	Error rate	Number of spots	Error rate
1–5,000	10	104	972	107	265	143	980	134	2227	122
5,000–10,000	33	58	617	68	101	72	170	68	921	68
10,000–20,000	166	29	511	47	87	66	122	69	886	49
20,000–30,000	232	26	241	47	31	63	39	65	543	42
30,000–40,000	137	22	86	48	18	51	16	62	257	37
40,000–50,000	114	23	59	44	4	43	10	74	187	35
50,000–60,000	69	31	29	47	2	128	3	48	103	38
>60,000	293	22	30	42	2	141	0	0	325	25
Total	1054	27.9	2545	71.5	510	106.2	1340	116.8	5449	54.5

Source: Internal data from the national traffic big data project team, The Korea Transport Institute (2018) [8]

4 New Approach Based on Mobile Base Station Data

The proposed procedure goes like this: first, a total trip O/D is created using mobile data. Then, a public transportation trip O/D is created out of information regarding railroads, buses, and subways by using the public transportation card data and other performance records. The O/D for non-public transportation trips, such as travel via passenger cars or other means, is created by deducting the public transport trip O/D from the total trip O/D. The purpose of travel for the national intercity O/D is divided into regular and irregular trips. Regular trips include commutes to work and school, which are performed regularly at certain times. Irregular trips include going shopping, for business, to travel, and or to transport cargo. The purpose of trip in the metropolitan area O/D is divided into from-home trips and not-from-home trips, and from-home trips is further sub-divided into regular and irregular trips. Not-from-home trips include business, leisure, and cargo transportation, and is not sub-divided into regular and irregular trips (Table 4).

Table 4. Classification of trip purpose based on mobile data

Item		Purpose of trip		Type of trip	Subject of trip
National intercity O/D	O/D approach	Regular		Commuting, commuting to school, etc.	Passengers and drivers of cargo trucks
		Irregular		Shopping, business, leisure, cargo transportation, etc.	
Metropolitan O/D	P/A approach	From-home	Regular	Commuting, commuting to school, etc.	
			Irregular	Shopping, leisure, etc.	
		Not-from-home		Business, leisure, cargo transportation, etc.	

Even though mobile data is collected from base stations, the reception range of each base station is different. Therefore, a spatial analysis unit should be set using a standardized method. In general, the spatial analysis unit to create the O/D trip volume is made consistent with administrative boundaries. Therefore, the mobile data sets an output area that has a spatial range similar to the reception range of the base station as a spatial analysis unit. The type of passer activity location should be estimated in order to understand information regarding the origin, destination, and waypoint based on their trip's purpose by using mobile data that has spatial uncertainties. The type of passer activity location is estimated considering travel time and visit frequency, as was done in the international examples, and classified into pass-by, stay area, and potential stay area. The purpose of trip is classified based on trips between the stay areas and potential stay areas. The mobile data is sample data (30% market share) that has collected trip information of mobile communications subscribers. Therefore, the population-based total trip O/D is forecasted by applying a data generating coefficient to the sample. The data generating coefficient is calculated with consideration for the present condition of a mobile service provider's subscribers and socioeconomic indicators (Table 5).

Trip information is classified into personal, public, or other transportation mode based on the predicted trip's purpose, determined using the mobile data. The personal transportation mode includes passenger cars (including taxis), walking, and bicycling whereas the public transportation mode includes railways, buses, subways, and so on. Other modes include trucks used to transport goods (Table 6).

Table 5. Method for forecasting trip purpose between activity locations

Item			Contents
Nationwide intercity O/D	Regular		Nighttime stay area (residential area) → Weekly stay area traffic Weekly stay area → Nighttime stay area (residential area) traffic
	Irregular		Daytime stay area → Potential stay area trip Nighttime stay area (residential area) → Potential stay area trip Potential stay area → Daytime stay area Potential stay area → Nighttime stay area (residential area) trip Potential stay area → Potential stay area trip
Metropolitan area O/D	From-home trip	Regular	Nighttime stay area (residential area) → Weekly stay area trip Weekly stay area → Nighttime stay area (residential area) trip
		Irregular	Nighttime stay area (residential area) → Potential stay area trip Potential stay area → Nighttime stay area (residential area) trip
	Not-from-home trip	Regular Irregular	Daytime stay area → Potential stay area trip Potential stay area → Daytime stay area Potential stay area → Potential stay area trip

Table 6. Mode classifications

Item	Means of traveling
Personal mode	Passenger cars (including taxis), walking, bicycling, chartered buses
Public transportation mode	Public transportation (High speed rail + general rail + subway + express bus/intercity bus + Red Bus + city bus + town bus + chartered bus + air + ocean)
Other modes	Trucks (for business, for non-business)

5 Application and Comparison with KTDB

In most cities and provinces, mobility data was estimated to a greater degree than the KTDB's estimation for both departure and arrival trip volume. KTDB comprehensively estimates O/D trips among districts considering the actual condition surveys for household trips, public transportation shipping records, highway TCS data, and the traffic volume of cordons and screen lines in pertinent cities and provinces. Therefore, KTDB may have errors resulting from uncertainty regarding final destinations after a trip has exited from highway toll gates and due to the absence of trip performance between adjacent cities and provinces. The results of analyzing O/D trip flow among regions in the 16 cities and provinces showed that trip volume to adjacent cities and provinces, which cannot be calculated using the highway TCS data or public transportation performance, is much more accurately accounted for in through the mobile data (Table 7 and Fig. 2).

Table 7. Comparison of departure and arrival traffic volume (only for intercity trips)

Departure zone	Mobile	KTDB	Error[1)]	Arrival zone	Mobile	KTDB	Error
Seoul	4,374,155	3,908,234	11.9%	Seoul	4,283,751	3,887,710	10.2%
Busan	786,883	589,977	33.4%	Busan	827,897	588,702	40.6%
Daegu	591,373	539,014	9.7%	Daegu	645,088	549,129	17.5%
Incheon	1,462,916	1,070,726	36.6%	Incheon	1,621,204	1,128,417	43.7%
Gwangju	414,425	428,363	−3.3%	Gwangju	415,491	252,528	64.5%
Daejeon	444,156	417,735	6.3%	Daejeon	448,413	427,015	5.0%
Ulsan	366,966	235,099	56.1%	Ulsan	331,092	224,580	47.4%
Gyeonggi	4,874,189	4,148,774	17.5%	Gyeonggi	4,817,221	4,105,482	17.3%
Gangwon	265,543	225,079	18.0%	Gangwon	238,719	223,015	7.0%
Chungbuk	415,340	389,640	6.6%	Chungbuk	429,503	379,058	13.3%
Chungnam	610,286	548,995	11.2%	Chungnam	582,768	541,726	7.6%
Jeonbuk	222,358	195,859	13.5%	Jeonbuk	230,656	199,458	15.6%
Jeonnam	474,001	301,592	57.2%	Jeonnam	475,304	478,371	−0.6%
Gyeongbuk	799,509	678,161	17.9%	Gyeongbuk	758,145	679,669	11.5%
Gyeongnam	835,848	625,618	33.6%	Gyeongnam	837,504	637,079	31.5%
Jeju	37,786	37,430	1.0%	Jeju	38,003	38,003	0.0%
Sejong	212,385	186,858	13.7%	Sejong	207,360	187,211	10.8%

1) Error = (Communication data-KTDB)/KTDB

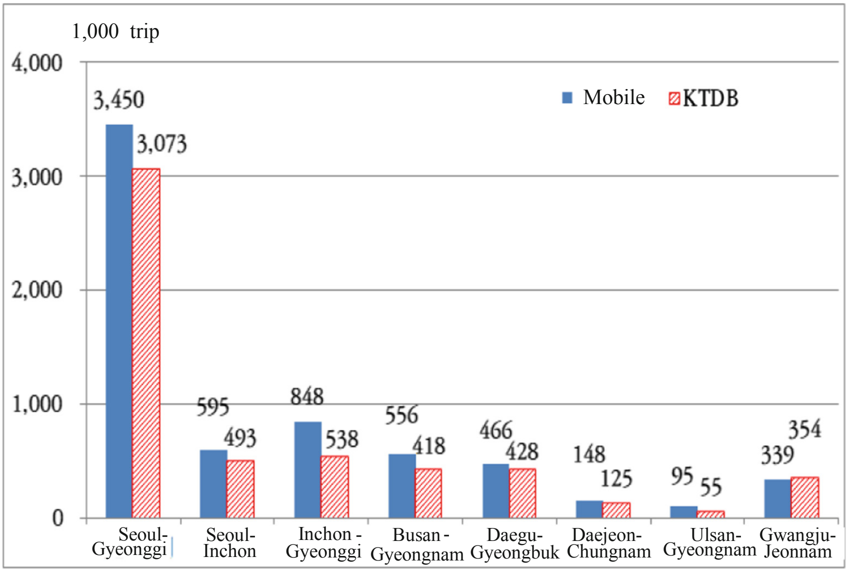


Fig. 2. Comparisons of trip volume between adjacent cities and provinces

When mobile data and KTDB were compared with trip volume inside the zones of cities and local provinces, research results showed that cities have more short-distance trips than local provinces. It was found that there was fewer mobile data in cities, compared to KTDB, when trips shorter than 3 km were excluded. Additionally, there was fewer mobile data in local provinces, compared to KTDB, when trips shorter than 4 km were excluded (Table 8).

Table 8. Change in traffic volume caused by excluding pedestrian traffic (distance constraint)

City and province	Full data generation	Pedestrian traffic classification scenario based on travel distance[1]					KTDB[2]
		≤ 500 m	≤ 1 km	≤ 2 km	≤ 3 km	≤ 5 km	
Urban area (trip volume)	123.11	107.52	78.72	45.74	32.58	20.06	36.26
Ratio (ver. KTDB)	3.4	3.0	2.2	1.3	0.9	0.6	1.0
Rural area (trip volume)	112.29	106.32	93.57	70.93	53.56	28.21	40.51
Ratio (ver. KTDB)	2.8	2.6	2.3	1.8	1.3	0.7	1.0
Total	235.40	213.84	172.29	116.67	86.14	48.27	76.77

1) Trips shorter than the criteria are defined as walking and excluded from the trip volume
2) Trip volume that excludes pedestrian trips

According to NTS (household trip survey) by KTDB, 0.08 and 0.06 trip units occur across the country and throughout Seoul respectively, during lunch time. It was found that only 4% (nationwide) and 3% (in Seoul) of trips occurred during lunch time, with consideration for the fact that the trips were for traveling back and forth. On the contrary, according to the mobile data, occurred trips, with basic units being those trips traveling more than 2 km, were 17% (nationwide) and 11% (in Seoul). Whereas trips of 5 km were 8% (nationwide) and 7% (in Seoul). The occurred trip unit for the evening shows similar results. The trip unit is 0.22 (nationwide) and 0.25 (in Seoul) according to the condition survey on household traffic, whereas the mobile data shows 0.48 (nationwide) and 0.38 (in Seoul) for trips longer than 2 km, which is the level that excludes "pedestrian traffic." However, if only trips longer than 5 km are taken into account, it is similar to the actual condition survey on households (0.21 nationwide and 0.24 in Seoul respectively). As described above, even though various trips, excluding major trips like commuting and travel for business, occur every day, it is often not taken into account in the household trip survey. This is a direct cause of the underestimation of O/D trip volume (Table 9).

Table 9. Trip unit comparison of NTS and mobile data

Time zone	Area	NTS	Mobile					
			Total	≤ 0.5 km	≤ 1 km	≤ 2 km	≤ 3 km	≤ 5 km
11:30 ~ 13:30 ~	Nationwide	0.08	0.64	0.59	0.48	0.35	0.27	0.17
	Seoul	0.06	0.71	0.53	0.37	0.22	0.18	0.14
18:00 ~ 21:59 ~	Nationwide	0.22	0.98	0.89	0.72	0.48	0.36	0.21
	Seoul	0.25	1.09	0.92	0.64	0.38	0.31	0.24

The KTDB showed that 0.6% zero cells were created in the O/D for nationwide intercity travel and 56.2% zero cells were created in the capital area O/D. On the other hand, mobile data showed that 0.1% zero cells were created in the O/D for nationwide intercity travel and 8.1% zero cells were created in the capital area O/D. It was analyzed that the rate of zero cell (no traffic volume between zones) occurrence could be reduced significantly by utilizing mobile data (Table 10).

Table 10. Comparison of zero cells between KTDB and mobile data

Item	KTDB		Mobile	
	Nationwide intercity O/D	Capital area O/D	Nationwide intercity O/D	Capital area O/D
Number of zero cells	392	723,688	62	104,120
Total	62,500	1,288,225	62,500	1,288,225
Zero cell ratio	0.6%	56.2%	0.1%	8.1%

When the volume ratio by trip length is reviewed using internal trips only that excludes inter-regional trips, the trip volume ratio below 2 km, which includes pedestrian trips, was found to be high in most of the metropolitan cities. In Seoul, the ratio of trips with a traveling distance of 5 km or more account for more than 16%, and about 7.8 million passages occur on a weekday basis. The ratio of trip volume of more than 10 km length versus the entire trip volume is 23.3% in Seoul, followed by Busan (17.5%), and Ulsan (14.4%), with trips shorter than 2 km excluded. This indicates that Seoul has the longest commute distance, though this may be influenced by the spatial size of the city (Table 11).

When traveling person km is analyzed by metropolitan city to compare mobile data with the KTDB data, it was found that the total trip distance of the mobile data, including short distance pedestrian trips, is more than the KTDB when pedestrian trips are excluded. When trips shorter than 1 km, 2 km, and 3 km are excluded sequentially to remove pedestrian trips, mobile data was found to be similar to the existing KTDB when trips shorter than 2 km and 3 km (Seoul and other metropolitan cities respectively) were excluded (Table 12).

Table 11. Trip volume ratio by trip length (intercity trips only)

City		≤ 2 km	2–5 km	5–10 km	10–30 km	30–50 km	≥ 50 km
Seoul	Trip volume	33,905,402	6,097,998	4,488,672	3,223,268	820	–
	Ratio	71.1%	12.8%	9.4%	6.8%	0.00%	–
Incheon	Trip volume	9,872,522	4,815,446	1,689,102	697,520	17,757	5,686
	Ratio	57.7%	28.2%	9.9%	4.1%	0.1%	0.03%
Busan	Trip volume	11,800,431	3,913,617	2,152,969	1,275,222	10,212	65
	Ratio	61.6%	20.4%	11.2%	6.7%	0.1%	0.0%
Ulsan	Trip volume	4,735,416	3,216,139	1,414,743	773,647	4,031	–
	Ratio	46.7%	31.7%	13.9%	7.6%	0.04%	–
Daegu	Trip volume	6,909,473	2,884,287	1,337,847	579,090	4,414	–
	Ratio	59.0%	24.6%	11.4%	4.9%	0.04%	–
Gwangju	Trip volume	5,136,273	2,145,605	996,029	197,236	–	–
	Ratio	60.6%	25.3%	11.8%	2.3%	–	–
Daejeon	Trip volume	5,011,820	2,607,342	958,753	229,546	–	–
	Ratio	56.9%	29.6%	10.9%	2.6%	–	–

Table 12. Comparisons of person/km by metropolitan city

		Seoul	Incheon	Busan	Daegu	Gwangju	Daejeon
KTDB		120,538,830	26,952,666	39,923,023	28,653,609	16,230,984	13,633,422
Mobile data	Full data generating	152,832,453	56,216,533	68,675,138	40,562,376	25,225,863	27,409,055
	Ratio (mobile/KTDB)	1.3	2.1	1.7	1.4	1.6	2.0
	Excluding trip shorter than 1 km	137,506,580	52,913,264	63,662,337	38,020,135	22,914,566	25,292,308
	Ratio (mobile/KTDB)	1.1	2.0	1.6	1.3	1.4	1.9
	Excluding trip shorter than 2 km	117,188,929	44,981,962	55,084,095	31,998,092	19,220,608	21,164,278
	Ratio (mobile/KTDB)	1.0	1.7	1.4	1.1	1.2	1.6
	Excluding trip shorter than 3 km	108,789,794	35,826,208	48,931,822	27,303,655	15,752,058	17,065,600
	Ratio (mobile/KTDB)	0.9	1.3	1.2	1.0	1.0	1.3

When the trip volume on weekdays and weekends by trip length was analyzed using the mobile data, it was analyzed that the trip volume on weekdays was larger than that on weekends if the trip length was less than 30 km. On the contrary, if the trip length was more than 30 km, the trip volume on weekends was larger than that on weekdays. In particular, it was found that the trip volume of trips with a travel distance of 50 km or more increased by more than 50% on weekends, compared to weekdays (Table 13).

Table 13. Ratio of trip volume by trip distance on weekdays versus weekends

Traffic distance	Weekday traffic volume	Weekend traffic volume	Weekend/weekday rate
Less than 5 km	91,590,746	81,272,646	88.7%
5–10 km	15,887,967	13,406,691	84.4%
10–30 km	12,131,170	10,701,583	88.2%
30–50 km	2,028,600	2,441,881	120.4%
50–70 km	666,270	1,032,428	155.0%
70–100 km	497,341	859,737	172.9%
100–200 km	518,837	955,535	184.2%
More than 200 km	243,324	374,522	153.9%
Total	123,564,255	111,045,022	89.9%

6 Conclusion

The results of this analysis can be summarized as follows: First, it was analyzed that the condition survey on household trips, because of its form as a survey, had significant failures in recording trip history. Second, the large disparity of observed data between the existing KTDB was caused by the diversity of trips collected through mobile data. There is a tendency for the existing KTDB, which is based on household trip surveys, to miss trip history. On the contrary, mobile data records all peoples' trips. As a result, there is a considerable disparity between the two data sets, especially for trips of distances less than 30 km. Third, the coefficient for full data generation that matches the characteristics of the population should be used to make a full set of data out of the samples and by using the mobile data. Fourth, various verifications are needed when using raw mobile data. The national traffic survey applies 25 min of pass-by time, which is classified as a discontinued trip when using mobile data. It is necessary to determine the appropriate pass-by time by analyzing the sensitivity of trip volume changes among different regions such as rural and urban areas, while adjusting the pass-by time every five minutes. It is also necessary to analyze the characteristics of from-home regular and irregular trips, and not-from-home trips, by analyzing daily patterns in the mobile data. For example, various verification processes should be performed regarding which trips will be defined as pedestrian trips when using mobile data. Based on such verifications, a connection with the existing KTDB can be established. Fifth, it appears that mobile big data, much like mobile data and public transportation card data, alone is not enough to establish trip purpose and means for the

O/D matrix. The condition survey on trips plays a supplementary role and is desirable, using the O/D development methodology centered on big data which has been differentiated from existing household trip surveys. Lastly, the reliability issue regarding the process of collecting and aggregating mobile big data should be reviewed. The data loss patterns should be reviewed exhaustively, when analyzing trips using mobile, navigation, and public transportation card data.

Acknowledgments. This research was supported by a grant (#20TLRP-B148659-03: Development of Future Transport Operation Technology based on Big Data & AI) from Transportation & Logistics Research Program (TLRP) funded by the Ministry of Land, Infrastructure and Transport of Korean government.

References

1. Caceres, N., Wideberg, J.P., Benitez, F.G.: Deriving origin destination data from a mobile phone network. IET Intell. Transp. Syst. **1**(1), 15–26 (2007)
2. Alexander, L., Jiang, S., Murga, M., González, M.C.: Origin–destination trips by purpose and time of day inferred from mobile phone data. Transp. Res. Part C Emerg. Technol. **58**, 240–250 (2015)
3. Çolak, S., Alexander, L.P., Alvim, B.G., Mehndiratta, S.R., González, M.C.: Analyzing cell phone location data for urban travel: current methods, limitations, and opportunities. Transp. Res. Rec. J. Transp. Res. Board **2526**, 126–135 (2015)
4. Iqbal, M.S., Choudhury, C.F., Wang, P., González, M.C.: Development of origin–destination matrices using mobile phone call data. Transp. Res. Part C Emerg. Technol. **40**, 63–74 (2014)
5. Maldeniya, D., Lokanathan, S., Kumarage, A.: Origin-destination matrix estimation in Sri Lanka using mobile network big data. In: Proceedings of the 13th International Conference on Social Implications of Computers in Developing Countries, Negombo, Sri Lanka (2015)
6. Wang, H., Calabrese, F., Di Lorenzo, G., Ratti, C.: Transportation mode inference from anonymized and aggregated mobile phone call detail records. In: 2010 13th International IEEE Conference Intelligent Transportation Systems (ITSC), pp. 318–323 (2010)
7. Larijani, A.N., Olteanu-Raimond, A.M., Perret, J., Brédif, M., Ziemlicki, C.: Investigating the mobile phone data to estimate the origin destination flow and analysis; case study: Paris region. Transp. Res. Procedia **6**, 64–78 (2015)
8. Kim, J.Y., Kim, D.H., Sung, H.M., Song, T.J.: A study on the reliability of traffic demand prediction based on big data. The Korea Transport Institute (2018)

Driver's Visual Attention Analysis in Smart Car with FHUD

Yanjun Zhang[1(✉)], Tian Yang[1], Xia Zhang[3], Yongjin Zhang[2], and Youchao Sun[3]

[1] College of Mechanical Engineering, Yangzhou University, Yangzhou 225127, China
zhangyj_204@163.com

[2] School of Mathematics and Physics, Anhui University of Technology, Maanshan 243002, China

[3] College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Abstract. This study explores the impact of cues provided by the full windshield head-up display (FHUD) on the visual attention allocation of the driver under different scenes. The full windshield highlights situational cues related to the driving task. In terms of the perception of situational cues, FHUD changed number of fixations and mean fixation durations of drivers. Number of fixations and mean fixation durations in the condition with FHUD were lower than that in the condition without FHUD. In the terms of percentage dwell times in area of interest (AOI), percentage dwell times in front of the field of vision of the driver in the condition with FHUD was larger than that of the control condition, percentage dwell times on both sides was smaller than that of the control condition. It indicated that FHUD could help drivers' more effectively perceive cues and improve drivers' visual attention allocation.

Keywords: Full Windshield Head-Up Display · Visual attention allocation · Number of fixations · Mean fixation durations · Percentage dwell times

1 Introduction

The task of driving is considered one of the complex perceptual–information–processing tasks requiring processes of perception, identification, selection, and response [1]. Drivers mainly obtain driving information via vision and audition, and it is vision that is used (over 90%) for perceiving information most of the time [2]. When the driver observes head-down display, there will be a contradiction that it is difficult to take into account the driving environment outside the vehicle [3]. HUD plays an auxiliary role in driver's cognitive decision-making [4]. The concept of full-windshield head-up display based on the ordinary vehicle head-up display has emerged. Full windshield head-up display can not only integrate and display system information, but also visually enhance environment information. It has the characteristics of large information capacity, diverse display modes and wide display range [5, 6]. It can assist the driver to obtain more driving information and reduce the frequency of the driver

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 68–73, 2020.
https://doi.org/10.1007/978-3-030-50943-9_9

looking down to observe the instrument in the vehicle [7]. Highlighted cues attract the driver's attention [8] and makes drivers respond more quickly.

This work will discuss whether there are differences in visual attention allocation when drivers see relevant information prompts on sunny, foggy, rainy, snowy and night simulation scenes when FHUD is on, as compared with the results when FHUD is off in the same weather.

2 Method

2.1 Participants

Forty-eight volunteers were recruited to participate in the study (20 women and 28 men, mean age = 29, max age = 36, min age = 24, SD = 3). 48 volunteers were randomly divided into two groups.

2.2 Materials

Vehicle driving visual environment building module is based on MultiGen and Vega Prime software to design different driving scenes (Sunny, rainy, foggy, snowy). Highlighted cues (Road-related clues and driving-related clues) were emphasized with semi-opaques shapes superimposed on the image. Area of Interests (AOI) were defined as Zone 1 to Zone 5 according to the driver's visual field and driving environment (Fig. 1).

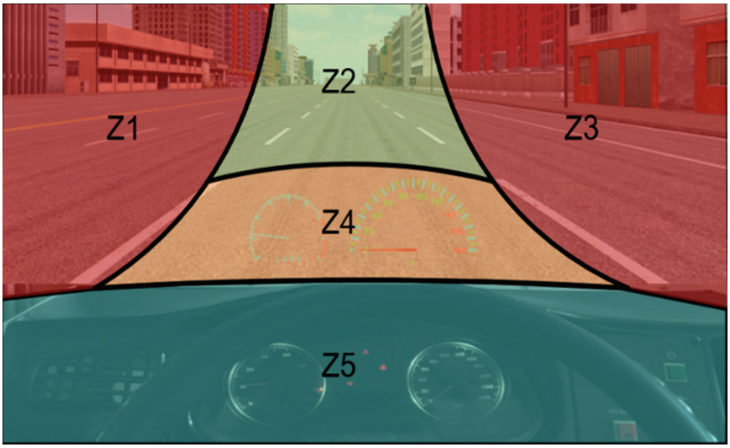


Fig. 1. The allocation of each area of interest

2.3 Experimental Setup

The head-up display interface was developed using GL Studio software which enabled the real-time creation of interactive geometries and driven by the real-time driving state data received from the data processing module of simulator. A 60 Hz SMI ETG 2 W head-mounted eye-tracking system that had a wider tracking range comparing to the desktop-fixed ones was used for recording drivers' eye movements synchronizing with the virtual dynamitic scenes, the accuracy of gaze tracking is 0.5° and the range of gaze tracking is 80° horizontally and 60° vertically. Stimulation was displayed on the curved surface simulation display (4700 mm × 2950 mm), resolution (1280 × 720). Begaze SMI software (Version 3.5) was used for analyzing the data on eye movements. Fixation was calculated with the velocity based algorithm and the minimum duration was set to 100 ms.

2.4 Experimental Design

In this study, we used a 4 (weather) × 2 (FHUD availability) experimental design. The driving routes were geographically the same. Participants had to encounter variable driving conditions due to the interactive nature of the simulator (e.g. some participants had to stop at a red light while others might have encountered a green light at the same point). Symbols including velocity of the vehicle, lane changing, augmented lanes, speed limitation, destination and safety distance keeping were set to display corresponding to the orientation of the vehicle. All participants drove a 10-min practice route different from the test route. After the test drive, drivers were asked to have a 5-min break. Then the participants wore and calibrate the eye tracker and began to perform the driving task in sequence. During the experiment, the data on the eye movements was recorded by eye-tracking system corresponded to the number of each participant

2.5 Dependent Variables

We selected three dependent variables to evaluate the effect of FHUD on visual attention: percentage of dwell time (PDT) on AOIS, number of fixations and mean fixation durations.

3 Result

With the aim to investigate the FHUD impact on the allocation of the visual attention accordingly to type of highlighted cues and type of weather, a 4 × 2 × 4 mixed ANOVA was carried out with one between-subject factor (condition of FHUD) and two within-subject factors (type of highlighted cues and type of weather). The ANOVA was used with the two dependent variables (number of fixations and mean fixation durations).

In order to study the effect of FHUD on the percentage of dwell time on AOIS. A chi square test was performed between the component factors (condition of FHUD) and the test factors (type of weather) by using a 4 × 2 chi square comparison.

3.1 Number of Fixations

In the condition with FHUD, we found that number of fixations on lane line and vehicle distance cues was higher than the other cues (Fig. 2).

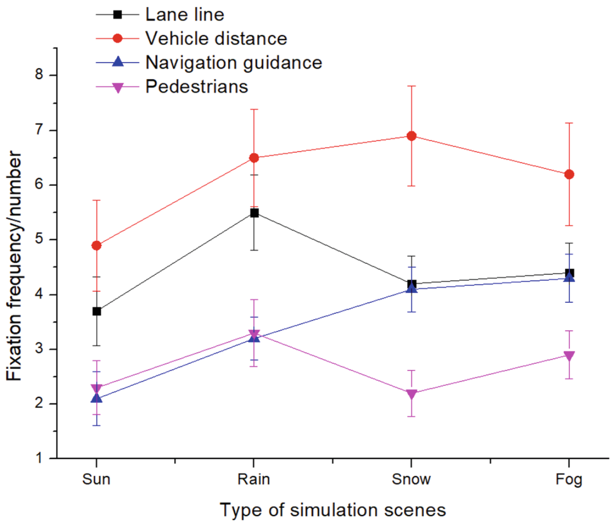


Fig. 2. Number of fixations on four different cues in four simulation scenes

3.2 Mean Fixation Durations

In the condition with FHUD, we also found that mean fixation durations on lane line and vehicle distance cues was higher than the other cues (Fig. 3).

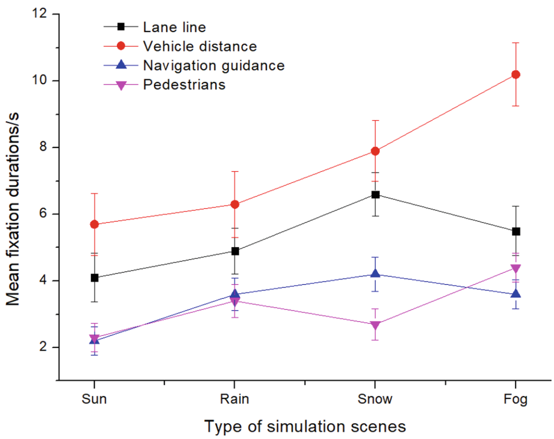


Fig. 3. Mean fixation durations on four different cues in four simulation scenes

3.3 Areas of Interest Analysis

In the four scenes, there were similarities in drivers' visual attention allocation, and the condition with FHUD took up a larger proportion (with FHUD vs without FHUD: 60% vs 72%; 59% vs 72%; 56% vs 68%; 49% vs 71% in the term of Zone 2 and Zone 4). Less percentage dwell time was allocated in the bilateral areas (Zone 1 and Zone 3). The proportion of Zone 1 and Zone 3 was less in the condition with FHUD than in the condition without FHUD. The percentage dwell time in Zone 5 was the least, which was especially obvious in the condition with FHUD, accounting for a small proportion (Fig. 4).

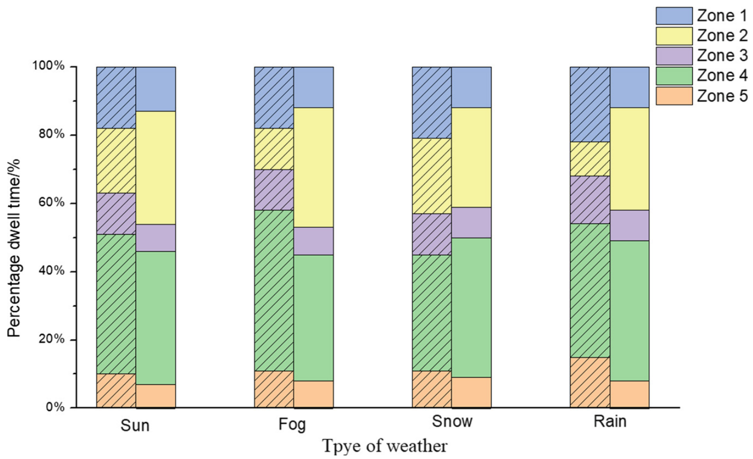


Fig. 4. Percentage dwell time in four scenes. Diagonal fill represents the condition without FHUD, Pure color fill represent the condition with FHUD

4 Discussion and Conclusion

In this paper, we studied effects of full windshield head-up display on visual attention allocation. Our experiments show that attention is a process of further processing of selected information. The shorter the average fixation duration is, the shorter the processing time is [9]. Compared with the control condition, drivers have more effective handling of cues in the condition with FHUD Yung-Ching Liu [10]. The number of fixation and the mean fixation durations of the four cues perceived by drivers in the condition with FHUD are lower than those in the control condition, which indicates that FHUD can help drivers reduce the response time to cues in low visibility weather. This result is similar to that of Vassilis charissis et al. [11], which were based on the experimental model of HUD. When the cues were more specific to driving, compared with the road-related cues, the driver increased the allocation of visual attention and the number of fixation [12], which led to the increase of the fixation durations of driving-related cues. In the terms of percentage dwell times in area of interest (AOI), percentage dwell times in front of the field of vision of the driver in the condition with FHUD was larger than that of the control condition, percentage dwell times on both sides was smaller than that of the control condition. FHUD improves the driver's visual attention allocation in four simulation scenes, which makes the driver's visual attention allocation more focused in the front of the field of vision.

Acknowledgments. The work is supported by the project of National Natural Science Foundation of China (51605424), Six Talent Peaks Project in Jiangsu Province of China (JXQC-018), Natural Science Foundation of Jiangsu Province of China (BK20150455), China Postdoctoral Science Foundation (2016M601898).

References

1. Kim, B.J., Bishu, R.R.: Cognitive abilities in driving: differences between normal and hazardous situations. Ergonomics **47**(10), 1037–1052 (2004)
2. Lansdown, T.C.: Visual allocation and the availability of driver information. In: Gothengater, T., Carbonell, E. (eds.) Traffic and Transport Psychology: Theory and Application, pp. 215–223. Pergamon, Amsterdam (1997)
3. Trivedi, M.M., Cheng, S.Y.: Holistic sensing and active displays for intelligent driver support systems. IEEE Comput. **40**(5), 60–68 (2007)
4. Eyraud, R., Zibetti, E., Baccino, T.: Allocation of visual attention while driving with simulated augmented reality. Transp. Res. Part F **32**, 46–55 (2015)
5. Park, H.S., Park, M.W., Won, K.H., et al.: In-vehicle AR-HUD system to provide driving-safety information. ETRI J. **35**(6), 1038–1047 (2013)
6. Park, M.W., June, S.K.: A projector-based full windshield HUD simulator to evaluate the visualization methods. In: 2014 Sixth International Conference on Ubiquitous and Future Networks, pp. 509–510. IEEE (2014)
7. Charissis, V., Papanastasiou, S.: Human-machine collaboration through vehicle head up display interface. Cogn. Technol. Work **12**(1), 41–50 (2010)
8. Johnson, M.B., Voas, R.B., Lacey, J.H., McKnight, A.S., Lange, J.E.: Living dangerously: driver distraction at high speed. Traffic Inj. Prev. **5**(1), 1–7 (2004)
9. Konstantopoulos, P., Chapman, P., Crundall, D.: Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. Accid. Anal. Prev. **42**(3), 827–834 (2010)
10. Liu, Y.-C., Wen, M.-H.: Comparison of head-up display (HUD) vs head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. Int. J. Hum Comput Stud. **61**, 679–697 (2004)
11. Charissis, V., Papanastasiou, S., Vlachos, G.: Comparative study of prototype automotive HUD vs. HDD: collision avoidance simulation and results. Society of Automotive Engineers World Congress (2008)
12. Triesch, J., Ballard, D., Hayhoe, M., Sullivan, M.: What you see is what you need. J. Vis. **3**, 86–94 (2003)

Human Machine Interaction

How Important is the Plausibility of Test Scenarios Within Usability Studies for AV HMI?

Nadja Schömig[1(✉)], Katharina Wiedemann[1], Frederik Naujoks[2], Sebastian Hergeth[2], Andreas Keinath[2], and Alexandra Neukum[1]

[1] Würzburg Institute for Traffic Sciences (WIVW GmbH), Robert-Bosch-Street 4, 97209 Veitshöchheim, Germany
{schoemig, wiedemann, neukum}@wivw.com

[2] BMW Group, Knorrstr. 147, 80788 Munich, Germany
{frederik.naujoks, sebastian.hergeth, andreas.keinath}@bmw.de

Abstract. We examined the necessity for plausibilization of test scenarios within usability studies for AV HMIs in driving simulator studies. One group of drivers experienced system-initiated transitions without any obvious reason, the other with plausible reasons (e.g. fog for L3 → L2 transition, broken-down vehicle for L3 TOR). The results showed that reaction times to TORs were not influenced by the plausibility while the type of reaction was. Drivers reported less system trust but still knew how to react to the transitions. Non-plausibility did not negatively affect system acceptance. It can be concluded that plausibilization is not necessarily required for all kinds of research questions.

Keywords: Plausibility · Test scenarios · HMI · Automated driving · Methodology

1 Introduction

The evaluation of the understandability of Human-Machine-Interface (HMI) of automated vehicles (AVs) [1] requires a standardized testing procedure and a profound consideration of several methodological issues such as the choice of the relevant evaluation criteria, the studied test sample and the way of instructing the sample (see [2]). Another consideration is the selection and implementation of relevant test scenarios. Especially for system-initiated transitions between automation levels, a major question is whether the driver must receive a plausible explanation for them. For example, must a driver get a reason for a control transition from L3 to a lower level in order to make valid statements whether he/she is able to understand the system and uses it in a meaningful way? The answer to this question can influence the internal and external validity of study results and the inferences that may be drawn from them.

Arguments for giving plausible reasons are that a realistic driving situation and its associated visual and cognitive demands account for a higher external validity of the results. In addition, a comprehensible transition is more likely to lead to the desired and

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 77–84, 2020.
https://doi.org/10.1007/978-3-030-50943-9_10

correct behavior, e.g. an obstacle on the ego-lane as reason for a take-over request (TOR) requires a lane change while the same situation with additional traffic on the neighboring lane requires a braking response. Altogether, it can be expected that the plausibility of transitions increases the traceability of the system behavior which in turn increases system acceptance and usability.

On the other hand, a detailed specification of a test scenario makes the observed behavior less generalizable as the driver shows it in the reaction to exactly this specific situation. It remains unclear whether the driver would have reacted in a similar way if the situation had been slightly different. In addition, it is difficult to differentiate between reactions that are based on the information displayed via the HMI and reactions to the driving scenario. Another problem arises when trying to create "realistic" reasons for system-initiated transitions in a standardized test protocol. They may require non-verifiable assumptions about the system architecture and system limits, which are not valid for every system to be tested (e.g. might fog really require a transition from L3 to L2? Would a system really shut off lateral control when entering a construction site?). This increases the risk of making wrong assumptions about system limits for a specific AV and the need to adapt the plausibility for each test scenario according to the investigated system.

In experimental studies with focus on the behavior of drivers in TOR situations, most often reasons for these TORs are obvious from the design of the scenarios (see e.g. [3, 4]). They are very often based on system limits that are supposed to be considered as realistic for current L3 automated systems (e.g. obstacles, high curvature, work zone, missing lane markings). Only few publications trigger TOR without any obvious reasons. Very often they were used to simulate a "sensor failure" scenario (e.g. [5]), or to create a "false alarm" scenario (see [6]). Sometimes an unexplained event is consciously included within a broad range of other explained TOR-scenarios ([7]).

Studies where this approach is explicitly explained by methodological considerations are very rare. Naujoks et al. [8] use it in real traffic environments in Wizard-of-Oz approaches in order to create a non-critical take-over situation as real TORs are not producible and would be too dangerous. Wintersberger et al. [9] adapted the standardized lane change task and therefore decided to trigger TORs on straight sections without any obvious reasons.

2 Method

In order to verify whether plausible test scenarios are necessary in order to differentiate a "good" from a "poor" HMI design (the goal of a classical usability study) we conducted a driving simulator study. As the relevant factors the quality of the HMI ("poor" = low compliance with design guidelines vs. "good" = high compliance with design guidelines) and the plausibility of the test scenarios (with vs. without plausibilization) were varied as between-subject factors.

2.1 Sample

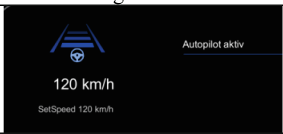
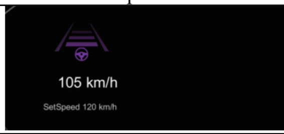
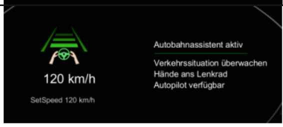
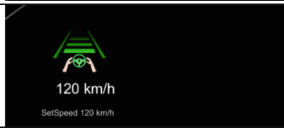
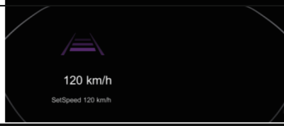
N = 24 subjects participated in the study. Twelve of them were female. They were recruited from the WIVW test driver panel, which ensures that all have experienced an extensive simulator training beforehand. Mean age was 40 years with a standard deviation of 14.8 years. Only subjects were selected who had previously participated in at least one study regarding L2 automation and one regarding L3 automation.

2.2 System and HMI

The study took place in the moving-based driving simulator of the WIVW GmbH (Würzburg Institute for Traffic Sciences). The conditionally automated driving system was able to control the longitudinal and lateral vehicle guidance on highways. It included driving on automation level 2 as well as automation level 3. The system maintained a constant speed of 120 km/h in free driving and adapted speed to limits (100/80 km/h).

Two different HMI variants were designed based on existing guidelines and recommendations that were collected in a checklist developed in [10].

Table 1. HMI variants “good” vs. “poor” in exemplary system states (L3 active, L2 active, take-over request)

modes	“good” HMI	“poor” HMI
L3 active		
L2 active		
Take-over request		

The design of the “good HMI” fulfilled a high number of required guidelines. It contained additional text for the non-standardized mode indicators, speech outputs, clear transition indications via additional acoustic feedback, the indication of the non-availability of the L3 system as well as consistent labelling of operation device and mode indicators.

The design of the “poor HMI” fulfilled only few required guidelines. This HMI used a heavily reduced mode indication, a worse contrast between background and visual indicators and no conventional colors. There was no additional explaining text, no speech outputs, the TOR was only issued by an acoustic warning, there was an unclear assignment between the operation device and the mode indicators, the fonts used for the few text messages were small and bad to read. Table 1 shows examples for the visual HMI variants in various system states. The system was operated via steering wheel buttons.

2.3 Test Scenarios

The driving course consisted of twelve scenarios on a 2-lane highway and lasted about 30 min. The presented results concentrate on the following four system-initiated transitions:

- TOR1: The first one of two TOR scenarios in L3
- L3 → L2: System-initiated transition from Level 3 to Level 2
- L2 System limit: System limit without warning, mere system-shutoff
- End of ODD: End of operational design domain of the system where a TOR is announced 1000 m before reaching the end of ODD

We varied the factor plausibility as a between factor in the four scenarios (the remaining scenarios were constant between the drives):

- *With*: Scenarios contain plausible reasons for transitions
  - TOR in L3: broken-down vehicle
  - System-initiated transition from L3 → L2: fog/bad weather conditions
  - System limit in L2: sharp curve
  - End of ODD: highway exit
- *Without:* the above transitions are triggered without any obvious reason

2.4 Measures

The study used a various set of measures. The paper however reports only those who are necessary for answering the study question. The analyses include reaction times in the TOR scenario, i.e. hands-on times and take-over times. The experimenter observed and recorded type of deactivation in TOR scenarios (using the off-button, gas pedal, brake pedal or steering) directly during the drive. Subjective measures were driver's ratings of the understandability of the system's behavior and the clarity of knowing what to do (directly asked after the respective scenarios). In order to check whether the non-plausibility of scenarios might have an effect on the general system acceptance, we analyzed the ratings on the van-der-Laan Scale.

3 Results

With regard to the traceability of the system's behavior (see Fig. 1) no significant effects of the HMI variant could be found for the four transition scenarios (TOR1: $F = 3.521$; $p = .075$; L3 → L2: $F = 1.738$; $p = .204$; L2 system limit: $F = 0.363$; $p = .554$; End of ODD: $F = 0.215$; $p = .648$). The fact whether a reason for the transition was given had a significant influence in the scenarios L3 → L2 ($F = 4.827$; $p = .041$) and first TOR scenario ($F = 5.070$; $p = .036$) but not in the scenario L2 system limit ($F = 1.498$, $p = .236$) and end of ODD ($F = .1570$; $p = .227$).

When drivers were asked about how clearly they knew what they had to do in the four analyzed transition scenarios (see Fig. 2) there was a clearly significant HMI effect in the scenarios TOR1 ($F = 17.658$; $p < .000$) and L3 → L2 ($F = 13.554$; $p = <.001$). In the scenarios L2 system limit and end of ODD the HMI variant had no significant influence (L2 system limit: $F = 3.160$, $p = .091$; end of ODD: $F = 1.688$; $p = .200$). The plausibility variation had no significant effect on that knowledge in all four scenarios (First TOR $F = .000$; $p = 1.000$; L3 → L2: $F = .542$; $p = .470$; L2 system limit: $F = .140$, $p = .712$; end of ODD: $F = .068$; $p = .798$).

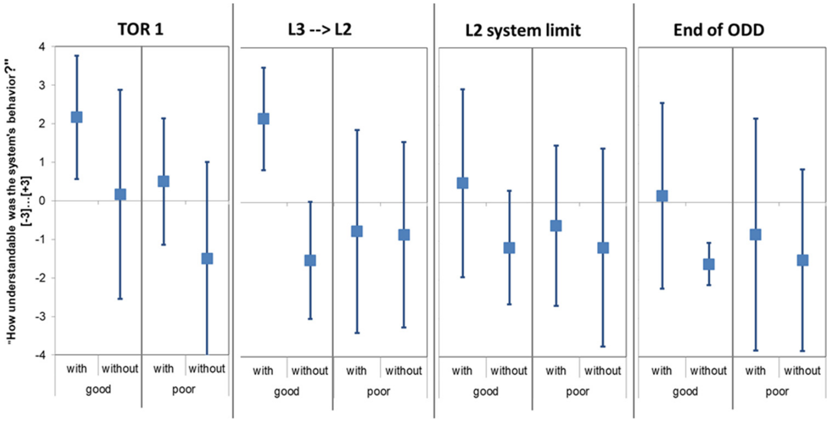


Fig. 1. Drivers' subjective rating of the possibility to trace the system's behavior in the respective scenarios, dependent from the HMI variant and the presence of a reason for the transition (with vs. without). Graph show means and standard deviations.

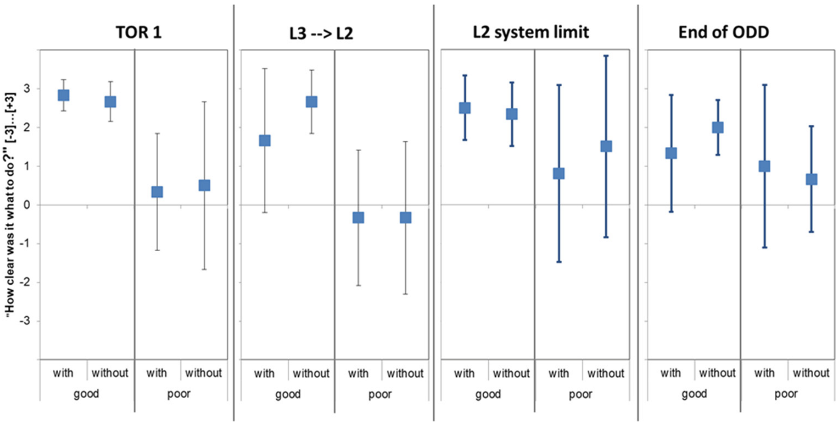


Fig. 2. Drivers' subjective rating of the clarity how to behave in the respective scenarios, dependent from the HMI variant and the presence of a reason for the transition (with vs. without). Graph show means and standard deviations.

Mean hands-on time (see Fig. 3 left) in the first TOR scenario was significantly shorter in the good HMI variant compared to the poor HMI condition (F = 12.557, p = .002). Plausibility had no effect (F = 1.921; p = .181). Mean take-over times did not differ significantly dependent from the HMI variant as well as from the plausibility condition (HMI variant: F = .238; p = .631; plausibility: F = .496; p = .490; see Fig. 3 right).

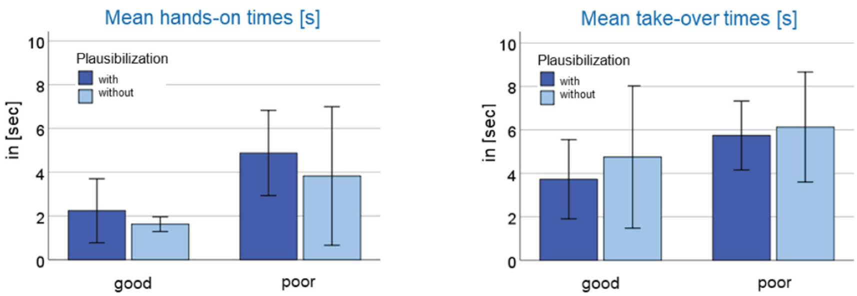


Fig. 3. Mean hands-on times (left) and mean-take-over times (right) in first TOR-scenario dependent from the HMI variant and the presence of a reason for the transition (with vs. without). Graph shows means and standard deviations.

The type of deactivation differed dependent from whether the TOR was plausible or not. If the reason was an obstacle on the road, more drivers used the brake pedal instead of the gas pedal. If no reason was obvious in the environment, a higher percentage of drivers used the gas pedal in order to maintain a constant speed. This effect seems to be independent from the HMI variant (see Fig. 4).

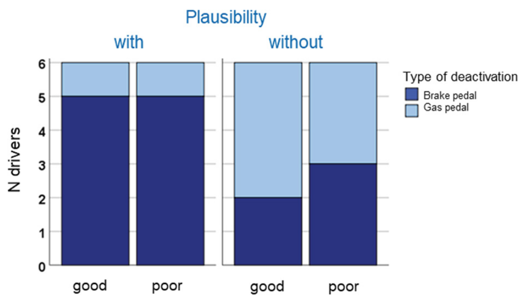


Fig. 4. Number of drivers who deactivated the system either by pressing the brake pedal or the gas pedal dependent from the HMI variant and the presence of a reason for the transition (with vs. without).

In the two subscales (usefulness and satisfying) of the van-der-Laan scale only HMI quality did show a significant influence (usefulness: F = 8.096; p = .009; satisfying: F = 10.682: p = .004). Plausibility had no significant effect (usefulness: F = .011; p = .019; satisfying: F = .006; p = .937).

4 Summary and Conclusion

The results showed that some drivers commented to feel somewhat insecure due to the absence of reasons of system-initiated transitions. The drivers also stated that receiving plausible reasons increases understandability of system's behavior. When analyzing TOR behavior, the results indicate that drivers differed in the type of reaction (i.e. accelerating in case of no visible reason for a TOR in order to maintain constant speed vs. braking in case of a broken-down vehicle as a reason for take-over in order to avoid a collision with that obstacle). On the other hand, it could be shown that take-over time was not negatively affected (i.e. no increase in hands-time and reaction time) by the absence of plausibility. In addition, non-plausibilization did not negatively affect the general system acceptance. The drivers stated that they know that they have to react in a certain way indicating that the HMI was sufficient to transfer the need to take over control from the automation. Furthermore, it got obvious that some ways of plausibilization did not reach the desired effect. Some causes for system limits that we tried to make them plausible (e.g. sharp curve for L2 systems, exiting the highway for L3 system) were not that self-explanatory for some of the drivers.

Based on the theoretical discussions and the presented results of the conducted simulator study it seems that – at least for the question at hand - it is not necessary to design test scenarios in a way that reasons for certain system behavior are given. In the worst case, it might even be counterproductive in case reasons are not perceived or understood correctly resulting in uncontrollable influences on study results. It seems sufficient to verify whether the HMI is able to transfer the necessary information to the drivers to enable them to adequately react to a change in the automation mode within a reasonable time and with a reasonable behavior.

References

1. NHTSA: Federal Automated Vehicles Policy, U.S. Department of Transportation NHTSA (updated version). Voluntary Guidance for Automated Driving Systems, U.S. Department of Transportation (2017)
2. Naujoks, F., Hergeth, S., Wiedemann, K., Schömig, N., Forster, Y., Keinath, A.: Test procedure for evaluating the human–machine interface of vehicles with automated driving systems. Traffic Inj. Prev. **20**(1), 146–151 (2019)
3. Radlmayr, J., Fischer, F.M., Bengler, K.: The influence of non-driving related tasks on driver availability in the context of conditionally automated driving. Chair of Ergonomics, Technical University of Munich (2019)
4. Hergeth, S., Lorenz, L., Krems, J.F.: Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. Hum. Factors **59**(3), 457–470 (2017)
5. Richardson, N., Flohr, L., Michel, B.: Takeover requests in highly automated truck driving: how do the amount and type of additional information influence the driver-automation interaction? Multimodal Technol. Interact. **2**(4), 68 (2018)
6. Morando, A., Victor, T., Bengler, K., Dozza, M.: Users' response to critical situations in automated driving: rear-ends, sideswipes, and false warnings (2019)

7. Molnar, L.J., Ryan, L.H., Pradhan, A.K., Eby, D.W., Louis, R.M., Zakrajsek, J.S.: Understanding trust and acceptance of automated vehicles: an exploratory simulator study of transfer of control between automated and manual driving. TRF **58**, 319–328 (2018)
8. Naujoks, F., Purucker, C., Wiedemann, K., Marberger, C.: Noncritical state transitions during conditionally automated driving on German freeways: effects of non–driving related tasks on takeover time and takeover quality. Hum. Factors **61**(4), 596–613 (2019)
9. Wintersberger, P., Schartmüller, C., Riener, A.: Attentive user interfaces to improve multitasking and take-over performance in automated driving: the auto-net of things. Int. J. Mob. Hum. Comput. Interact. (IJMHCI) **11**(3), 40–58 (2019)
10. Naujoks, F., Hergeth, S., Keinath, A., Wiedemann, K., Schömig, N.: Development and application of an expert based assessment for evaluating the usability of SAE Level 3 ADS HMIs. In: ESV Conference 2019, Eindhoven, Netherlands (2019)

Employing Natural Finger Positioning Strategy for Improving Blind-Positioning of Steering Wheel Mounted Switches

T. K. Philip Hwang[1(✉)], Yao-Tin Huang[2], and Pin-Chieh Kuo[2]

[1] Da-Yeh University, Changhua, Taiwan
phwang@mail.dyu.edu.tw
[2] National Taipei University of Technology, Taipei, Taiwan
i0936223817@gmail.com, kpj0819@gmail.com

Abstract. The purpose of this study was to avoid off-road glances of drivers while operating steering wheel mounted switches. However, a large number of switches on multi-functional steering wheel give rise to difficulties in memorability and finger-reaching and require hand-eye coordination. The ideation and prototyping process is as follows. Firstly, we integrated three directional entry sets into one, accompany with a function mode switching button for circulating selections of "radio control mode", "cruise control mode" and "digital dashboard display control mode". Secondly, we employed natural finger positioning (NFP) strategy and introduced multi-finger pull gesture NFP device for more effective blind-positioning movement of controlling directional entry keysets. Thirdly, to avoid function mode confusion, the "voice prompt upon key operation" is provided and proved to be much effective in the feasibility test. It allowed subjects to acknowledge the status of function mode timely and reduce operation errors.

Keywords: Natural Finger Positioning · Multi-functional Steering Wheel

1 Introduction

Over the past decades, more and more functions have been migrating from their traditional locations on the console or dashboard onto the steering wheel. Radio controls, cruise control operation and several others have been integrated into steering wheel technology. The manufacturers advocated that Steering Wheel Remote Controller greatly improve the convenience and safety of drivers.

A driver's eyes and hands are extremely important tools for driving. It is true that keeping a driver's eyes on the road and hands on the Steering wheel can increase auto safety. However, we would argue that does Steering Wheel Remote Controller really keep a driver's eyes on the road? Especially when dealing with so many functions and key buttons.

Our exploratory research remarked that users who used Steering Wheel Remote Control constantly employed hand-eye coordination and glanced over the Steering Wheel Remote Controller during switch operation.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 85–91, 2020.
https://doi.org/10.1007/978-3-030-50943-9_11

2 Problem Discovery

Good control display improves usability, allowing users to quickly access to the control devices they are looking for. Avoid presenting too many selections at one time. However, in our survey, the existing steering wheel mounted switches are up to 18 buttons.

As a part of problem discovery which is an important step in the creative process, we conducted an observation and employed the theory of switch operation for identifying explicit and implicit problems during the process of operating steering wheel mounted switches.

The existing steering wheel mounted switches (18 buttons) are divided into 4 button sets base on diversity of functions including “radio control”, “cruise control”, “digital dashboard display control” and “voice command”. Except for voice command function, the rest are provided with directional entry for adjustment of respective function (Fig. 1).

Exploratory Research. An exploratory research was carried to observe the usage pattern and problems in the process of operating existing steering wheel mounted switches. Five subjects respectively took four experimental tasks, each of which contains 9 function chores, including volume control, radio seeking, cruise control, cruise speed, voice command, telephone receiver operating and dashboard selection. During the test, subjects were required to set their eyes on the scenario screen and remark at the road conditions frequently while doing his/her experimental tasks.

Each experimental task contains various function chores, including radio seeking, voice command, cruise control, cruise speed, volume control, telephone receiver operating and dashboard selection. control button glance rate. operation time and error rate were recorded. The results are as follows.

Control Button Glance Rate was high, 3 subjects (subject A, B and C) of novices frequently glanced over the control buttons in 81% of the function chores tests, while 2 subjects (subject D and E) of experienced users glanced over the control buttons in 46% of the function chores tests.

The control button glances rate (N = 5) insignificantly decreased following the experiment task 1 through task 3 (Fig. 2). We concluded that drivers constantly employ hand-eye coordination during switch operation. The effect of learnability on multi-function steering wheel controller did not apply to the subjects, although some report noted that the users will learn quickly and gain a fast sense of mastery.

Operation time of four different experimental tasks that testing by 5 subjects respectively. The result remarked the learning effect of experimental task 1 through task 3 was insignificant. However, In the experimental task 4, the operation time was longer than previous tasks in most cases, owing to that the steering wheel mounted switches were covered in task 4 for operation without hand-eye coordination. It remarked the importance of hand-eye coordination in operating existing steering wheel mounted switches.

The evaluation indicated that although the total number of control keys was significantly decreased, the operation of function selection keys still demands hand-eye coordination. Meanwhile, the location of directional entry keyset needs to be carefully adjusted to meet thumb reach area.

3 Exploring Possibilities

Natural Finger Positioning Device. In violin playing the pivoting thumb is on the opposite side of the violin neck from the fingers. A pivot finger can assist complex movements in the same position using multiple fingers. This concept of natural finger positioning has been broadly applied to the design of sophisticate finger operation situations like the design of mouse, joysticks and flute.

The finger placement skill in natural finger positioning emphasized that thumb or palm pivoting allowing the steering wheel to pivot across the thumb providing accurate positioning operation of the rest of the fingers.

Space Exploration Around Steering Wheel. The design space of a steering wheel can be divided into three areas: the front of the steering wheel, the back of the steering wheel, and the steering wheel rim. At the front side of the steering wheel various interaction elements can be mounted including buttons, thumbwheels, and touchpads. The rim of the steering wheel can be used to provide haptic feedback and may allow tab or swipe gestures. The back of the steering wheel often contains pull elements such as shift paddles [1].

Interaction with the back of the steering wheel is usually done with fingers located behind the steering wheel. As Fig. 2 illustrate, multi-finger pull gesture (a) seem to be promising especially combined with chorded keys; touch gestures (b) may be performed in the xy-plane but only in the area reachable by fingers (ibid.).

In contrast to the front of the steering wheel, interaction with the back of the steering wheel is done with the fingers (usually 3 fingers) located behind the steering wheel and supports bi-manual and multi-finger operation (ibid.) (Figure 2).

Fig. 1. Control button sets of various functions on existing multi-function steering wheel.

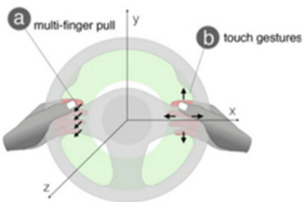


Fig. 2. Application of "Multi-Finger Pull" and "Touch Gestures" of natural finger positioning. Source: Meschtscherjakov 2014.

4 Idea Development of Prototyping

The concept of natural finger positioning (NFP) was introduced to fulfill the goal of blind-positioning operation of the driver. A multi-finger pull gesture NFP device, that is commonly used in car radio remote control, was selected as a substitute for existing directional entry keys.

4.1 Multi-finger Pull Gesture NFP Device of Directional Entry Keys

The device of multi-finger pull gesture NFP directional entry sets (Fig. 3) at the back of the steering wheel was accompany with a function mode switching button. The function mode switch at the front of the steering wheel simply provides users with a push-button switch for circulating selections of function mode including: “radio control mode”, “cruise control mode” and “digital dashboard display control mode”. The design of push button switch for circulating selections of function mode followed touch gestures of the thumb (Fig. 4(b)).

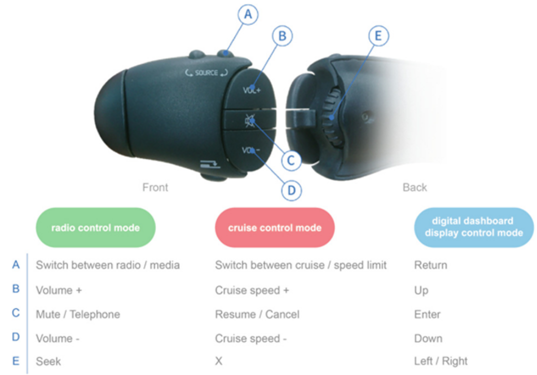


Fig. 3. The multi-finger pull gesture NFP device (installed at the back of the steering wheel).

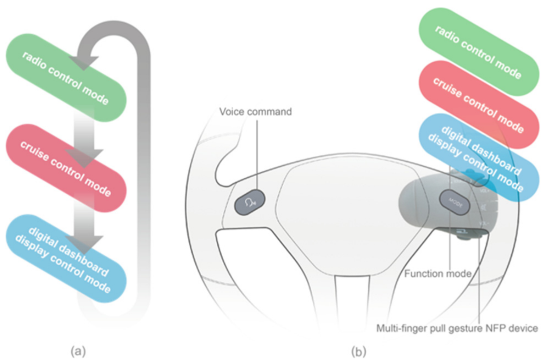


Fig. 4. (a) The process of circulating function mode (b) Push-button switch for circulating selections of three function modes of Prototype.

4.2 Feasibility Test

Five subjects respectively took four experimental tasks, each of which contains 9 function chores, including volume control, radio seeking, cruise control, cruise speed, voice command, telephone receiver operating and dashboard selection. Nine function chores in Task 1 were set to be successive with no more than 10 s pauses, while function chores in Task 2 through Task 4 were set to be intermittent with 60 s pauses. The application of "Return to Default Function Mode by Timing" were initiated with voice message if the function chores in the task were not successively operated within 30 s. In the experimental task 4, the steering wheel mounted switches were blindfolded for avoiding hand-eye coordination during function chores operation.

During the test, subjects were required to set their eyes on the scenario screen and remark at the road conditions from time to time while doing his/her experimental tasks. Data were recorded including operation time, error rate and control button glance rate.

5 Results

In the Exploratory Research of existing steering wheel mounted switches, the Control Button Glance Rate was 81% of the 27 function chores tests in novice subjects, 46% of the 18 function chores tests in experienced users. Comparatively, by employing Natural Finger Positioning (NFP) Device of Directional Entry Set, the Control Button Glance Rate decreased to 7% in novice subjects, and surprisingly, none in experienced users. The result suggested that the multi-finger pull gesture NFP Device with 3 fingers located behind the steering wheel, could enable the drivers to operate steering wheel mounted switches without off-road glances when driving. On the other hand, task operation time remained stable even with the switches blindfolded in task 4.

Errors happen when you know what you are doing - that is while carrying out a routine task in familiar circumstances - but the actions don't go as planned. These errors take the form of systematic 'absent-minded' action slips and memory lapses. They can also appear as trips, fumbles and stumbles [2].

The total number of fumbles in prototype of multi-finger pull gesture NFP device tasks was zero while in the tasks of existing multi-function steering wheel the subjects frequently fumbled, totally 29 times, with keys for target positioning.

The subjects lapsed 8 times totally in the tasks of prototype NFP device.

Compare to 12 times in existing multi-function steering wheel, we assumed that memory of key position was improved. However, we noticed that three mistakes derived from function mode switching were observed. The derivative problem was defined as mode confusion which has to be improved (Table 1).

Table 1. The number of errors in existing multi-function steering wheel and in prototype

Number of errors	Slip	Lapzse	Fumble	Mistake
Existing multi-function steering wheel	1	12	29	2
Prototype of NFP device	2	8	0	5

Mode Confusion. The process of cycling function mode was set while standpoint of function mode may confuse the user in the intermittent operation. There are alternatives for function mode standpoint control: (1) remain in place (2) return to default function mode by a 30 s timer. Unfortunately, our user test revealed that function mode confusion happened from time to time in both settings of standpoint control. Working memory is limited resource of human mind. It is not feasible demanding a driver to recall the standpoint of function mode while executing driving tasks.

6 Employing Voice Prompt on Intermittent Key Operation

Effective interfaces instill a sense of control in their users. Keeping users in control makes them comfortable; they will learn quickly and gain a fast sense of mastery [3]. Voice prompt reminds the user of the existing state of function mode when he/she needs, in this case, when he/she touch any NFP key button at the intermittence operation occasion. Voice prompt indicates the current state of the task (i.e. function mode). The participants (drivers) were aware of the state of the function mode without having to remove the visual focus from their driving task.

6.1 Visual and Auditory Prompt for Car Driver

One of the Jakob Nielsen's 10 usability heuristics advises promoting recognition over recall in user interface design. Recognizing something is much easier than recalling it because recognition involves more cues in our brain (cues spread activation to related information in memory, and those cues help us remember information) [4]. In prototype of our design, the current function mode is clearly displayed in the digital dashboard. However, in driving context, checking the current function mode by setting eyes on digital dashboard can never be a proper resolution. Alternatively, the effect of timely message is to provide user with recognition of current function mode so that neither recall nor visual recognition is needed.

The addition of auditory prompt of current situation allowed the drivers to confirm the current function mode on keyset operation while driving.

6.2 Feasibility Test of Voice Prompt on Intermittent Key Operation

While the interface design of "Return to Default Function Mode by Timing" consequently lead to mode confusion in the intermittence operation of steering wheel control, it had to be improved. An alternative design of auditory prompt at the opportune moment was raised. A timely auditory prompt of current situation was given at the intermittence operation occasion.

Following previous experiment settings, the application of "Voice Prompt on intermittent Key Operation" were initiated if the function chores in the tasks were not successively operated within 30 s. In other words, Voice Prompts of current function mode were given whilst the subjects intermittently operated any steering wheel mounted switches for avoiding hand-eye coordination during operation.

6.3 Results

The operation time increased in Task 2 through Task 4 that provided the users with Voice Prompts of current function mode in the intermittence operation occasion. Nevertheless, the average subjective measures of satisfaction remain high.

Neither error operation nor control button glance were observed. The implication of the experiment findings is that voice prompt initiated by intermittent key operation is feasible to remind the drivers of current function mode and effectively eliminate function mode confusion in the context of intermittent control.

If a situation (function mode) is to be reminded just before keyset operation while driving, vocal prompt is applicable.

7 Conclusions

Natural Finger Positioning (NFP) strategy effectively enhanced blind-positioning operation of the driver and significantly avoided drivers' off-road glances while operating steering wheel mounted switches.

To integrate three directional entry sets into one single set of controls did reduce control key number, but may cause function mode confusion. The idea of "voice prompt upon key operation" could effectively release working memory load but relatively less efficient, that should be tolerable comparing to the demanding memory effort until next operation.

The results suggest that the integrated navigation key-set with function mode voice prompt is feasible in conducting steering wheel remote control and could outperform the prevailing Multi-functional Steering Wheel system for achieving a safer and more convenient way of blind positioning key operation while driving, and finally, enhance the driving safety.

References

1. Meschtscherjakov, A., Wilfinger, D., Murer, M., Osswald, S., Tscheligi, M.: Hands-on-the-wheel: Exploring the design space on the back side of a steering wheel. European Conference on Ambient Intelligence, pp. 299–314. Springer, Cham (2014)
2. Reason, James: The Nature and Varieties of Human Error, in The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries, vol. 3. Ashgate, Burlington (2008)
3. Babich, N.: The 4 golden rules of UI design (2018). https://xd.adobe.com/ideas/process/ui-design/4-golden-rules-ui-design/
4. Nielsen, J.: 10 Usability Heuristics for User Interface Design (1994). https://www.nngroup.com/articles/ten-usability-heuristics/

Towards a Truly Cooperative Guidance and Control: Generic Architecture for Intuitive Human-Machine Cooperation

Marcel Usai[1(✉)], Ronald Meyer[1], Hiroshi Nagahara[2], Yusaku Takeda[2], and Frank Flemisch[1,3]

[1] IAW of RWTH Aachen University, Bergdriesch 27, 52062 Aachen, Germany
{m.usai, r.meyer, f.flemisch}@iaw.rwth-aachen.de
[2] Mazda Motor Corporation, Technical Research Center, 3-1, Shinchi, Fuchu-cho, Aki-gun, Hiroshima 730-8670, Japan
{nagahara.hi, takeda.yus}@mazda.co.jp
[3] Fraunhofer Institute for Communication, Information, Processing and Ergonomics (FKIE), 53343 Wachtberg, Germany
frank.flemisch@fkie.fraunhofer.de

Abstract. Human-machine cooperation (HMC) is often still rigid and unintuitive. However, with more ability transferred to machines, the need for intuitive cooperation rises. To achieve this, new concepts need to arise and be implemented for machines to get a better understanding of their cooperation partner and to be able to act as expected. This includes adapted cooperation schemes based on actual dimension of control, e.g. conscious or subconscious HMC. In this paper, we give an overview on a generic architecture designed to achieve intuitive HMC and introduction to an example application.

Keywords: Human-machine cooperation · Shared control · Intuitive control · Cooperative guidance

1 Introduction

Cooperation seems to be a key tool to the evolution of humankind [1], especially in interplay with its counterpart, competition or its extreme form, war [2].

In today's world, we develop ever-powerful tools to which we hand over not only physical tasks, but also cognitive tasks or functions that they complete autonomously. However, interaction with machines is often limited or follows strict, unintuitive rules. Moreover, looking only at interaction leaves out important aspects of the game, e.g. that not only outer compatibility between human and machines is important, but also a minimum amount of inner compatibility [3]. To overcome strict interaction, as a next step, these tools should be able to truly cooperate with the human and with each other.

To create a successful human-machine cooperation (HMC) accessible for everyone, it is obvious to relate HMC to existing cooperation schemes between humans [4]. Beyond human-human cooperation, human-animal cooperation like the relationship between rider and horse can also be used as a blue-print for the cooperation between a human and a non-human agent [5]. The metaphor of rider and horse or coachman and

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 92–98, 2020.
https://doi.org/10.1007/978-3-030-50943-9_12

carriage has been used by car manufacturers since many years, e.g. Mazda (1987ff), Volvo (2001ff), Daimler (2016ff). In addition to increased availability, implementing easily understandable schemes may help prevent a range of negative effects, i.a. unsuccessful interaction or overtrust in machine's abilities resulting from an uncanny valley of automation [6].

1.1 Human-Machine Cooperation (Framework)

Former research on mental models and human machine interaction indicated that a human's decent understanding of the system's actions leads to improved interaction [7].

This paradigm can also be used vice versa: A proper model of the human's actions on the machine's level can help to predict the human's actions and involvement. Hence, the machine's ability to adapt to human's behavior can be increased.

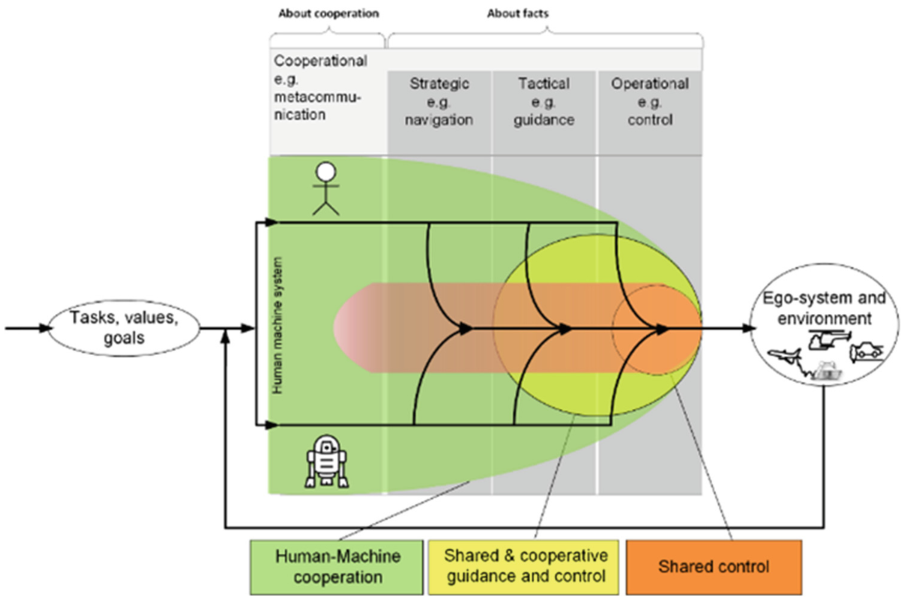


Fig. 1. Framework of human-machine cooperation proposed by Flemisch et al. [2]. Cooperation happens on three levels of cooperation and a transversal cooperational layer. Human and machine merge their inputs.

Complex models of human behavior can be represented in an abstract way by framing them into abstract image schemas [8]. To be able to implement such interaction schemes, the machine's architecture must be capable of enabling these schemes. Historically, architectures of human-machine cooperation involve discrete states of control or authority distribution [9]. In recent work the idea of isolated control states has more and more been extended to transitions between states to manage different isolated situations, e.g. [10, 11].

Flemisch et al. describe an approach to a four layer model of shared and cooperative control [2] and proposed a general layered approach of a system design on how to achieve a cooperative guidance and shared control between human and machine considering multiple layers of cooperation dimensions which are the strategic, tactical and operational layer (see Fig. 1) [2]. These three layers are subordinated to a cooperational layer which enables meta-communication which can e.g. include communication about cooperation and is transversal to the three other layers of cooperation.

In addition to the layers of cooperation proposed by [2], the horizontal layers of cooperation (HLC), we propose that HMC arises in additional vertical layers of cooperation (VLC), i.e. in different dimensions. They reflect different forms of cooperation characterized by diversion in tasks (e.g. movement and shovel control of an excavator) or by diversion of effect in a single task (e.g. conscious and subconscious cooperation). As the same human and machine are cooperating, this dimension may not be modelled as new, unconnected frameworks, because the layers influence and may depend on each other.

2 System Architecture for Adaption Based on Human Involvement

The proposed general architecture describes cooperation between multiple entities, e.g. between human and a machine. Both share a task or goal, but take different roles while cooperating. All entities produce their own normalized actions (response), which are then blended into a single joint action that is imposed onto the environment or processed on the next-deeper HLC. There is a single distribution coordination role, which is done by a human-machine mediator. This mediator facilitates the distribution of control for each HLC between itself and other entities based on its own models and observations of environment and other entities. For each VLC the same mechanism applies, potentially changing the role of mediator.

This leaves the mediator with ability, and with that responsibility to facilitate the distribution influence on the environment. In case the mediator is not human, we advise adding a possibility to the conscious layer for the human to immediately take over full control on request even when the system is malfunctioning to prevent failures in the human-machine interaction relationship [4] that would occur per system design.

This conceptual frame applied to the case of intuitive human-machine cooperation results in a system of two entities of which the machine is considered a mediator and supporting human as a supporting entity. The mediator can be equipped with an involvement and intention prediction (IIP) unit on each HLC to fill gaps in the control distribution while human is not involved on a given HLC. I.e. human controls the control distribution with its own involvement, which eventually provides intuitive cooperation from at least the human perspective. The HMC framework for this case is depicted in Fig. 2. Not shown is the conscious part of this framework, which includes feedback paths from mediator to human about machine state and desires as well as human observing machine behavior. In this case, human and machine cooperate on all three HLC and possibly multiple VLCs. To mediate the correct control distribution, the IIP unit of the mediator on a given layer needs information on the observed human action and intention of its own layer and the next lower layer, the own intention of given and next layer and joint action. Given this information, involvement and intended involvement of human is estimated. When the mediator has the authority to take over control in certain situations, it may do so based on its understanding of the situation given by internal models on its layer of cooperation. These models are fed the current perceived environment and directly or derived observed values provided by IIP unit. To achieve a less hierarchical setup, the machine may be allowed to voice its suggestions or even debate with other entity to reach a more optimized influence on the environment.

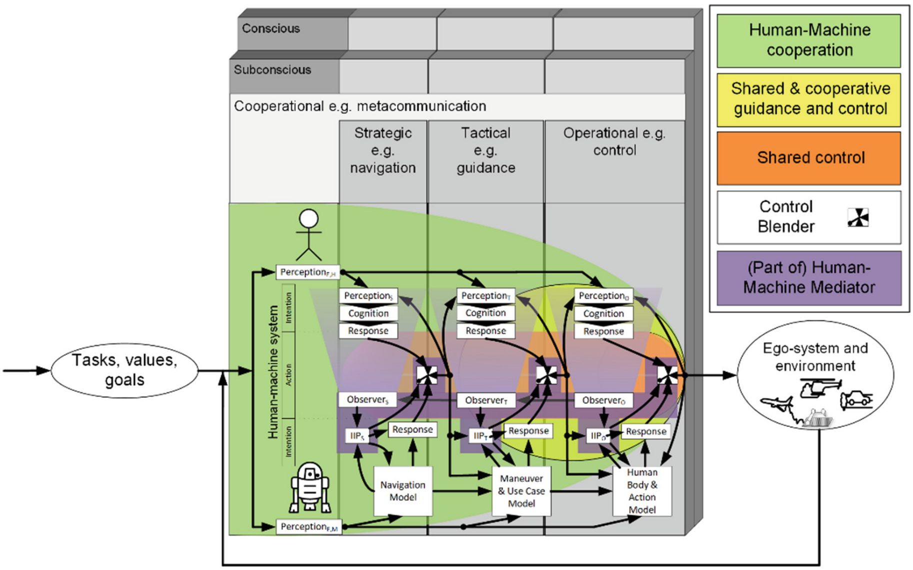


Fig. 2. Proposed general concept of intuitive human-machine interaction in subconscious dimension. Conscious dimension is not shown but coupled to subconscious dimension. Fundamental to intuitive cooperation is the observation and involvement prediction of the human by the machine. The mediator arbitrates between human and machine. (Possible direct observation of the machine's response by the human is not shown here.) Based on [2] and [11].

2.1 Intuitive Control Distribution

Distribution of control is very dependent on the HLC. On the operational layer, a separate blended shared controller (see [10] for example application) for longitudinal and lateral control may manage the distribution of control. In this case, the blending factor is continuously varied by the mediator, dependent also on the input of the human.

Sharing control on the tactical and strategical layers is not as straightforward.

On the tactical layer, decision of driving style and maneuvers is made. These decisions may not be a simple mix of the inputs of human and machine. Depending on the situation this output may not be part of the set of all inputs or a variant of at least one input. In the case of the tactical layer, the mediator has to contain an arbitration unit, which does not decide on rules based on a fixed automation level design as in [11] but favors a single output based on the inputs of all involved entities based on their individual involvement determined by IIP unit. This results in variable rules based on a fluid automation level design, as already hinted by [5] and discussed by [9] as Fluidity of automation. Based on the remaining time to execute a given input safely, the arbitration unit in the mediator gives a time frame for all entities to still negotiate the final decision by either changing their input or involvement on the tactical layer.

On the strategical layer, goals and sub goals of the navigation are decided. These decisions are less time-critical as decisions on the lower layers. The constant involvement of an entity is not of importance. On this layer, the human or entity with highest involvement while goal decision takes place decides goals and possible sub goals. These goals will leave a set of other goals still possible to reach. In this so-called null space, all other entities may set additional goals. E.g., the vehicle decides to fit in an extra stop while still driving the human to an appointment on time.

3 Intuitive Assistance in Operational Layer for Movement Control

Figure 3 depicts a human-machine cooperation concept proposed to provide intuitive assistance, which has a conscious and subconscious dimension. Our assistance changes relationship through two dimensions depending on two different purposes as follows:

1. The conscious dimension affects consciousness by a sense of incongruity, when it makes the human be conscious of an aim of assist.
2. The subconscious dimension affects sub consciousness without a sense of incongruity, when it supports the human's operations as if it was the human's own shadow supporting.

The assist of human's operation in the subconscious dimension (in Purpose 2) is more difficult than Purpose 1, as its assist needs to operate the human movements directly without a sense of incongruity. The assist by direct force will lead to a sense of incongruity while subconscious motion, if it is not provided in short time before the stretch reflex of muscles [12]. Therefore, our study proposes the overview of assist without the sense of incongruity. Its overview is that our assistance adjusts the environmental field in operational layer, when there is the difference between the result of actual motion and the result of prediction based on combined motion tactics and observed human intention.

The traditional researches of human arm movements reported that human can learn the ideal trajectory of a task in sub consciousness without a direct input of ideal trajectory, by using a motion environmental field from which human can derive his motion tactics easily [13, 14]. These reports suggested that mechanical impedance of environmental field (e.g. viscoelasticity) can be influenced to modify a human motion trajectory. Furthermore, Shadmehr et al. report that the environmental field suitable for the human internal model inspires the ideal movement [15].

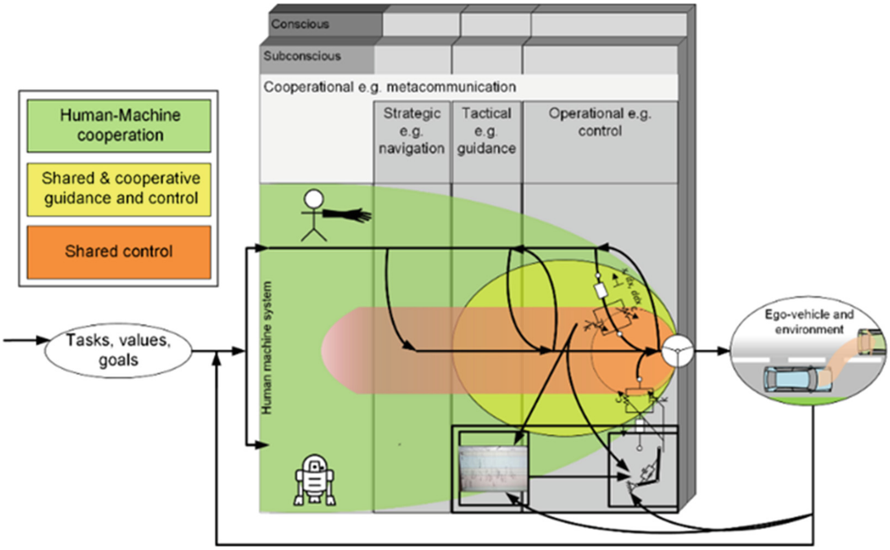


Fig. 3. The proposed relationship between shared and cooperative guidance and control, and human-machine cooperation while the intuitive assistance. Based on the framework from [2].

Based on the above traditional researches, our proposed assistance affects the human sub consciousness action by the operational layer's assist, which regulates the mechanical impedance of an environmental field. Additionally, an output from the tactical layer to operational layer (machine side) is the ideal trajectory corresponding to the task. Therefore, it is necessary to clarify how to decide the motion tactics in each environment, because the motion tactics depend on the personal internal model. To predict the motion tactics in a way to serve Purpose 2, the progress of researches of the human motion trajectory [16] and the ideal vehicle dynamic trajectory [17] are very important to define an output from the tactical layer to operational layer. For example, many researches about human arm movement report that the ideal trajectory in the complex voluntary movement is expressed by the output of the feedforward motor command from the central nervous system [15]. In addition to trajectories, the concept of maneuvers and maneuver models (e.g. [8]) could contribute to a better balance between conscious and subconscious cooperation of human and machines.

4 Conclusion

In this work a framework of human-machine cooperation for explicitly modelling conscious and subconscious behavior has been created. Therefore, an existing model providing three layers of cooperation was extended by conscious and a subconscious layers as a supplementary dimension. Hence, human-machine cooperation is levered into a new multi-dimensional space. It places a fundament to create truly intuitive human-machine systems. The next step is to implement each HLC in a way that it systematically resolves any incongruity between human and machine. The key components to be designed are mediator, arbitration, control blender and IIP units.

References

1. Nowak, M.A.: Five rules for the evolution of cooperation. Science **314**, 1560–1563 (2006)
2. Flemisch, F., Abbink, D.A., Itoh, M., Pacaux-Lemoine, M.-P., Weßel, G.: Joining the blunt and the pointy end of the spear: towards a common framework of joint action, human–machine cooperation, cooperative guidance and control, shared, traded and supervisory control. Cogn. Technol. Work **21**(4), 555–568 (2019). https://doi.org/10.1007/s10111-019-00576-1
3. Flemisch, F., Adams, C.A., Conway, S.R., Goodrich, K.H., Palmer, M.T., Schutte, P.C.: The H-Metaphor as a guideline for vehicle automation and interaction; NASA/TM—2003-212672 (2003)
4. Hoc, J.-M.: From human-machine interaction to human-machine cooperation. Ergonomics **43**(7), 833–843 (2000)
5. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. IEEE SMC – Part A Syst. Hum. **30**(3), 286–297 (2000)
6. Flemisch, F., Altendorf, E., Canpolat, Y., Weßel, G., Baltzer, M., Lopez, D., Voß, G. M. I., Schwalm, M., Schutte, P.: Uncanny and unsafe valley of assistance and automation: first sketch and application to vehicle automation. In: Advances in Ergonomic Design of Systems, Products and Processes, pp. 319–334 (2017)
7. Rasmussen, J.: Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models; IEEE-SMC, May 1983
8. Flemisch, F., Schieben, A., Kelsch, J., Löper, C.: Automation spectrum, inner/ outer compatibility and other potentially useful human factors concepts for assistance and automation. Human Factors for Assistance and Automation (2008)
9. Flemisch, F., Schutte, P.: Fluidity in control distributions of automation. NASA Langley 2003
10. Allain, M., Konduri, S., Maske, H., Pagilla, P.R., Chowdhary, G.: Blended shared control of a hydraulic excavator. IFAC-PapersOnLine **50**(1), 14928–14933 (2017)
11. Baltzer, M., Flemisch, F., Altendorf, E., Meier, S.: Mediating the interaction between human and automation during the arbitration processes in cooperative guidance and control of highly automated vehicles. In: Proceedings of the 5th International Conference on AHFE, pp. 2107–2118 (2014)
12. Bennett, D.J.: Stretch reflex responses in the human elbow joint during a voluntary movement. J. Physiol. **474**(2), 339–351 (1994)
13. Burdet, E., Osu, R., Franklin, D.W., Milner, T.E., Kawato, M.: The central nervous system stabilizes unstable dynamics by learning optimal impedance. Nature **414**, 446–449 (2001)
14. Shadmehr, R., Mussa-Ivaldi, F.A.: Adaptive representation of dynamics during learning of a motor task. J. Neurosci. **14**, 3208–3224 (1994)
15. Kikuuwe, R., Yamamoto, T., Fujimoto, H.: A guideline for low-force robotic guidance for enhancing human performance of positioning and trajectory tracking: it should be stiff and appropriately slow. IEEE SMC - Part A Syst. Hum. **38**(4), 945–957 (2008)
16. Domen, K., Latash, M.L., Zatsiorsky, V.M.: Reconstruction of equilibruim trajectories during whole-body movements. Biol. Cybern. **80**, 195–204 (1999)
17. Ishio, J., Ichikawa, H., Kano, Y., Abe, M.: Vehicle-handling quality evaluation through model-based driver steering behavior. Veh. Syst. Dyn. **46**, 549–560 (2008)

Researchers and Public Views on Electronic Sideview Mirror System (ESMS) in the 21st Century Cars

Bankole K. Fasanya(✉), Skandip Anand, and Guna Sreeja Kallepalli

College of Technology, Purdue University Northwest, 2200, 169th Street, Hammond, IN, USA
fbankole@pnw.edu

Abstract. Each year, there are thousands of road accidents in the United States. The issue has led to additional safety adaptive features in recent cars of which many complained about the safety and the cognitive workload impacts on decision-making processes of some of the features, particularly the ESMS. The study aimed at increasing public awareness on the potential safety of ESMS support in cars. Nine websites (710 public comments) and two articles were reviewed, including. The comments were categorized into positive and negative impacts. Of the 710 comments, 73% were negative and only 27% were positive. Findings highlighted possible hazards associated with ESMS in cars such as image obscuration, high display brightness in dark driving conditions, poor visibility during raining, distraction, etc. Manufacturer improvement on the ESMS was also suggested. Findings from this study would add to the available knowledge on ESMS for the National Highway Traffic Safety Administration (NHTSA).

Keywords: Redundancy · Cognitive overload · Clustering · Trust · Electronic Sideview Mirror System

1 Introduction

Recently, the automotive industry has been working together with other major technology companies in order to deliver the most advanced, safest and most comfortable vehicles to the public. Cars are becoming large keen devices with advanced emergency braking capabilities, mapping technology for autonomous driving, better fuel efficiency and cars as a form of transportation [1]. Technology evolving perhaps has the largest impact on the automotive sector in the arena of safety. No doubt, science and technology have shifted the focus of the automobile manufacturers from safety features and user comfortability to fancifulness of the products. Driving is a complex information-processing task and is one of the most challenging activities people engage daily. The roadway information a driver must process (such as signs, signals, pavement markings, road curvature, position, and distance of other vehicles as well as pedestrians) make driving a dynamic task and changes constantly as a driver proceeds along his or her path. According to Fasanya et al. [2] driving a car comes with few challenges such as distraction from on-coming cars, stress, nervousness, psychology problems, etc. some

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 99–106, 2020.
https://doi.org/10.1007/978-3-030-50943-9_13

of these factors could lead to an accident, if not handled properly. Therefore, fancifulness cannot be an option for products design particularly for automobile. Furthermore, at high speeds, driver information processes require few second to be completed, thus, driver's ability to respond to situations appropriately depends on how he or she processes the information within the limited time [3].

Today, cars are equipped with many semi-autonomous features, like assisted parking and self-braking systems, geographical position system (GPS) etc [2, 4]. Every decade, automobile companies are progressively moving away from mechanical structure production of cars toward electronics production [5]. The replacement of the mechanical parts with electronic components according to vehicle manufacturers is to increase vehicle power and fuel consumption efficiency. The safety of the drivers and the vehicle occupants have been undermined in the recent car productions. Most safety measures are focused on drivers/human errors as the primary cause of accidents. Safety in terms of sideview visibility is evolving as an important focus for customers and vehicle manufacturers [6]. Therefore, what should customers and the vehicle manufacturers do to ensure adequate safety while driving a car, Fasanya et al. [2] concluded in their findings that highly automated cars reduce human tasks, while operating the car. Contrary to Fasanya et al.'s findings, Bainbridge [7] findings revealed that with a highly automated system, the output is very vital when humans interacted with a machine for its operation. Additionally, Gordon-Becker et al. [8] study revealed that a highly automated car reduces driver's attention and situation awareness, as the driver might not be actively involved in the selection of the actions recommended by the automation. Yang and Coughlin [9] findings revealed that the term "self-driving" or "autonomous" could easily mislead people into thinking that the driver's role in vehicle operation will become insignificant with the arrival of advanced vehicles. Inbuilt car technology is gradually increasing human misconception that driver interactions with the system would be eliminated. However, people forget that distraction and inattention are important driving safety issues. As the use of in-vehicle technologies becomes more popular, there are concerns about a related increase in driver distraction from their usage. The introduction of automated devices is intended to reduce distraction and not to complicate driving processes as majority of the road accidents are known to be associated with human errors [10]. However, a major part of the distraction associated with the use of the automated devices has been reported to increase drivers' cognitive workload during use, and not from the manual manipulation of the devices [11]. Additionally, NHTSA [9] reported that distracted driving in 2017 claimed 37, 133 lives in US highways. Distracted driving is any activity that diverts attention from driving, including talking or texting on your phone, eating and drinking, talking to people in the vehicle, interacting with the stereo, entertainment or use of navigation system etc. [12]. According to a recent report from the U.S. Department of Transportation, over 17% of police-reported crashes involved a distracted driver [13]. Another study showed that over 65% of near crashes and 78% of crashes resulted from distracted driving [14].

Harbluck et al. [11] concluded that under conditions of increased cognitive load, drivers made fewer saccades, spent more time looking centrally and spent less time looking to the right periphery in visual behavior. The same author further concluded in the study that increase in cognitive load reflected in drivers' increased rate of workload and distraction as well as reduced safe driving. Recently, electronic sideview system was introduced in cars as a technological advancement for drivers to change lanes. Fasanya et al. [2] found in their study that as drivers are habituated to regular sideview mirrors, it is difficult to adapt to the new technology quickly which might indirectly increase driver's stress levels by adding more information into their cognitive processing. The same authors concluded that switching back and forth with the two-side view systems (ESMS and side manual mirrors) increased drivers' cognitive workload and reduced decision-making processes of the drivers on highways. However, few drivers included in Fasanya, et al.'s [2] study confirmed to preferred cars equipped ESMS. Therefore, it is imperative to investigate more on the effectiveness of the ESMS in the cars for drivers' safety. The purpose of this study is to document people's concerns and comments on driving cars with ESMS.

The objectives of the study are as follow:

- Review articles on cars equipped with ESMS.
- Review online drivers' comments on cars equipped with ESMS.
- Analyze comments and draw conclusion.

2 Methodology

The study is a cross sectional review of public comments on the pros and cons of ESMS from nine different websites. We also review two scholarly articles on the drivers' perceptions on ESMS. The website visited are Auto News, Ars Technica, The Truth About Cars, Electrek, Car Magazine, Road Show, ExtremeTech, LeftLane News, and Edmunds. Years of comments review ranged from 2013 to 2019. Auto News comments were reviewed in 2015 as well as in 2018, users comments on The Truth About Cars were reviewed in 2013, 2017 and in 2018. Users comments on LeftLane News were reviewed in 2013, but in 2019 on Road Show site. Comments on car magazine are reviewed in 2019, Extreme Tech and Edmund in 2016. The sources of the articles are international conference on Applied Human Factors and Ergonomics proceeding and international journal for automotive.

This is a research base study, whereby secondary data were used for analysis. Data clustering analysis approach was used to find the goodness of fit on public comments where risk and poor safety as feedback were combined to form "risk" during the data analysis. Clustering in this study is define as the process of combining similar comments (comments with the same characteristics/attributes) in one group. Similarly, Trust and Reliable were combined as "Reliable" and Easy to Use, Stress Free and Convenient were combined to form "Convenient". The feedback recorded were segregated into two groups, the positive and the negative. Comments documentation were based on criteria on clustering formation.

3 Data Analysis and Results

The public comments were sorted and categorized into thirteen factors based on the key points emphasized in the comments. Factors include reliable, safety, convenient, effective, helpful, supportive, cost, indispensable, distraction, poor visibility, not reliable, and repetition. The sum of the comments listed in Table 1 does not match the sum of the public comments, because some of the comments did not fall under the inclusion criteria, therefore, were not included in the analysis.

Data were compiled and analyzed using Excel spreadsheet version 2016 (Microsoft, Redmont, WA, USA). In total, there were seven hundred and ten (710) public comments retrieved from online search and only three hundred and ten (310) public comments were included in the analysis, which amount to 44% of the entire online comments retrieved. Results show that 73% of the analyzed comments were negative and only 27% were positive about the ESMS in the cars. Of the 27% positive comments, only 14% acknowledged that the ESMS is a good feature for driver safety, 28% emphasized on the convenience, approximately, 18% affirmed on the helpfulness, 33% believe it is reliable, 1% said it is supportive when parking and 7% concluded that it is supportive while driving to consciously aware of the situation in the surrounding.

Table 1. Variables categorization based on year of comments

Factors	2013	2014	2015	2016	2017	2018	2019	Total	Res. type
Reliable	1	6	4	5	0	10	26	26	Positive
Safety	1	9	1	1	0	0	12	12	Positive
Convenient	1	2	2	0	4	14	23	23	Positive
Effective	3	0	0	0	0	3	6	6	Positive
Helpful	1	0	1	0	8	5	15	15	Positive
Supportive	0	0	0	0	0	1	1	1	Positive
Cost	0	0	0	0	13	12	25	25	Negative
Indispensable	0	1	0	0	0	0	1	1	Negative
Distraction	3	1	1	3	2	38	48	48	Negative
Poor Visibility	0	0	0	0	2	0	2	2	Negative
Not Reliable	2	0	2	3	0	12	19	19	Negative
Repetition	0	1	0	0	0	0	1	1	Negative
Risk	1	1	0	3	8	118	131	131	Negative

Of the 73% negative comments, 58% complaints on the problem associated with using the ESMS and the manual mirror simultaneously, approximately, 21% emphasized on the risk for the drivers, 11% of the negative comments only focused on the cost of repair and purchase cars equipped with the ESMS technology. Additionally, 8% of the comments concluded that the ESMS is not reliable for the purpose it's designed for. Analysis by year is detailed using pictorial representation. Figure 1 shows the public concerns by factors in 2013 and 2015.

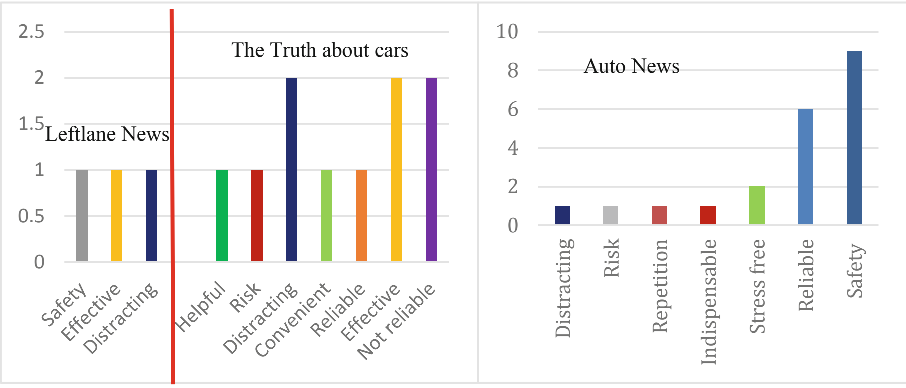


Fig. 1. Comments pictorial representation for year 2013 and 2015, left to right.

In 2013, reviewers were not specific on their concerns about the ESMS, but generalized their statement; therefore, this makes the number of comments analyzed in 2013 smaller compared with that of 2015. Findings from the two scholarly articles on the effectiveness of the ESMS also revealed significant negative impacts of the technology on the drivers' visibility, cognitive workload, physical health etc. Table 2 shows the factors identified in the findings from the articles. Table 2 indicates that poor visibility is one of the major factors noticed in driving cars equipped with ESMS in the articles findings. Fasanya, et al. argued that weather condition sometimes degrades the clarity of the ESMS camera display and thereby, reduced drivers' visibility.

Table 2. Findings on in-vehicle Technology from literatures

Source	Year	Factors
AHFE Proceeding Fasanya et al. [2]	2019	Distraction
		Stress
		Better image
		Cognitive workload
		Confusion
		Poor visibility
Int. journal for automotive Yang and Coughlin [9]	2014	Distraction
		Confusion
		Poor integration
		Frustration
		Cognitive workload

4 Discussion

It is discovered from the analysis that the public felt that factors such as risk/poor safety is a major area of concern with respect to the ESMS usage in the cars. The findings from the articles revealed poor visibility, distraction, and cognitive workload as the major factors identified by the public on the effectiveness of the ESMS. The major feedback from such a high proportion of concerns is the possible hazards associated with the inclusion of ESMS in cars. One of the article's findings clearly discussed about the image obscuration, high display brightness in dark driving conditions, which causes glare and affect drivers' visibility during raining periods. On the contrary, 8% of the public comments indicated that users believe that ESMS in cars is reliable, 4% believe that it ensures driving safe, 2% like the effectiveness, and 5% believe that having ESMS in the car is helpful to navigate successfully in the traffic or tight roads. The findings from the study also made us understand that inclusion of a new technology in cars can make the user to go through a sudden shock and hence create a fear factor that can lead to an accident while driving on a busy road. The findings from the public comments also agreed with the findings from the articles that distraction, risk, time for decision-making etc are the issues of concerns. Findings further revealed that the negative impacts of ESMS on drivers over weigh the positive impacts.

5 Conclusion

The dynamic changes in car inbuilt technology would continuously evolve as new devices are introducing to the market daily to ensure users comfortability and to move with the technology wind. The issue remains, how the car manufacturers would effectively introduce user safety features and ensure comfortability in the new cars inbuilt technology. Findings from this study revealed that the negative impacts of the ESMS on drivers over weigh the positive impacts. Therefore, following the design rules, the research ethics and principles, ESMS technology should not be allowed in the market for human use at the present design condition. Products design is aimed to be on the user capability and not fitting the users to the functionalities of the products. This is main reason why every industry now adopt ergonomics principle in their design is to ensure that products usability fit the capability of the users [15]. Distraction is one of the drivers' complaints, and obviously driving requires full concentration, so little distraction as cited in the public comments can end in a great loss of both lives and properties. Additional concerns pointed out by the public are the reliability and risk associated with the use of the ESMS in the car for safe driving. Risk as defined in the literatures as the degree to which a particular material can cause harm. Therefore, having ESMS in the car has higher probability of causing accident compared with having the regular/manual side mirror. The articles reviewed emphasized on the difficulties, which the drivers experienced using the ESMS for lane changing while driving in the direction of sunlight ray.

Further, it was also discovered from the reviewed articles that cognitive workload is another concern for using the ESMS for lane changing and driver maneuvers. Increased in drivers cognitive workload occurred as mentioned in the articles when drivers tried to compare image seen on the camera display with the image on the manual mirrors. Overall, ESMS has failed to perform as expected by the public and therefore, car manufacturers proper review on the existing ESMS in the cars is encouraged. This research suggests for additional studies on ESMS technology in cars to increase the available knowledge on the benefits versus the disadvantages of having ESMS in cars. In the future, a simulated study will be appropriate to investigate the effectiveness of the ESMS in cars.

References

1. Utermohlen, K.: The Future of Technology in the Automotive Industry [internet], May 2018. https://towardsdatascience.com/the-future-of-technology-in-the-automotive-industry-11081c8a1999?gi=331c37c6eeb4. Accessed 30 Jan 2020
2. Fasanya, B.K., Avula, Y., Keshavula, S., Aragattu, S., Kurra, S., Kummari, B.: Human factors concerns: drivers' perception on electronic sideview system in 21st century cars. In: International Conference on Applied Human Factors and Ergonomics, pp. 556–563. Springer, Cham (2019)
3. Dewar, R.E., Olson, P.L., Alexander, G.J.: Human Factors in Traffic Safety, 2nd edn. Lawyers & Judges Publishing Company, Tucson, Arizona (2007)
4. Takacs, A., Rudas, I., Bösl, D., Haidegger, T.: Highly automated vehicles and self-driving cars [industry tutorial]. IEEE Robot. Autom. Mag. **25**(4), 106–112 (2018)
5. Oswald, E.: How and why electric vehicles will change the way cars look. [Internet], 26 July 2019. https://www.digitaltrends.com/cars/how-electric-cars-will-change-car-design/. Accessed 31 Jan 2020
6. Nancy Gupton, N.: The Science of Self Driving Cars [INTERNET- The Franklin Institute] (2016). https://www.fi.edu/science-of-selfdriving-cars. Accessed 20 Dec 2019
7. Bainbridge, L.: Ironies of automation. In: Analysis, Design and Evaluation of Man–Machine Systems, pp. 129–135. Pergamon (1983)
8. Gordon-Becker, S.E., Lee, J.D., Liu, Y., Wickens, C.D.: An Introduction to Human Factors Engineering. Google Scholar (2004)
9. Yang, J., Coughlin, J.F.: In-vehicle technology for self-driving cars: advantages and challenges for aging drivers. Int. J. Automot. Technol. **15**(2), 333–340 (2014)
10. NHTSA, 2017. Drivers assistant technologies (internet). https://www.nhtsa.gov/equipment/driver-assistance-technologies. Accessed 30 Jan 2020
11. Harbluk, J.L., Noy, Y.I., Trbovich, P.L., Eizenman, M.: An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. Accid. Anal. Prev. **39**(2), 372–379 (2007)
12. Li, N., Busso, C.: Driver mirror-checking action detection using multi-modal signals. In: Proceedings of the 6th Biennial Workshop in Digital Signal Processing for In-Vehicle Systems and Safety, pp. 101–108, September 2013
13. NHTSA. Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. Technical Report NHTSA-2010-0053, National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT), Washington, D.C., USA, April 2012

14. Harbluk, J.L., Noy, Y.I., Eizenman, M.: The impact of cognitive distraction on driver visual behaviour and vehicle control (No. TP# 13889 E) (2002)
15. Fasanya, B.K., Shofoluwe, M.: Occupational ergonomics: emerging approaches toward improved worker productivity and injury reduction. In: International Conference on Applied Human Factors and Ergonomics, pp. 385–395. Springer, Cham (2018)

Future Transportation Service Technology Platform System Based on Internet of Vehicles

Jia-xin Liu[1(✉)], Zi-yun Li[2], and Ying Cao[1]

[1] Huazhong University of Science and Technology, Luoyu Road 1037, Wuhan, China
2840731796@qq.com

[2] Hubei University of Technology, Nanli Road 28, Wuhan, China

Abstract. In the era of 5G and intelligent development, the layout and architecture of the Internet of Vehicles are slowly being established. In the future, the urban travel system will gradually become a research hotspot in the automotive field. Mobile mobility services are an important part of smart cities. The combination of multi-modal transportation, shared travel, electric vehicles, and autonomous driving is an important direction for automobile industrial transformation. In order to study the future transportation service system, This article will build a new type of transportation service system from the perspective of social supply and demand, improving user experience and promoting trilateral win-win situation. The combined market users cognitive model theory of the present system, logic and feasibility.

Keywords: Smart city · Transportation service system · Automobile industry restructuring · Multimodal transportation · Shared travel electric car · Autonomous driving

1 Introduction

Cars from Germany Bailun root user experience studio after an investigation found. Today, the life cycle of each car only accounts for 5% of its life cycle, and a large amount of time has not been used reasonably. In the time that is not being used reasonably, Most cars are parked in private garages or community parking lots. On the one hand, carrelated consumption is increased, and on the other hand, the size of the parking lot and the development cost are increased. At the same time, in today's society, the influence of the circle has become greater and greater. Circle life is a rising trend culture that can bring more value to everyone who joins the circle. The same concept of life will bring us greater vision pattern, more career opportunities, In common philosophy of life, we can find a sense of identity. Based on the above problems and social pain points, the author came up with the idea of creating an automotive platform community.

In the future, the development of transportation vehicles will no longer be a single provision of goods, but more services will be added and service-oriented providers will be added [1]. It is no longer for a single user, but a unified collection of car companies. Future transportation can meet the needs of private cars and services, and also increase

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 107–111, 2020.
https://doi.org/10.1007/978-3-030-50943-9_14

the utilization rate of cars, so that cars are no longer a consumable, It has even become a revenue-generating and investment product, and this platform community can provide car companies with a huge user data network and popularity, and carry out service transformation. This system will have a strong information chain and function chain, build an intelligent site led by the OEMs, and form an online and offline integration platform for automobiles based on the Internet of Vehicles.

2 Development Trend of the Autonomous Driving Industry

In the next ten to twenty years, self-driving cars will gradually replace the simple manual labor of human beings. It will change people's travel and lifestyle, and will also slowly impact the traditional transportation industry. The society attaches importance to the service industry The degree will increase.

Google commissioned VTTI to compare "use of natural data for automated vehicle collision comparison" data, which shows that the comfort and safety of driverless cars will exceed the average level of human drivers for several reasons: A. Autonomous vehicles will be equipped with keen sensing elements (hardware), which can surpass human perception, so autonomous driving can make decisions and respond earlier than humans; B. Fatigue driving has become a cause of traffic safety accidents. One of the important reasons, but autonomous vehicles can avoid this problem; C. Machines are more logical and more rational than humans, and will not make wrong decisions due to emotions (such as emotional problems such as road rage). Although there are still some defects in the hardware of automobiles at this stage, with the continuous improvement of hardware technology, autonomous driving will have complete advantages, making human travel safer. The future autonomous driving industry has the following trends in technology, products and competitors (see Table 1).

Table 1. Future development trends of the autonomous driving industry

Involved in the field	Technology	Applications	Competitive subject
Future trends	Human-computer interaction solutions for future models, R & D of V2X and 5G technologies	Commercialization of ADAS, urban travel plans, and strategic layout of autonomous driving	Intervention of Internet companies, traditional car companies, and service industries

3 Blockchain in the Automotive Industry

3.1 Disruption of the Automotive Industry

With the current transformation of the automotive industry, the degree of integration of the automotive market will increase significantly in the future. In the future, automobiles may be able to provide personalized services according to user needs. The importance of blockchain will increase significantly. Blockchain technology can be widely used in the entire automotive industry to promote product-oriented and service-oriented transformation. Blockchain has the advantages of accuracy and extensiveness [2], which can prompt enterprises to formulate a more detailed future travel strategy layout.

With the advent of the sharing era, the consumer behavior of users is undergoing drastic changes. Shared mobility, new energy transportation, and on-demand services are increasing [3]. At the same time, users are more willing to share data and use new technologies than in the past. The accelerated transformation has promoted the transformation of the automotive industry. The digitization of users is constantly increasing, indicating that they are also attaching greater importance to user experience. The increase of digitalization and the user's demand for data usage (optimizing the user experience) will prompt enterprises to provide more travel-oriented service experiences. There is a large amount of data and user information that has not been effectively used today. By analyzing these consumer data, customer insights can be obtained for commercialization purposes. Analysis of vehicle use and experience can bring new commercial opportunities. Autonomous and unmanned driving are setting new paradigms for driving and transportation.

3.2 The Practical Application of the Service System

In the automotive industry, blockchain can be used to securely share data of vehicles, owners, manufacturers, and dealers; improve vehicle utilization, and sell unused space through vehicle sharing, thereby commercializing the above data and mining more Great value; realize data communication between various links such as manufacturers, dealers, auto financial services institutions and insurance companies; save costs for consumers by "pay as you go", and provide data based on the use of cars and accurate data More affordable insurance products [4].

In this solution, the vehicle platform community is composed of a virtual community, a vehicle owner, and a vehicle module. During the non-use time period, the vehicle owner makes income-generating investments through the APP community platform provided by the platform, signs up for community activities, and the vehicle will be used in non-use vehicles. Participate in activities integrated by the platform during the time period, and automatically go to the event location to generate revenue; the car interconnection platform can capture effective information and data through the app background, more targeted research and development of future vehicle models, and increase the visibility of related auto companies (see Fig. 1).

Fig. 1. Three major hub layouts

4 Construction of Urban Travel System

In the future, the car will arrive at the doorstep of the owner's home on time through the "owner call appointment" function of the travel platform app and pick up the owner to go to work. Immediately on the way to the company, car owners can send cars to participate in work activities through the car community platform and confirm the time and place of the recalled car. After successfully sending the owner to the company building, the car will receive the information communicated to the 5G car networking site through the integrated car platform, and organize the car with other vehicles that are on the way to the event (work) site, making the vehicle more efficient and convenient. Function division, quickly go to the scene. After the vehicle arrives at the event (work) site, it automatically parks into the designated functional area of the car during the event through the automatic driving assistance system, and uses the background information of the automobile community to organize the vehicle in an orderly and efficient manner. The parking area of each car in the venue belongs to the wireless faster charging module. On the one hand, it is used for the power supply of the functions required by the vehicle at the event site. On the other hand, the owner can recall the vehicle at any time and keep the vehicle's power sufficient. At the same time, the car of the future will switch modeling modes to maximize the use of interior space and meet the application needs of different scenarios in the same event [5]. After the work is over, the car arrives at the agreed place on time to pick up the car owner, and at the same time, the day's revenue-generating settlement and activity vehicle infrastructure loss review information will be pushed to the car owner's APP, so that the car owner has no worries (Fig. 2).

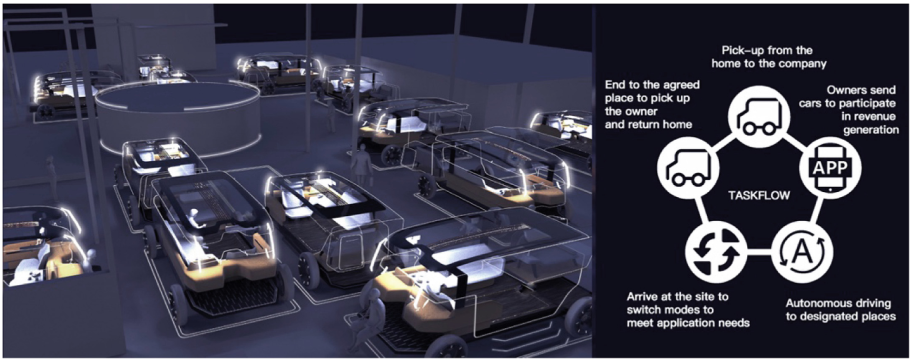


Fig. 2. System application process

5 Summarize

In the future, the development of transportation vehicles will no longer be a single provision of goods, but more services will be added, and service-oriented travel providers will no longer be oriented to a single user, but a unified collection under the control of automobile companies [6]. Future transportation can meet the needs of private cars and services, and can increase the utilization rate of vehicles, so that cars are no longer a consumable, or even a revenue-generating, investment product, and this platform community can also serve car companies. Provide a huge user data network and popularity, and carry out service-oriented transformation. This system will have a strong information chain and function chain, build intelligent sites dominated by OEMs, and form an online and offline integration platform for automobiles based on the Internet of Vehicles.

References

1. China Autonomous Driving Industry Development Research Report, Prospective Industry Research Institute (2019)
2. Gartner, Forecast: Blockchain Business Value, Worldwide, 2017-2030, March 2017
3. Deloitte, Blockchain Auto Finance – How blockchain can enable the Future of Mobility, Blockchain Institute (2017)
4. Ginther, A.J.: Smart city: The city of Columbus, E & E News, Accessed 15 Dec 2017
5. Corwin, S., Vitale, J., Kelly, E., Cathles, E.: The future of mobility: How trans portation technology and social trends are creating a new business ecosystem, Deloitte University Press, 24 September 2015
6. Corwin, S., Jameson, N., Pankratz, D.M., Willig-mann, P.: The future of mobility: What's next ? Deloitte University Press, 14 September 2016

Impact of Speedometer Forms on Integration Task Performance for Train Driving

Chenchen Gao[1(✉)], Weining Fang[1], Ke Niu[2], and Jianxin Wang[1]

[1] State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China
464369568@qq.com

[2] School of Locomotive and Rolling Stock, Zhengzhou Railway Vocational & Technical College, No. 56 Pengcheng Avenue, Zhengdong New District, Zhengzhou 451460, China

Abstract. Accurate speed control can effectively improve the efficiency of train operation. The advantages and disadvantages of the speed display form directly impact the interaction efficiency and the driver's correct situational awareness of the train running state. Three kinds of train speedometer forms were designed, and the existing interface was compared with the three types. 28 subjects were selected to carry out four speed control experiments. The results showed that, the response time and accuracy performance were the highest in the case of color highlighting; the response time performance was higher but the accuracy performance was the lowest in the case of graph without numbers; the accuracy was higher but the response time was the lowest in the case of scale sparsity. This study shows that the speed display design should be give priority to color highlighting, while weighting the influences of number highlighting and scale sparsity.

Keywords: Train driving · Speedometer form · Integration task · Eye movement

1 Introduction

Unlike other vehicles, there is no lateral control for a train driver. To guarantee the safety and punctuality of a train, the only required skill of a train driver is to control the speed along the railroad tracks. However, with the rapid development of metros, the driving mode has gradually changed from the original manual operation to fully automated operation. The fully automated operation mode realizes intelligent and automatic operation and is not equivalent to unmanned driving. In the event of an emergency, the driver's intervention is still required [1]. In an emergency, the only thing that the driver can control is the train's speed.

Boles and Wickens noted that with many tasks involving multiple-element arrays, information from several sources must be combined, and a single response must be generated. They referred to these types of tasks as integration tasks [2, 3]. The train driver provides a single-operation response—acceleration, deceleration or maintenance of speed according to the information provided by the elements in the speedometer.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 112–119, 2020.
https://doi.org/10.1007/978-3-030-50943-9_15

Therefore, the speed control of the train driver is an integration task. The train driver's correct perception of the information from the speedometer is the decisive factor in ensuring the performance of the integrated task. The advantages and disadvantages of the speed display form directly impact the interaction efficiency and the driver's correct situational awareness of the train running state.

Studies have been performed on the effects of elements in car speedometers on people's performance. The factors that have a significant impact on drivers' speed recognition performance can be summarized as color highlighting [4] and the density of the dial scales [5]. However, there is no research on the effects of the elements in train speedometers on train drivers' performance, and the speed recognition of the train driver, which is the integration task, is not the same as a car driver's speed recognition. Therefore, it is necessary to conduct research on the influence of factors in train speedometers on human performance.

2 Method

2.1 Participants

In this speed recognition experiment, 30 students at Beijing Jiaotong University were recruited by online advertisements. The 28 valid samples in total (M = 22.3 years of age, SD = 2.12 years) included 15 males and 13 females. In the eye movement experiment, 14 students at Beijing Jiaotong University were recruited by online advertisements. The 10 valid samples in total (M = 21.9 years of age, SD = 2.13 years) included six males and four females. No color blindness, color weakness or other eye diseases were observed. The distance between the monitor and the eyes of the participants was 66 cm. The experiments were conducted in a bright and quiet environment.

2.2 Apparatus

The speed recognition experiment was conducted on a computer with a 24-inch monitor and a standard keyboard. The eye movement experiments used a 22-inch screen, standard keyboard, and SMI RED500 eye tracker.

2.3 Procedure

In the speed recognition experiment, first, the participants learned the train driving rules, which were derived from the train driving video analysis. After learning, the training program was started.

Training process: A total of 48 different combinations of allowable speed, train speed and target speed were used as training materials. Each training material picture was randomly presented once, and a total of 48 judgment operations were required by the participants. The experimental program was compiled in the E-prime software [6]. First, a white "+" appeared at the center of a black screen, and the presentation time was 1000 ms. Then, the training picture was presented, and the presentation time was

10,000 ms. When the picture was presented, the subjects selected fast acceleration, slow acceleration, hold, slow deceleration, or fast deceleration and pressed the corresponding key, where "x" corresponds to fast acceleration, "c" corresponds to slow acceleration, "v" corresponds to hold, "b" corresponds to slow deceleration, and "n" corresponds to fast deceleration. The response time and correct rate were displayed on the interface after the test operation ended. The participants could run the training program multiple times until the accuracy rate reached more than 90%, and then, they could start the formal experiment. The training experiment process is show in Fig. 1.

Formal experiment process: The formal experiment was divided into four groups. Each group used 188 different combinations of allowable speed, train speed and target speed as experimental materials. The formal experimental speedometer formats were different from the training format. Each picture from the experimental materials was randomly presented once, and a total of 188 judgment operations according to the 188 pictures were required. Examples of each group of experimental material are shown in Fig. 4. The participants performed four sets of experiments in random order. The experimental program was compiled in the E-prime software. First, a white "+" appeared on a black screen, the presentation time was 2000 ms, the material was presented, and the presentation time was 8000 ms. When the picture was presented, the subjects selected fast-acceleration, slow-acceleration, hold, slow-deceleration, or fast-deceleration and pressed the corresponding key, as in the training process. The formal experiment process is show in Fig. 2. The experimental scenario is shown in Fig. 3.

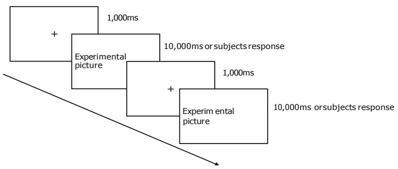


Fig. 1. The training experiment process

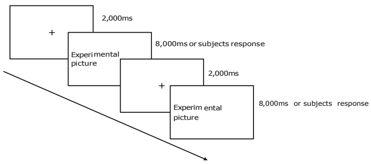


Fig. 2. The formal experiment process

The eye movement experiment was divided into four groups. Each group used 60 different combinations of allowed speed, train speed, and target speed as the experimental materials. After the participants thought that they had obtained the information, they pressed the space bar. Then, a picture that describes the ten types of speed information was presented on the monitor. At the same time, the subject chose the type of speed information and indicated this to the main testing system. Examples of the experimental materials for each group are the same as in the formal experiment. The participants also performed four sets of experiments in random order.

Fig. 3. Examples of each group of experimental materials

Fig. 4. The experimental situation

2.4 Data Analysis

The speed recognition experimental program recorded the correct rate and response of the test participants when each picture was presented, calculated the average and standard deviation, and performed a significance analysis. The eye movement experiment program recorded the eye movement data. If the P-value was less than 0.1, a significant difference was considered to exist.

3 Results

The speedometer form of the Group 1 experiment has both numbers and graphics and no color highlighting; the small scale represents 5 km/h, and the large scale represents 10 km/h. The speedometer form of the Group 2 experiment has no numbers. The speedometer form of the Group 3 experiment is a speedometer form with color highlighting. The speedometer form of Group 4 is a speedometer form with only a large scale representing increments of 10 km/h, and there is no small scale. Therefore, from the comparison of the Group 1 and Group 2 experimental data, the impact of the numbers on the integration task of the train driver is obtained; from the comparison of the Group 1 and Group 3 experimental data, the effect of the color highlighting on the integration task of the train driver is obtained; the comparison of the Group 1 and Group 4 experimental data shows the influence of the dial scale density on the integration task of the train driver. According to the comparison between the four sets of experimental data, the degree of influence of different factors on the integration task of the train drivers is obtained.

3.1 Reaction Time

The reaction time is an indicator of the efficiency of the train driver in the integration task. The shorter the reaction time is, the faster the reaction operation made by the train driver based on the speed information, and the more efficient the train can be. The average reaction times of the four groups of experiments are shown in Table 1.

Table 1. The mean and variance of the reaction time of the four groups experiments.

Group of experiments	Average response time (ms)	Variance (ms^2)
1	2827.12	727.14
2	2804.15	640.68
3	2731.70	741.97
4	2905.34	632.64

Analysis of the significance of the experimental reaction time data of the Group 2 and Group 3, F (1,26) = 4.03, p = 0.00018 < 0.1. Analysis of the significance of the experimental reaction time data of the Group 2 and Group 4, F (1,26) = 4.56, p = 5.66E−05 < 0.1, Analysis of the significance of the experimental reaction time data of the Group 3 and Group 4, F (1,26) = 4.55, p = 5.8E−05 < 0.1. Therefore, each two of the four groups of experimental reaction time data were significantly different. In terms of reaction time performance, the speed display form with color highlighting is better than the speed display form with only figures and no numbers, and the speedometer form with sparse dial scales.

The analysis of the Group 1 and Group 2 experimental data, F (1,26) = 5.87, p = 4.38E−06 < 0.1, shows that, in terms of response performance, pure graphics speedometer forms are obviously better than the speed display form of mixed graphics and numbers, which is consistent with the literature [3, 4]. The analysis of the experimental data of Group 1 and Group 3, F (1,26) = 7.07, p = 5.85E−07 < 0.1, shows that, regarding reaction performance, speedometer forms highlighted in color are significantly better than the speed display form without color highlighting, which is consistent with the literature [5]. The analysis of the experimental data of Group 1 and Group 4, F (1,26) = 6.71, p = 1.04E−06 < 0.1, indicates the performance with respect to response time. The speed display form of the sparse dial scale is obviously worse than the speed display form of the dense dial scale, which is inconsistent with the literature [6]. Therefore, eye movement experiments were carried out.

Analysis of the significance of the experimental reaction time data of Group 2 and Group 3, F (1,26) = 4.03, p = 0.00018 < 0.1. Analysis of the significance of the experimental reaction time data of Group 2 and Group 4, F (1,26) = 4.56, p = 5.66E−05 < 0.1. Analysis of the significance of the experimental reaction time data of Group 3 and Group 4, F (1,26) = 4.55, p = 5.8E−05 < 0.1. Therefore, each of the four groups of experimental reaction time data were significantly different. In terms of reaction time performance.

3.2 Correct Rate

The accuracy rate is an indicator that reflects the safe operation of the train driver in the integration task. The higher the accuracy rate is, the higher the probability that the operation corresponding to the correct response by the train driver is based on the speedometer form, and the safety of the train can be guaranteed. The average correct rates of the four groups of the experiments are shown in Table 2.

Table 2. The mean and variance of the correct rate of the four groups of experiments.

Group of experiments	Average correct rate	Variance
1	0.89	0.074
2	0.88	0.095
3	0.91	0.066
4	0.90	0.070

The analysis of the experimental data of Group 1 and Group 2, F (1,26) = 2.40, p = 0.012, indicates that the speed display with the pure graphics in terms of accuracy is obviously worse than the speed display with a mixture of graphics and numbers. The analysis of the Group 1 and Group 3 experimental data, F (1,26) = 5.02, p = 3E−05, indicates that, in terms of accuracy, the speedometer with color highlighting is significantly better than the speedometer form without color highlighting. The analysis of the experimental data of Group 1 and Group 4, F (1,26) = 7.70, p = 3.61E−07, shows that, in terms of the correct rate performance, the speedometer form with the sparse dial scale is obviously better than the speedometer form with the dense dial scale. The experimental results of the correct rate are consistent with the literature [2–5].

Significance analysis of the accuracy data of Group 2 and Group 3, F (1,26) = 1.70, p = 0.08) = 3.17, p = 0.0014. For the significance analysis of the experimental data of Group 2 and Group 4, F (1,26) = 5.56, p = 7.68E−06. Therefore, two of the four groups of experimental correct rate data were significantly different.

3.3 Eye Movement Experiment Data Analysis and Results

The ratio of the fixation count and saccade count was selected as the index of the information search efficiency [7, 8], and the ratio of Group 3 was 1.09, the ratio of Group 2 was 1.08, the ratio of Group 1 was 1.07, and the ratio of Group 4 was 1.02. This is consistent with the response performance ranking. Therefore, the factor that affects the performance of the reaction time may be the efficiency of the information search.

4 Discussion

In the experiment, the participants made a single response based on the information presented by the pointer, scale, and number combination of the speedometer interface and made the choices of fast acceleration, slow acceleration, hold, slow deceleration, or fast deceleration. Therefore, train driving is an integration task. According to the results of the experimental data, when the speedometer has color high-lighting, the reaction time and correct rate of the integration task are the best. The color highlighting significantly improves the efficiency of train driving and helps to ensure the safe and efficient operation of the train. When the speedometer interface has only graphics and no numbers, the performance of the integration task response is better, and the accuracy rate decreases, indicating that when there are no numbers, the speedometer interface information is obtained faster but with poor accuracy. When the scale density of the speedometer interface is sparse, the performance of the integration task is poor, and the accuracy rate is better, indicating that the speed of obtaining the speedometer interface information is lower when the scale density of the dial is sparse but is performed with higher accuracy.

According to the results of the eye movement experiment data, the higher the ratio between the fixation counts and the glance counts is, the better the performance level of the response of the train driver during the integration task. According to the literature, the ratio between the number of fixations and the number of glances represents the efficiency of the information search, that is, the higher the efficiency of the information search is, the better the reaction time. This shows that color highlighting is conducive to improving the efficiency of information search and thus improving the speed of the train driver's acquisition of interface information, thereby improving the performance level of the integration task response [9–11].

According to the experimental results, in the design of the train speedometer, the effect of color highlighting should be considered first, and the effects of different color highlighting or color combinations on the integration task of train drivers can be studied in depth. Second, the degree of digital highlighting and the scale density should be considered carefully. Based on the degree of influence of the above factors on the integration task of train drivers, a design plan or guidance for the train speedometer can be given.

5 Conclusion

Regarding the reaction time and correct rate performance, the speedometer forms with color highlighting are the best. Therefore, the effects of color should be considered first in the design of speedometers. The response time was shorter, but the accuracy was the lowest in the case of graphs without numbers; the accuracy was higher in the case of scale sparsity, but the response time was the longest. Therefore, there should be a trade-off between reaction time and accuracy when considering the use of numbers and scale density in a speedometer to guarantee safety and efficiency.

References

1. Vuchic, V.R.: Maintaining Performance with Full Automation, vol. 3, pp. 36–39 (2014)
2. Boles, D.B., Wickens, C.D.: Display formatting in information integration and nonintegration tasks. J. Hum. Factors Ergon. Soc. **29**(4), 395–406 (1987)
3. Payne, D.G., Lang, V.A., Blackwell, J.M.: Mixed versus pure display format in integration and nonintegration visual display monitoring tasks. J. Hum. Factors Ergon. Soc. **37**(3), 507–527 (1995)
4. Juanfang, Xu: Car dashboard interface design based on SEEV attention model. J. Mach. Design **12**, 119–122 (2016)
5. Xu, J.: Analysis and Research on Cognitive Errors of Digital Interface Information by Eye Tracking. School of Mechanical Engineering Southeast University (2015)
6. Macwhinney, B., James, J.S.S., Schunn, C., et al.: TEP–a system for teaching experimental psychology using E-Prime. J. Behav. Res. Methods, Instrum. Comput. J. Psychon. Soc., Inc. **33**(2), 287–296 (2001)
7. Card, S.K., Moran, T.P., Newell, A. (eds.) The Psychology of Human-Computer Interaction. CRC Press, Boca Raton (1983)
8. Wu, X., Xue, C., Zhou, F.: An experimental study on visual search factors of information features in a task monitor-ing interface. In: 17th International Conference on Human-Computer Interaction, pp. 525–536. Springer Press, Los Angeles (2005)
9. Kruger, J.L., Hefer, E., Matthew, G.: Measuring the impact of subtitles on cognitive load: eye tracking and dynamic audio visual texts. In: Conference on Eye Tracking South Africa. ACM, pp. 67–71. ACM Press, South Africa (2013)
10. Goldberg, J.H., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. **24**(6), 631–645 (1999)
11. Mayr, U., Kuhns, D., Rieter, M.: Eye movements reveal dynamics of task control. J. Exp. Psychol. Genera. **142**(2), 489–509 (2013)

New Technology Implementation in High-Risk Organizations - The Application of HRO Principles in New Technology Implementation in Railroad Industry

Yalda Khashe(✉) and Najmedin Meshkati

Viterbi School of Engineering, University of Southern California, Los Angeles, USA
{Khashe,Meshkati}@usc.edu

Abstract. High-risk organizations are inherently complex and depend on the latest technologies to survive and function properly. Therefore, introducing new technology to such an organization is inevitable. Studies show that the installation of new technology always involves some changes to the organization and its members. The railroad industry, as an example of a high-risk and safety-critical organization, strives to avoid catastrophic events, while performing dynamic tasks under strict time constraints, operating technology posing large-scale physical hazards. High Reliability Organizations (HROs) are a subset of high-risk organizations designed and managed to avoid such accidents. This paper discusses the adaptation of HRO principals as part of the implementation process for the Positive Train Control (PTC) technology in a safety-sensitive railroad organization.

Keywords: Human-systems integration · New technology implementation · System safety · High reliability organizations · Transportation · Railroad

1 Introduction

High-risk organizations are inherently complex and depend on the latest technologies to survive and function properly. High-reliability Organizations (HROs) are a subset of high-risk organizations designed and managed to avoid such accidents.

The railroad industry, as an example of a high-risk and safety-critical organization, strives to avoid catastrophic events, while performing dynamic tasks under strict time constraints, operating technology posing large-scale physical hazards. Failures in these systems are rare but they are highly visible, making the consequence of such failures disastrous. The fact that trains are major means of transportation around the world is another factor that makes these failures highly significant.

One of the accidents that had a significant impact on the US railroad industry was the 2008 collision between Metrolink commuter train and Union Pacific freight near Chatsworth, California. The catastrophic result of the accident was the loss of 25 lives, 135 injuries and millions of dollars in damages. Shortly after the accident, US Congress passed the Rail Safety Improvement Act that requires Class I railroads to install

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 120–126, 2020.
https://doi.org/10.1007/978-3-030-50943-9_16

Positive Train Control (PTC) systems on their tracks, which carry passengers or toxic-by-inhalation materials by the end of 2015. PTC is a generic term referring to a range of fully integrated technologies that overlay existing safety systems to prevent train-to-train collision and improve worker safety. One of the challenges that railroad industry is facing for implementing PTC is the complications of introducing this new technology to an already existing system.

This study discusses the adaptation of HRO principals as part of the implementation process for the PTC technology in a safety-sensitive railroad organization.

2 New Technology Implementation in High-Risk Organizations

High-risk organizations are inherently complex and depend on the latest technologies to survive and function properly. Therefore, introducing new technology to such an organization is inevitable. Studies show that the installation of new technology always involves some changes to the organization and its members. Therefore, organizational and technological change must be considered simultaneously. Technological change often requires some organizational changes, and it will fail if both organizational and technological changes are not effectively integrated and managed to achieve alignment [1].

As Bea et al. [2] pointed out, the system modeling approach should be interdisciplinary, especially in the case of a high-risk organization. The same applies to efforts taken by the organization to implement new technology. Sometimes, the technical aspect of the implementation could overshadow the human, social, and organizational elements. However, studies show that this approach could negatively affect the outcome of the implementation. In a study of 2000, U.S. firms implementing new office systems, less than 10% of the failures were attributed to technical failures; the majority of the reasons given were human and organizational in nature [3]. Failed technology investment would cost an organization a lot of time, effort, and considerable financial loss, especially if we want to consider a high-reliable organization, which would require sophisticated and expensive tools.

The purpose of implementing new technology is to improve the effectiveness of the work done in the organization. However, to achieve that goal, organizations need to reestablish boundaries between different working groups in the organization and rethink the way they interact with one another [4]. Adaptive Structuration Theory (AST) suggests that technology has structures in its own right but that social practices moderate their effects on behavior [5]. In other words, causality runs in both directions: technology influences the patterns of human activity and the technology changes as it is modified in the course of daily activity. Therefore, it is critical to understand the relationship between technological and social factors over time. As Barley [6] proves "The complexity and uncertainty are functions of how the machine merged with the social system; they are not attributes of the machine itself."

The way technology is presented to the organization is also very important. An implementer's positively biased presentation of technology makes negative surprises inevitable; the paradox of negative experience is that these negative surprises if managed well, become valuable positive learning experiences for users [7]. Studies show that new technological systems do not impact precisely as they were designed. The designer misses the unexpected features and behaviors; therefore, they fail to account for events that fall outside the pre-established scope of work. Technologies interact with both the structural and institutional arrangements of an organization and the assumptions, frames, and mental images of its people. Most of these change processes, because of their cognitive and social embeddedness, are largely unnoticed in organizational and design settings [8].

This gap could be further explained by two concepts of technology affordance and technology constraints as explained by the Technology Affordances and Constraints Theory (TACT), which shows how people and organizations use information systems and how the use of information systems affects individuals, organizations, and their performance. Technology affordance is an action potential to what an individual or organization with a particular purpose can do with a technology or information system. Technology constraint is the ways in which an individual or organization can be held back from accomplishing a particular goal when using technology or a system [9].

Although many studies argue that technological adaptation is a gradual and continuous process, Tyre and Orlikowski [10] believe it is highly discontinuous and it is only through experience with a new technology that users would discover its implications. The Authors future argue that there is a relatively brief window of opportunity to explore and modify new process technology following initial implementation.

These are real concerns, and they are not likely to go away with new management, engineering skills, or bleeding edge technologies. To manage these difficulties, the designers should understand the role of cultures in alignment, recognize the breadth of factors and their relationships, and focus alignment on business purpose, not as a technology fix to a localized problem [1].

3 Hallmarks of High Reliability Organizations

The two key attributes that mark high reliability organizations are a chronic sense of unease and strong responses to weak signals. According to Weick and Sutcliffe [11], "hallmarks of high reliability", or major characteristics of HRO while "anticipating and becoming aware of the unexpected", include 1) Preoccupation with failure, 2) Reluctance to simplify interpretations, 3) Sensitivity to operations. In addition, when the "unexpected occurs", HROs attempt to contain it by 4) Commitment to resilience and 5) Deference to expertise. Not all of these characteristics apply to all the organizational processes.

The application of HRO starts with a deep knowledge of the technologies that are used in the organization, and clear understanding roles and responsibilities. These organizations foster a culture that values diversity and encourages productive collaboration. The concept of complex adaptive systems enables the organizational change in the system and it is internalized in an HRO through systems, processes, culture, and education [12]. HROs aim to empower expert people closest to a problem by distributing expertise and shifting leadership to people at the "sharp-end" who have the answer to the problem at hand [13].

HROs are process rather than goal-oriented, and failure as the outcome of eroded processes that ignored precursor events identified by an operator because the culture stressed other operational goals at the expense of safety [14]. Over time, organizations learn how to approach and eliminate visible and routine problems, and the positive feedback that they receive creates a culture that directly influences organizational performance. The same concept does not apply to high-risk operations since in these organizations risks are not clear. Studies show that serious events are often the result of systemic failures, human errors, or organizational weaknesses. Some of these factors may seem inconsequential when evaluated in isolation [15].

HROs strive to minimize the gap between "work-as-imagined" versus "work-as-done". These organizations develop detailed operation procedures, and due to their safety-sensitive nature strive to perform within the boundaries and safety limits. However, it as a known fact that there are discrepancies between plans and how the actual performance of the sub-systems (ΔW_g). The ΔW_g represents "what" is not working that puts us out of the physics-based safety basis, and it is determined using Causal Factors Analysis (CFA). The "why" in ΔW_g is determined using some of the organizational and culture investigative CFA tools. The goal is to design, implement, and manage work and processes in a way that minimizes this gap [16].

4 Positive Train Control

Positive Train Control (PTC) is a generic term referring to a range of fully integrated technologies that overlay existing safety systems to prevent train-to-train collision and improve worker safety. The current PTC system gives the notice of impending penalty brake application if the train approaches a speed-limiting with full speed or if the train is traveling beyond speed restrictions. If the engineer does not take the necessary action, the system brings the train to a stop with a full-service brake application. It also prevents the train to move beyond the speed restrictions. The conventional safety approach used signal systems with colored signs along the track, and daily bulletin reports to manage the speed of the train. If the train travels through unsignaled (dark) or automatic signal territories, movement authorities are transmitted to and confirmed with train crews over the analog voice radio system [17].

The current state of the train operation in most railroad organizations is Centralized Train Control (CTC). In CTC territory, authority for train movements and track occupancy is verbally exchanged between the dispatcher and train crew over the radio. A movement authority consists of determining a safe point to which a train can travel, for example, an absolute signal displaying a stop aspect [18]. General Code of Operating Rules (GCOR) governs these exchanges of information [19]. CTC uses electrical track circuits to determine the train location. In PTC systems in addition to electrical track system and wayside unites, a Global Positioning System (GPS) is used for tracking train movements [18].

5 Incorporation of HRO Principles in PTC Implementation

High-risk organizations must be able to flexibly reconfigure and synchronize their system elements to meet the challenges arising from the introduction of new technology. They need to use a mechanism to reinforce the weak interconnections to address such a complex integration [20]. Research on organizational culture and safety outline four HRO practices to deliver high reliability in an organization: 1) manage the system not the parts, 2) reduce variability in HRO system, 3) foster a strong culture of reliability, and 4) learn and adapt as an organization [21]. We briefly address these HRO practices in the context of PTC implementation.

Managing the System not the Parts. PTC is a collision avoidance technology that uses a communications network of digital data links to coordinate the activities of locomotives, wayside units, and dispatch centers. The PTC human subsystem includes the engineer, the dispatcher, and maintenance of the way who are interacting with the technological subsystems. Since railroad organizations share at least segments of the track, they need to develop an interoperable network to have a fully operational PTC system.

Reduce Variability in the HRO System. System reliability is defined as "the lack of unwanted, unanticipated, and unexplainable variance in performance" [22]. In other words, minimizing ΔW_g. Variances in performance of the train operation under PTC can be defined as an uninterrupted/successful PTC performance during a train run, and train delays due to PTC related issues. The results of our study show that the factor that has caused the most delay was "consist error". Train consist is the lineup or sequence of train cars. In order for the PTC system to operate the crew needs to confirm that the PTC system is displaying the most current train consist. If there are any discrepancies, the train crew need to get authorization from the dispatcher to update the information. The consist error is mainly a human and organizational issue, and the only approach for resolving this issue is through communication between the crew and the dispatcher, and well-established organizational procedures.

Foster Strong Culture of Reliability. Research on organizational culture and safety has outlined developing a system of process checks to spot expected and unexpected safety problems as one of the main processes that are useful in developing HROs [23]. To integrate the HRO principle into PTC operations, we adopted the hallmarks of high reliability [24], as a guide for each HRO principle in PTC operations. In developing the checklists, we incorporated the findings of the interviews and site visits that conducted as part of the PTC implementation project, as well as the findings of the research studies to customize the checklist for PTC operations [25].

Learn and Adapt as an Organization. Most practical learning is by trial and error; however, it is not feasible in railroad operations due to their complexity, tight coupling, and the dangerous outcome. Organizations often learn as much about themselves and their internal relationships as they learn about critical events and near misses. HROs actively seek to learn what they do not know, design systems to disseminate relevant knowledge relating to a problem available to everyone in the organization, learn rapidly and efficiently, train staff to recognize and respond to system abnormalities, empower staff to act, and design redundant (sub-)systems to anticipate problems. This step highlights the importance of continuous learning in the implementation of new technology.

6 Conclusion

Railroad organizations are mandated to implement PTC to improve the reliability of the railroad operations. However, implementing the technology alone will not guarantee such results. On April 3, 2016, Amtrak train 89 struck a backhoe with a worker inside while traveling 99 mph near Chester, Pennsylvania. Two Amtrak employees were killed and 39 passengers were injured as the result of this accident [26]. PTC is designed to prevent the incursion of the train into the work zone, and that section of the track was also equipped with PTC [27]. However, the accident was not prevented because the equipment used for maintenance, as the backhoe in this accident, is not detectable by the PTC system and needs to be registered in the systems and the tracks should be manually shunted. This accident highlights the importance of human and organizational factors in the successful implementation of the PTC system.

Research shows the technical change is always accompanied by organizational change and to ensure a successful implementation, these two must be aligned. It is a proven fact that the HRO characteristics constitute the "secret of success" for a safe, sustainable, and result-oriented system, which must operate in a safety-sensitive, non-routine, and rapidly changing environment. Implementing these principals from the design stage of the system, would reinforce the pillars of the organization and enhance resiliency.

References

1. Majchrzak, A., Meshkat, N.: Aligning technological and organizational change. In: Handbook of Industrial Engineering, Technology and Operations Management, 3rd edn., ch. 36, pp. 948–974. Wiley-Interscience Publication (2001)
2. Bea, R., Mitroff, I., Farber, D., Foster, H., Roberts, K.H.: A new approach to risk: the implications of E3. Risk Manag. **11**(1), 30–43 (2009)
3. Long, R.J.: Human issues in new office technology. In: Computers in the Human Context: Information Technology, Productivity, and People, T. Forester. MIT Press, Cambridge (1989)
4. Black, L.J., Carlile, P.R., Repenning, N.P.: A dynamic theory of expertise and occupational boundaries in new technology implementation: building on barley's study of CT scanning. Adm. Sci. Q. **49**(4), 572–607 (2004)
5. DeSanctis, G., Poole, M.S.: Capturing the complexity in advanced technology use: adaptive structuratio theory. Organ. Sci. **5**(2), 121–147 (1994)
6. Barley, S.R.: Technology as an occasion for structuring: evidence from observations of CT scanners and the social order of radiology departments. Adm. Sci. Q. **31**, 78–108 (1986)
7. Griffith, T.L., Northcraft, G.B.: Cognitive elements in the implementation of new technology: can less information provide more benefits? MIS Q. **20**(1), 99–110 (1996)

8. Ciborra, C.U., Lanzara, G.F.: Formative contexts and information technology: understanding the dynamics of innovation in organizations. Acc. Manag. Inf. Technol. **4**(2), 61–86 (1994)
9. Majchrzak, A., Markus, M.L.: Technology Affordances and Constraints in Management Information Systems (MIS). Encyclopedia of Management Theory (2012)
10. Tyre, M.J., Orlikowski, W.J.: Windows of opportunity: temporal patterns of technological adaptation in organizations. Organ. Sci. **5**(1), 98–118 (1994)
11. Weick, K.E., Sutcliffe, K.M.: Managing the Unexpected: Assuring High Performance in an Age of Complexity, Jossey-Bass (2001)
12. Carnes, E.: Highly reliable governance of complex socio-technical systems. In: The Macondo Blowout 3rd Progress Report, pp. 135–165 (2010)
13. Meshkati, N., Khashe, Y.: Operators' improvisation in complex technological systems: successfully tackling ambiguity, enhancing resiliency and the last resort to averting disaster. J. Contingencies Crisis Manag. **23**(2), 90–96 (2015)
14. Placencia, G., Meshkati, N., Moore, I.J., Khashe, Y.: Technology and high reliability organizations in railroad operations safety: a case study of Metrolink/SCRRA and Positive Train Control (PTC) Implementation. In: ASME/IEEE Joint Rail Conference, Colorado Springs, Colorado, USA (2014)
15. IAEA, Safety Culture in Pre-Operational Phases of Nuclear Power Plant Projects - Safety Reports Series 74. International Atomic Energy Agancy (2012)
16. Hartley, R.S.: High reliability organizations and practical approach. In: CCRM HRO Conference (2011)
17. Ditmeyer, S.: Confused about PTC yet? Trains **71**(10), 24–31 (2011)
18. Southern California Regional Rail Authority (Metrolink), SCRRA Positive Train Control System Concept of Operations (2010)
19. Federal Railroad Administration, Human Reliability Analysis in Support of Risk Assessment for Positive Train Control, U.S. Department of Transportation, Cambridge, MA (2003)
20. Khashe, Y., Meshkati, N.: Evaluation of human-systems interaction of the PTC technology in railroad, a preliminary study. In: Proceedings of the 2015 Industrial and Systems Engineering Research Conference (2015)
21. Hartley, R.S.: High reliability organization implementation. In: 8th Annual HRO Conference, Fort Worth, TX (2014)
22. Hollnagel, E.: The four cornerstones of resilience engineering. In: Resilience Engineering Perspective, Preperation and Restoration ed., vol. 2, Farnham, Surrey: Ashgate, pp. 117–133 (2009)
23. Wong, D.S., Desai, V.M., Madsen, P., Roberts, K.H., Ciavarelli, A.: Measuring organizational safety and effectiveness at NASA. Eng. Manag. J. **17**(4), 59–62 (2005)
24. Weick, K., Sutcliffe, K.: Managing the Unexpected - Sustained Performance in a Complex World, 3 edn. Wiley, Hoboken (2015)
25. Khashe, Y., Meshkati, N.: Human and organizational factors of positive train control safety system - the application of high reliability organizing in railroad. In: Proceedings of the Human Factors and Ergonomics Society 2019 Annual Meeting, Seattle, WA (2019)
26. National Transportation Safety Board, Amtrak Train Collision with Maintenance-of-Way Equipment Chester, Pennsylvania April 3, 2016, NTSB (2017)
27. Laughlin, J.: Safety took a back seat at Amtrak, feds say. Then workers died, 14 Nov 2017. http://www.philly.com/philly/business/transportation/ntsb-report-amtrak-derailment-april-2016-chester-pennsylvania-20171114.html?mobi=true

Eye Movement Analysis of Interactive Interface of CRH High-Speed Train Braking Test

Jun Li(✉) and Jinyi Zhi

School of Architecture and Design, Southwest Jiaotong University, Chengdu, Sichuan, China
{li.jun,zhijinyi}@swjtu.edu.cn

Abstract. Playing a crucial role in the start of high-speed train driving, braking test is a necessary task to be conducted before train departure. In the train braking test, a driver needs to make a quick response according to a hint displayed on the human-machine interactive interface screen and then operate a driver's master handle to the designated position. This paper makes a braking test for the driver's operations of a Chinese CRH train based on the application of an eye tracker. Then, through analysis of eye movement data involved during the interaction, vision characteristics of the interaction are illustrated. The research result is of referential values to design optimization of man-machine interaction interfaces.

Keywords: High-speed train · Braking test · Eye movement

1 Introduction

Playing a crucial role in the start of high-speed train driving, braking test is a necessary task to be completed before train departure. In general, braking test is initiated by a driver and realizes interactions with a braking system by a screen and a handle, and it is a typical course of interactions between people and an automatic system.

Based on a lot of previous studies in fields of experimental psychology [1–3] and neuropsychology [4, 5], a theoretical basis has been established for relations between human eye movement and attention. As shown by many studies, the attribute of obtaining more attention plays a more important role in decision making [6, 7]. Such relationship is called as the "utility effect" and has been embodied in many scenes [8]. As found in the studies of the man-machine interaction field, the analysis of eye movements can help know about how a person distributes the attention in manipulation and how to design scene, interface, control lever, instruments, exteriors and the like for optimization of human-machine interactions.

There is a large inventory of CRH380 trains in China, and they play an important role in current actual operation [9]. In this paper, a human-machine interaction task of the braking test for a CRH380 train was researched. Drivers' braking test task was studied with an eye tracker. In this way, the research result displays vision characteristics of

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 127–132, 2020.
https://doi.org/10.1007/978-3-030-50943-9_17

interactions and provides reference for design optimization of the human-machine interaction interface.

As indicated in the operation instructions, operation steps and instructions of the braking test of CRH380D can be divided into 6 steps. A driver will see instructions 1–6 on IDU (IDU2 is close to the driver's master handle in general) in succession and then make quick responses to the instructions. The steps are given in Table 1.

Table 1. Steps of braking test for CRH380D train

Steps	Instructions on IDU Screen	Actions requested for driver
1	The driver's master handle is placed at Shift 0	The driver's master handle is checked and switched to Shift 0 if it is not put in place
2	The driver's master handle is placed at Shift 7	The driver's master handle is moved down to Shift 7
3	The driver's master handle is placed at Shift 0	The driver's master handle is moved up to Shift 0
4	The driver's master handle is placed at Shift 8	The driver's master handle is moved down to Shift 8 (emergency braking shift)
5	The driver's master handle is placed at Shift 0	The driver's master handle is moved up to Shift 0
6	Braking test is completed	The test result given by IDU is waited; and other steps will be continued if the test is passed; otherwise, failures shall be checked and another test is needed

2 Experiment

The experiment in this research was conducted on a CRH380D driving simulator which is generally used by drivers to make daily training of driving skills. In addition to the driving simulator, equipment such as an eye tracker (The model of eye tracker is Tobii Pro Glasses 2) and a video camera was used. Experiment system composition displayed in Fig. 1. In the experiment, 14 on-duty drivers aged 25–45 from a high-speed train operation station were selected as samples, with subjective reporting of good health. Those drivers were asked to put on the eye tracker glasses and complete a task on the simulator by picking up the train non-electrically from a station and parking the train and opening the door at next station according to standard regulations of train driving.

During task execution of those drivers, a research assistant would make assistant recording for the driving course by a video camera. Braking test is a must operation to be conducted after train picking at the departure station, so it is involved in drivers' operations. During this task, those drivers must quickly responded to instructions displayed on the human-machine interactive interface screen IDU2 and switched the driver master handle to a designated position.

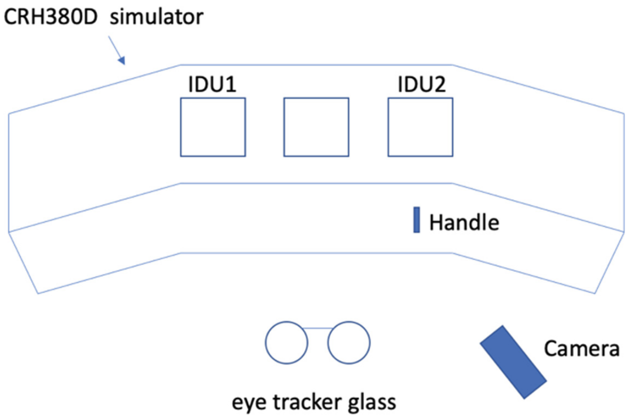


Fig. 1. Eye movement experiment system of CRH380D braking test based on a CRH380D driving simulator

3 Data

After the test, those drivers' eye movement data were obtained from the eye tracker, involving drivers' eye movement information and sounds of videos and environments in their visual field. Import the data into the eye tracker data analysis software (Tobii Pro Lab trial 1.130), and set some settings on the record by setting, the software can automatically analyze the blinking, gaze and other events.

3.1 Settings of TOI (Times of Intersest)

At first, start and stop events of the train braking test were determined according to videos and sounds contained in the eye movement data. Then, events of braking start and braking end were generated through event customization. TOI was set according to these two events, so data concerning the braking test were selected from eye movement data of the overall driving task.

Next, TOI was set for instructions. Each instruction (Listed in Table 1) was analyzed one by one again according to videos and sounds in eye movement data of the overall braking test. Through customized setting of events completed according to the instructions, eye movement data of the braking test were divided. In this way, a corresponding TOI was set for each instruction. In some experiments, the initial position of the handle is different. If it is placed in gear 0, it will cause instruction 1 to skip to instruction 2. This problem requires attention when setting a custom event.

3.2 Visualization

Eye tracker data analysis software provides data virtualization analysis tools, through which you can visually understand the behavior of eye movements in the experiment.

In the end, gaze event sequences of TOI corresponding to the tasks were overlaid. Therefore, through visualization analysis of data, a gaze plot and a attention heat map can be generated for operation eye movement data corresponding to each instruction. Then the image output of the eye-tracking data analysis software was combined according to the 6 steps and instructions in Table 1 to form Fig. 2 and Fig. 3 for further analysis.

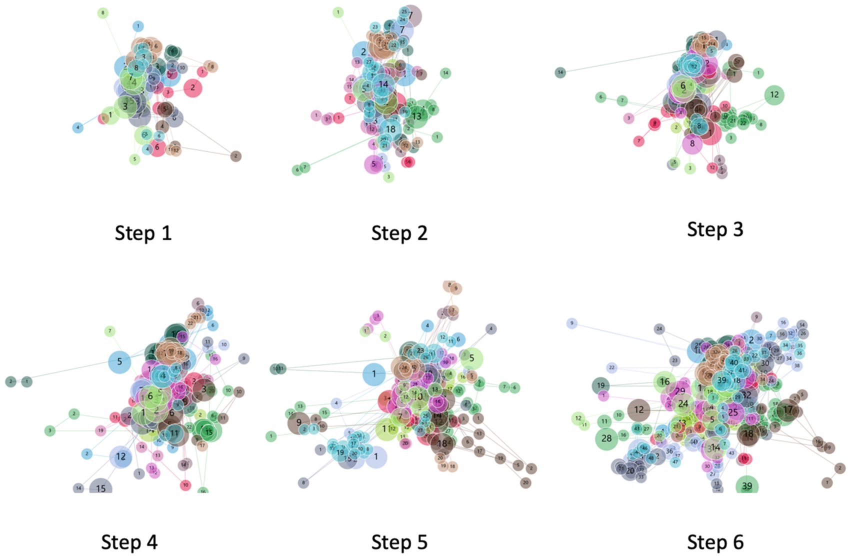


Fig. 2. Gaze stacking chart for each step in CRH380D braking test. The six steps in the picture correspond to the six instructions listed in Table 1.

4 Analysis

As found through comparison of the gaze stacking charts (Fig. 2) and attention heat maps (Fig. 3), Step 1, Step 3 and Step 5 were similar and corresponding to similar tasks (Table 1), namely driver's master handle was put at Shift 0 according to IDU instructions. Step 2 and Step 4 were similar probably because of the close distance between Shift 7 and Shift 8 of the driver's master handle. At Step 6, drivers who were waiting for results of the braking test would focus their visual field randomly.

Gaze stacking charts and visual pictures of the braking test captured by the glass were compared. The upper half of the pictures represents the IDU2 screen. The lower half is corresponding to the driver's master handle. During the whole braking course, drivers needed to keep scanning from the IDU2 screen to the driver's master handle. The scanned distances were long at Steps 2 and 4, while those at Steps 1, 3 and 5 were short. Scanning at Step 6 focused on IDU2.

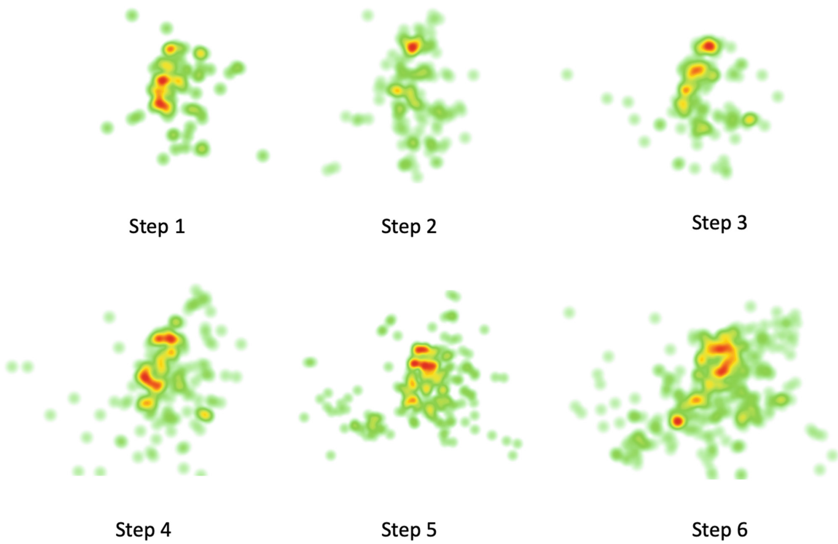


Fig. 3. Attention heatmap in CRH380D braking test. The six steps in the picture correspond to the six instructions listed in Table 1.

5 Conclusion

With regard to the layout, the interaction efficiency of braking test can be increased by the combined design of IDU2 and driver's master handle, which enables more precise and explicit display of accurate handle position information on IDU2 or additional position display beside the handle aiming at phenomena reflected by eye movements.

As for design of screen information, it is feasible to design more striking instruction information displayed on IDU2 and provide transitional hints during instruction shifting. In this way, drivers do not need to divert visual attention to other places to find data concerning system working states.

References

1. Deubel, H., Schneider, W.X.: Saccade target selection and object recognition: evidence for a common attentional mechanism. Vis. Res. **36**(12), 1827–1837 (1996)
2. Hoffman, J.E., Subramaniam, B.: The role of visual attention in saccadic eye movements. Atten. Percept. Psychophys. **57**(6), 787–795 (1995)
3. Kowler, E., Anderson, E., Dosher, B.A., Blaser, E.: The role of attention in the programming of saccades. Vis. Res. **35**(13), 1897–1916 (1995)
4. Kustov, A.A., Robinson, D.L.: Shared neural control of attentional shifts and eye movements. Nature **384**(6604), 74–77 (1996)
5. Mohler, C.W., Wurtz, R.H.: Organization of monkey superior colliculus: intermediate layer cells discharging before eye movements. J. Neurophysiol. **39**(4), 722–744 (1976)

6. Wedell, D.H., Senter, S.M.: Looking and weighting in judgment and choice. Organ. Behav. Hum. Decis. Process. **70**(1), 41–64 (1997)
7. Glockner, A., Herbold, A.: An eye-tracking study on information processing in risky decisions: evidence for compensatory strategies based on automatic processes. J. Behav. Decis. Making **24**(1), 71–98 (2011)
8. Orquin, J.L., Loose, S.M.: Attention and choice: a review on eye movements in decision making. Acta Physiol. (Oxf) **144**(1), 190–206 (2013)
9. High-speed Rail in China. https://en.wikipedia.org/wiki/High-speed_rail_in_China

Research on Optimal Design of Metro Driving Interface Based on Driver's Operating Characteristics

Yang Du, Jin-Yi Zhi(✉), Ze-Rui Xiang, and Jing Kang

School of Architecture and Design, Southwest Jiaotong University,
Chengdu 610031, China
zhijinyi@swjtu.edu.cn

Abstract. For the purposes of relief of drivers' working strength and spiritual stress and increase of working efficiency, this paper studies design of the metro train's driving interface and comes up with an optimization proposal aiming. Firstly, investigation survey is conducted on drivers' driving satisfaction; secondly, importance and use frequency of equipment are surveyed; thirdly, drivers' driving operation processes are recorded with videos, the tracks of drivers' operation processes are analyzed; fourthly, based on the above survey analysis, an optimization design proposal for driving interface layout is proposed, while man-machine verification of the proposal is conducted with the JACK virtual simulation technology; and finally, drivers' overall satisfaction with the optimization proposal is investigated. As shown in the results, the optimized proposal is advantageous in rational layout, convenient operations and clear tracks of operation processes, which can help relieve drivers' driving strength effectively and increase their satisfaction significantly.

Keywords: Metro train · Driving interface · Layout optimization · Operation process · Virtual simulation

1 Introduction

Driving interface of a train is the operation space for a driver to obtain information, make decisions, conduct instruction-oriented control of relevant systems and drive a train to complete various tasks [1]. It comprises an interface for output of equipment information to a person, an interface for equipment to obtain the information output by the person and a man-machine-environment interface. Person, console and seat are three key factors in the driving interface. The interactions between person and console are closely associated with driving tasks, so compatibility of man-machine interactions has great impacts on drivers' feelings in driving operations and travelling efficiency [2].

As early as 1980s, Swedish work physiology researchers [3] and Japanese scholars [4] already started the exploration in train cab man-machine interfaces, and found that drivers would suffer fatigue in shoulder, neck and lumbar muscle strain during long-time driving due to irrational size design of the consoles. Several years later, Chinese scholars started focusing on the research of train cabs [5–7], and found that current train driving interfaces were subject to many problems such as irrational layout of

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 133–138, 2020.
https://doi.org/10.1007/978-3-030-50943-9_18

consoles and inconformity of their structure and size with ergonomics. Most of these studies pay attention to cabs in high-speed trains, but rarely concern those in metros. Cabs in high-speed trains and metros are quite different in aspects such as layout, functions and operations. These studies mainly focus on specific problems rather than make systematic consideration of overall layout optimization of driving interfaces [8, 9]. This paper aims to propose an optimization design proposal for driving interfaces, which conform to drivers' operation features and task processes, expecting to support subsequent design from theoretical perspectives.

2 Method

2.1 Participant

After obtaining approval from Chengdu Metro Operation Co., Ltd., we made an investigation of 28 metro drivers, including 27 males and 1 female with the age of 18-40 years old and the right hand habitually used in practice. With 3 years' driving experience on average, all the participants have normal visual acuity or normal corrected visual acuity.

2.2 Design Process

Satisfaction Survey. During driving, drivers' information output behaviors are generally manifested by operation motion of effector organs (such as hand and foot), namely those drivers interact with the console in a relatively continuous and concurrent manner. Therefore, the matching degree between equipment attributes of the console and driver features can direct influence drivers' driving feelings and operation efficiency. As investigated in the survey of Chengdu metros, there are 40 console devices which are classified into 8 types based on functions: display screen, operation console, knee storage space, vehicle-mounted radio station, control devices, layout, driver controller and others. Driver features can be classified into physiological, visual and operational characteristics. Through cross analysis between equipment's physical indexes and operation modes and driver features, the *Questionnaire for Man-Machine Relationship Satisfaction between Metro Drivers and Console* was finally determined based on discussion with two industrial experts. The questionnaire is subject to a Likert scale form. Drivers were asked to give a mark for equipment according to their feelings in daily driving (1 point for dissatisfied, 2 points for generally satisfied, 3 points for satisfied, 4 points for relatively satisfied and 5 points for very satisfied). As verified, the questionnaire is of high reliability and structural validity.

Questionnaires were issued to drivers joining the experiment. As shown in the questionnaire results, those drivers are more satisfied with size and visual effects of consoles, while remain relatively unsatisfied with operations. Specifically, drivers are satisfied with the display screen, operation console and lighting, which can be maintained in optimization design of the driving interface; and in view of low scores, optimization shall be conducted on knee storage space, color and shape, vehicle-mounted radio stations and other auxiliary equipment.

Importance of Equipment – Investigation on Use Frequency. Equipment layout optimization is an important link in optimization design of metro driving interfaces. Equipment layout is correlated with criticality of equipment. Two indexes including “importance” and “use frequency” are commonly used to judge whether a device is crucial. In general, these two aspects are taken into account comprehensively though definition of “relation chain value = importance*frequency”, where the higher a chain value is, the more crucial the device will be. Through the method similar with satisfaction investigation, the *Questionnaire for Importance and Use Frequency of Metro Cab Equipment* was compiled for survey of importance and use frequency of equipment. Through screening of chain values, equipment with high chain values, such as driver controller, TCMS and door switch, are preferentially allocated within drivers' optimum operation scope; then, equipment with relatively low chain values is successively allocated; and in the end, equipment buttons with low chain values, such as window shade, car wash and sprinkling, are taken into account.

Procedure Model of Driving Operation Tasks. Metro operation is characterized by high density, high efficiency, high safety standards, etc. Metro drivers bear heavy tasks. During a complete travelling course of the metro vehicle, operation processes are quite different under different tasks. In order to analyze drivers' driving operation characteristics under different situations, a complete travelling process is divided into five tasks, namely delivery from carport, departing, main track running, returning and entry into carport. Detailed processes of these five tasks are decomposed and analyzed in combination with video records, so the decomposed process charts of each task motion are obtained. According to the decomposed task charts, operation process track charts of drivers at different task stages are drawn. During driving, drivers make visual cognition and conduct operations nearly at the same time. They are recorded separately for clearer observation. A driver visual cognition track and an operation process track are drawn respectively. The process track of delivery from carport is shown in Fig. 1, where equipment is numbered according to operation process sequences. It is thus clear that under the task of delivery from carport, drivers' visual track is centralized, but the operation process is chaotic. The frequent switching of left-hand and right-hand operations brings up drivers' workloads. It is necessary to make a new layout for the equipment in the driving interface design, so as to optimize drivers' operation process track.

3 Results and Discussion

3.1 Optimization Design of Driving Interface

A new layout is made for console equipment according to console man-machine matching satisfaction, importance and use frequency of equipment obtained in the above research as well as driving task operation processes. In combination with metro technologies and standard requirements in the region where the target train is located [10, 11], the console size is optimized based on physical data of Chinese people. A proposal is preliminarily designed for the console. The optimized driver operation process track under train delivery from carport is shown in Fig. 2.

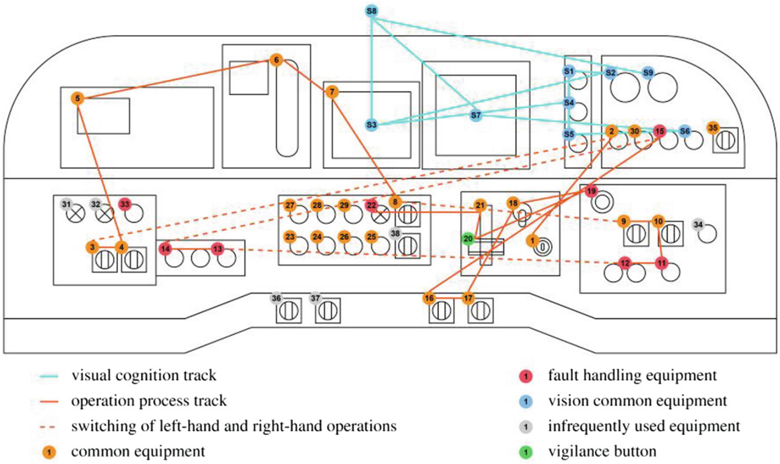


Fig. 1. Track chart for driver operation process of train delivery from carport.

Then, the man-machine size is verified with JACK visual simulation software. The physical size of the 5% males is taken as the upper limit for size design of upper limb reachable scope; and the physical size of the 95% males is taken as the lower limit for size design of limb motion scope in the operation space. Obviously, all the console devices are reachable for fingers; and the passing size of knee storage space is rational. As shown in the virtual verification result, the man-machine size in driving interface optimization design is rational and feasible.

In the end, route representative colors of the target vehicle are extracted and integrated into the concept design. On this basis, a metro driving interface design proposal which satisfies route characteristics and conforms to orientations of Chengdu Metro Line 3 is proposed, and the proposal is presented in the form of computer aided design. The final proposal is shown in Fig. 3.

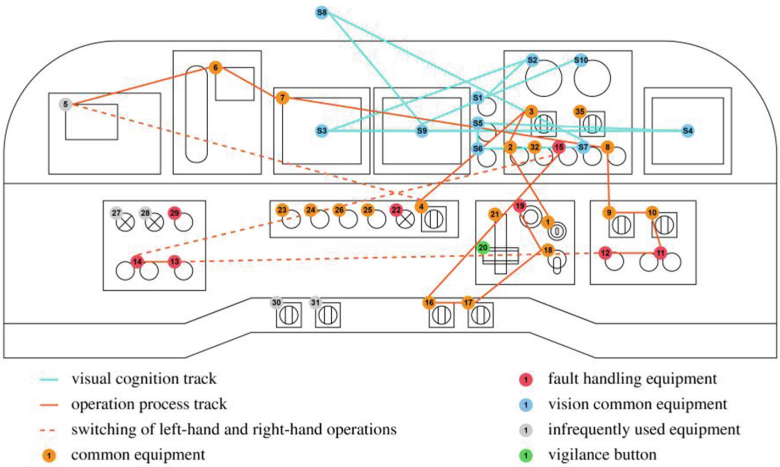


Fig. 2. Diagram of driver operation process tracks under optimized interface.

Fig. 3. Design proposal for metro driving interface.

The optimized driving interface is advantageous in the following aspects: 1. Devices with high degree of importance and low use frequency are placed on the inclined panel, so they can be quickly reached and bring no influence on daily operations; 2. Slope radian is increased, and equipment is controlled within an effective operation scope; 3. Some equipment is integrated based on the prior art (For example, lighting in guest rooms is integrated in the touch display screen) to simplify the console; 4. Driver operation processes are more concise after optimization.

3.2 Satisfaction Survey

Online questionnaires were issued to metro drivers via WeChat to investigate drivers' satisfaction with the new proposal. See the online questionnaire via the website: https://wj.qq.com/s2/5084436/0d30/. The questionnaire was issued on the Internet at 11:26 of November 29, 2019 to collect public votes. Till 17:00 of November 29, 2019, 111 drivers from metro companies in Lanzhou, Guangzhou, Zhengzhou and Wenzhou participated in the voting. As shown in the questionnaire research, 90% drivers think the optimized driving interface proposal is better than the current metro driving interface proposal.

4 Conclusion

For the purpose of improving metro drivers' operation efficiency and driving comfort, we carried out an investigation survey on man-machine matching between drivers and the driving interface under a series of conditions. A model for drivers' driving operation task processes was established through survey on man-machine matching satisfaction, investigation of equipment importance-use frequency as well as on-site data collection. Many problems existing in the current driving interface were improved properly. An optimization design proposal for the driving interface was formed in the end. As verified, this proposal can help improve drivers' overall satisfaction.

It's worth noting that due to limitations in conditions and scopes of investigation, this paper fails to make sufficient consideration on emergency settings in the metro driving interface layout. Therefore, the following outlooks are suggested:

1. It is necessary to conduct a driving course in new interactive forms, reduce multifarious types of control devices, and design a more concise and integral console.
2. Degree of freedom of the console display and seat can be enhanced. In this way, drivers can make less motion during driving. More humanized design shall be taken into account.

References

1. Guo, B.Y.: Research on the Human Compatibility Design Theory and Method of High-Speed Train Driving Interface. Beijing Jiaotong University. Beijing (2010)
2. Chen, D.J.: Ergonomic Design Method for Human-Machine Interface of Metro Train Driver's Desk. Beijing Jiaotong University. Beijing (2014)
3. Hedberg, G., Bjorkten, M., Ouchterlony-Jonsson, E., Jonsson, B.: Rheumatic complaints among swedish engine drivers in relation to the dimensions of the driver's cab in the rcengine. Appl. Ergon. **12**(2), 93–97 (1981)
4. Kogi, K., Onishi, N., Saki, K.: Requirements of a driving cab seat. In: 10th Proceedings of the Asian Conference on Occupational Health, Singapore, pp. 262–270 (1982)
5. Zhang, Q.H., Liu, D.M.: Discussion on Ergonomics of ND_5 Locomotive Driver's Cab Console. Diesel Locomotive, pp. 21–26 (1986)
6. Chen, X.G.: Discussion on optimal design of locomotive driver's cab. Diesel Locomotive **39**(12), 24–28 (2005)
7. Li, Y.: Study on the Optimized Design of Locomotive Driving Interface Based on Trainman's Properties. Southwest Jiaotong University. Chengdu (2012)
8. Zwolinski, P., Sagot, J.C.: A co-operative simulation approach to the design and evaluation of auditory interfaces in a high speed driving cab. In: The Fifth International Conference on Auditory Display. University of Glasgow, Glasgow (1998)
9. Angelos, A., Aris, P., Luisa, A., Evangelos, B.: Communication and interaction strategies in automotive adaptive interfaces. Cogn. Technol. Work **8**(3), 193–199 (2006)
10. Chen, Y.R., Tong, S.Z.: The compilation and research of "Guidelines for Ergonomic Design of Control Centers". Chin. J. Ergon. **6**(3), 19–24 (2000)
11. General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. GB10000-88, Human Dimensions of Chinese Adults. Standards Press of China. Beijing (1989)

Effect of Intercity Train Vehicle Layout on Boarding and Alighting

Chen Wang[1], Weining Fang[1(✉)], Yueyuan Chen[1], and Caifeng Li[2]

[1] State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China
{18121354,wnfang}@bjtu.edu.cn
[2] Shanghai Jiues Information Technology Co., Ltd, Shanghai 201206, China

Abstract. To achieve the operational objectives of intercity trains, in particular, reducing the station dwell time, it is necessary to improve the boarding and alighting efficiency. The vehicle layout considerably influences the boarding and alighting efficiency. In this paper, a method to simulate pedestrian dynamics is proposed. The effects of three layout factors on the efficiency are investigated; furthermore, the regression curves of different factors and cumulative high density maps are considered for analysis. The results show that when the door width ranges from 800–1900 mm, the relationship between this width and the boarding and alighting efficiency is a quadratic function. When the hall width ranges from 1300–2100 mm, the relationship between this width and the efficiency is a power function. When the aisle width ranges from 650–950 mm, the relationship between this width and the efficiency is a cubic function.

Keywords: Intercity train · Vehicle layout · Boarding and alighting efficiency · Pedestrian simulation

1 Introduction

The operating purpose of intercity trains is to satisfy the increase in the quantity and quality of passenger transport demand. Under a situation involving a large capacity and multisite stops, the key to achieve the operational objectives is to shorten the time taken by the station dwell time. Because the dwell time is mainly affected the time taken by the passengers to board and alight, improving the boarding and alighting efficiency is crucial to improve the traffic capacity and service quality. Bernhard Ruger once indicated that by optimizing the vehicle layout design of trains, the dwell time can be reduced by 1/3rd [1], thereby demonstrating the importance of the train vehicle layout. Thus, this paper focuses on the impact of the passenger room layout on the efficiency.

To effectively reduce the train dwell time, an increasing number of scholars have focused on the boarding and alighting efficiency from the perspective of the train vehicle layout. In 2013, Nigel G. Harris [2] found that the door width showed a nonlinear change at approximately 1.4 m. In 2014, Taku Fujiyama [3] noted that a wider door, corresponded to a lower efficiency. In 2014, Jungtai Kim et al. [4] observed that a door width of 1.6 to 2 m was the most suitable. In 2017, Roselle Thoreau [5] proposed that the optimal door width should between 1.7 and 1.8 m. These researchers

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 139–146, 2020.
https://doi.org/10.1007/978-3-030-50943-9_19

used actual experimental platforms to study the effects of the vehicle layout on the boarding and alighting efficiency. In 2019, Qiu H [6] used the simulation method for this investigation and observed that a wide aisle and reasonable seat pitch can promote a proper evacuation. In addition, a normal train door width has no effect on an evacuation. Qiu Hanzhao [7] proposed an experimental method to verify the effectiveness of using LEGION to simulate the evacuation process of train passengers. It shows that the simulation results are not different from the real environment. This paper is based on this research, using the method to simulate the pedestrian dynamics, which the boarding and alighting efficiency can be investigated. The findings can help reduce the labor, material and financial resources.

This paper mainly includes five parts: the first part is the introduction section; the second part describes the simulation experiment scenario of the passenger boarding and alighting based on LEGION; the third part describes the research on the influence of the factors on the boarding and alighting efficiency, and the last two parts correspond to the discussion and conclusions.

2 Passenger Boarding and Alighting Simulation Scene Developed Using LEGION

2.1 Simulation Model Construction and Parameter Input

In this paper, the TP03 vehicle of the intercity emu designed by TANGSHAN Co. Ltd. was taken as the basic research object, and the experimental simulation model was built based on the control variables such as the platform and passenger attributes of the Beijing Tianjin intercity railway operation background. To analyze the most significant passenger congestion effect, an experimental scenario with a 1:1 passenger ratio was selected to study the boarding and alighting efficiency under the rated passenger load scenario. The initial number of passengers in the vehicle was set as 96.

The input parameters of the passenger sex, age and speed were defined according to the survey results of Liu Dongdong [8] (Table 1).

Table 1. Pedestrian sex, age and velocity distribution

Sex	Age	Percentage	Average speed (m/s)	Actual input (m/s)
Male	Juvenile	0.61%	–	1.3 (10%), 1.2 (90%)
	Youth	45.28%	1.21	1.3 (10%), 1.2 (90%)
	Middle aged	13.46%	1.11	1.2 (10%), 1.3 (90%)
	Senior	1.84%	1.04	1.1 (40%), 1.0 (60%)
Female	Juvenile	0.39%	–	1.2 (50%), 1.1 (50%)
	Youth	28.72%	1.15	1.2 (50%), 1.1 (50%)
	Middle aged	8.54%	1.08	1.1 (80%), 1.0 (20%)
	Senior	1.16%	0.99	1.0 (10%), 1.1 (90%)

The boarding and alighting behavior of the passengers involves not only the instantaneous boarding and deboarding, but the entire process. According to the passenger boarding and alighting process, the interior space is divided into four areas, including the entrance, stagnant zone, interior walking space and personal space of the passengers. To board a bus, the passengers pass through these areas in turn, and they deboard in reverse. The main facilities in this process include the door, hall and aisle. Based on the ergonomic standards of human factors, the layout parameters of the mainstream intercity train passenger vehicle and the research scope of the domestic and foreign scholars, the specific research scope of the three factors is determined, as shown in Table 2.

Table 2. Space factor research scope

Variable	Range mm	Level
Door width	800,900,1000,1100,1200,1300,1400,1500,1600,1700,1800,1900	12
Hall width	1300,1400,1500,1600,1700,1800,1900,2000,2100	9
Aisle width	450,500,550,600,650,700,750,800,850,900,950	11

2.2 Establishment of Analysis Area and Determination of Analysis Indicators

To analyze the experimental results, three pairs of upper and lower analysis lines were set 130 mm from the door (thickness of the vehicle body), and the movement time of the passengers was determined. According to the definition of the passenger boarding and alighting provided by Buchmueller [9], the passenger boarding and alighting time range in this study was defined as starting with the first passenger passing through the analysis line and ending with the last passenger passing through the analysis line.

As an important part of the train dwell time, we selected the total boarding and alighting time T, mean boarding time $\overline{T_b}$, and mean alighting time $\overline{T_a}$ as the experimental indexes to measure the efficiency of the multiplication and reduction.

The total time can be calculated as

$$T = T^{last} - T^{first} \tag{1}$$

T^{last} is the last time point, determined by analyzing the time point of a line passenger; T^{first} is the first time point, determined by analyzing the time points of the line passengers.

The mean disembarkation time is calculated as

$$\overline{T_a} = \left(T_a^{last} - T_a^{first}\right)/N_a. \tag{2}$$

T_a^{last} is the last time point obtained using the downlink analysis of the passengers; T_a^{first} is the first time point obtained by considering the downlink passenger time point; N_a denotes the total number of alighting passengers.

The mean boarding time can be calculated as

$$\overline{T_b} = \left(T_b^{last} - T_b^{first}\right)/N_b \tag{3}$$

T_b^{last} is the last time point through the upline of the analysis passenger; T_b^{first} is the first passenger time point through the upline analysis; N_b denotes the total number of passengers on board.

3 Influence of the Passenger Vehicle Layout Factors on the Boarding and Alighting Efficiency

3.1 Analysis of the Influence of the Door Width on the Boarding and Alighting Efficiency

The quadratic curve fits the relationship between the door width and $\overline{T_a}$, $\overline{T_b}$ and T (Fig. 1). The left figure shows the quadratic curve of the door width, $\overline{T_a}$($R^2 = 0.9499$), and $\overline{T_b}$($R^2 = 0.9185$), and the right figure shows the quadratic curve of the door width and T($R^2 = 0.9741$).

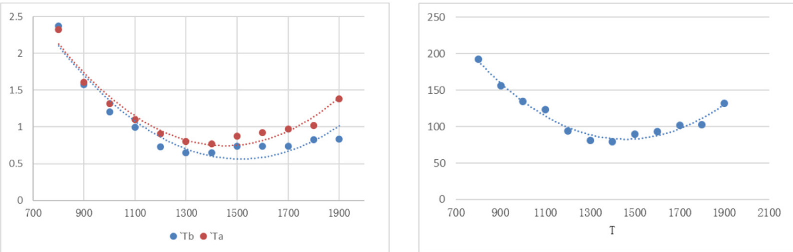


Fig. 1. Boarding and alighting efficiency curve of the door

When the door width is 800 mm, the efficiency is the lowest; with the widening of the door, the efficiency is constantly improved, and when the door width ranges from 1300 to 1400 mm, the efficiency is the highest thereby reducing gradually. To explain the phenomenon, we selected the cumulative high density maps of the door width of 800, 1300 and 1900 mm for analysis, as shown in Fig. 2. As shown in the figure, the relatively high density (red area) inside the train is 1900 mm, and the lowest is 800 mm (blue area); the highest density at the entrance is 800 mm, and the lowest is 1900 mm.

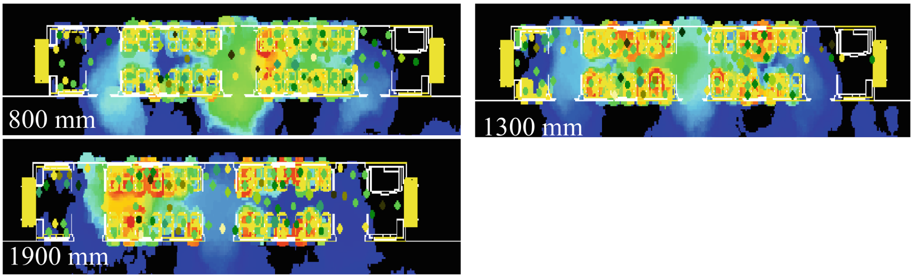


Fig. 2. Cumulative density maps for door widths of 800, 1300 and 1900 mm

3.2 Analysis of the Influence of the Hall Width on the Boarding and Alighting Efficiency

The power curve fits the relationship between the hall width and $\overline{T_a}$, $\overline{T_b}$ and T. The left figure of Fig. 3 shows the power curves of the hall width, $\overline{T_a}(\mathrm{R}^2 = 0.9374)$ and $\overline{T_b}(\mathrm{R}^2 = 0.9155)$, and the right figure shows the power curves of the hall width and $T(\mathrm{R}^2 = 0.7199)$.

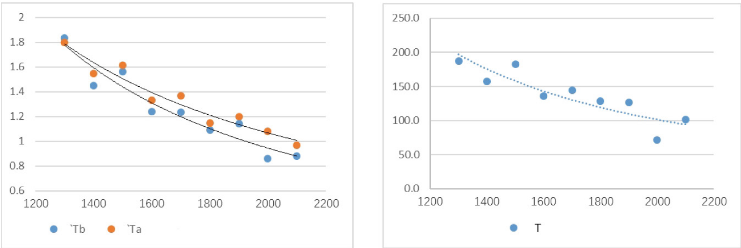


Fig. 3. Boarding and alighting efficiency curve of the hall

When the hall width is 1300–2100 mm, the three curves present a continuous downward trend. We selected the cumulative high density map of the extreme and intermediate values for analysis, as shown in Fig. 4. It is observed that when the hall width is 1300 mm, the density of the detention area is slightly high; when the hall width is 2100 mm, the density of the detention area is the lowest.

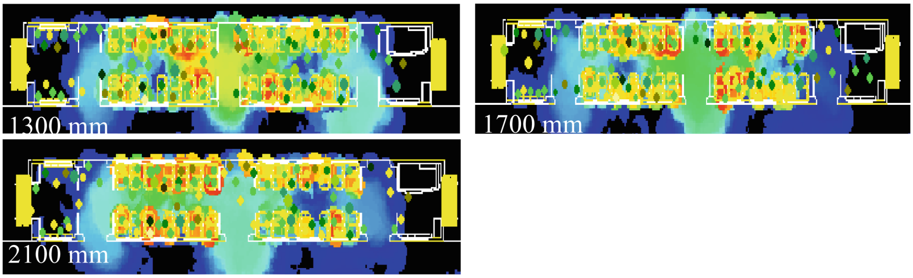


Fig. 4. Cumulative high density maps for the hallway widths of 1300, 1700 and 2100 mm

3.3 Analysis of the Influence of the Aisle Width on the Boarding and Alighting Efficiency

The cubic curve fits the relationship between the aisle width and $\overline{T_a}$, $\overline{T_b}$ and T. The left part of Fig. 5 shows the cubic curve of the aisle width, $\overline{T_a}(\mathrm{R}^2 = 0.9308)$ and $\overline{T_b}(\mathrm{R}^2 = 0.8619)$, and the right figure shows the cubic curve of the hall width and $T(\mathrm{R}^2 = 0.7994)$.

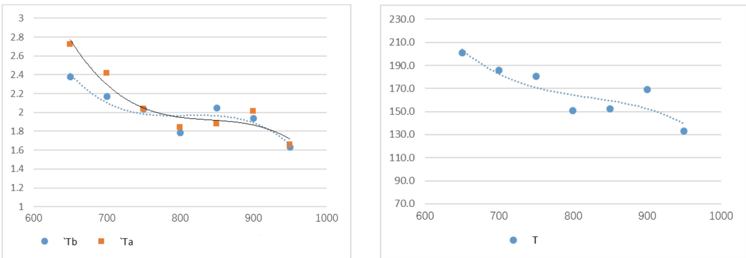


Fig. 5. Regression analysis results for the experimental data of the aisle width

The aisle width is within the range of 650–950 mm, and the three curves exhibit a downward trend later becoming stable. To explain the phenomenon, we selected the cumulative high density maps for the aisle widths of 450, 650, 800 and 950 mm for the analysis, as shown in Fig. 6. As can be seen from the figure, as the aisle widens, the density in the train decreases continuously.

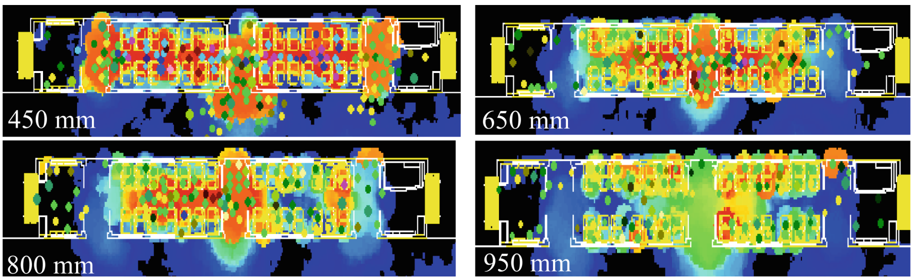


Fig. 6. Cumulative high density maps for the aisle widths of 450, 650, 800 and 950 mm

4 Discussion

The door width is affected by the hall and aisle widths. When the door width is 800 mm, it plays a leading role in limiting, while when the door width is 1900 mm, the capacity of the aisle plays a leading role in the limiting. A door width of approximately 1300 mm exhibits good coordination with the other facilities (see Fig. 2).

The detention area, which is the buffer zone for the boarding and alighting, affects the total boarding and alighting time. In the range of 1300–2100 mm, a wider hall corresponds to a higher boarding and alighting efficiency (see Fig. 4).

The possible reasons for the findings shown in Fig. 6 are as follows. When the aisle width is 450 mm, only one person can pass through. The disembarking passengers enter the aisle while coordinating within the aisle. Because the passage is extremely narrow, people cannot effectively avoid one another, and the interior of the carriage is critically blocked. In this case, the densities of the passenger personal space area and passenger walking area are high. When a large number of passengers get off the bus, the aisle with a width of 450 mm cannot meet the passengers' boarding and alighting requirements. When the aisle width is 650 mm, a conflict occurs in the middle of the aisle, and although the passengers can avoid it, the coordination time is extremely large. When the aisle width is 800 mm, the downlink passenger flow may collide with the uplink passenger flow when it arrives at the joint of the hall and aisle, resulting in a high density in the hall and low boarding and alighting efficiency. When the aisle width is 950 mm, the passengers can exit the aisle and enter the hall smoothly. At this point, the aisle capacity matches that of the door and the hall, and the passengers can promptly complete the boarding and alighting process.

5 Conclusion

To study the influence of the three factors on the boarding and alighting efficiency, we built a passenger boarding and alighting simulation model. By changing the values of the variables, the passenger boarding and alighting time at different levels of each factor was output. This study provides a reference for the train vehicle layout design in door width, hall width and aisle width. The optimal value of the door width ranges from 1300 to 1400 mm; the optimal value of the hall width is 2100 mm; and the optimal value of the aisle width is 950 mm.

Further studies will be aimed at analyzing more train vehicle layout factors and more values of the passenger ratio.

References

1. Rüger, B., Tuna, D.: Optimizing railway vehicles in order to reducing passenger change over time. J. XIII naučno-stručna konferencija o železnici. **9**(10), 25–28 (2008)
2. Harris, N.G., Graham, D.J., Anderson, R.J., Li, H.: The impact of urban rail boarding and alighting factors. In: Transportation Research Boarding Meeting, London, pp. 1–13 (2014)

3. Fujiyama, T., Thoreau, R., Tyler, N.: The effects of the design factors of the train-platform interface on pedestrian flow rates, pp. 9–95. Springer International Publishing (2014)
4. Jungtai, K., Moo, S.K., Jae, S.H., Yong, H.C., Taesik, K.: An analysis of boarding and alighting times for urban railway vehicles. J. Korean **17**(3), 210–215 (2014)
5. Thoreau, R., Holloway, C., Bansal, G.: Train design features affecting boarding and alighting of passengers: Train design affecting boarding and alighting. J. Adv. Transp. **50**(8), 1–10 (2017)
6. Qiu, H., Fang, W.: Train vehicle structure design from the perspective of evacuation. J. Chin. Mech. Eng. **32**(1), 88–100 (2019)
7. Qiu, H., Fang, W.: Effect of high-speed train interior space on passenger evacuation using simulation methods. J. Phys. A. **528**, 378–392 (2019)
8. Liu, D., Zhao, B., Li, L.: Investigation and analysis of pedestrian characteristic parameters at Beijing South Railway Station. J. Architectural Sci. **27**(5), 61–66 (2011)
9. Buchmueller, S., Weidmann, U., Nash, A.: Development of a dwell time calculation model for timetable planning. WIT Trans.State Art Sci. Eng. **1**, 105–114 (2010). G. Version 1

Vehicle-to-Infrastructure and Human-to-Infrastructure Models for Smart Civil Infrastructure Systems

Sara Mostowfi(✉) and William Glen Buttlar

University of Missouri, 65201 Columbia, MO, USA
{sara.mostowfi, Butttlarw}@mail.missouri.com

Abstract. Vehicle-to-infrastructure (V2I) technologies are known as the next generation of intelligent transportation systems (ITS), enabling data transmission between vehicles and road infrastructure. V2I communication models offer a variety of safety, mobility, and environmental applications in smart cities. Another innovative concept within the ITS is the human-to-infrastructure (H2I) communication model. H2I deals with the vulnerability of the users to damages and deteriorations of civil infrastructure. While many studies deal with V2I, limited research has been conducted in the area of interaction between humans and transportation infrastructure systems. This study presents a comprehensive review of the current studies in the arena of V2I and proposes the vital basics for establishing a robust H2I model for the next generation of the smart, civil infrastructure systems.

Keywords: Smart city · Smart infrastructure · Vehicle-to-infrastructure interaction · Human-to-infrastructure interaction · Intelligent transportation systems

1 Introduction

Nowadays, more than half of the world's population live in cities. Ten percent of this segment lives in the top 30 metropolises. The world's urbanization speed is at an unprecedented level. The problems caused by this rapid urbanization are: "waste management difficulty", "traffic congestion", "limited resources", "air pollution", "human health concerns", "deteriorated and aging infrastructures", "social and political complexity". Making cities smart is a new approach to solve problems associated with rapid urbanization [1]. A smart city may be defined in one of several ways, depending on one's perspective. One of the well-accepted definitions denotes it as a city that connects "the physical infrastructure, the information-technology infrastructure, the social infrastructure, and the business infrastructure to leverage the collective intelligence of the city" [2]. Smart civil infrastructure systems as a central component of the smart cities form the backbone of our future intelligent cities. These systems are based on the integration of necessary infrastructures like roads, bridges, tunnels, railways, and other constructed facilities with distributed smart sensing systems, information and communication technology, and the Internet of Things (IoT). Smart infrastructure

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 147–155, 2020.
https://doi.org/10.1007/978-3-030-50943-9_20

development demands developing innovative solutions such as vehicle to infrastructure (V2I) communication models. Recently, notable research has been conducted in the area of deployment of V2I technologies that can be used to alleviate congestion and improve safety without the need for large-scale, disruptive roadworks [3]. Another innovative intelligent transportation system (ITS) solution is the human-to-infrastructure (H2I) commutation model. H2I deals with several parameters that can affect the communication between human and infrastructure systems.

An example is "People as Sensors," in which data are not only collected via calibrated hardware sensors but also people can contribute by their sensation and observation in collecting data of infrastructures (air quality impressions, street damages) via smartphone [4]. Research has shown that infrastructure defects can have a significant impact on the public health and surrounding living environment of citizens. For instance, road surface anomalies such as surface roughness, unevenness, potholes, cracks, and road humps are among the top causes of car and motorcycle accidents. Besides, highway noise can contribute to many health problems such as hearing loss and physical health issues typically associated with stress (e.g., cardiovascular problems) [5]. The primary goal of this study is to present a comprehensive review of the current studies in the arena of V2I and H2I models for smart civil infrastructure systems.

2 Materials and Methods

A content analysis approach recommended by [6] is used, which allows for qualitative and quantitative operations and, therefore, providing an inclusive disclosure of the V2I and H2I applications in civil engineering. The articles were collected from well-accepted academic databases such as Web of Science, Scopus, Science Direct, ASCE Library, Engineering Village, Wiley Online Library, Sage, and Emerald. Different keywords were used to assure the related research have been included. Some keyword examples are "smart city", "V2I", "I2V", "H2I", "ITS", "smart transportation", "smart cities", "smart infrastructure", etc. The period under review is from 2008 to 2018, which led to the identification of approximately 180 candidate articles. Subsequently, a two-round article selection technique is used in this study. Accordingly, titles, abstracts, and keywords of the noted articles are checked in the first round to assure that they fall within the scope of the current literature review. Then the second round consisted of reading and analyzing the entire article was done to ensure that all the selected papers are closely related to the review objective [7]. Finally, 15 articles are selected and used for the present review study.

3 Results

The most prominent applications of V2I, V2V, and V2X with respect to smart city has been reviewed (see Table 1).

Table 1. Collected articles from mentioned search engines

Author	Focus	Research area/Contribution
1. Benzaman and Sharma [8]	Integration of V2V into a simulation model for checking traffic flow performance in intersections which shows the efficiency of traffic light availability model in decreasing the total waiting time in comparison to space availability approach	**Traffic performance, V2I, V2V** **Contribution:** Opening more opportunities in the field of traffic simulation modelling in terms of utilizing road segment resources as a constraint
2. Djahel et al. [9]	Introduction an approach based on V2I communication in which the connection between road-side facilities, and traffic light cycle information is available for approaching vehicles to determine their optimal speeds, and any other necessary actions to crossroad intersections with the lowest delay and no stops	**Traffic performance, roadside communication, V2I** **Contribution:** Providing the Belief-Desire-Intension architecture modelling for vehicle decision making
3. Aldegheishem et al. [10]	Developing a new protocol, named as the Traffic Accidents Reduction Strategy (TARS), for Vehicular Ad-hoc NET works (VANETs) to minimize the number of road accidents, decrease the death rate caused by road accidents, and for the successful deployment of the Intelligent Transportation System (ITS)	**Traffic congestion** **Contribution:** Forecasting the probability of the occurence of an accident in advance before it occurs. It also re-rotes vehicle traffic to prevent traffic jams on the road that may cause accidents
4. Parrado and Donoso [11]	Introducing an algorithm which has been developed based on the global congestion in vehcular networks. In this network, the vehicle's route assignment and recalculation service are done in nodes	**Traffic congestion** **Contribution:** Efficiency and fairness
5. Lingala et al. [12]	Introducing a new "Geo broadcasting" based strategy which makes traffic control possible through V2V and V2I communications by suing vehicular Ad-hoc Networks (VANET)	**Traffic control, V2I V2V** **Contribution:** Broadcasting the prior information about the status of road for controlling the traffic congestion

(*continued*)

Table 1. (*continued*)

Author	Focus	Research area/Contribution
6. Eder and Wolf [13]	V2I has been introduced as a contributor to the better understanding of the V2X which can be combined with further traffic light system to develop more complex traffic scenarios	**V2X, Traffic performance** **Contribution:** Providing an overview over V2X in general and related technologies
7. Olaverri-Monreal et al. [14]	A traffic light assistance system (TLA) has been developed to recommend an optimal velocity based on the acceleration and speed for a vehicle approaching an intersection	**Traffic control** **Contribution:** Provides speed advice regarding optimal velocity when approaching an intersection controlled by a traffic light
8. Rakha et al. [15]	The "Eco-cooperative Adaptive Cruise (EcoCACC)" is a system which calculates the fuel-efficient speed of the received data in V2I from the traffic signal controller based on "Signal Phasing and Timing (SPaT)"	**Traffic control** **Contribution:** Minimizing the total fuel consuming in vehicles
9. Dey et al. [16]	Introducing an optimal method which utilizes "Het-Net communication" for two CVT including: "traffic data collection", and "forward collision warning"	**Traffic data collection** **Contribution:** Providing additional connectivity beyond DSRC range to collect traffic data for a larger traffic network, which is required for many CVT applications. This study is the first of its kind to evaluate the performance of Het-Net for a seamless V2I and V2V communication for CVT applications
10. Bilgin et al. [17]	Smart meter has been suggested as a new solution for getting data about busses daily commuting by combining Vahicular Ad Hoc Network (VANET) and smart grid communication	**Vehicular Ad Hoc Network (VANET)** **Contribution:** The evaluation show that the proposed data collection mechanism achieves high delivery ratio and low end-to-end delay when the DSR routing protocol is used
11. Li and Shahidehpour [18]	The framework has been designed to safety faciliate the cooperation among drives and between drives and the traffic management authority	**Traffic management** **Contribution:** Analyze the prevailing traffic congestion circumstances with the objective of ensuring the traffic efficiency and safety at the presence of cyber incidents

(*continued*)

Table 1. (*continued*)

Author	Focus	Research area/Contribution
12. Ferdowsi et al. [19]	Suggesting an analytics for ITS in which data is transmitted and exchanged in the vehicle or roadside smart sensor level	**V2I** **Contribution:** Deep learning techniques to solve a plethora of edge analysis challenges in ITSs that include measurements heterogeneity, path planning, autonomous vehicle control, platon control, semi-autonomous ITSs, and cyber-physical security
13. Richter et al. [20]	Developing the creation of virtual models that correspond to interaction hot spots on the Austrian road network-from digitizing the infrastructure, to calibrating a simulation scenario with congruent traffic measurements-while it concludes with the evaluation of scenario simulation results	**Digital infrastructures** **Contribution:** This research demonstrates that properly scaled adaptions of the road infrastructure can ease the transitition towards increasing automation of vehicles
14. Karaaslan, E. et al. [21]	Creating a smart, human-centered method that offers significant contributions to infrastructure inspection, maintenance, management practice, and safety for the bridge owners. By developing a smart mixed reality (MR) framework, which can be integrated into a wearnable holographic headset device, a bridge inspector, for example, can automatically analyze a certain defect such as a crack that he or she sees on a element, and display its dimension information in real-time along with the condition state	**Infrastructure management** **Contribution:** Explaining how a framework for collective human-AI intelligence can be created and how it can outperform the conventional or fully automated concrete inspections
15. Whyte [22]	Providing a background to systems integration; articulates the challenge of civil infrastructure in the 21st century; and reviews the state-of-the-art in the research on systems integration in the delivery and operation of civil infrastructure	**Civil infrastructure operation** **Contribution:** Suggesting a new agenda for research on system integration within civil infrastructure

4 Discussion

The literature revealed that there is an increasing interest in developing a more human-centered civil infrastructure in which communications is not only possible between vehicle and infrastructure but also between human and infrastructures. The focus of most articles in the literature was on controlling traffic congestion, improving traffic performance, and minimizing road accidence through V2I, V2V, and V2X.

Road deterioration not only has a negative effect on ride quality but also affect vehicle operating costs (VOCs), travel times, and traffic congestion [23]. On the other hand, health issues arise from infrastructure defects contribute to many health problems. There is not also any specific statistic that illustrates the health issues cost on human health. A research done by [24] in Sweden revealed that back problems costs in Sweden had exceeded SEK 20 billion per year. Furthermore, whole-body vibration (WBV) as a leading source of back problems is lacked in general health statistics. Similarly, the total cost of low back pain (LBP) had exceeded $100 billion per year in the US [25]. Figure 1 is the estimated cost on users, which will be incurred due to pavement damage and deterioration of health (Fig. 2).

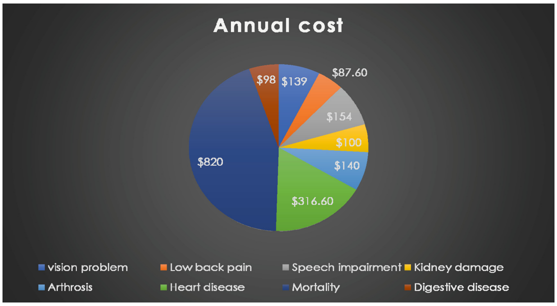


Fig. 1. Approximate annual costs of civil infrastructure defects on human health in billion annually in US (developed by the author)

Among different road features, road roughness and textures have been known as the indirect source of the noise, infrasonic sounds, whole-body vibration, and stress on road users. About road roughness and associated health effects, there isn't any specific research to reveal their correlation. However, [26] investigation revealed that the whole-body vibration (WBV) exposure differs across various road types. In other words, road and vehicle types are two essential factors in transmitted vibration to the human body [27]. The reason WBV is having an adverse impact on human health is that each body organ has its natural frequencies, the vibration makes organs to excite their natural frequencies and allowing the free oscillation [28]. Table 2 represents the vibration frequency and related health issues.

Table 2. Disease and their resonance frequency range adopted from [29]

Resonance frequency	Human body part	Associated symptomatology
1–4 Hz	Respiratory system	Dyspnea symptomatology
1–10 Hz	Ocular area	Vision ability decreasing
4–6 Hz	Brain	Sleeping and attention loss
4–8 Hz	Inner ear, Heart	Equilibrium disorder and health disease
20–30 Hz	Vertebral Column	Cervical and low back pain
20–40 Hz	Ocular area	Decrease of image focusing ability

Overall, the vibration effects on the human body can range from pure motion sickness to severe discomfort, organ failure, or tissue degeneration. Among them, low back pain is the most pronounced and common effect of the WBV, leading to the degeneration of the little cartilage (intervertebral) discs, allowing tissues and nerves to be strained and pinched.

The extent to which the human body is affected by WBV depends on the vibration frequency, duration time, body part exposed to vibration [29]. Furthermore, pavement condition is among the most prominent risk factors contributing to accidents. Statistics have introduced pavement as a contributing factor in:

– 31% of total crashes;
– 52% of fatalities (resulting from accidents), and
– 38% of non-fatal injuries (resulting from accidents) [30].

5 Conclusion

In this study, we first highlighted the applications of the V2I and I2V technology within the smart city concept. Results indicated that most V2I works had been dedicated to solving the traffic congestion issues, which can result in the reduction of CO2 emission and less oil and gas consumption and saving citizens time. Although traffic congestion and other applications of the V2I are prominent topics to be done for current cities and future cities, the important role of citizens in a smart city has been underestimated. Our literature review revealed that there is not currently any comprehensive model that addresses the interaction of humans with infrastructure. Cities are in the needs of developing such models to ensure the health of their citizens and infrastructure. Therefore, by promoting the human infrastructure model, we will be able to qualify the health of our citizens.

In this study, a novel, human-infrastructure model was developed, which will possibly be benefitting planners and citizens of smart cities. The collected data from both infrastructure and human can be utilized in the following areas:

- Clarifying deterioration effect on human health;
- Health monitoring of the infrastructures and human;
- Collect and publish online pertinent data on traffic congestion, and weather, and
- Data exchange for ITS applications, as well as connected and automated vehicles.

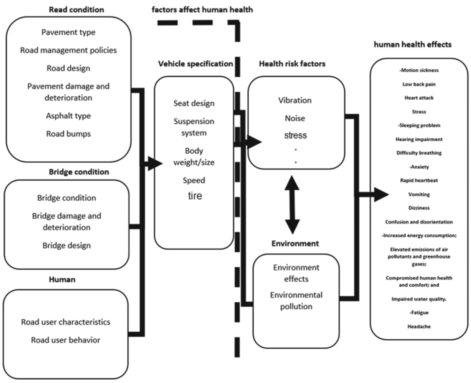


Fig. 2. Human-infrastructure interaction (developed by the author)

Therefore, making our transportation infrastructure systems (roads and bridges) smarter will allow countries to run government services more efficiently and cost-effectively. Besides, the establishment of a robust H2I interaction model paves the road for future theory expansion and practical developments.

References

1. Nam, T., Pardo, T.: Smart city as urban innovation. In: Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance, ICEGOV 2011 (2011)
2. Harrison, C., Eckman, B., Hamilton, R., Hartswick, P., Kalagnanam, J., Paraszczak, J., Williams, P.: Foundations for smarter cities. IBM J. Res. Dev. **54**(4), 1–16 (2010)
3. Devilly, O., Margaret, O.: Assessing the benefits of installing v2i communications on an 1 urban orbital motorway in a medium sized european city 2 (2018)
4. Resch, B.: People as Sensors and Collective Sensing-Contextual Observations Complementing Geo-Sensor Network Measurements. Springer, Heidelberg (2013)
5. Evans, G.W., Bullinger, M., Hygge, S.: Chronic noise exposure and physiological response: a prospective study of children living under environmental stress. Psychol. Sci. **9**, 75–77 (1998)
6. Krippendorff, K.: Content Analysis: An Introduction to its Methodology. Sage Publications, Thousand Oaks (2018)
7. Salehi, H., Burgueño, R.: Emerging artificial intelligence methods in structural engineering. Eng. Struct. **171**, 170–189 (2018)
8. Benzaman, B., Sharma, D.: Discrete event simulation of a road intersection integrating V2V and V2I features to improve traffic flow. In: 2017 Winter Simulation Conference (WSC) (2017)
9. Djahel, S., Jabeur, N., Barrett, R., Murphy, J.: Toward V2I communication technology-based solution for reducing road traffic congestion in smart cities. In: 2015 International Symposium on Networks, Computers and Communications (ISNCC) (2015)
10. Aldegheishem, A., Yasmeen, H., Maryam, H., Shah, M.A., Mehmood, A., Alrajeh, N., Song, H.: Smart road traffic accidents reduction strategy based on intelligent transportation systems (tars). Sensors **18**(7), 1983 (2018)

11. Parrado, N., Donoso, Y.: Congestion based mechanism for route discovery in a V2I-V2V system applying smart devices and IoT. Sensors **15**(4), 7768–7806 (2015)
12. Lingala, T., Galipelli, A., Thanneru, M.: Traffic congestion control through vehicle-to-vehicle and vehicle to infrastructure communication. Int. J. Comput. Sci. Inf. Technol. (IJCSIT) **5**(4), 5081–5084 (2018)
13. Eder, M., Wolf, M.: V2X communication overview and V2I traffic light demonstrator (2017)
14. Olaverri-Monreal, C., Errea-Moreno, J., Díaz-Álvarez, A.: Implementation and evaluation of a traffic light assistance system based on V2I communication in a simulation framework. J. Adv. Transp. **2018**, 1–11 (2018)
15. Rakha, H., Chen, H., Almannaa, M., Kamalanathsharma, R., El-Shawarby, I., Loulizi, A.: Field testing of eco-speed control using V2I communication. http://hdl.handle.net/10919/72260. Accessed 30 Jul 2018
16. Dey, K., Rayamajhi, A., Chowdhury, M., Bhavsar, P., Martin, J.: Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication in a heterogeneous wireless network – Performance evaluation. Transp. Res. Part C Emerg. Technol. **68**, 168–184 (2016)
17. Bilgin, B., Baktir, S., Gungor, V.: A novel data collection mechanism for smart grids using public transportation buses. Comput. Stan. Interfaces **48**, 19–29 (2016)
18. Li, Z., Shahidehpour, M.: Deployment of cybersecurity for managing traffic efficiency and safety in smart cities. Electr. J. **30**(4), 52–61 (2017)
19. Ferdowsi, A., Challita, U., & Saad, W. (2017). Deep learning for reliable mobile edge analytics in intelligent transportation systems. arXiv preprint arXiv:1712.04135
20. Richter, G., Grohmann, L., Nitsche, P., Lenz, G.: Anticipating automated vehicle presence and the effects on interactions with conventional traffic and infrastructure. In: EPiC Series in Computing, vol. 62, pp. 230–243 (2019)
21. Karaaslan, E., Bagci, U., Catbas, F.N.: Artificial intelligence assisted infrastructure assessment using mixed reality systems. Transp. Res. Rec. **2673**(12), 413–424 (2019)
22. Whyte, J.: The future of systems integration within civil infrastructure: a review and directions for research. In: INCOSE International Symposium, vol. 26, no. 1, pp. 1541–1555, July 2016
23. Egan, M., Petticrew, M., Ogilvie, D., Hamilton, V.: New roads and human health: a systematic review. Am. J. Public Health **93**(9), 1463–1471 (2003)
24. Múčka, P., Granlund, J.: Is the road quality still better? J. Transp. Eng. **138**(12), 1520–1529 (2012)
25. Crow, W.T., Willis, D.R.: Estimating cost of care for patients with acute low back pain: a retrospective review of patient records. J. Am. Osteopath. Assoc. **109**(4), 229–233 (2009)
26. Thamsuwan, O., Blood, R.P., Lewis, C., Rynell, P.W., Johnson, P.W.: Whole body vibration exposure and seat effective amplitude transmissibility of air suspension seat in different bus designs. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, no. 1, pp. 1218–1222, September 2012. SAGE Publications, Los Angeles (2012)
27. Burdzik, R., Konieczny, Ł., Adamczyk, B.: Automatic control systems and control of vibrations in vehicles car. In: International Conference on Transport Systems Telematics, pp. 120–129. Springer, Heidelberg, October 2014
28. Mohideen, A.: Reduction of whole-body vibration transmitted to test drivers performing complete vehicle durability testing at Hällered Proving Ground. Master, Chalmers University of Technology (2018)
29. Archer, W.: Whole-body vibration: causes, effects and cures. In: NAPIT (2018)
30. BloodyPavement, khafagy (2018). https://www.slideshare.net/MonaKhafagyMScPE/bloody-pavement. Accessed 3 Aug 2018

Driving Automation

Online Feedback Control for Driver-Vehicle Interaction in Automated Driving

Khazar Dargahi Nobari(✉), Franz Albers, Katharina Bartsch, and Torsten Bertram

TU Dortmund University, Otto-Hahn-Str. 8, 44227 Dortmund, Germany
{khazar.dargahinobari, franz.albers, katharina2.bartsch, torsten.bertram}@tu-dortmund.de

Abstract. Driver assistance systems have been in use for a decade and the automated vehicles are expected to hit the market soon. Collaboration between drivers and assistance systems, especially in SAE Level 3 of driving automation, plays a significant role as it is directly related to driving safety and the acceptance of automated vehicles. This contribution proposes driver state feedback control as a possible method for taking the driver state into account in driver-vehicle interaction. The feedback control creates a loop in which the takeover request affects the driver state and the driver state adjusts the takeover request. The functionality of the proposed interaction method is examined in an exemplary experiment on a driving simulator with twenty participants in which the gaze direction of drivers acts as a sensory state. The results indicate an improvement in the performance of drivers during the takeover situation by involving driver state in the design of the takeover request.

Keywords: Takeover request · Driver state · Driver-vehicle-interface · Selective attention · Workload · Human factors

1 Introduction

Driver-vehicle interaction (DVI) plays a significant role especially in the first generation of automated vehicles in which the driving task is assigned to both the human driver and the automated system. Efficient collaboration between the human driver and the automated system depends on having appropriate communication to gain a mutual understanding of the driving situation, the state, and the intention of both sides.

Depending on the automation level, the task distribution between driver and system changes. Consequently, the aim of the DVI is changed and its design is reformed. The interaction in SAE Level 1 and 2 aims to keep the driver attentive to the driving situation during the whole drive, whereas in SAE Level 3, the interaction tries to bring the driver back into the loop when the system is unable to continue driving utilizing a takeover request (TOR). In SAE Level 4 and 5, the main idea of interaction is to provide a comfortable drive for the driver (passenger) and increase the trust of the driver on the automated vehicle. Generally, in the higher automation levels, DVI aims at increasing the comfort of the drivers; however, in the lower automation levels, DVI

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 159–165, 2020.
https://doi.org/10.1007/978-3-030-50943-9_21

is assumed to provide higher safety and hence a careful design of interfaces is crucial in lower levels.

DVI consists of obtaining data from the driver and exposing stimuli to the driver. Several factors are involved in designing proper stimuli. Modality, direction, time, and type of stimuli are some of the influencing factors that affect the performance of drivers [1, 2]. A survey done by Bazilinskyy et al. [3] indicates the necessity for considering situation criticality in designing TORs. According to this survey, urgent situations require multimodal TORs, while auditory warnings are preferred for low-urgent situations.

Several researchers have investigated the driver state and its influence on driving performance during takeover scenarios [4, 5]. Lapoehn et al. [6] have recommended integrating the driver state assessment in the interaction design. In [7], drivers receive feedback on their performance after driving and then the effect of feedback on their perceived ease of use is considered. However, in this study, the feedback is not online and it does not improve driver performance. The present contribution proposes a general framework for an online feedback control that aims to improve driver performance by providing a comprehensive interaction between driver and automated system.

In the next section, the structure of feedback control for the DVI concept is investigated in detail. The application of the proposed approach is explained in Sect. 3 within an exemplary scenario on a driving simulator and the results are depicted. Section 4 includes a discussion of results, limitations, and possible next steps.

2 Structure of Feedback Control in Interaction

The driver state is a description of the driver's physical and mental condition with respect to the driving task. Marberger et al. [8] split the driver state into sensory, motoric, and cognitive states. The performance of the driver depends on the current driver state. Hence, to interact with human drivers effectively and to increase situation safety in critical driving situations, in which task transition between the automated system and the human driver is unavoidable, a comprehensive understanding and a constructive influence on the driver state is of crucial importance for the automated system.

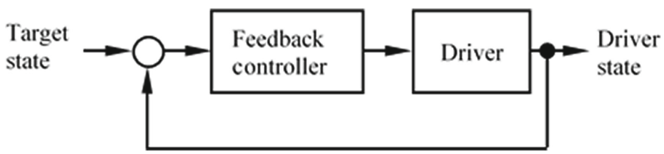


Fig. 1. Driver state feedback control

The goal of the driver state feedback controller is to include the driver state in the generation of stimuli that are exposed to the driver. Figure 1 depicts the regulation process. The controller detects the driver state via driver sensors (e.g., eye-tracker, physiological sensors) and compares it with a target state that is the desired state defined according to the current driving situation and the usual characteristics of a human driver. The difference between the driver state and the target state is then implemented in combination with the situation criticality to regulate the stimuli. The regulation loop is repeated at each time step to meet the desired driver state. The regulation intends to prescribe or proscribe behaviors to prevent driving failures. However, the driver state is not always controllable and a backup strategy should be planned for such situations.

Situation Criticality. A TOR should be able to reflect the situation criticality at each moment until the driver perceives the driving situation correctly. Situation criticality can be influenced by several factors, such as weather conditions, traffic density, and speed. Distance to an obstacle, which is capable of being a danger, is a popular measure for determining situation criticality. A shorter distance indicates higher criticality. While in some researches this relationship is determined linearly, Wang et al. [9] have assumed a power function form for the criticality with respect to the distance between ego-vehicle and hazardous obstacle.

In general, hazardous objects can be divided into dynamic and static obstacles. Althoff et al. [10] have grouped dynamic obstacles due to their behavior into the classes car, truck, motorcycle, bicycle, and pedestrian and treated them in a separate manner. Static obstacles can cause collision such as a parked vehicle, or only impose limitations on traffic behavior such as traffic signs and lane markings [9].

Target State. The target state determines the best driver state according to the current driving situation and the characteristics of the driver. Within the driving context, the sensory state of the driver is mainly defined as the driver's gaze direction and is situation dependent. In critical situations, the target sensory state of the driver is to look in the direction of an upcoming hazard, which indicates the attraction of selective attention and gradually increases situation awareness. The target motoric state can be identified as hands on steering wheel or feet on pedals concerning the driving situation and individual's preferences. In the target cognitive state, several variables must be taken into account, the determination of which is more complex. The preferred target cognitive state is to make the driver attentive towards the driving scene, increase the driver's situation awareness, and remove any irrelevant workload on the driver.

The measurement of sensory and motoric states is done directly by eye-trackers and cameras. The cognitive state, however, has only indirect indicators such as electroencephalography activities, pupil diameter, heart rate, skin conductance, and respiration rate. In this contribution, it is assumed that the driver provides the safest and most comfortable performance when the driver state tends towards the target state.

TOR. The goal of the controller is to generate and adapt the TOR. Various parameters and features of a TOR can be adjusted. TOR time, modality, position, color, direction, and frequency are some of the adjustable parameters. TOR intensity or position are two possible modifications for matching the TOR with the situation criticality. TOR should also be adapted to the current driver state. Drowsiness, emotional state, and workload of the driver should be considered when generating TOR by variations in auditory and visual modality or addition of ambient light and vibrotactile signals.

3 Exemplary Experimental Implementation

To illustrate the implementation of the proposed method, a driving scenario in SAE Level 3 of automation was designed. The critical driving situation took place on a two-lane highway when the drivers were distracted by reading a text displayed on the center console during automated driving. The drivers were asked to overtake the driving task and exit the highway. The TOR was issued 18 s before the beginning of the exit lane, where the ego-vehicle was in the left lane and the right lane was occupied by other vehicles (Fig. 2.a). The drivers could choose between different maneuvers: increase speed and exit the highway before arrival of other vehicles, pass through the vehicles and exit the highway, wait until all of the other vehicles are passed and then exit the highway, and go straight on the highway without exiting.

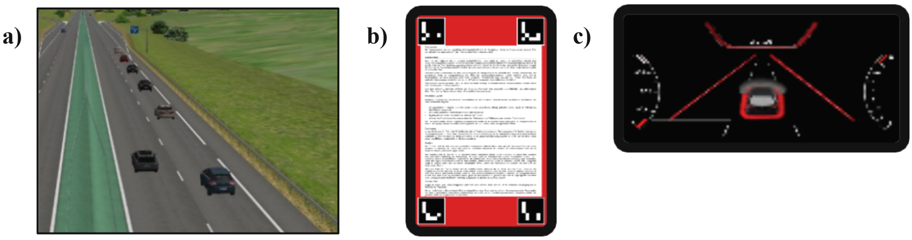


Fig. 2. a) Critical situation caused by occupied right lane by traffic vehicles, b) TOR on the dashboard, c) TOR on the center console

Experiment Design. In total, 22 students of the TU Dortmund University (8 females, 14 males, aged M = 22.4, SD = 2.13) took place in the experiment. All had normal or corrected to normal vision. Trials of three participants experiencing technical problems or motion sickness were removed from the data analysis. The experiment was done on a static driving simulator and all eye-related data was collected by a Tobii Pro Glasses 2 eye-tracker [11]. All the participants signed a consent form after their entrance. At the beginning of the experiment, an introductory drive of twenty minutes manual and automated driving was planned for participants to get familiar with driving on the simulator. The takeover scenario was driven after the introductory drive. Time-to-exit, maximum absolute lateral acceleration, and average lateral shift after the takeover on the exit lane were observed as safety and performance measures. The data of five participants who did not exit the highway was excluded from the reaction data analysis.

Controller Components. The gaze direction was employed as driver state and the situation criticality was identified by time-to-exit during the experiment. The TOR had visual and auditory modalities. The auditory modality was used to inform the driver about the current situation criticality and consisted of beep pulses with pulse frequency adapted to the current time-to-exit. As a visual modality of the TOR, the color of the dashboard turned red (Fig. 2.c) and additionally if the driver was looking at the console (was reading the text) at TOR time, the console's background started blinking in red

(Fig. 2.b; feedback from driver state). Thus, the target was to inform the driver about the upcoming takeover directly in his field of vision.

Results. To investigate the effect of feedback in the takeover situation, drivers' performance with feedback (N = 11) and without feedback (N = 9) was compared. The dependent variables for measuring driver performance were takeover time (time interval from TOR issuance to the first reaction), time-to-exit (time budget until missing highway exit at takeover time), average lateral shift, maximum and minimum longitudinal acceleration, maximum lateral acceleration, and average speed on the exit lane. All the results were analyzed by ANOVA to check the differences. Gender and sequence of scenarios did not show any effect on dependent variables (Table 1).

Table 1. Results of data analysis for scenarios with and without feedback

Dependent variables	Without feedback		With feedback	
Success	C = 2 NE = 2	NC = 7 E = 7	C = 0 NE = 3	NC = 11 E = 8
Takeover time (s)	M = 5.26	SD = 3.16	M = 4.21	SD = 1.86
Time-to-exit (s)	M = 19.92	SD = 2.06	M = 20.36	SD = 1.76
Avg. lat. shift (m)	M = 0.64	SD = 0.02	M = 0.43	SD = 0.10
Min. long. acc. (m/s^2)	M = −1.51	SD = 1.20	M = −1.55	SD = 1.78
Max. long. acc. (m/s^2)	M = 0.98	SD = 0.78	M = 0.94	SD = 0.81
Max. lat. acc. (m/s^2)	M = 4.49	SD = 3.79	M = 1.48	SD = 1.03
Avg. speed (km/h)	M = 26.80	SD = 7.96	M = 28.24	SD = 4.45
SART	M = 23.33	SD = 7.82	M = 17.4	SD = 4.99

C: crashed, NC: not crashed, NE: not exited, E: exited, avg.: average, lat.: lateral, long.: longitudinal, acc: acceleration, M: mean, SD: standard deviation

Due to the results, without feedback two out of nine participants experienced an accident with other traffic vehicles when they intended to exit the highway. However, adding feedback reduced this statistic to zero. The ANOVA analysis did not show any significant difference in takeover time, time-to-exit, maximum and minimum longitudinal acceleration, and average speed on exit lane. The average lateral shift of the drivers on exit lane is significantly less for drivers who got additional TOR on the console as driver state feedback ($p = 0.02$). The visual feedback has also decreased the maximum lateral acceleration on the exit lane ($p = 0.02$). In addition to performance measures, a subjective questionnaire (SART) was also filled by participants after the experiment to evaluate the effect of visual feedback on their situation awareness. The analysis of the results revealed no significance in self-rated awareness of drivers.

4 Discussion

The present study has proposed a control framework for DVI in automated driving, in which the driver state is propelled to the desired state. The controller adopts the TOR to the situation criticality and the driver state. An exemplary driving scenario is designed to present the implementation of the suggested approach in a driving situation. The results show an improvement in driver performance in terms of reduced lateral shift and acceleration by implementing the driver state feedback. In the case of getting feedback, no accident is reported, as well. The reaction type of drivers was also examined at the takeover time. Without feedback, all of the drivers reacted with braking as their first action. However, for the drivers who received additional visual feedback on the console, the reaction type statistics did not show any clear difference between braking and steering. This result can be interpreted as the drivers without feedback react immediately with braking before considering other possible reactions; however, drivers with feedback react more consciously. The designed experiment included only feedback from the sensory driver state. The next step would be to cover more driver state variables in the feedback loop to create a comprehensive TOR. The search for a suitable strategy to incorporate the situation criticality and the driver state into the TOR design is still an open point for which further investigations are required.

Acknowledgments. The present work is supported by the German Federal Ministry of Transport and Digital Infrastructure as result of the research project “Moffa – Holistic model to describe the allocation and the transfer of tasks between the human driver and advanced driver assistance systems during automated and interconnected driving” [grant number 16AVF2005A].

References

1. Yoon, S.H., Kim, Y.W., Ji, Y.G.: The effects of takeover request modalities on highly automated car control transitions. Accid. Anal. Prev. **123**, 150–158 (2019)
2. Roche, F., Somieski, A., Brandenburg, S.: Behavioral changes to repeated takeovers in highly automated driving: effects of the takeover request design and the non-driving related task modality. Hum. Factors **61**(5), 839–849 (2019)
3. Bazilinskyy, P., Petermeijer, S.M., Petrovych, V., Dodou, D., de Winter, J.C.: Takeover requests in highly automated driving: a crowdsourcing survey on auditory, vibrotactile, and visual displays. Transp. Res. Part F Traffic Psychol. Behav. **56**, 82–98 (2018)
4. Hecht, T., Feldhuetter, A., Radlmayer, J., Nakano, Y., Miki, Y., Henle, C., Bengler, K.: A review of driver state monitoring systems in the context of automated driving. In: Congress of the International Ergonomics Association, pp. 398–408. Springer, Cham (2018)
5. De La Fuente, H.L., Berthelon, C., Fort, A., Etienne, V., De Weser, M., Ambeck, J., Jallais, C.: Electrophysiological and performance variations following driving events involving an increase in mental workload. Eur. Transp. Res. Rev. **11**(1), 42 (2019)
6. Lapoehn, S., Dziennus, M., Utesch, F., Kelsch, J., Schieben, A., Dotzauer, M., Hesse, T., Köster, F.: Interaction design for nomadic devices in highly automated vehicles. In: Weyers, B., Dittmar, A. (eds.) Human and Computer 2016 – Workshop volume (in German), pp. 1–8. German Informatics Society, Aachen (2016)
7. Forster, Y., Hergeth, S., Naujoks, F., Krems, J., Keinath, A.: Tell them how they did: feedback on operator performance helps calibrate perceived ease of use in automated driving. Multimodal Technol. Interac. **3**(2), 29 (2019)
8. Marberger, C., Mielenz, H., Naujoks, F., Radlmayr, J., Bengler, K., Wandtner, B.: Understanding and applying the concept of “driver availability” in automated driving. In: International Conference on Applied Human Factors and Ergonomics, pp. 595–605. Springer, Cham (2017)

9. Wang, J., Wu, J., Li, Y.: The driving safety field based on driver-vehicle-road interaction. IEEE Trans. Intell. Transp. Syst. **16**(4), 2203–2214 (2015)
10. Althoff, M., Stursberg, O., Buss, M.: Safety assessment of driving behavior in multi-lane traffic for autonomous vehicles. In: 2009 IEEE Intelligent Vehicles Symposium, pp. 893–900. IEEE Press (2009)
11. Tobii Pro Glasses 2 Wearable Eye-Tracker (2020). https://www.tobiipro.com/product-listing/tobii-pro-glasses-2

Automated Driving on the Motorway: A Users' Perspective on Conditional Versus High Automation

Johanna Wörle(✉), Barbara Metz, Aaron Lutz, and Marcus Schmitt

Wuerzburg Institute for Traffic Sciences, Robert-Bosch-Straße 4, 97209 Veitshöchheim, Germany
{woerle,metz,schmitt}@wivw.de, aaronlutz@gmx.de

Abstract. The development of driving automation is one of the major topics in automotive engineering and is progressing at a fast pace. One key to the success of these systems is the acceptance by users and the actual usage. Surveys on the potential usage and acceptance have shown that the higher the automation level, the higher the acceptance and willingness to engage in side-activities. A driving simulator study with N = 61 was conducted with a between-subject design to compare the acceptance and usage of a conditionally (L3) and a highly (L4) automated driving system (ADS). Drivers were free to use the ADS as they liked. Results show that when using the L3 ADS, drivers spent more time on non-driving related activities and drivers using the L4 ADS spent more time with their eyes closed. The overall acceptance of the L4 ADS was higher.

Keywords: Automated driving · User acceptance · SAE automation levels · Driving simulator

1 Introduction

The development of driving automation has progressed to a point that allows the testing of conditionally automated driving systems (ADS, SAE level 3) on open roads as it is currently done in the EU-funded L3 Pilot project. Conditional automation is the first level of automation that allows the driver to withdraw her or his attention from the driving task and execute other activities. However, in case of system limits the driver is still the fall-back option and has to be available to take back the driving task at any point during the automated drive.

The taxonomy of driving automation by the Society of Automobile Engineers (SAE [1]) defines 6 levels of driving automation; from level 0 (L0), which is purely manual driving up to level 5 (L5), full automation. ADS of level 2 (L2) are already available on the market (e.g. Tesla Autopilot, General Motors Super Cruise, Mercedes-Benz Distronic Plus). Those ADS execute longitudinal and lateral guidance; however, the driver is required to supervise the system and pay attention constantly. The main expected benefit of level 3 (L3) from a user's perspective is that she or he does not have to pay attention while the ADS is operating and they can engage in non-driving related activities (NDRAs) such as reading or watching a movie. However, if the ADS reaches

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 166–172, 2020.
https://doi.org/10.1007/978-3-030-50943-9_22

a system limit, the driver has to be available to take back the vehicle guidance at any point and is responsible for the drive. This changes for highly ADS, level 4 (L4) ADS. The driver is not required at all while the system is operating in AD mode. In contrast to L5, the operational design domain (ODD), that is the circumstances under which the ADS is working, is limited (e.g. to motorways, clear weather). The driver still has the option to drive her- or himself.

A survey on 5000 respondents [2] asked the drivers' willingness to engage in NDRAs for different automation levels. Not surprisingly, they found that the higher the automation level, the more drivers wanted to engage in other activities. This was especially true for the activities "rest/sleep", "watch movies" and "read". For the comparison of drivers' evaluation of L3 and L4 AD, especially the activity "rest/sleep" might be relevant since drivers are not allowed to sleep in L3, but in L4 AD. In the survey, drivers were also willing to pay more for higher automated ADS.

A recent survey by Becker and colleagues [3] found that potential users had most interest in spending their automated drive with "sleeping and relaxing" followed by "working and being productive", "eating and drinking", "entertainment" and "beauty, wellness and fitness". The most desired type of usage "sleeping and relaxing" is only allowed in L4 automation. We could thus expect that drivers see more benefits of L4 AD than of L3 AD. We have to consider however, that respondents of these surveys had not experienced AD and therefore the answers are rather hypothetical.

However, while driving with the ADS activated, drivers might not recognize a difference between an L3 ADS an L4 ADS. Only the instruction/legal constraints and potentially the frequency of system limits might differ from a user's perspective. It is thus questionable whether after usage drivers actually see an increased benefit of an L4 ADS compared to an L3 ADS.

The aim of the study was to investigate whether the usage and acceptance of automated driving was different for an L3 ADS and an L4 ADS when drivers had the opportunity to experience them.

2 Method

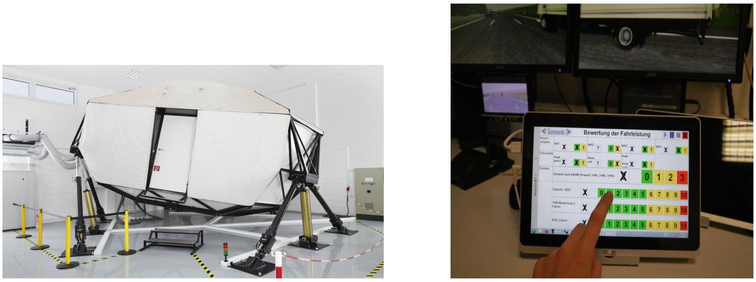
We conducted a study in a high fidelity moving-base driving simulator that has a 240° surround view and that displays realistic dynamics (see Fig. 1a). The simulator runs with the simulation software SILAB 6.0® (WIVW GmbH, Veitshöchheim).

N = 61 drivers (29 female, mean age = 38 years, SD = 12 years) participated in the study and were randomly assigned to one of two conditions: the L3 condition (N = 31 drivers) and the L4 condition (N = 30 drivers). All drivers experienced six drives on a motorway in six different driving sessions which took place on different days. During the drives, they were free to use a motorway chauffeur of L3 (L3MC) resp. L4 (L4MC). Drive 1, 2, 4 and 6 had a duration of 30 min and drives 3 and 5 had a duration of 90 min. One drive (drive 3 or 5) took place at 6 a.m. and drivers had the instruction to sleep no more than 4 h the night before the session in order to investigate their usage of the ADS in the state of fatigue. Both groups were instructed in all drives that they were free to use the ADS as they would like to use it in real life.

Both ADS implementations were able to execute longitudinal and lateral guidance as well as automated lane changes in case the ADS detected a slower lead vehicle, had a speed range of 0–130 km/h and issued a take-over request (TOR) in case system limits were reached. Differences between the ADS were: System limits of the L3MC were motorway exits and entrances, construction sites, missing lane markings and heavy rain. The L4MC had the same system limits except that it was able to operate when lane markings were missing. The take-over times also differed between the two ADS implementations: The L3MC had a take-over time of 15 s while the L4MC had a take-over time of 45 s.

Another main difference between the two conditions was the instruction. Participants in the L3 condition were instructed with the actual wording of §1b of the German Road Transport Law [4] that regulates the role of the driver while using an ADS. The paragraph obliges the driver to be receptive to TORs at any time during the drive and to respond appropriately in case the ADS issues a TOR. In contrast, in the L4 condition, the instruction said that the vehicle was able to solve all situations and that they were not responsible for the drive.

In the experiment, questionnaires developed in the L3Pilot were used [5]. These questionnaires are planned to be used within all studies of L3Pilot to allow a pooling of the data later in the project. We administered the L3Pilot pre-questionnaire before the 1st drive and the L3 Pilot post drive questionnaire after the 6th drive. The L3Pilot questionnaire consists of standardized scales (like, e.g., the Acceptance scale of [6]) and customized items that were specifically designed for the purposes of L3Pilot (for a full description of the L3Pilot questionnaire see [5], Annex 3). Before and after every other drive, participants filled in a short version of the questionnaire. For this paper, we only evaluate the pre-questionnaire of the first session and the post-questionnaire of the 6th session. After the 6th drive, participants had experienced the systems for a total of 5 h. During the drives, an experimenter constantly coded NDRA engagement of the drivers via a tablet application and saved it together with the data from the driving simulator in SILAB®. For an overview of the sample description, see Table 1.



(a) (b)

Fig. 1. Figure a (left) shows the high-fidelity driving simulator. The experimenter constantly coded NDRAs via a tablet application throughout all drives, Figure b (right).

Analysed parameters were:

- The 9-item Acceptance scale [6] administered after the 6th drive
- Items about the frequency of several NDRAs during AD from the L3 Pilot questionnaire [5]
- Total activation time: Percentage of time the system was activated of the time it was available averaged for all 6 drives
- Duration and type of NDRA and time with closed eyes coded throughout all 6 drives via a tablet application (see Fig. 1b)

Independent t-tests were conducted to compare the results for the L3 group and the L4 group.

Table 1. Study sample: means and standard deviations (in brackets) for the description of the two sub-samples

	L3 condition	L4 condition
N	31	30
Age	37.1 (11.9)	39.0 (12)
Gender	13 female	16 female

3 Results

We compared the total activation time for the L3 and the L4 condition throughout all six drives calculating independent t-tests. No difference was found in the overall activation time between the L3 ($M = 92.9\%$, $SD = 9.7\%$) and L4 condition ($M = 93.2\%$, $SD = 8.0\%$; $t(362) = -0.34$, $p = .733$).

However, drivers using the L3 ADS used the time with the system activated differently than drivers using the L4 ADS. Drivers in the L3 condition spent 73.2% ($SD = 18.3$) of the time on NDRAs while drivers in the L4 condition spent 54.7% ($SD = 25.2\%$) of the time on NDRAs ($t(59) = 3.367$, $p = .002$).

The coding scheme for coding NDRAs consisted of 9 categories: “Smartphone”, “Reading magazines or newspapers”, “Reading books”, “Reading an eBook”, “Eat and drink”, “Other technical devices”, “Hygiene and cosmetics”, “Work” and “Others”. The most frequent side-task in both conditions was “Smartphone” with a mean percentage of 52,3% of the time the system was available in the L3 condition ($SD = 37\%$) and 33,0% in the L4 condition ($SD = 30\%$). An independent t-test revealed a significant difference for smartphone usage between the two conditions ($t(59) = 2.24$, $p = .029$). Other popular NDRAs were “Reading magazines or newspapers”, “Reading books, “Reading an eBook” and “Eat and drink”. However, there was no difference between the conditions.

A significant difference, however, was evident for the time drivers spend with their eyes closed. In the L3 condition, drivers spend 10.5% ($SD = 13.1\%$) of the time the system was available with eyes closed, but 19.5% ($SD = 18.7\%$) of the time in the L4 condition ($t(59) = -2.17$, $p = .034$).

For the analysis of the Acceptance scale [6] the inversely coded items 3, 6 and 8 were recoded according to the instructions in the original paper. That means that the resulting scale ranges from −2 to 2 with −2 being the worst evaluation and 2 the best. Table 2 shows the results. Except for the item *"Raising Alertness – Sleep inducing"* all statements were answered in favour of the L4MC.

Table 2. Means and standard deviations are shown for the Acceptance scale [6] of the L3 system and the L4 system

	M (L3)	SD (L3)	M (L4)	SD (L4)	p
Useful – Useless	1.3	0.8	1.8	0.4	.00*
Pleasant – Unpleasant	1.2	0.7	1.8	0.4	.00*
Good - Bad	1.1	0.9	1.7	0.4	.00*
Nice – Annoying	1.0	0.9	1.5	0.9	.00*
Effective – Superfluous	0.6	1.2	1.5	0.7	.00*
Likeable - Irritating	0.9	0.9	1.8	0.6	.00*
Assisting – Worthless	1.2	1.0	1.9	0.3	.00*
Desirable - Undesirable	1.0	1.1	1.6	0.6	.00*
Raising Alertness – Sleep-inducing	−0.7	0.9	−0.4	1.1	.29

*significant on the .005 level

After the 6th driving session drivers were asked how frequently they would do the following activities if they had the L3/L4 function in their cars: *"Texting", "Music, radio, audiobooks", "Interact with a passenger", "Eating or drinking", "Calling", "Smoking", "Personal hygiene/Cosmetics", "Smart phone apps", "Social media", "Navigation", "Browsing the internet", "Sleeping", "Watching movies", "Office/work tasks", "None"*. The answer scale ranged from 1 = "Never" to 6 = "Very frequently".

Activities that were chosen across both conditions as the most frequent were "Listening to music, radio or audio book" (M = 5.7, SD = 0.6), "Interacting with a passenger" (M = 5.6, SD = 0.6) and "Writing text messages" (M = 5.1, SD = 0.9). There was only a difference in the stated frequency for the activities "Sleeping" and "Watching movies". Participants of the L4 condition indicated that they would sleep more frequently (M = 3.6, SD = 1.4) than participants in the L3 condition (M = 2.2, SD = 1.4; p = .000). In the L4 condition participants also reported that they would watch movies more frequently (M = 3.7, SD = 1.2) than participants in the L3 condition (M = 3.0, SD = 1.5; p = .041).

4 Conclusions

The comparison of usage and acceptance of a conditionally and a highly automated motorway ADS yielded differences in the usage of time with the ADS as well as in acceptance of the ADS. There was no difference in the overall activation, the drivers drove with both systems activated more than 90% of the time the ADS was available.

However, they spent the time with the systems activated slightly differently. In the L3 condition, the drivers spent a higher proportion of the driving time with ADS active on NDRAs. This effect can be explained by four drivers in the L4 condition who only very rarely engaged in NDRAs and the fact that with the L4 system drivers spent significantly more time with eyes closed. The finding that drivers spent more time with eyes closed in the L4 condition is not surprising since sleep is one of the use cases of L4 AD. The rather high percentage of time with eyes closed of 10.5% with an L3 ADS is concerning, since with L3 AD drivers have to be receptive to TORs at any point during the drive.

The results of the Acceptance scale [6] administered after the 6th drive show a high overall acceptance of both tested automated driving systems. However, drivers evaluated the L4 system more positively than the L3 system on almost all scales. Both systems were evaluated negatively on the scale "Raising alertness – sleep inducing" indicating that both the L3 and the L4 function increased drowsiness in drivers. This finding is also supported by other studies on automated driving. Vogelpohl et al. [7] found that drivers became drowsy faster and showed higher levels of drowsiness when driving with automation than when driving manually. This finding might be especially critical for the release of L3 ADS where drivers have to be receptive to TORs at any point during the automated drive.

Since not all NDRAs were available in the simulator study, e.g., there was no radio or navigation system in the simulator and no passenger was present, after the 6th driving session participants were asked how often they would execute certain NDRAs while driving with an L3 or L4 ADS. Stated NDRAs were similar for both ADSs, drivers of the L4 condition only stated the activities "Sleeping" and "Watching movies" more frequently. This results corresponds to findings of surveys on potential usage of AD [2, 3]. Especially the NDRA "Sleeping" is expected to occur more often in L4 than L3 AD since it is clearly forbidden in L3. It is therefore concerning that half of the participants in the L3 condition stated that they would sleep at least "very infrequently" and 3 out of 31 participants stated that they would sleep "frequently" when driving with the ADS. However, sleep is a clear misuse of L3 AD.

The study showed an overall high acceptance and positive evaluation of AD with a slightly better evaluation of the L4 ADS. Drivers used the ADSs differently with drivers in the L3 condition spending more time on NDRAs and drivers in the L4 condition spending more time with eyes closed. It is however concerning that in the L3 condition drivers also spend about 10% of the time with the ADS active with eyes closed and that drivers in the L3 condition stated that they would use the time with the ADS for sleeping. The overall very positive evaluation of both ADS and the high acceptance are a positive signal for market success. However, our results support the concern of a potential misuse of L3 ADS by drivers becoming drowsy and sleeping behind the wheel.

Acknowledgments. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723051. The sole responsibility of this publication lies with the authors. The authors would like to thank all partners within L3Pilot for their cooperation and valuable contribution.

References

1. SAE: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (2018)
2. Kyriakidis, M., Happee, R., de Winter, J.C.: Public opinion on automated driving: results of an international questionnaire among 5000 respondents. Transp. Res. Part F Traffic Psychol. Behav. **32**, 127–140 (2015)
3. Becker, T., Herrmann, F., Duwe, D., Stegmüller, S., Röckle, F., Unger, N.: Enabling the Value of Time. Fraunhofer Institute for Industrial Engineering IAO, Stuttgart, Germany, pp. 1–27 (2018)
4. Straßenverkehrsgesetz, D.: § 1b Rechte und Pflichten des Fahrzeugführers bei Nutzung hoch- oder vollautomatisierter Fahrfunktionen, Bundesministerium der Justiz und für Verbraucherschutz (2019)
5. Metz, B., Rösener, C., Louw, T., Aittoniemi, E., Bjorvatn, A., Wörle, J., et al.: L3 Pilot Deliverable D3.3. Evaluation methods, L3 Pilot (2020)
6. Van Der Laan, J.D., Heino, A., De Waard, D.: A simple procedure for the assessment of acceptance of advanced transport telematics. Transp. Res. Part C Emerg. Technol. **5**(1), 1–10 (1997)
7. Vogelpohl, T., Kühn, M., Hummel, T., Vollrath, M.: Asleep at the automated wheel—sleepiness and fatigue during highly automated driving. Accid. Emerg. Med. **126**, 70–84 (2018)

On the Road Again - Explanatory Factors for the Users' Willingness to Replace Private Cars by Autonomous on-Demand Shuttle Services

Ralf Philipsen(✉), Teresa Brell, Hannah Biermann, and Martina Ziefle

Human-Computer Interaction Center (HCIC), RWTH Aachen University, Campus-Boulevard 57, 52074 Aachen, Germany
{philipsen, brell, biermann, ziefle}@comm.rwth-aachen.de

Abstract. Increasing urbanization and the simultaneous growth of car-based individual transport pose major challenges for many cities. On the one hand, streets and parking facilities are reaching their capacity limits, and on the other hand, car emissions are a burden for both residents and environment. A transformation from one-person car use to emission-free shared mobility could be part of the solution. Small autonomous, electrified shuttle buses, which can carry 10–15 people on-demand, are already in the technical development stage. At present, however, many questions remain unanswered regarding the design of mobility services around the technological platform. This paper therefore explores which factors influence the willingness to share rides and to abandon private car ownership. While for the willingness to use ridesharing, especially the familiarity with the passengers and, to a small extent, the purpose of the trip has an influence, the expected advantages in terms of flexibility and costs are particularly important for the willingness to abandon a car. A positive attitude towards conventional car driving counteracts both measures surveyed.

Keywords: Ridesharing · Ridehailing · Autonomous shuttles · On-demand mobility · Private car ownership · Usage intention

1 Introduction

The increasing use of motorized private transport creates major challenges for modern industrial societies [1, 2]. Especially in growing cities, the available road infrastructure is reaching its limits, which can lead to increasing congestion and accident rates. One possible solution, in addition to switching to alternative means of transport, would be to share rides with cars, so that the rate of people per car increases and the total number of vehicles and trips decreases [3]. However, all existing offers and constraints, such as Car-Sharing services or the restriction of lanes to carpool use, have so far not led to any significant changes in the traffic mixture and quantity. Even autonomous driving, which is associated with many hopes of increasing efficiency and safety [4], would initially even lead to more (empty) trips in cities if private conventional cars would be replaced one-to-one by private autonomous vehicles.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 173–185, 2020.
https://doi.org/10.1007/978-3-030-50943-9_23

Autonomous on-demand shuttle services, which transport several people to the same or a similar destination and are more flexible than existing public transport, could, however, make a real contribution to easing the traffic situation if there is a large-scale shift from private car use to this kind of shared mobility. From the perspective of manufacturers, urban planners, and mobility service providers, many questions arise, but the most important is probably under which conditions users are willing to forego the use of a private car for such a mobility service, or ideally not to buy a private car at all.

2 State of Research and Resulting Research Questions

Even though fully autonomous driving is not yet available in practice and thus hands-on experience is not yet available to most people, there is already initial research on the acceptance of self-driving vehicles. Analogue to traditional technology acceptance models (TAM) [5], the focus here is on the fundamental intention to use. For example, Lee et al. were able to show that perceived usefulness, affordability, social support, lifestyle fit, and conceptual compatibility are the most important predictors for the willingness to use self-driving cars [6]. They and other researchers also identified age as a negative influence factor: The older a user, the less willing he or she is to use the technology and the more critical his or her attitude towards it [6–8]. Also, gender influences the attitude towards self-driving cars. Men are somewhat more open to this technology than women [7–10]. In addition to demographic characteristics, personality traits also play an important role in the acceptance of autonomous vehicles. Choi & Ji, for example, have shown that technical competence has a positive influence on trust in such vehicles and that trust in turn has a positive effect on the usage intention [11]. What all presented sources have in common is that they address the one-to-one replacement of private conventional cars by vehicles with automatic driving functions and do not address on-demand ridesharing concepts.

Regarding bookable autonomous shuttles, Motak et al. found that a classical TAM for this context cannot explain sufficient variance in willingness to use and therefore other factors, such as social, motivational, and experience-related, should be taken into account [12]. A step in this direction is taken by Stoiber et al., who in particular define cost, comfort, and speed as distinguishing features between modes of autonomous transport. The results for their scenario indicate that users would not completely abandon private autonomous cars and autonomous taxis in single use in favor of autonomous shared shuttles [13]. However, there is still a lack of linkage to conventional means of transport, such as human-controlled cars or the existing public transport system. A differentiation according to the purpose of use of the means of transport is also open.

Another relevant factor influencing the acceptance of autonomous shuttles is the immersion of the user experience [14]. Preliminary user surveys during prototype operations of autonomous shuttles in Germany – albeit on fixed routes and not on-demand – have shown that users are basically open to the technology and have partly idealized expectations prior to the first hands-on experience [15], but then, when dealing with the prototypes particularly criticize the lower speed and effectiveness of the current state of the technology compared with the means of transport used to date [16]. It can therefore be assumed that the speed of travel will be a major trigger for a change of transport mode [17], e.g., away from the private car. User evaluations during field tests in Finland and Norway followed a similar pattern. Here, too, users did not see any fundamental usefulness, particularly because of security concerns [18, 19]. Regarding user characteristics, openness towards and trust in technology in general could be identified as positive influencing factors, while privacy concerns and personal need for control had a negative effect [20]. All previous studies with hands-on experience have two limitations due to the state of the technology. On the one hand, the vehicles used are very slow (often <= 15 km/h) compared to previous motorized means of transport, and on the other hand, fixed routes and stations are used. An on-demand service that can drive flexible routes is not yet available in practice, which is why an evaluation of the transferability of previous findings is still pending, although.

To summarize the state of research, it can be said that there are already some approaches that consider a general willingness to use and which have identified both user factors (as for instance age, gender, personal need for privacy, and control or technology commitment) and service parameters (like costs, speed, or security) as predictors. Due to the already, partly very differentiated models of general willingness to use [20], the focus of this paper will be a different one. Two aspects have so far mostly been excluded: On the one hand, the context-dependency of the usage decision, on the other hand, the concrete willingness to change in the sense of a complete replacement of private, human-controlled cars by autonomous shuttle services. For the former, there is still a lack of understanding, in particular, regarding the willingness to share the vehicle depending on the purpose of the trip, such as travel to work, shopping, leisure activities, or other errands [21], and on the type and number of passengers. Therefore, based on the existing knowledge gaps, the following research questions were derived:

1. *How do user, technology, and service factors influence the general willingness to share a ride?* The latter is the basic prerequisite for autonomous shuttles to be able to reduce the volume of traffic, which would not be possible with individual vehicle use [3].
2. *Which factors influence the willingness to abandon a private car in favor of an autonomous shuttle service?* By giving up own cars in favor of shared mobility services, the utilization of road infrastructure, but especially the parking load in inner cities could be reduced.

3 Methodology

To answer the research questions, a consecutive methodological approach was chosen. First, qualitative interviews ($N = 40$) were conducted with car owners to identify their requirements for a switch to an autonomous shuttle service. In a second step, a questionnaire was developed as a measurement tool to quantify those requirements and to relate them to the willingness to use and switch to an autonomous shuttle service, which is the focus of the present paper. Due to the limited knowledge available, no specific hypotheses were developed, but instead an explorative, structure-discovering approach was chosen.

3.1 Qualitative Preliminary Work

As a basis for the item construction in the quantitative study, at first interviews with car owners (*15* women, *25* men, M_{AGE} = *36.8 years,* SD_{AGE} = *16.1*) were conducted to identify requirements and attitudes towards conventional driving and the new autonomous shuttle service. The interviews were transcribed and evaluated by content analysis. Therefore, information was structured and transferred into a category system. In general, most respondents were open to the new mobility service. However, it was mentioned as a basic requirement for changing existing mobility behavior that the new service must provide at least the same performance, ideally improvements. Possible areas for improvement in this respect were *cost*, *travel time*, *flexibility*, *environmental protection*, and *noise*. Furthermore, the interviews revealed, as expected, that the question of who the co-passengers are, whether the passenger has an influence on the choice of these and for what purpose the trip is made, could have an influence on the usage intention.

3.2 Questionnaire Design

The quantitative study was designed as a conventional questionnaire in order to measure the users' attitudes and underlying influencing factors that had been identified previously, as well as their interrelationships. The questionnaire consisted of 5 thematic blocks (see Fig. 1) containing questions on socio-demography (including age, gender, educational level, income, and place of residence), mobility behavior – in terms of the frequency of use of different means of transport –, personality traits, attitudes towards conventional driving, and the autonomous shuttle service itself. Regarding the personality traits, the technology commitment [22], trust in automation, the willingness to take risks [23], and the need for privacy were surveyed. For the latter, an influence on the willingness to share trips was assumed; while for the first three traits, an effect on the general willingness to use the service was expected. Referring to conventional driving, the participants were asked both whether they are dependent on their cars and what their general attitude to driving itself is, in terms of a positive or negative connotation.

Since autonomous shuttle services are not yet available in practice, a scenario-based approach was used. The scenario outlined a mobility offer based on small autonomous shuttle buses with an electric drive and a capacity of up to 15 persons. The user should be able to book the service on-demand, e.g., via mobile phone app. Both hop-on and drop-off are not bound to fixed stops. No information on speed or costs but a level of safety comparable to other means of transport was defined.

After presenting the scenario, questions were asked about the perceived usefulness of the service, the intention to use it, the willingness to replace a private car with the service, and the general willingness to share trips. The latter was additionally queried depending on the purpose [21] and the length of the trip. The question of whether the vehicle is shared with known or foreign passengers was designed as a between-group factor with random allocation of participants.

Unless stated otherwise, self-developed item sets were used to measure all attitudes and characteristics. 6-point Likert scales were used for most questions regarding the autonomous shuttle service with min = 0 *full disagreement/not important/no*

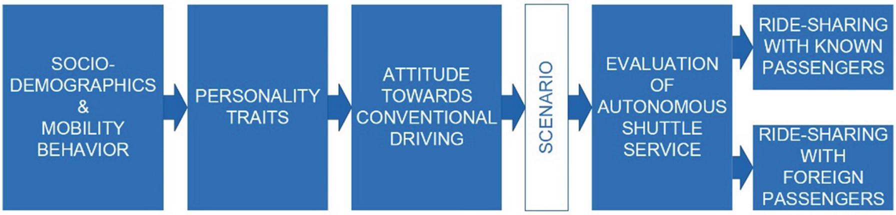


Fig. 1. Thematic and sequential structure of the developed questionnaire

willingness and max = 5 *full agreement/very important/full willingness*. In contrast, the probability of abandoning a private car was measured as a self-reported percentage value between 0 and 100.

3.3 Data Collection, Processing, and Analysis

The study was conducted as an online survey and participants were acquired both in the university environment and mainly in specialist Internet forums on the subject of (auto)mobility. Participation was voluntary and there was no form of incentive. Possession of a driving license and private car ownership were defined as screening criteria. The survey was conducted in German and was therefore limited to the corresponding language region.

Following the completion of data collection, a quality check of the complete responses was carried out. Due to conflicting response behavior and speeding (processing time less than half the median processing time of the sample), twelve responses were excluded from the analysis.

The analysis itself was then carried out using both univariate and multivariate inferential statistical methods after checking the respective distributions and other prerequisites. The level of significance was set to $\alpha = .05$ and η_p^2 and Cohen's *d* were used for effect sizes, whereas adjusted R^2 and Root Mean Square Error (*RMSE*) were used for the prognosis quality of the models developed. Furthermore, Cronbach's Alpha (α) was used to measure the scales' internal consistency with $\alpha > .7$ interpreted as good reliability [24]. To improve readability, all measures of central tendency and dispersion were rounded to one decimal place.

4 Sample

In total, the sample that was included in the analysis consisted of *545* participants. A slight majority were male (*55.8%, n = 304*), the remaining female (*44.2%, n = 241*). The participants ranged from *18* to *84* years with an average age of $M = 40.3$ years ($SD = 16.00$). The average age in the sample was thus slightly lower than the mean age of private car owners in Germany (*45.5* years [25]), while the gender ratio is even too balanced (male share 64.5% [25]). With *44.9%* ($n = 240$) of the participants, the high-school diploma was the most frequently mentioned highest educational level, followed by a university degree (*39.5%, n = 212*). The remaining participants reported lower levels of educational qualifications. Overall, the sample can therefore be regarded as rather educated. The median monthly net household income was between 3,000 and 4,000€. More than half of the participants were employed (*60.6%, n = 326*), while the rest were distributed among persons in education, retired, unemployed, and housewives/men. *41.6%* (*n = 224*) said they lived in a suburban area, while the remaining participants were distributed almost equally between city centers and rural areas.

Regarding mobility behavior and in line with the screening criteria, all participants stated that they hold a car driving license and were owners of a private car. *77.1%* (*n = 420*) of them stated that they were dependent on their car. A more than one in two used the private car daily (*55.5%, n = 302*), *157* participants used it several times a week (*28.9%*), and *59* at least once a week (*10.8%*).

5 Results

In the following, the willingness to share a ride in an autonomous shuttle with others, depending on different travel purposes, will be presented first, followed by the willingness to give up one's own private car in the presence of an autonomous on-demand mobility service.

5.1 Willingness to Share Rides

The willingness to share rides differs significantly depending on whether the passengers are known to the user or are foreigners. While there is generally a willingness to ride with known passengers ($M = 3.9,\ SE = 0.08$), users tend to be undecided regarding foreign passengers, with an average willingness close to the center of the scale ($M = 2.8,\ SE = 0.09$). This effect of passenger familiarity was significant with $t(533) = 10.2,\ p < .001,\ d = 0.865$.

When differentiating more precisely between the trip purposes, the effect of the passengers' familiarity is preserved. As can be seen in Fig. 2, the willingness to share the trip is higher for all questioned trip purposes if known passengers are on board. The highest levels of willingness were observed for trips to work, to the place of education, and for leisure trips, while rides for shopping or other errands were less likely to be shared.

The effect of the trip purpose itself on the willingness to share rides was significant, albeit of little strength: $F(3,1629) = 44.08,\ p < .001,\ \eta_p^2 = .075$. Pairwise comparisons of trip purposes revealed that there was no significant difference between shopping and other errands in terms of willingness ($p > .05$), while all other purposes differed from each other ($p < .001$ for all comparisons). Furthermore, a weak interaction effect between the purpose of the trip and the familiarity of the passengers could be found: $F(3,1629) = 6.64,\ p < .001,\ \eta_p^2 = .012$. This is reflected in the fact that work and leisure as trip purposes do not differ significantly for sharing with known passengers, while in the case of foreign passengers the willingness for ridesharing is significantly higher for work than for all other purposes (see Fig. 2).

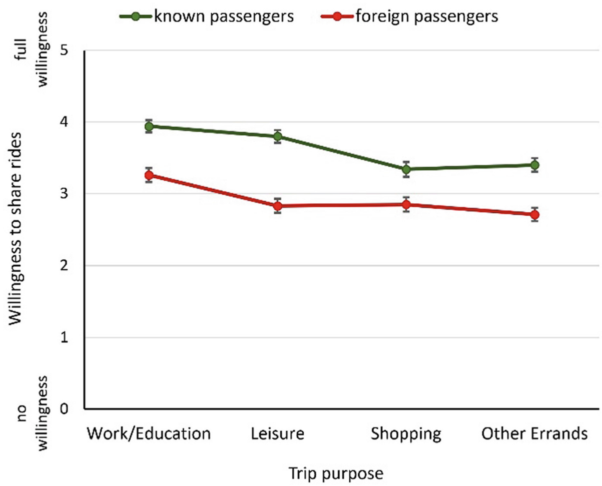


Fig. 2. Willingness to share rides depending on trip purpose and familiarity of passengers (Means and related Standard Errors; scale min = 0, max = 5)

Regardless of the purpose of the journey, also the length of the trip influences the willingness to share it with other passengers. As can be seen in Fig. 3, the willingness is significantly higher for short trips ($F(1,543) = 95.2, p < .001, \eta_p^2 = .149$). If the familiarity of passengers is taken into account, the difference in users' willingness for ridesharing between known and foreign co-passengers on short trips is less than on long trips, with $F(1,543) = 18.7, p < .001, \eta_p^2 = .033$ for the interaction effect. Therefore, there is a weak rejection of foreign passengers on long-distance trips, while on short-distance trips there is an approval to share the vehicle for both groups of passengers.

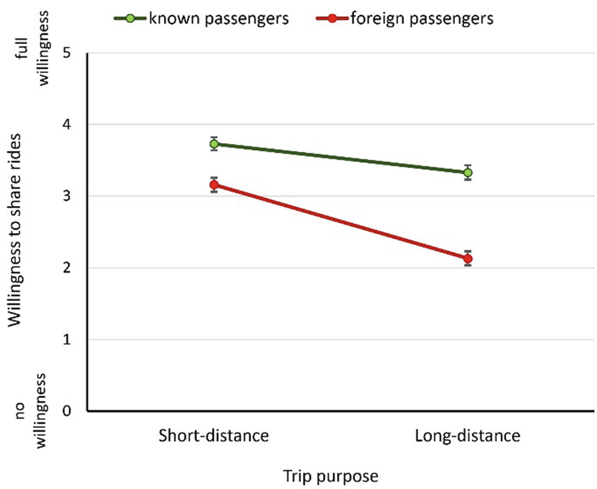


Fig. 3. Willingness to share rides depending on trip length and familiarity of passengers (Means and related Standard Errors; scale min = 0, max = 5)

Based on the previously presented findings and the factors surveyed in the questionnaire, a regression model was developed to explain the general willingness for autonomous ridesharing. A significant regression equation was found ($F(4,528) = 38.6$, $p < .001$), with adjusted $R^2 = .221$, $RMSE = 1.31$. The willingness to share rides equals to *2.64 − 0.28 × attitude towards conventional driving + 0.15 × trust in automation + 0.13 × usage intention + 1.14 × familiarity of passengers* (see Table 1), with a scale of min = 0 to 5 = max. Therefore, the familiarity of the passengers was the most powerful predictor in the present dataset.

Table 1. Significant predictors for willingness to share a ride (Unstandardized Beta (B), Standard Errors for Unstandardized Beta (SE B), Standardized Beta (β), Test statistic (t), and Probability (p)

Predictor	*B*	*SE B*	*β*	*t*	*p*
(Intercept)	2.64	0.24		10.80	< .001
Attitude towards conventional driving	−0.28	0.06	−0.18	−4.75	< .001
Trust in automation	0.15	0.07	0.09	2.12	.034
usage intention	0.13	0.04	0.13	3.10	<.001
Familiarity of passengers	1.14	0.11	0.77	10.03	< .001

Further predictors could not be identified. Neither socio-demographic characteristics (e.g., gender, age, education, income) nor personality traits (e.g., need for privacy, willingness to take risks, commitment to technology) could explain further variance. Similarly, mobility behavior in terms of the current frequency of use of various means of transport (car, public transport, etc.) was not a significant predictor of the model.

5.2 Willingness to Replace Private Cars by Autonomous on-Demand Shuttle Services

Based on the general willingness for ridesharing, the next step now focuses on the usage intention of the autonomous on-demand shuttle service and on the willingness to give up one's own private car.

First, it could be revealed that the shuttle service was basically considered useful, with $M = 3.2$ ($SE = 0.06$) out of 5 max. Regarding a concrete intention to use, however, the agreement levels were slightly lower, closer to the undecided middle of the scale with $M = 3.0$ ($SE = 0.06$). The self-reported probability of replacing the private car with the presented shuttle service was on average $M = 44.2\%$ ($SE = 1.34$). In the following the aim is to better understand and explain the variance in this probability, which varied from 0 to 100%, meaning that there are users who would not give up their car under any circumstances and others who are very open to replacing their private cars by the new mobility service.

Since such new service or technology, if it is to replace an existing one, must entail benefits for the user, the first step was to consider the perceived importance of various possible improvements over conventional motorized private transport. As can be seen in Fig. 4, improvements in flexibility and costs are of paramount importance for users. Next come, by some margin, improved environmental friendliness, shorter travel times, and less noise.

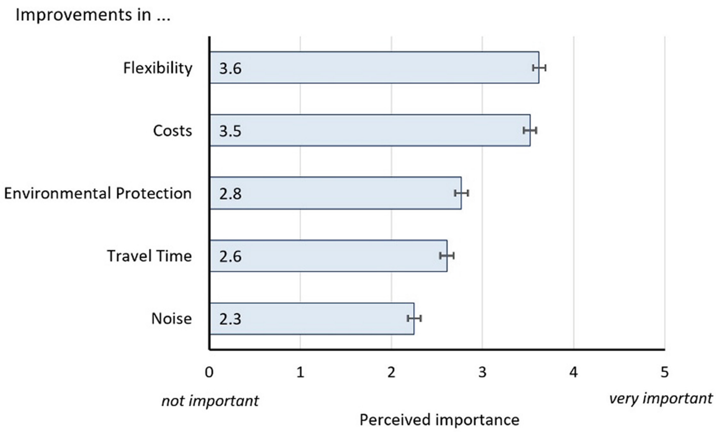


Fig. 4. Perceived importance for replacing private cars of improvements by an autonomous shuttle service (Means and Standard Errors; scale min = 0, max = 5)

There was a small significant main effect of the improvement type, with $F(4,2156) = 129$, $p < .001$, $\eta_p^2 = .193$. In line with the previously reported gradations, pairwise comparisons revealed that improvements in costs and flexibility, as well as in environmental protection and travel time, did not differ significantly in their respective perceived importance ($p > .05$), resulting in three importance classes for improvements: *High:* flexibility and cost, *Medium:* environmental protection and travel time, *Low:* noise.

Finally, a model was developed to explain the probability of abandoning a private car. This model was significant with $F(6,475) = 64.7$, $p < .001$), with adjusted $R^2 = .443$, $RMSE = 22.0$. The probability for car abandonment equals to $-0.13 - 3.66 \times$ *attitude towards conventional driving* $+ 6.7 \times$ *usage intention* $+ 3.4 \times$ *importance of noise improvement* $+ 2.9 \times$ *importance of flexibility improvement* $+ 2.57 \times$ *importance of costs improvement* $+ 0.15 \times$ *age* (see Table 2). Thus, the attitude towards conventional driving has the strongest negative influence on the abandonment probability, whereas the usage intention has the most positive effect.

Table 2. Significant predictors for willingness to replace a private car by an autonomous on-demand shuttle service (Unstandardized Beta (B), Standard Errors for Unstandardized Beta (SE B), Standardized Beta (β), Test statistic (t), and Probability (p)

Predictor	*B*	*SE B*	*β*	*t*	*p*
(Intercept)	−0.13	5.16		−0.02	.981
Attitude towards conventional driving	−3.66	1.11	−0.12	−3.30	.001
Usage intention	6.70	0.76	0.35	9.09	<.001
Importance of noise improvement	3.40	0.72	0.19	4.72	<.001
Importance of flexibility improvement	2.90	0.89	0.15	3.24	.001
Importance of costs improvement	2.57	0.85	0.14	3.04	.003
Age	0.15	0.06	0.08	2.33	.020

Similar to the explanatory model for ridesharing willingness, no further predictors based on socio-demographics, personality traits, or mobility behavior were found. Even the fundamental willingness to share rides (as presented in Sect. 5.1) did not become significant within the model.

6 Discussion

The present study was able to show that the attitude of the participants towards autonomous shuttle services is not in general rejective in terms of willingness to share trips, perceived usefulness and intention of use, which is in line with previous research [e.g., 15]. Nevertheless, the results on willingness to switch means of transport suggest that a complete replacement of private cars is unrealistic, at least in the short term. It is therefore even more important that the design of the mobility service takes user needs into account. The model described in this paper to explain the switchover probability could only explain a small part of the variance in the user attitudes with an adjusted $R^2 < .5$. This corresponds to existing models of general willingness to use [12] and makes it necessary to take a more differentiated view of the user, including a profound consideration of the attitude towards conventional driving, as this revealed a strong impact on the changeover. To this, the research focus could be on the perceived joy of driving and well-being of users in their role as driver or passenger to find out how the autonomous driving environment should be designed to make change attractive. Interestingly, earlier findings regarding effects of gender [7–10], technology commitment, or privacy needs [20] could not be confirmed in this paper. Furthermore, in contrast to the state of research [6–8], age also played only a minor role. However, it must be considered that, in contrast to the above-mentioned studies, the general intention of use was not the variable to be explained here, but the usage intention itself was incorporated into the models as an explanatory factor. If the usage intention is not included in the regression models in the present paper, for example gender becomes a significant predictor, but with an overall lower variance explanation for the model. Consequently, gender may indeed be a factor in user profiling by mobility service providers, but – as presented – other user or service factors are more relevant to the willingness to share or use autonomous shuttles. In particular, costs and flexibility are factors that determine whether a means of transport is changed or not and on which the mobility provider has an influence. Regarding the attitude towards conventional driving, a direct adjustment is not possible. The only feasible approach is that autonomous driving or the mobility service becomes so attractive in terms of other factors that the personal driving experience loses relative importance for the user. Methodologically, further in-depth exploratory research projects are recommended, which have proved successful in this paper, but also in other studies on technology acceptance in this context [9], including both qualitative and quantitative methods next to increased hands-on experiments that can be implemented on test tracks as technology matures.

Regarding ridesharing itself, it could be shown that the purpose of the trip has only a small influence on the willingness to share an autonomous shuttle. An exception is the journey to/from work, where the willingness to take along foreign passengers is higher than for other purposes. In the future, it would have to be researched whether this remains the case if the shuttle service not only takes passengers to work but also enables users to work during the journey, e.g., if the distances are longer, which according to the results also leads to a reduction in willingness to share. The purpose of the autonomous bringing and collecting of other persons, e.g., children, which has been excluded in the present study, must also be in focus in the future to reveal whether the familiarity of co-passengers makes a difference regarding the usage intention in this context. So far it can be stated that the willingness to share an autonomous shuttle service was basically present for all purposes. Compensation of foreign passengers by lower fares or, conversely, surcharges for journeys with known passengers therefore do not seem necessary or possible on a larger scale. But, based on the qualitative results *("On the bus I am not alone, too, and I don't know all my fellow passengers")*, it can be assumed that the users' mental models of the mobility service plays an important role in this and that those models currently correspond more to existing public transport than to a car replacement. If this is the case, and autonomous shuttles will not supplement existing public transport but partially replace it, this could reverse the hoped-for positive effects regarding the relief of traffic infrastructure. Further research is necessary in this context, because the model presented can only explain very little variance in user opinions and accordingly there must be other yet unidentified factors regarding the user, technology, service, or context of the journey that have an influence on the willingness to share.

7 Limitations and Outlook

The limitations of the present study, in addition to the low variance clarification in the regression models already discussed, lie particularly in the local composition of the sample. In general, the sample had a sufficient size to identify and quantify significant influencing factors even with small effect sizes. However, an examination of the indications of origin (postcodes) showed that a large proportion of the participants came from West German conurbations. Even if participants indicated that they come from rural areas, these can vary substantially in terms of population density and distance to city centers. Although the self-classification of the place of residence had no influence on the willingness to abandon a car or to use the mobility service, it would nevertheless make sense to draw a locally more diverse sample in follow-up research in order to exclude further influence at this point. Furthermore, the same applies to the cultural background of the respondents. It is an all-German sample and the Germans are at least said to have a special relationship with their cars. Whether this is the case and how this affects the replacement of private cars by a shuttle service needs to be further explored in a cross-cultural approach.

Based on the results themselves and the limitations presented, it is apparent that the present study is only a first step towards a better understanding of the conditions necessary for the widespread introduction of a mobility service such as the one discussed here. In particular, a holistic model for transport mode selection, which integrates a wider range of user decision criteria and new means of transport and services, is necessary in order to simulate more precisely how the introduction of a new mobility service, such as the one discussed in this paper, affects the transport mix within a region. In this context, the identification of the service parameters that must be satisfied in order to ensure that private individual transport is actually replaced by the users – and not the existing public transport system – is of great importance for the design and implementation of mobility services.

Acknowledgements. This research was funded by Federal Ministry of Transport and Digital Infrastructure (BMVI) within the funding guideline "Automated and Connected Driving" under the grant number 16AVF2134B. The authors thank all participants for sharing opinions on a novel technology. Furthermore, thanks go to Susanne Gohr for research assistance, to Dr. Johanna Kluge, Dr. Simon Himmel, and Julian Hildebrandt for valuable research advice, as well as to the icon of outlaw country for inspiring the paper title.

References

1. Briggs, D., de Hoogh, K., Morris, C., Gulliver, J.: Effects of travel mode on exposures to particulate air pollution. Environ. Int. **34**, 12–22 (2008)
2. European Commission: Annual Accident Report 2018. European Road Safety Observatory. https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/statistics/dacota/asr2018.pdf, last accessed 06 Feb 2020
3. Shaheen, S., Cohen, A., Chan, N., Bansal, A.: Sharing strategies: carsharing, shared micromobility (bikesharing and scooter sharing), transportation network companies, microtransit, and other innovative mobility modes. In: Transportation, Land Use, and Environmental Planning, pp. 237–262 (2020)
4. Michałowska, M., Ogłoziński, M.: Autonomous vehicles and road safety. In: Communications in Computer and Information Science, pp. 191–202 (2017)
5. Davis, F.: A Technology Acceptance Model For Empirically Testing New End-User Information Systems. Massachusetts Institute of Technology, Boston (1985)
6. Lee, C., Ward, C., Raue, M., D'Ambrosio, L., Coughlin, J.: Age differences in acceptance of self-driving cars: a survey of perceptions and attitudes. In: Human Aspects of IT for the Aged Population. Aging, Design and User Experience, pp. 3–13 (2017)

7. König, M., Neumayr, L.: Users' resistance towards radical innovations: the case of the self-driving car. Transp. Res. Part F Traffic Psychol. Behav. **44**, 42–52 (2017)
8. Hulse, L., Xie, H., Galea, E.: Perceptions of autonomous vehicles: relationships with road users, risk, gender and age. Saf. Sci. **102**, 1–13 (2018)
9. Philipsen, R., Brell, T., Ziefle, M.: Carriage without a driver – user requirements for intelligent autonomous mobility services. In: Advances in Intelligent Systems and Computing, pp. 339–350 (2018)
10. Hohenberger, C., Spörrle, M., Welpe, I.: How and why do men and women differ in their willingness to use automated cars? The influence of emotions across different age groups. Transp. Res. Part A Policy Pract. **94**, 374–385 (2016)
11. Choi, J., Ji, Y.: Investigating the importance of trust on adopting an autonomous vehicle. Int. J. Hum. Comput. Interact. **31**, 692–702 (2015)
12. Moták, L., Neuville, E., Chambres, P., Marmoiton, F., Monéger, F., Coutarel, F., Izaute, M.: Antecedent variables of intentions to use an autonomous shuttle: moving beyond TAM and TPB? Eur. Rev. Appl. Psychol. **67**, 269–278 (2017)
13. Stoiber, T., Schubert, I., Hoerler, R., Burger, P.: Will consumers prefer shared and pooled-use autonomous vehicles? A stated choice experiment with Swiss households. Transp. Res. Part D Transp. Environ. **71**, 265–282 (2019)
14. Distler, V., Lallemand, C., Bellet, T.: Acceptability and acceptance of autonomous mobility on demand. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI 2018 (2018)
15. Nordhoff, S., de Winter, J., Payre, W., van Arem, B., Happee, R.: What impressions do users have after a ride in an automated shuttle? An interview study. Transp. Res. Part F Traffic Psychol. Behav. **63**, 252–269 (2019)
16. Nordhoff, S., de Winter, J., Madigan, R., Merat, N., van Arem, B., Happee, R.: User acceptance of automated shuttles in Berlin-Schöneberg: a questionnaire study. Transp. Res. Part F Traffic Psychol. Behav. **58**, 843–854 (2018)
17. Wintersberger, P., Frison, A., Riener, A.: Man vs. Machine. In: Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI 2018 (2018)
18. Roche-Cerasi, I.: Public acceptance of driverless shuttles in Norway. Transp. Res. Part F Traffic Psychol. Behav. **66**, 162–183 (2019)
19. Salonen, A.: Passenger's subjective traffic safety, in-vehicle security and emergency management in the driverless shuttle bus in Finland. Transp. Policy **61**, 106–110 (2018)
20. Herrenkind, B., Brendel, A., Nastjuk, I., Greve, M., Kolbe, L.: Investigating end-user acceptance of autonomous electric buses to accelerate diffusion. Transp. Res. Part D Transp. Environ. **74**, 255–276 (2019)
21. Nobis, C., Kuhnimhof, T.: Mobilität in Deutschland - MiD Ergebnisbericht. In: Studie von infas, DLR, IVT und infas 360 im Auftrag des Bundesministeriums für Verkehr und digitale Infrastruktur (FE-Nr. 70.904/15). Bonn, Berlin (2018)
22. Neyer, F., Felber, J., Gebhardt, C.: Entwicklung und Validierung einer Kurzskala zur Erfassung von Technikbereitschaft. Diagnostica. **58**, 87–99 (2012)
23. Beierlein, C., Kovaleva, A., Kemper, C.J., Rammstedt, B.: Kurzskala zur Erfassung der Risikobereitschaft (R-1). In: Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS) (2015)
24. Cortina, J.: What is coefficient alpha? An examination of theory and applications. J. Appl. Psychol. **78**, 98–104 (1993)
25. Kraftfahrt-Bundesamt: Private Halter von Personenkraftwagen (2014). https://www.kba.de/DE/Statistik/Fahrzeuge/Neuzulassungen/Halter/2014/2014_kurzbericht_halter_n_pdf.pdf?__blob=publicationFile&v=6. Accessed 06 Feb 2020

How Visual Cues on Steering Wheel Improve Users' Trust, Experience, and Acceptance in Automated Vehicles

Arun Muthumani[1](✉), Frederik Diederichs[2], Melanie Galle[4], Sebastian Schmid-Lorch[3], Christian Forsberg[1], Harald Widlroither[2], Alexander Feierle[4], and Klaus Bengler[4]

[1] Autoliv Research, Wallentinsvägen 22, 447 37 Vårgårda, Sweden
{arun.muthumani, christian.forsberg}@autoliv.com
[2] Fraunhofer Institute for Industrial Engineering, Stuttgart, Germany
{frederik.diederichs, harald.widlroither}@iao.fraunhofer.de
[3] University of Stuttgart, Stuttgart, Germany
sebastian.schmidlorch@gmail.com
[4] Technical University of Munich, Munich, Germany
{melanie.galle, alexander.feierle, klaus.bengler}@tum.de

Abstract. With the introduction of ADAS systems and vehicle automation, an interface informing the driver of the automation state is required. This study evaluates the suitability of a visual interface comprising up to 64 LEDs on the steering wheel perimeter; it displays continuous visual feedback about the automation state—including notifications of an unscheduled hand-over due to sudden system failure. Three HMI (Human Machine Interface) designs were evaluated: two versions with visual cues on the steering wheel and one without (baseline). We implemented the designs in a driving simulator and compared the subjective responses of 38 participants to questionnaires measuring user experience, trust, and acceptance. The designs with visual cues improved the participants' user experience, as well as their trust in, and acceptance of, automated vehicles. Moreover, both designs were well perceived by participants.

Keywords: Human machine interaction · Visual cues on steering wheel · Visual interface · Human factors in automation · Trust in automated driving

1 Introduction

Vehicle automation changes the nature of driving tasks. Recent years have seen a significant rise in the integration of driving assistance systems aimed at supporting drivers with functions such as lane keep assistance, collision avoidance, route planning and navigation assistance [1, 2]. However, the driver is still responsible for object and event detection, classification, and response (OEDR), and must intervene if the system's performance is not appropriate. This level of automation is Level 2, or *Assisted Driving (ASD)*, according to the Society of Automotive Engineers (SAE) [3]. In Level

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 186–192, 2020.
https://doi.org/10.1007/978-3-030-50943-9_24

3, *Conditional Automation*, drivers can engage in non-driving related tasks although they must be fallback-ready, able to resume the driving task upon receipt of a take-over request if there is a sudden failure of the automated system. In Level 4 or *Automated Driving (AD)*, the driver is no longer considered a fallback option; however, driver-initiated transitions are possible. Hence, the OEDR requirements for the driver depend on the level of automation. A vehicle capable of supporting all automation levels is a complex system that needs to distinctly communicate the responsibilities and tasks of the user and the system at each level [4]. The right level of trust (neither too much nor too little) in automated systems is vital, not only for their usage and acceptance, but also for their safety and performance. Driver trust evolves over time, as the system successfully copes with actual driving situations (e.g., at road junctions, pedestrian crossings, etc.) without failure [5]. Like trust, user experience and acceptance can be evaluated using standardized questionnaires [6]. The results from an online survey indicate that vehicles with the currently available levels of automation receive high ratings on user experience and acceptance, but as the degree of autonomy increases this may change [7].

As automation increases, so does the driver's need for new information about the system, which requires additional interactive surfaces [8]. Visual interfaces, including ambient interior lighting, have been tested for their effectiveness at communicating the automated vehicle's state and limitations to the driver. For example, an LED strip near the windshield, which used a configurable lighting sequence to provide system information, enhanced the user's trust in the system [9]. Because the steering wheel is right in front of drivers, with good visibility and reachability, it is a prime location for vehicle interfaces. In fact, this concept is already implemented in some production and concept vehicles (Cadillac Super Cruise, Daimler Pilot, and Daimler Safety Car). Autoliv's research prototype z-force steering wheel [10], pioneering this concept, is investigated in this study in terms of user experience, trust, and acceptance.

2 Human Machine Interface

There are three HMI designs in this study: the baseline design provides auditory and visual cues about the automation system, while the other two, using the Autoliv prototype, provide additional visual cues generated by 64 multicolored LEDs on the steering wheel perimeter. The LEDs can be modulated to provide the desired luminance, color, and pattern. The touchpad interfaces on the left and right side of the steering wheel spoke allow the driver to physically interact with the system using thumb presses. Two different automated modes were realized in this study: the AD mode is activated by a thumb press on the left touchpad and the ASD mode is activated by simultaneous thumb presses on both left and right touchpads. In the Manual mode, the driver has sole responsibility for the driving task. The events signaled by the systems in all three HMIs were: mode availability, mode activation, and unscheduled transition to manual mode. (Note that all text in the interfaces was originally in German).

2.1 Baseline HMI Design

In the Baseline design, the availability of AD mode is conveyed by a “gong” sound and a voice message, with “*hands off*” icon and the text “*Automation available*” are displayed on the cluster (Fig. 1a). The successful activation of AD mode is communicated by a voice message, along with a “*hands off*” icon and “*Automation activated*” displayed on the cluster (Fig. 1b). The availability of ASD mode is conveyed by a voice message only; its successful activation is communicated with a voice message and the icons “*hands off*” and “*eyes on road*” on the cluster, along with the text “*Assisted driving activated*” (Fig. 1c). In the event of an unscheduled transition (in either mode), the vehicle makes a continuous “gong” sound, and the cluster displays the “*hands back on wheel*” icon (Fig. 1d). After the driver takes over control, the cluster once again displays conventional vehicle-related information until another automated system request is made. Concepts A and B both also use all the cues from the Baseline HMI design.

2.2 Concept A

The availability of AD mode is demonstrated when the top 14 LEDs turn blue (Fig. 1e). On activation, all the remaining LEDs illuminate in a circular pattern of blue around the steering wheel (Fig. 1f). The illumination level is slightly decreased after 5 s. The availability of ASD is conveyed when the top 14 LEDs change from blue to amber (Fig. 1g) and the illumination level is increased again. The LEDs are individually modulated in order to attract the driver’s attention. On activation, the moving pattern disappears and turns into static illuminance in amber. In case of sudden system failure, all LEDs start to pulsate in red at 1 Hz to get the driver’s attention (Fig. 1h). After driver take-over, all the mode-related cues are switched off.

2.3 Concept B

The availability of AD is conveyed when 28 LEDs (14 on each side of the steering wheel) start pulsing in turquoise (Fig. 1i). On activation, more LEDs on each side start to light up, forming a turquoise ring (Fig. 1j). The illumination level of the ring was slightly dimmed after 5 s. The availability of ASD is conveyed by switching off the top and bottom 14 LEDs, creating a dynamic pattern with normal illumination level to attract the driver’s attention. When ASD is activated, the top 14 LEDs remain switched off while the remaining 50 LEDs turn amber (Fig. 1k). In the event of sudden system failure, 46 LEDs (18 on top and 28 on the bottom) light up in a pulsating pattern (Fig. 1l), while the remaining 18 LEDs (nine on the left and nine on the right) are switched off. This pattern encourages the driver to grab the steering wheel at the 10 o’clock and 2 o’clock hand positions, which provide the best maneuvrability. When the driver grabs the steering wheel and takes over manual control of the vehicle, all LEDs are switched off.

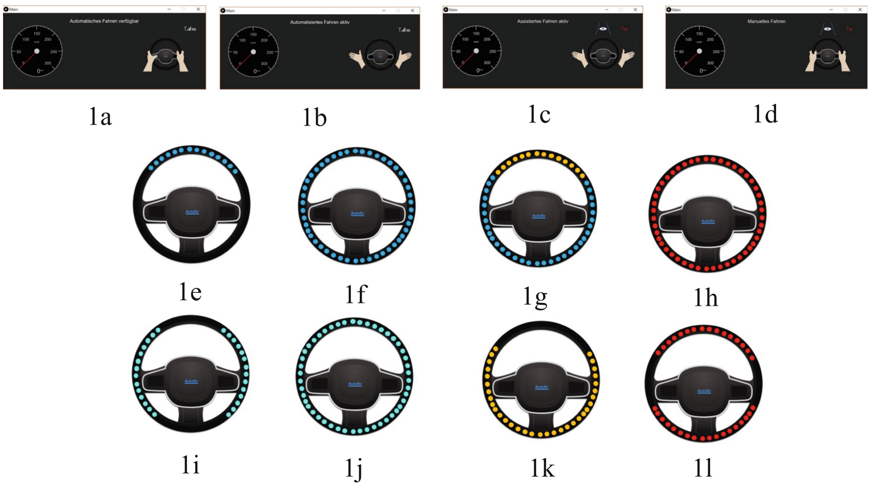


Fig. 1. All three HMI designs. 1a (Baseline-Availability of AD), 1b (Baseline AD activated), 1c (Baseline-ASD activated), 1d (Baseline-System failure); 1e (Concept A-Availability of AD), 1f (Concept A-AD activated), 1g (Concept A-ASD activated), 1h (System failure); and 1i (Concept B-Availability of AD), 1j (Concept B-AD activated), 1k (Concept B-ASD activated), 1l (Concept B-System failure).

3 Methods

This study was conducted with a Porsche Maccan vehicle mockup in the Fraunhofer IAO, immersive driving simulator. There were six projection screens: three in the front covering 180° of the driver's forward field of view, one on each side projecting the side mirror views, and one at the back of the mockup for the rearview mirror. The driver`s activities were recorded using a USB web camera in the interior. The simulation software SILAB version 5.0. collected data on the driving task at a rate of 60 Hz. A ten-inch tablet mounted on the center stack of the interior was used by the subjects to perform the Surrogate Reference Task (SuRT) when the vehicle was in AD or ASD mode. The data from the steering wheel, simulator, and SuRT were transmitted to SILAB for recording and synchronization. Thirty-eight subjects participated in the study. The data from five subjects were omitted for technical reasons; the remaining 33 consist of 18 female (55%) and 15 male (45%) drivers, between 25 and 61 years old (mean = 39.2 years; SD = 12.0 years).

3.1 Procedures and Data Collection

Subjects were informed about the study in a briefing session and then filled in the demographic information and consent forms. They had a five-minute practice session with the driving simulator, during which they became familiar with the different automation modes (including their activation procedures), the unscheduled transitions, and the SuRT. Following the practice session, the experimental drive was started, which lasted for nearly 18 min. The driving scenario consists of first exiting a parking space and then merging onto a European two-lane highway (speed limit 130 km/h). A few seconds later, the vehicle gets connected to the 5G network and initiates an automation availability request. After activation of the AD mode, subjects perform the SuRT task until a further request is received. Less than a minute after the start of automated driving, the vehicle loses the network connection and initiates an ASD mode request. In this mode, subjects are requested to monitor the vehicle while they are performing the SuRT task. Within a minute of driving in this mode, a system failure occurs along a curved section of the road in one of six possible locations, which were predefined—but unknown to the subjects (Fig. 2). Failure to respond to the take-over request results in a collision with the guardrail at the side of the road. All subjects drove the test scenario twice for each of the three HMI design concepts, so each participant drove six times. The trial order was randomized. The experiment was a within-subject design; the independent variables were the HMI design (Concept A, Concept B, and Baseline). The dependent variables were the subjective questionnaire results (user experience, trust, and acceptance) with Likert scale metrics. The collected data were tested for normality. For analyzing parametric datasets, ANOVA and Bonferroni-adjusted post-hoc analysis with t-test were used. For non-parametric datasets, Friedman T-test and Bonferroni-adjusted post-hoc analysis with Wilcoxon test were used.

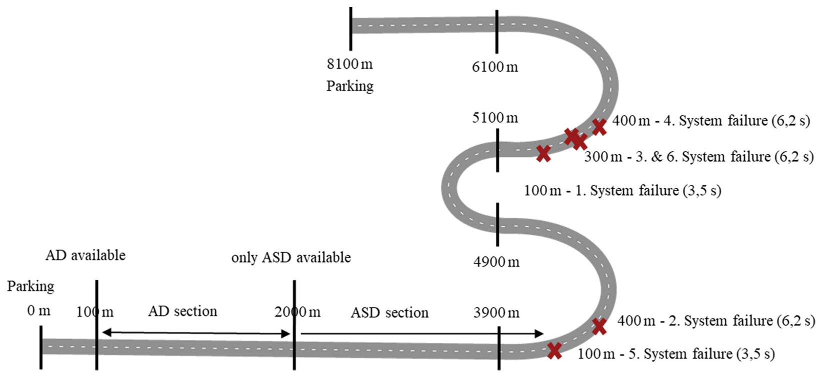


Fig. 2. Driving scenario showing the AD and ASD mode activations and the six possible system failure locations (red x's)

4 Results

The self-reported measure of *trust* used a two-item Likert scale. The results indicate that Baseline was significantly less trusted than either Concept A (p = .042) or Concept B (p = .006): see Fig. 3a. The *user experience* questionnaire consisted of six items (Fig. 3b): for *attractiveness,* Concept A and Concept B both scored higher than Baseline (p < .001 for both); for *efficiency*, there was no difference between Concept A and Baseline, but Concept B scored higher than Baseline (p < .17 and p = .002, respectively); for the item *perspicuity*, Concept A and Concept B both scored higher than Baseline (p = .006 for both); for *dependability*, both Concept A and Concept B scored higher than Baseline (p = .001 for both); similarly, for *stimulation* and *novelty*, Concept A and Concept B both scored higher than Baseline (p < .001 for all four tests). Results from the *user acceptance* questionnaire (Fig. 3c) show that Concept B scored significantly higher than Baseline (p = .001) and Concept A (p = .003). There was also a significant difference observed between Concept A and Baseline (p < .027).

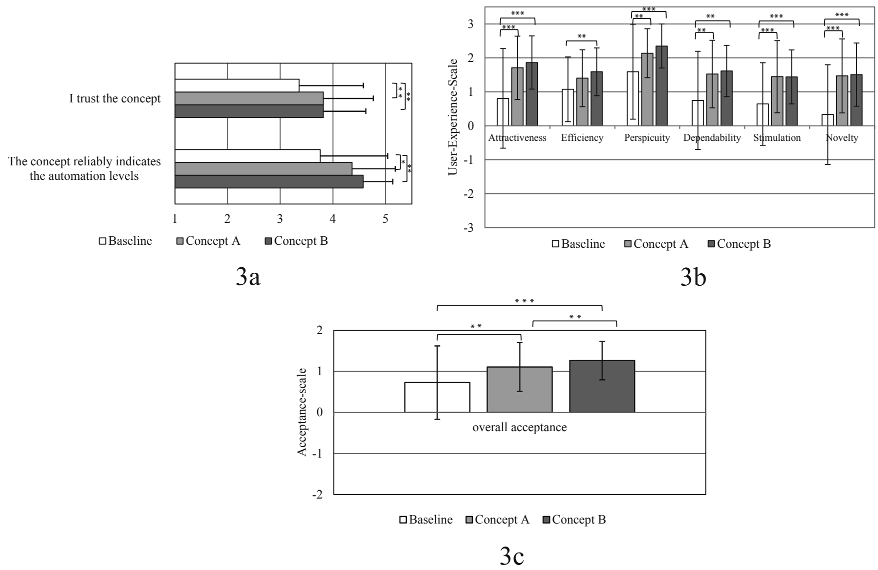


Fig. 3. Subjective responses: (3a) *Trust in automation*; (3b) *User experience*; (3c) *Acceptance* (***p < 0.001, **p < 0.05, p* < 0.1)

5 Discussion

In this study, we evaluated three HMI designs. While the baseline design provided both auditory and visual cues about the automated vehicle's state and limitations, the other two provided additional visual cues generated by up to 64 multicolored LEDs on the steering wheel perimeter, which led to a significant increase in drivers' trust in the system. This finding is in line with a study reporting that transparency of automation could build trust [9]. In terms of user experience, there was a significant difference between the baseline and the designs with visual cues on steering wheel. Except for *stimulation* and *novelty*, Concept B scored relatively higher than Concept A in all categories. One reason could be the distinctive colors differentiating AD (turquoise) and ASD (amber) modes only in Concept B, which might have enabled the drivers to perceive the vehicle status more easily. The results from the acceptance questionnaire further illustrate Concept B's superior performance, possibly for the same reason. Based on these findings, we conclude that visual cues on the steering wheel are important for increasing user experience, trust, and acceptance—ultimately resulting in an enhanced human-vehicle experience in the context of automated driving.

One limitation associated with this study had to do with the use of a driving simulator and the ambient illumination. Even though Fraunhofer IAO driving simulator has the possibility to simulate interior daytime lighting, it is not equivalent to bright sunlight condition. Another limitation is that the two HMI designs with visual cues were very similar to each other. Future investigations should examine more distinct designs under day time conditions. Besides that, the investigations should also consider the light properties (color, intensity) and its effect on driver state (sleep) in vehicle automation [11].

Acknowledgments. This study was a part of the ADAS & ME project, funded from European Union's Horizon 2020 research and innovation programme under grant agreement No. 688900. The authors would like to thank Kristina Mayberry for language revisions.

References

1. Melcher, V., Rauh, S., Diederichs, F., Widlroither, H., Bauer, W.: Take-over requests for automated driving. Procedia Manuf. **3**, 2867–2873 (2015)
2. Yang, Z., Shi, J., Zhang, Y., Wang, D., Li, H., Wu, C., Zhang, Y., Wan, J.: Head-up display graphic warning system facilitates simulated driving performance. Int. J. Hum. Comput. Interact. **35**, 796–803 (2019)
3. SAE: Surface Vehicle. SAE International, vol. 30 (2016)
4. Radlmayr, J., Ratter, M., Feldhütter, A., Körber, M.: Take-overs in level 3 automated driving – proposal of the take-over performance score (TOPS): Volume VI : Transport Ergonomics and Human Factors (TEHF). In: Presented at the Aerospace Human Factors and … Take-overs in Level 3 automated driving – Proposal of the take-over (2019)
5. Carsten, O., Martens, M.H.: How can humans understand their automated cars? HMI principles, problems and solutions. Cogn. Technol. Work **21**, 3–20 (2019)
6. Helgath, J., Braun, P., Pritschet, A., Böhm, P., Schubert, M., Isemann, D.: Investigating the effect of different autonomy levels on user acceptance and user experience in self-driving cars with a VR driving simulator. In: Design, User Experience, and Usability: Users, Contexts and Case Studies (2018)
7. Rödel, C., Stadler, S., Meschtscherjakov, A., Tscheligi, M.: Towards autonomous cars: the effect of autonomy levels on acceptance and user experience. In: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 1–8 (2014)
8. Strömberg Helena, K., Lars-Ola, B., Marianne, K.: HMI of autonomous vehicles - more than meets the eye. In: Proceedings of the 20th Congress of the International Ergonomics Association, pp. 359–368 (2018)
9. Yang, Y., Karakaya, B., Dominioni, G.C., Kawabe, K., Bengler, K.: An HMI concept to improve driver's visual behavior and situation awareness in automated vehicle. In: IEEE Conference Intelligent Transportation Systems Proceedings, ITSC, November 2018, pp. 650–655 (2018)
10. Intutive Steering Wheel (2014). https://www.autoliv.com/products/electronics/human-machine-interaction
11. Hirsch, M., Diederichs, F., Widlroither, H., Graf, R., Bischoff, S.: Sleep and take-over in automated driving. Int. J. Transp. Sci. Technol. **9**(1), 42–51 (2019)

The Impact of a Biological Driver State Monitoring System on Visual Attention During Partially Automated Driving

Alice Stephenson[1(✉)], Iveta Eimontaite[2], Praminda Caleb-Solly[1], and Chris Alford[1]

[1] University of the West of England, Bristol, UK
{Alice.Stephenson,Praminda.Caleb-Solly, Chris.Alford}@uwe.ac.uk
[2] Cranfield University, Bedford, UK
Iveta.Eimontaite@cranfield.ac.uk

Abstract. As the shift from manual to automated driving occurs, the driver will be required to take a supervisory role in monitoring the driving environment and system parameters. Driver State Monitoring Systems (DSMS) have been proposed to evaluate the state of the driver and provide support for driver engagement. However, it is not clear how a DSMS may impact attentional mechanisms. Nineteen young adults (mean ± SD age = 19.58 ± 0.94 years) experienced a simulated semi-autonomous driving journey. Participants' visual attention via eye tracking fixation and visit metrics were compared before, during, and after two distinct notifications designed to enhance driver engagement. The first notification displayed biofeedback changes in physiological state; the second notification provided speed limit changes. Results revealed participants spent longer attending to the outside driving environment during biofeedback. The results suggest the potential for feedback based on relevant physiological parameters to enhance global visual processing strategies during semi-autonomous driving.

Keywords: Semi-autonomous vehicle · Notifications · Biofeedback · Visual attention

1 Introduction

The process of selective attention enables us to gather relevant information in the visual environment. This process is largely successful across complex and dynamic environments where sustained visual attention is distributed over a large region of the visual field. During real-world attention tasks such as driving, accurately scanning and responding appropriately to the situation is critical for safety. Driver inattention has been shown to predict driving performance [1] and has been long recognised as a leading factor contributing to road crashes [2]. Although fully autonomous vehicles aim to eradicate all human input, as the shift from manual to automated driving ensues, the driver will be required to take a supervisory role monitoring the driving environment and system parameters [3]. Although the role of the driver will become more

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 193–200, 2020.
https://doi.org/10.1007/978-3-030-50943-9_25

passive, it is vital that they still engage in successful visual processing of the driving environment to prepare for intervention when required [4].

Insufficient monitoring of an automated system and the environment has been found to be a contributing factor in major transportation incidents. Research has found that incidents occur due to an insufficient response by the operator due to over-reliance on the system, delayed responses, and a failure to act [5]. Deficit attentional strategies have consistently been shown to be implemented. Self-reported over-reliance of an automated system has been negatively related to frequency of monitoring of automation [6]. Individual differences in sustained attention have been related to delayed responses [7]. Furthermore, automation complacency is commonly found in multi-tasking environments, where attention allocation favours other tasks over automation monitoring [5]. This has important implications for semi-autonomous vehicles, where passive drivers will engage in non-driving related activities such as reading.

Subsequently, Driver State Monitoring Systems (DSMS) have been proposed to monitor driver status and provide necessary assistance [8]. A main function of a DSMS is to encourage driver engagement to assist with perception of the driving environment to detect potential risks. Utilising biophysical sensors to monitor the drivers' physiological state enables the system to detect potentially distracted drivers. The advanced system will invoke action to prevent critical situations. In some instances, this may be haptic, visual or auditory feedback; in other instances, the system may take the necessary steps for vehicle safety. Yet, a DSMS may have the reverse anticipated effect on attention. Visual and auditory notifications encourage divided attention which may make it more difficult to process multiple stimuli inside and outside the vehicle [9]. Feedback of changes in physiology may induce anxiety and negatively impact alerting and orienting attentional networks [10]. Therefore, it is important to understand how notifications of changes in driver state might impact driver attention during monitoring of a semi-autonomous vehicle.

An experiment was designed to investigate attentional allocation during notifications designed to enhance driver engagement. Using a Wizard of Oz within-subject protocol, participants experienced a simulated semi-autonomous journey. During the journey, two notifications were presented; the first notification applied biofeedback data of changes in the participants' physiology, and the second notification of speed limit changes. Overt visual attention shifts evoke certain patterns of responses in eye movements and therefore visual attention was measured via eye tracking fixation and visit metrics before, during, and after the presentation of a notification.

2 Method

2.1 Participants

Thirty-three healthy young adults (21 females, mean age ± SD = 20.06 ± 1.48 years, range 18–25 years) took part in the study. Nineteen participants data were used in the final analyses (12 females, mean age ± SD = 19.58 ± 0.94 years, range 18–20 years) due to five participants experiencing technical failures and a further eight participants having different experiences with the notifications (varying in duration or appearing in different journeys with different instructions) as a result of simulator traffic and participants not following the script for journey set up.

Individuals with severe motion sickness, severe health conditions (i.e. epilepsy, neurological impairments) and uncorrected vision or hearing were excluded. Ethical approval was obtained by the Faculty of Health and Applied Sciences University of the West of England Research Ethics Committee.

2.2 Design and Materials

Design. This study used a within-subject design to investigate the impact of notifications (pre-notification vs. during notification vs. post-notification) and type of notification (biofeedback notification vs. speed limit change notification) on attentional mechanisms measured via eye tracking.

Simulated Environment. The semi-autonomous simulation environment consisted of a static Lutz Pathfinder pod and three large forward projector screens with a display resolution of 1280 × 1024 providing a 210° horizontal forward field of view. The driving simulation was generated by the SCANeR II® software (OKTAL Sydac, France) and consisted of a two-lane road passing through a neighbourhood with traffic passing in the opposite direction. The Lutz Pathfinder pod was a two-seater vehicle with two rear-hinged doors. The participant was told to interact with the vehicle using the in-vehicle Human-Machine Interface (HMI; see Fig. 1).

Human-Machine Interface (HMI). The HMI was presented on HannsG HT161HNB 15.6″ Multi-Touch Screen connected to a Kodlix GN41 Mini PC (Windows 10, Intel Celeron processor, 8/4 GB RAM, 64 GB). The main design and development were informed by a literature review [11] and feedback from previous iterations of the HMI used in earlier research studies [12, 13].

Journeys. Participants underwent eight randomised journeys. During one journey, participants were told they might be required to take control by stopping the vehicle. As we were interested in participants' monitoring behaviour, only this journey is described and subsequently analysed for this paper.

At the beginning of the journey, participants were provided with a scenario that suggested they might need to intervene with the vehicle behaviour by pressing the 'Safe Stop' button on the HMI. They were told that their physiological state was being monitored to assist vehicle safety, via the Empatica E4 wristband (Empatica Inc., Cambridge, MA, USA and Milan, Italy). During the journey, two distinct auditory and visual HMI notifications were presented: "Your physiology suggests that your reaction times are slowed. The vehicle will slow down"; and, "Approaching the town centre. Speed limit change". These notifications were pre-programmed and occurred at the same time during each journey. Journey start and endpoints were randomised so the message to appear first was different between participants.

Once the messages appeared, the vehicle reduced its speed. The participant was not provided with an opportunity to turn the notifications off, nor could they increase the speed of the vehicle. After approximately 30 s (depending on traffic and traffic lights), the notifications disappeared, and the vehicle returned to the appropriate speed limit. Overall, the journey lasted approximately ten minutes.

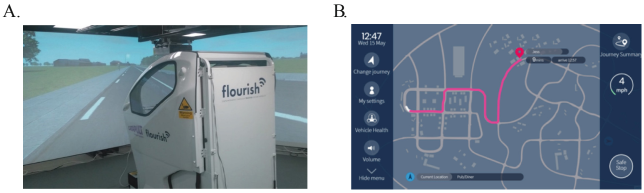


Fig. 1. *(A)* Exterior of the Lutz pod in front of three large forward projector screens. *(B)* The HMI default screen during the journeys.

2.3 Measurements and Pre-processing

Tobii Pro Glasses 2 was used to collect fixation and visit metrics (Tobii Technology, Stockholm, Sweden). The Tobii Glasses are a lightweight, wearable eye tracker. Sampling rate was 100 Hz. The manufacturer's calibration procedure was followed.

Eye tracking analysis was undertaken using Tobii Pro Lab software (Tobii Technology, Stockholm, Sweden). Eye tracking glasses captured a mean of 80% (SD = 0.12%) of gaze samples. Times of Interests (TOIs) were defined as three separate periods representing the 30 s before, during, and after presentation of the notification. Gaze data from the recording were manually mapped onto an image best depicting the overall visual view of the participant. Two Areas of Interests (AOIs) were defined to represent the HMI and the external driving environment. To finish, the raw gaze filter was applied to the data. Average visit duration, average fixation duration, time to first fixation metrics were exported (see Table 1).

Table 1. Visual attention eye tracking measures.

Measure	Description	Interpretation
Average visit duration	Average time spent inside an AOI. Can include many fixations	How long are they attending to different stimuli in the AOI?
Average fixation duration	Average time of a fixation within an AOI	How long are they attending to specific stimuli in an AOI?
Time to first fixation	The time it takes to the first fixation in an AOI	How quickly does an AOI capture attention?

2.4 Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 26 (IBM Corp., Armonk, N.Y., USA). Shapiro-Wilk test of normality and visualisation of QQ plots of the unstandardized residuals indicated that the data were not normally distributed. Therefore, data were normalised via the natural logarithm.

A mixed model with the maximum likelihood method was utilised as the dataset included missing observations due to loss of gaze points and absence of gaze data within an AOI for some participants. Fixed effects were Message (biofeedback vs. speed limit change), Time (pre vs. during vs. post), and AOI (HMI vs. driving environment). Post hoc Bonferroni-adjusted t-tests were performed.

3 Results

3.1 Eye Tracking

Visit Duration. The mixed model revealed significant main effect for AOI ($F(1, 173) = 271.88$, $p < .001$). Significant two-way interaction effects were found for Message × AOI ($F(1, 173) = 17.26$, $p < .001$), Message × Time ($F(2, 173) = 17.29$, $p = .02$), AOI × Time ($F(2, 173) = 10.54$, $p < .001$), and for a three-way interaction ($F(2, 173) = 233.41$, $p < .001$). Three-way interaction post-hoc comparisons revealed visit duration of the driving environment was greater during the biological feedback message compared to the speed limit message ($p = .049$; Fig. 2); and visit duration of the HMI was greater after the speed limit message compared to biofeedback message ($p < .001$). Visit duration was longer in the driving environment compared to the HMI (all $ps \leq .001$). In addition, visit duration of the driving environment was greater before both messages (all $ps < .02$). During the biofeedback message, visit duration of the HMI was greater compared to before and after the message (all $ps < .03$).

Fixation Duration. The mixed model on average fixation duration found a significant main effect for AOI ($F(1, 172) = 30.32$, $p < .001$) as well as a three-way interaction effect ($F(2, 172) = 3.17$, $p = .04$; Fig. 3). Three-way interaction post-hoc comparisons revealed fixation duration on HMI was marginally significantly greater before the speed limit notification, compared to the biological feedback notification ($p = .052$). In addition, fixation duration was significantly greatest on the driving environment compared to the HMI before and after the biological feedback notification; and during the speed notification (all $ps \leq .001$).

Time to First Fixation. The mixed model revealed a significant main effect for AOI ($F(1, 174) = 48.92$, $p < .001$) and a two-way interaction effect between AOI × Time ($F(2, 174) = 4.01$, $p = .02$). Pairwise comparisons of the interaction effect revealed that time to first fixation was significantly quicker for the driving environment when compared to the HMI before, during, and after both notifications (all $ps \leq .01$).

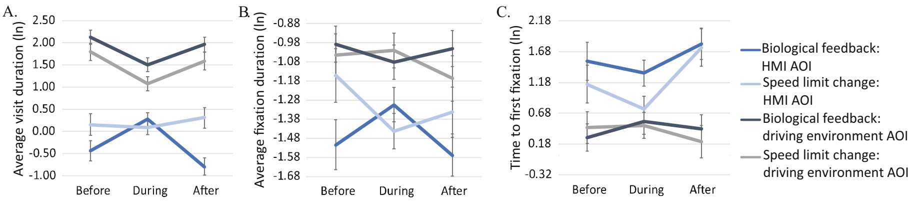


Fig. 2. Eye tracking measures (normalised) before, during, and after the biological feedback and speed limit change messages. *(A)* Average visit duration; *(B)* Average fixation duration; *(C)* Time to first fixation. Error bars represent ± SEM.

Fig. 3. Gaze plots represent visual behaviour before, during, and after the *(A)* Biological feedback notification and *(B)* Speed limit change notification for one participant only.

4 Discussion

The current study investigated the effect of a biological feedback notification compared to a speed limit change notification on visual attention of participants who underwent a simulated semi-autonomous journey. Visual attention was measured via eye tracking visit and fixation metrics of the driving environment and HMI.

In terms of visit duration, participants spent longer looking at the driving environment during the biological feedback notification, compared to the speed limit notification. As visit duration is defined as the time when a participant focuses on an AOI until they look away, this suggests that participants were engaged with monitoring the overall driving environment for longer during biological feedback.

However, visits to the driving environment were longer before either notification. This is logical, as in some cases, participants did not look at the HMI before the notification (as seen in Fig. 2). The scan paths, however, reveal that ocular attention was focused centrally, rather than globally. Future research should consider utilising metrics such as gaze entropy to capture such changes in visual scanning behaviour [14].

Participants demonstrated shorter visits to the HMI (compared to the driving environment) before, during, and after both notifications. This suggests that notifications did not decrease attentional resources. Time to first fixation also supports this notion as participants were quicker to attend to the driving environment compared to the HMI. Therefore, notifications did not trigger significant distraction. Although eye tracking is unable to detect the periphery of a participants' visual gaze, stimuli can be perceived preattentively in peripheral vision, and it is well known that attention can occur without a change of fixation [15]. Participants may have therefore discerned the notification and inhibited saccadic movement for further processing.

Fixation duration revealed fixations on the driving environment were longer (compared to the HMI) before and after the biological feedback notification, and during the speed notification. Therefore, during biological feedback, the decrease in fixation duration and the increase in visit duration suggests that once stimuli were detected and encoded, the focus of attention shifted away to another external element in the scenario.

Taken together, these results reveal that a biological feedback notification encouraged efficient global visual processing of a dynamic driving environment. The eye tracking data suggest a minimal distracting effect of the notifications with either biological feedback or speed limit change. This shows the beneficial potential for utilising HMI notifications for biofeedback based on relevant physiological parameters to enhance improved attentional strategies during partially automated driving.

References

1. Kircher, K., Ahlstrom, C.: Predicting visual distraction using driving performance data. Ann. Adv. Automot. Med. **54**, 333–342 (2010)
2. Klauer, S., Dingus, T., Neale, V., Sudweeks, J., Ramsey, D.: The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data [NHTSA Report No. DOT HS 810 594]. Virginia Tech Transportation Institute, Blacksburg (2006)
3. SAE International: U.S. Department of transportation's new policy on automated vehicles adopts SAE International levels of automation for defining driving automation in on-road motor vehicles (2018)
4. Braunagel, E., Kasneci, W., Stolzmann, W.: Driver-activity recognition in the context of conditionally autonomous driving. In: IEEE Conference on Intelligent Transportation Systems, pp. 1652–1657. IEEE Computer Society, Washington DC (2015)
5. Parasuraman, R., Manzey, D.H.: Complacency and bias in human use of automation: an attentional integration. Hum. Factors **52**, 381–410 (2010)
6. Hergeth, S., Lorenz, L., Vilimek, R., Krems, J.F.: Keep your scanners peeled: gaze behavior as a measure of automation trust during highly automated driving. Hum. Factors **58**, 509–519 (2016)
7. Jipp, M.: Reaction times to consecutive automation failures: a function of working memory and sustained attention. Hum. Factors **58**, 1248–1261 (2016)
8. Khan, M.Q., Lee, S.: A comprehensive survey of driving monitoring and assistance systems. Sensors **19**, 2574 (2019)
9. Lee, J.D.: Driving Safety. Rev. Hum. Factors Ergon. **1**, 172–218 (2005)
10. Pacheco-Unguetti, A.P., Acosta, A., Callejas, A., Lupiáñez, J.: Attention and anxiety: different attentional functioning under state and trait anxiety. Psychol. Sci. **21**, 298–304 (2010)

11. Morgan, P., Caleb-Solly, P., Voinescu, A., Williams, C.: Literature review: human-machine interface. Proj. rep. UWE Bristol (2016)
12. Morgan, P.L., Voinescu, A., Williams, C., Caleb-Solly, P., Alford, C., Shergold, I., Parkhurst, G., Pipe, A.: An emerging framework to inform effective design of human-machine interfaces for older adults using connected autonomous vehicles. In: Stanton, N.A. (ed.) Advances in Human Aspects of Transportation, pp. 325–334. Springer International Publishing (2018)
13. Eimontaite, I., Voinescu, A., Alford, C., Caleb-Solly, P., Morgan, P.: The impact of different human-machine interface feedback modalities on older participants' user experience of CAVs in a simulator environment. In: Stanton, N.A. (ed.) Advances in Human Factors of Transportation, pp. 120–132. Springer International Publishing (2019)
14. Shiferaw, B., Downey, L., Crewther, D.: A review of gaze entropy as a measure of visual scanning efficiency. Neurosci. Biobehav. Rev. **96**, 353–366 (2019)
15. Posner, M.I.: Orienting of attention. Q. J. Exp. Psychol. **32**, 3–25 (1980)

Analysis of Public Transport Ridership During a Heavy Snowfall in Seoul

Seonyeong Lee, Minsu Won(✉), and Seunghoon Cheon

The Korea Transport Institute, 370 Sicheong-Daero, Sejong-Si, Republic of Korea
{sylee708,minsuwon,sh1000}@koti.re.kr

Abstract. Severe weather conditions, such as heavy snowfall, rain, heatwave, etc., may affect travel behaviors of people and finally change traffic patterns in transportation networks. Hence, this study has focused on the impacts of a weather condition on travel patterns of public transportations, especially when a heavy snowfall which is one of the most critical weather conditions. First, this study has figured out the most significant weather condition affecting changes of public transport ridership using weather information, card data for public transportation, mobile phone data; and then, developed a decision tree model to determine complex inter-relations between various factors such as socio-economic indicators, transportation-related information. As a result, the trip generation of public transportations in Seoul during a heavy snowfall is mostly related to average access times to subway stations by walk and the number of available parking lots and spaces. Meanwhile, the trip attraction is more related to business and employment densities in that destination.

Keywords: Weather condition · Public transport ridership · Card data · Mobile phone data · Decision-Tree model

1 Introduction

The pattern of public transportation is related to various factors such as the number of cars owned, valuables, tolls, service levels, road conditions, weather conditions, etc., of which the weather conditions may affect decision making in a short time. It is one of the important variables [1]. In particular, global warming and rapidly changing weather conditions and the associated uncertainties in daily traffic patterns are one of the most important issues in the field of transportation, where a predictable and stable system must be maintained. Nevertheless, many existing studies fail to effectively analyze the relationship between public transport use patterns and weather conditions due to regional specialities, lack of data, etc.

Therefore, this study analyzes the change pattern of public transportation usage according to various weather conditions such as heavy rain, heavy snow, intense heat, and cold wave in Seoul City, and tries to understand the relationship more clearly.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 201–210, 2020.
https://doi.org/10.1007/978-3-030-50943-9_26

2 Preceding Research

Research on traffic characteristics, changes in demand, and usage patterns of public transportation has been extensively studied in Korea and overseas using various influence variables such as weather conditions. First, Choi, Rhee et al. In [1], Categorized into city buses, maul buses and subway means, confirming the differences, discomfort index and sensible temperature weather conditions appear to reduce demand for public transportation by about 2 to 7%, buses have higher correlation than subway Was confirmed. Shin and Choe [2] analyzed the effect of precipitation on changes in the number of passengers in public transportation by applying the SUR model. The study started with the hypothesis that traffic volume by means of travel would be affected. The main variables are the number of passengers on buses, urban railways, and village buses on a day, and work precipitation. The analysis showed that when the precipitation was more than 10 mm, the number of public transportation passengers decreased.

Stover and McCormack [3] analyzed the relationship between weather conditions and demand for public transport by season using a linear regression model. The analysis confirmed that strong winds have a negative effect on public transport demand in spring, autumn and winter, and that lower temperatures (below −7 °C) reduce public transport demand in winter. Kashfi, Lee et al. [4] analyzed the relationship between bus utilization and precipitation using a linear regression model. At peak times in the morning, weekend passengers confirmed that as rainfall increased, bus traffic decreased. In particular, it was found that the effects on precipitation and bus traffic were more sensitive in summer compared to other seasons. Arana, Cabezudo et al. [5], we analyzed the effect on bus traffic in consideration of the purpose of traffic using a multiple regression model in order to analyze the user traffic pattern of public transportation according to the weather conditions. When using the bus for leisure purposes, it was confirmed that as the wind and rainfall increased, the number of bus passes decreased, and that a rise in temperature increased the number of passes. Zhou, Wang et al. [6] analyzed the effects of weather conditions on regular public transport users. The analysis confirmed that increasing humidity, wind and rainfall had a reduced effect on public transport utilization at certain levels.

Existing research has shown that weather conditions also play an important role in the variety of variables affecting public transport paths and usage patterns.

In particular, previous studies have analyzed changes in public transportation using only materials that can estimate public transportation usage. For this reason, there is a limit in grasping whether a change in the traffic pattern of public transportation according to each weather condition is a potential change or a change in the means of traffic. Therefore, in this study, switching between pass means and potential I want to clarify changes in demand.

3 Data

In order to analyze changes in public transportation usage due to weather, weather report information provided by the Korea Meteorological Administration (Weather report), card data for public transportation in Seoul, and mobile phone data, socio-economic indicators [14], and other transportation-related information (Table 1).

Table 1. Descriptions of datasets

Dataset	Temporal scope	Spatial scope	Note
Weather report	2016.1.1 ~ 2017.12.31	Korea	Weather information, per every day
Card data for public transportation	2016.4.1 ~ 2017.03.31	Seoul	# of public transportation trip information, per every day
Mobile phone data	2016.4.1 ~ 2017.03.31	Korea	# of trip information, per every 15 min
Socio economic indicators	2016.1.1 ~ 2017.12.31	Korea	Business density Employment density, etc.
Other transportation related information	2016.1.1 ~ 2017.12.31	Korea	#of parking lots, access time to a subway, etc.

Mobile phone data has a different spatio-temporal unit from other data, it is necessary to standardize using each spatio-temporal unit for analysis. For this reason, a polygon for each base station was constructed using the Voronoi diagram technique.

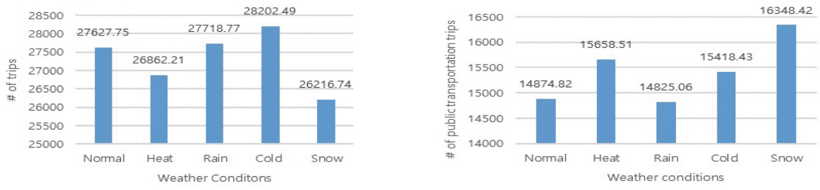
A breakdown of public transport card usage is collected for each event and includes ride and drop-off station ID information. These pieces of information are matched with bus stop information to calculate position information, and are configured in a polygon unit which is a collection unit of a mobile phone data. In this study, the Seoul Data Center (2018) and the Korea Transport Institute internal data (2015) were used. After establishing various socio-economic and traffic-related indicators, such as the number of employees, the number of businesses, the number of subway stations and bus stops, and the approach times of access to subway stations, it was used to analyze changes in public transportation use by weather conditions.

4 Analysis and Result

4.1 Analysis of Public Transport Ridership

The Korea Meteorological Administration uses the temperature, humidity, wind speed, and precipitation information of this and surrounding work to ferment weather information in four formats, including intense heat, heavy snow, heavy rain, and cold waves [7].

Under the assumption that such weather information will change the decision of traffic between vehicles and public transport users, we will verify using public transport card data and mobile phone data.



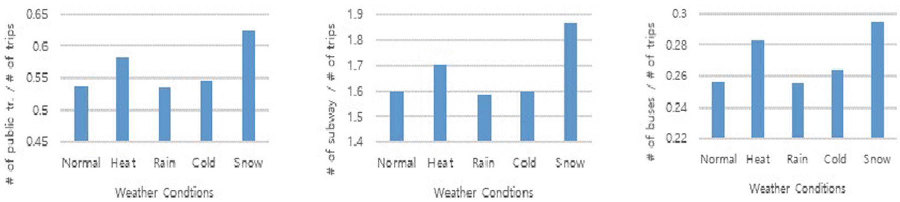
(a) mobile phone data (b) card data

Fig. 1. # of trips according to weather conditions

Table 2. Unpaired two-samples t-test for # of trips based on weather conditions

Weather conditions (Normal vs.)	Heat	Rain	Cold	Snow
(a) # of trips by mobile phone data	0.0009	0.7497	0.1533	0.0101
(b) # of public transportation trips by card data	0.0128	0.8917	0.2965	0.0679

Figure 1 (a) shows the difference in the traffic volume (departure+arrival) of the entire city of Seoul estimated by using mobile communication data by weather condition, and Table 2 shows the difference of these traffic volumes. Figure 1 (b) and the second part of Table 2 show the difference in public transport usage rates across Seoul City estimated by public transport card usage data by weather condition. Looking at the results, it can be confirmed that the usage rate of public transportation has increased significantly during the heat and heavy snowfalls compared to usual.



(a) Ratio of public transportation trips (b) Ratio of subway trips (c) Ratio of bus trips

Fig. 2. Ratio of public transportation trips by modes and weather conditions

In Fig. 2, it is necessary to confirm that the use rate of public transport during heavy snowfall increased relatively most. Subways and buses also show the highest usage during heavy snowfall. However, it was found that there was no special difference in the pattern of changes in subway and bus utilization. The use of public transport in Seoul changes significantly during intense heat and heavy snow. This phenomenon means that during intense heat and heavy snow, some people abandon traffic due to bad weather, or people who used existing cars and walking use public transportation. Is shown. This means that the potential demand decreases and the means of public transport change at the same time for the whole traffic.

4.2 Changes in Public Transport Usage by Region

In order to analyze the changes in public transport use estimated using public transport card data on a regional basis, the changes in public transport use in Seoul during heavy snowfall and the social The characteristics of economic indicators were investigated.

Fig. 3. Classification of area

Figure 3 shows the eight service areas of Seoul City classified according to the characteristics of public transport paths [11]. Table 3 shows changes in the use of public transport in these areas and various socio-economic and transport-related indicators. It can be confirmed that the public transport utilization rate in Gangdong and Songpa (area-4), which are located in the east of Seoul, increased the most during heavy snowfall. On the other hand, it can be seen that the use rate of public transportation is relatively small and increased in Gwanak, Kinsen, and operation (area-6) areas. The use rate of public transportation is relatively small and the area of increase-6 is relatively small in the worker density and the average number of subway stations per polygon (1.1). Average subway access times using transportation were relatively long. In other words, when the heavy snowfall changes the use of public transportation in Seoul by region, it can be confirmed that the density of workers, the number of subway stations, and the access time have a significant effect.

Table 3. Trip changes according to regions and their socio-economic conditions

Classification of area	1	2	3	4	5	6	7	8
Origin								
Trips under normal conditions	896,062	818,737	750,595	516,898	1,032,450	730,494	976,339	768,644
Trips under heavy snowfall	982,217	903,123	831,257	603,145	1,128,231	794,579	1,097,499	843,999
Rate of change (%)	**9.6%**	**10.3%**	**10.7%**	**16.7%**	**9.3%**	**8.8%**	**12.4%**	**9.8%**
Destination								
Trips under normal conditions	898,449	803,913	755,128	530,133	1,023,043	725,879	989,043	764,601
Trips under heavy snowfall	983,691	889,184	833,025	613,424	1,117,362	789,332	1,111,490	843,692
Rate of change (%)	**9.5%**	**10.6%**	**10.3%**	**15.7%**	**9.2%**	**8.7%**	**12.4%**	**10.3%**
Business density	1077.8	346.3	468.7	480.2	338.9	405.8	545.6	462.5
Employment density	5712	1404.1	2171.4	2827.1	3083	**1993.3**	2807.1	2335.7
Average number of bus stations by polygon[a]	64.9	87.2	64.1	**50.6**	65.7	76.8	85.7	91.7
Average number of subway stations by polygon	3.5	1.3	2.4	1.8	2.1	**1.1**	1.7	2.2
Average number of parking lots by polygon	118.7	153.7	178.1	209.6	148.8	232.4	204.1	279.6
Average number of parking spaces by polygon	1902.7	2070.3	2008	2875.9	2894	2175.3	3088.9	2803.9
Average access time to subway stations by the other transit (minutes)	9.8	13	10.6	9.6	10.7	**13.3**	11.6	10.9
Average access time to subway stations by walk (minutes)	7.2	9.8	7	7.2	7.8	**10.4**	8	7.7

[a]Polygon: a standard unit space to collect mobile phone data in this study

4.3 Analysis of the Main Effect of Change Public Transport

The various socio-economic and transport-related indicators that can be used as explanatory variables are highly correlated with each other [12] and cannot guarantee a linear relationship with changes in public transport use. In particular, a Classification Model is advantageous for finding relatively noticeable changes in public transport usage that are extreme in distribution, rather than changes in average public transport usage [8]. Classification models such as the decision-tree model are free to assume the distribution of each variable and are more convenient than explaining extreme values of the distribution [8, 9] Therefore, in this study, a classification model was constructed using the CART algorithm [10], which is a typical algorithm of the decision tree model [13].

Step-1: Divide the rate of increase or decrease in the use of public transportation "Origin" and "Destination", and set the rate of decrease/decrease, Increase 0 ~ 5, 5 ~ 10, 10 ~ 15, 15 ~ 20, 20 or more increase".

Step-2: The decision tree is constructed using the CART algorithm based on the rate of increase or decrease in the use of the public transportation used in Step-1 and various socioeconomic and basic traffic indicators.

Step-3: Deriving the final result of each terminal node by combining it with surrounding categories to ensure the reliability of each terminal node of the derived decision tree about 60% or more I do (Table 4 and Fig. 4).

Table 4. Frequency of each category of trip

Category of trip changes	Origin	Destination
Decreased	178 (11.3%)	152 (9.6%)
0–5% increased	165 (10.4%)	162 (10.2%)
5–10% increased	319 (20.2%)	348 (22.0%)
10–15% increased	317 (20.1%)	329 (20.8%)
15–20% increased	223 (14.1%)	244 (15.4%)
$\geq$ 20% increased	379 (24.0%)	349 (22.0%)
Total	1581 (100%)	1584 (100%)

The results show that during heavy snowfalls, the change in the usage rate public transport is as follows: "Average bus stops (bus)", "Average number of subway stations (subway)", "Average subway station walking access time (walk minutes)", " average subway station transit times, minutes", " average parking lots", " average parking spaces", "business density", It can be confirmed that it is described in "employee density (employment)" and its condition combination (Combination of Conditions). For example, in Fig. 5, if the average subway station walking time is 3.7 minutes or more and the average number of parking lots is 485 or more, public transport ridership during heavy snowfall will increase by about 15% or more.

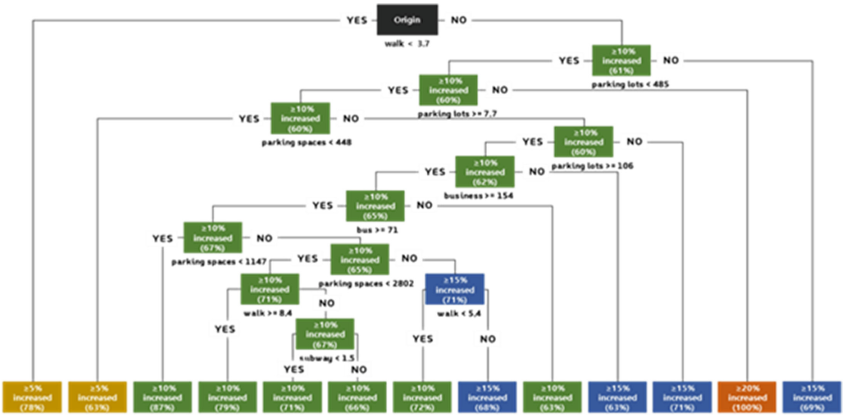


Fig. 4. Classification rules for trip changes of public transportation during heavy snowfalls (based on origin)

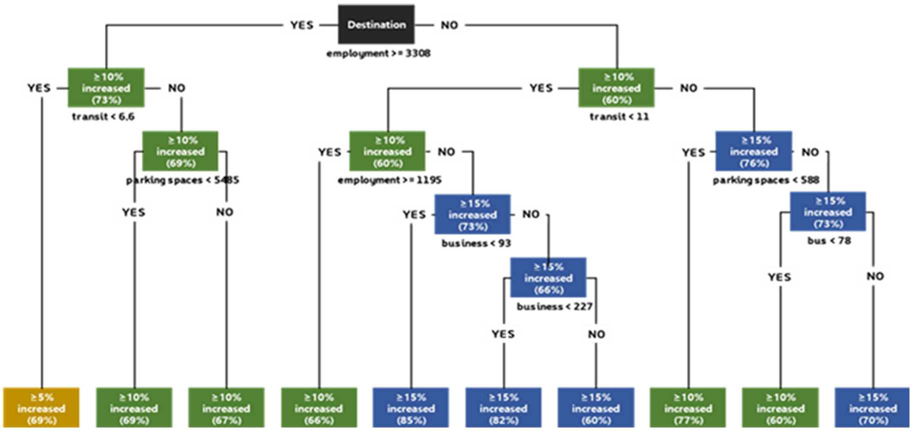


Fig. 5. Classification rules for trip changes of public transportation during heavy snowfalls (based on destination)

In particular, the first condition that affects the change in starting point reference public transport usage is the walk time for access by public transport. On the other hand, the first condition that affects the change in the use of the end point standard public transportation is the number of workers at the end point.

5 Conclusion

In this study, we analyzed the changes in public transportation usage according to weather conditions in Seoul City. First, using the weather, mobile phone communication, and public transportation card data, we look at the changes in traffic patterns and public transportation usage by each weather condition (hot weather, heavy rain, cold waves, and heavy snow). Set the situation. In the case of intense heat and heavy snowfall, the traffic volume in Seoul, which was estimated as mobile phone communication data, decreased, and the use of public transport, which was estimated as public transport card data, increased. In particular, it has been confirmed that such a phenomenon is conspicuous in heavy snow. Next, we analyzed the eight areas of Seoul. In the Gangdong and Songpa areas in the east of Seoul, and in the Gangseon, Kuro, Yangcheon and Yeongdeungpo areas in the west, public transportation usage has increased, and Gwanak, Gimcheon and Dongjak areas have confirmed not a large changes relatively. Also, the Gangdong and Songpa areas were relatively easy to access by public transportation. The Yangcheon and Yeongdeungpo districts are not easily accessible by these public transportations, but it is likely that the existing passengers have switched from the Gimpo and Incheon areas west of Seoul.

Finally, we constructed a decision tree that can analyze and estimate changes in Seoul public transport usage during heavy snowfall using polygon units and socio-economic transport related indices. Analysis shows that changes in public transport departures during heavy snowfall play important roles, such as pedestrian access pass times and parking availability near bus stops, and changes in arrivals depend on the arrival workers and businesses. It turned out to be closely related to density. In other words, it could be inferred that despite the decrease in overall traffic volume during heavy snowfall, the target path for commuting switched means mainly in areas easily accessible by public transport.

In this study, using the weather information, public transport card data, mobile phone data, socio-economic and traffic-related indicators, we analyzed the changes in public transport use in Seoul by weather conditions, it makes sense to clarify the vague and complex relationships of the main causes more clearly by using classification models. However, it is hoped that the findings of this study can be used in the future to create more accurate and accurate models of public transport usage patterns by weather conditions.

Acknowledgment. This study was funded by the Korea Meteorological Administration, Korea Meteorological Institute (Grant No. KMI2018-04910).

References

1. Choi, S.G., et al.: The effect of weather conditions on transit ridership. J. Korean Soc. Civil Eng. **33**(6), 2447–2453 (2013)
2. Shin, K., Choe, G.-J.: Analyzing the relationship between precipitation and transit ridership through a seemingly unrelated regression model. J. Korean Soc. Transp. **32**(2), 83–92 (2014)
3. Stover, V.W., McCormack, E.D.: The impact of weather on bus ridership in Pierce County. Washington. J. Public Transp. **15**(1), 6 (2012)
4. Kashfi, S., et al.: Impact of rain on daily bus ridership: a Brisbane case study. In: Australasian Transport Research Forum 2013 Proceedings. Australian Transportation Research Forum (2013)
5. Arana, P., et al.: Influence of weather conditions on transit ridership: a statistical study using data from smartcards. Transp. Res. Part A Policy Pract. **59**, 1–12 (2014)

6. Zhou, M., et al.: Impacts of weather on public transport ridership: results from mining data from different sources. Transp. Res. Part C Emerg. Technol. **75**, 17–29 (2017)
7. Chung, W.-Y., et al.: A study on local area weather condition monitoring system in WSN and CDMA. J. Korea Inst. Marit. Inf. Commun. Sci. **13**(8), 1712–1720 (2009)
8. Aggarwal, C.C.: Outlier Analysis (2013)
9. Han, J., et al.: Data mining concepts and techniques third edition. In: The Morgan Kaufmann Series in Data Management Systems, pp. 83–124 (2011)
10. Breiman, L.: Classification and Regression Trees. Routledge (2017)
11. Sung, H., Kim, T.-H.: A study on categorizing subway station areas in Seoul by rail use pattern. J. Korean Soc. Transp. **49**, 455–464 (2005)
12. Kim, J.: An analysis of the changes in the cause-and-effect relationships between socio-economic indicators and the road network of Seoul using structural equation model. J. Korean Geogr. Soc. **44**(6), 797–812 (2009)
13. Won, M., et al.: Knowledge-based system for estimating incident clearance duration for Maryland I-95. Transp. Res. Rec. J. Transp. Res. Board **2672**(14), 61–72 (2018)
14. Seoul Data Center. http://data.seoul.go.kr
15. The Korea Transport Institute (2015). Unpublished report (in Korean)

The Role of Attentional Networks in Secondary Task Engagement in the Context of Partially Automated Driving

Rui Lin[1(✉)], Yuchen Xu[2], and Wei Zhang[1]

[1] Department of Industrial Engineering, Tsinghua University, Beijing 100084, China
ruilin336699@163.com

[2] Department of Bioengineering, Jacobs School of Engineering, UC San Diego, La Jolla, CA 92093, USA

Abstract. The usage of partially automated systems has transformed certain traditional driving into monitoring tasks. Simultaneously, drivers may voluntarily switch attention to engage in a secondary visual task. The main purpose of this study was to explore how individuals' attentional networks (i.e., alerting, orienting, and executive control) influence their visual attention allocation between monitoring and visual secondary tasks. In this study, we simulated specific dual-task condition by a well-controlled monitoring task combined with a visual secondary task. Participants' visual attention allocation behaviour and task performances were compared with their attentional functions measured by standard Attention Network Test. Two attentional components' significant effects were found: First, participants with more efficient alerting network spent less time on the monitoring task and exhibited more frequent switching behaviour. Second, participants with more efficient orienting networks showed longer reaction times in the secondary task and conducted less consecutive secondary task trials in once switching.

Keywords: Dual-task · Multitasking · Automation system · Alerting · Orienting · Executive control

1 Introduction

The usage of partially automated systems has transformed certain traditional driving into monitoring tasks. It was recently found that in automated driving, while engaging in the secondary tasks, drivers' percentage of time looking front road was less and glance duration to the secondary task was longer than the manual driving condition [1]. The classical distraction problem in manual driving has gradually changed to secondary task engagement [2]. This shift means that these drivers may not be merely the passive recipients of irrelevant stimuli. Instead, they may also perform monitoring tasks within a dual-task context and switch their visual attention between the road monitoring task and secondary tasks voluntarily but strategically.

In the driving area, improper visual attention allocation characterized by gaze behaviour could decline drivers' situation awareness and have the potential to cause

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 211–218, 2020.
https://doi.org/10.1007/978-3-030-50943-9_27

serious accidents [3–6]. However, from the cognition perspective, attention is a more general and primeval concept [7]. According to an attentional networks model, individuals' attentional functions can be realized through three relatively independent anatomical networks, i.e., the alerting, orienting, and executive control networks [8]. The three attentional networks are defined as follows:

Alerting is defined as maintaining a state of high sensitivity to incoming stimuli [9]. Alerting here indicates phasic alerting, which is separate from the definition of tonic alerting (an increased and general arousal) [9]. The alerting mentioned in this study refers specifically to phasic alerting. Orienting indicates the ability to select information from the candidate's sensory inputs [9]. Orienting can be divided into endogenous and exogenous orienting according to whether such selection is guided by an internal goal or an external stimulus [10]. Executive control is responsible for monitoring and resolving the conflict in error detection and novel or not well-learned responses [9].

Two versions of the Attention Network Test are commonly used to assess the performance of these three networks. One is the original version (referred to here as the ANT) [11], and the other is the modified version, called the Attention Network Test for Interactions (ANT-I) [12].

This study will investigate the influence of different attentional functions on individuals' visual attention allocation between a critical monitoring task and a separate secondary task. To achieve this objective, multiple measures of attentional function were obtained via the standard ANT. The dual-task condition was simplified using a well-controlled monitoring task, which was combined with a separate secondary task. Then, the switching-related parameters, monitoring and secondary task performances were compared with the measures of the ANT.

2 Method

2.1 ANT

This study used the standard ANT (see [11] for more details). The participants were asked to judge whether an arrow in the centre of a screen was pointing to the left or right using a two-button mouse. The cue type and flanker type were manipulated based on characteristics of the three attentional networks. There was a training session of 12 practice trials with accuracy feedback, followed by three blocks of 96 trials without feedback. The cue (no cue, centre cue, double cue, and spatial cue) and target (neutral, congruent, and incongruent) conditions acted as within factors in analysing participants' reaction time (RT) and accuracy.

2.2 Dual Task

In the dual task, the automation failure monitoring was simplified via a well-controlled monitoring task. At the beginning of each trial, participants were shown a green balloon on the screen, the colour of which would change to red after a random duration (green balloon duration, GBD). The red balloon represented an automation failure. Participants needed to respond to the red balloon with any key within a time limit (red balloon duration, RBD); otherwise, the balloon would explode. The secondary task was a one-digit odd-or-even judgement task. The trial's GBD was pseudo-randomly selected from a duration set (16 s, 32 s, 48 s, 64 s, 80 s), and the five values occupied the same number of trials. Thus, the signal rate was approximately 72 per hour. The dual tasks were separated into two conditions according to the RBD levels. The value of the RBD was either 2 s or 4 s (see [13] for more details). The dependent variables included the following three parts (Fig. 1): (1) The attention-switching parameters within the trials' GBD: the percentage of time spent on the monitoring task, the frequency of switching to the digit task and the number of consecutive digit task trials once participant switched to the digit task. (2) The digit task performance: digit task RTs in the repletion trials (all the trials after removing the switching trials). (3) The monitoring task performance: the red balloon detection rate. The correlation analysis was conducted between the three attentional networks' scores and these behavioural parameters in the dual tasks.

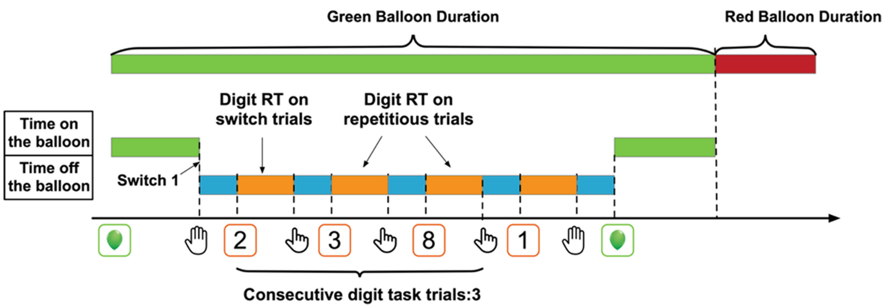


Fig. 1. Dual-task process and parameters

2.3 Apparatus and Participants

A 19-in. monitor with a resolution of 1280 × 1024 pixels displayed the stimulus. The stimulus was located 60 cm in front of the participants. Both tasks ran the E-Prime 2.0 software. Then, the 15-min session of the ANT proceeded, followed by a 5-min break, and finally, the 45-min session of the dual task. The experiment participants were 40 male graduates and undergraduates aged 20–31 (mean = 25.4, standard deviation = 3.3). Participants were recruited from a university population through an online bulletin board. All participants were right-handed and had normal or corrected-to-normal visual ability.

3 Results

3.1 ANT

For each participant, data was calculated after eliminating extreme values (less than 200 ms or greater than 1,700 ms). Data from one participant was removed from the final analysis due to a global accuracy of less than 90% on the ANT (72.9%). Table 1 summarized the correct RT and accuracy. Repeated-measures ANOVA showed that the main effect of cue conditions was significant on the correct RTs ($F(3,114) = 151.93$, $p < .001$, $\eta_p^2 = .138$), and the main effect of target conditions ($F(2,76) = 335.10$, $p < .001$, $\eta_p^2 = .250$). Longer RTs were found in no-cue conditions or incongruent target conditions. As Fig. 2a shows, there was a significant interaction between the cue and target conditions ($F(6,228) = 3.76$, $p = .001$, $\eta_p^2 = .005$). Repeated-measures ANOVA also showed that the main effect of cue conditions was significant for the accuracy ($F(3,114) = 3.02$, $p = .03$, $\eta_p^2 = .002$), and the main effect of target conditions was also significant ($F(3,114) = 20.71$, $p < .001$, $\eta_p^2 = .026$). The accuracy values were higher in spatial cue conditions than others and were lower in incongruent conditions than others. As Fig. 2b shows, there was a significant interaction between the cue and target conditions ($F(6,228) = 2.19$, $p = .045$, $\eta_p^2 = .002$). These results suggested some lack of independence among the three network scores.

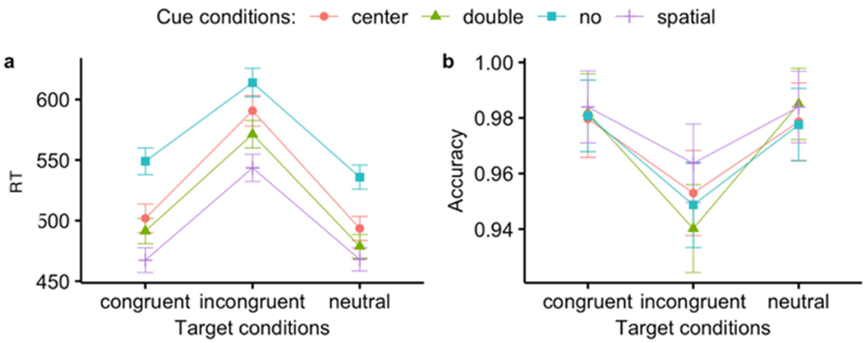


Fig. 2. The average in RT (a) and accuracy (b) under different cue and target conditions. Error bars represent the standard errors of the means

Table 1. Means and SDs in RT and accuracy under each condition

		Centre	Double	No cue	Spatial
RT	Congruent	501.90 (73.80)	491.35 (64.98)	548.99 (68.89)	467.33 (64.00)
	Incongruent	590.67 (78.83)	571.29 (70.55)	613.92 (74.15)	543.52 (70.35)
	Neutral	493.49 (62.19)	478.75 (61.20)	535.89 (62.17)	467.93 (59.64)
ACC	Congruent	0.98 (0.09)	0.98 (0.09)	0.98 (0.08)	0.98 (0.08)
	Incongruent	0.95 (0.10)	0.94 (0.10)	0.95 (0.10)	0.96 (0.09)
	Neutral	0.98 (0.09)	0.99 (0.08)	0.98 (0.08)	0.98 (0.08)

Table 1 showed that the means (SD) of the three attentional network scores were obtained individually based on the ANT: alerting was 52.47 ms (27.72), orienting was 35.76 ms (18.58), and executive control was 77.46 ms (26.57). The global average RT was 525.42 ms (82.02), whereas the global accuracy was 0.97 ms (0.09).

3.2 Dual-Task Performance

Table 2 summarized the chosen dependent variables under shorter RBD and longer RBD conditions. The results showed that the percentage of time spent on the monitoring task was significantly longer in the shorter RBD condition than the longer RBD condition ($p < .001$). Moreover, while engaging in the digit tasks, participants in the shorter RBD condition tended to conduct more consecutive digit trials ($p = .014$). While focusing attention on the monitoring task, participants were found to switch less frequently to the digit task than in the longer RBD condition ($p = .007$). No significant difference was found on the digit task performance characterized by the reaction times. It was also found that participants in the shorter RBD condition achieved less red balloon detection than in the longer RBD condition ($p = .026$).

Table 2. Means and analysis for dual-task performances under different RBD conditions

Dependent variables	Mean (SD)		t (38)	p	Cohen's d
	Shorter RBD	Longer RBD			
Percentage of time on monitoring task (%)	26.74 (13.81)	18.76 (14.32)	3.62	<.001	0.579
Frequency of switching to digit task (times/s)	3.43 (2.04)	4.46 (2.17)	−2.84	.007	0.454
Number of consecutive digit task trials	8.46 (4.03)	6.77 (3.67)	2.59	.014	0.414
Digit task RT (ms)	492.73 (61.67)	495.11 (65.18)	−0.31	.759	0.050
Red balloon detection rate (%)	61.03 (18.92)	68.21 (16.56)	−2.32	.026	0.372

3.3 Correlation Analyses

Table 3 showed that the alerting score negatively correlated with the percentage of time spent on the monitoring task in both conditions (shorter RBD: $p = .017$; longer RBD: $p = .003$): participants with greater alerting effects spent less time on the monitoring task in both conditions. The alerting score and the frequency of switching to the digit task were positively correlated in the longer RBD condition ($p = .046$), such that participants with greater alerting effects switched more frequently from the monitoring task to the digit task in the longer RBD condition. Additionally, the alerting score and the red balloon detection rate were negatively correlated ($p = .028$): participants with greater alerting effects detected more red balloons in the shorter RBD condition.

Table 3. Correlations between dual-task performance and attentional networks

Dependent variables	Alerting				Orienting				Executive control			
	Shorter RBD		Longer RBD		Shorter RBD		Longer RBD		Shorter RBD		Longer RBD	
	r	p	r	p	r	p	r	p	r	p	r	p
Percentage of time on monitoring task (%)	**−0.38**	**.017**	**−0.46**	**.003**	−0.17	.307	−0.24	.150	−0.09	.572	−0.02	.888
Frequency of switching to digit task (times/s)	0.27	.091	0.32	.046	0.23	.166	0.14	.411	0.02	.919	−0.15	.363
Number of consecutive digit task trials	−0.15	.361	−0.09	.57	**−0.32**	**.049**	−0.15	.374	−0.05	.773	0.06	.716
Digit task RT (ms)	0.07	.670	−0.02	.926	**0.35**	**.031**	**0.33**	**.038**	−0.08	.624	0.16	.327
Red balloon detection rate (%)	**−0.35**	**.028**	−0.27	.097	−0.22	.187	−0.24	.141	−0.06	.700	−0.20	.224

As shown in Table 3, the orienting score positively correlated with the digit task reaction times in the consecutive trials in both conditions (shorter RBD: p = .031; longer RBD: p = .038), such that participants with greater orienting effects responded more slowly to those consecutive trials' digit tasks. Moreover, the orienting score and the number of consecutive digit task trials were negatively correlated in the shorter RBD condition (p = .049), such that participants with greater orienting effects tended to achieve fewer digit task trials per switch from the monitoring task. No other significant correlations were found.

4 Discussion and Conclusion

This study first investigated the influence of different attentional functions underlying the attention-switching process in the context of performing a monitoring task in a dual-task condition. For this purpose, a well-controlled monitoring task combined with a secondary digit task was used to simulate the specific dual-task condition, and the behaviour parameters in the dual tasks were compared with three attentional functions scored by the standard ANT. The results indicated that the alerting and orienting networks played critical roles in influencing participants' visual attention allocation.

First, consistent with previous literature on attention lapse, the results in this study showed that participants with more efficient alerting networks might have difficulties in maintaining attention. A previous study has also shown that individuals with higher alerting efficiency tended to overestimate interval durations [14]. Therefore, after switching back to the monitoring task, participants with higher alerting efficiency might benefit more from the onset of the green balloon to increase and maintain a certain arousal level for the red balloon. Under such effects, these higher alerting-efficiency participants might also overestimate their time spent focused on the monitoring task and thus switch to the digit task more quickly.

Second, the participants with more effective orienting networks were found to exhibit longer RTs in the consecutive digit task trials and produced relatively fewer digit task trials once they switched from the monitoring task. In previous laboratory studies, such task interference (i.e., longer reaction time) was deemed as the cost for strategic monitoring for the prospective memory-related cue [15]. It should be highlighted that the orienting efficiency assessed by the ANT mainly focuses on endogenous orienting [12]. Thus, it is possible that participants with higher endogenous orienting efficiency relied more on strategic monitoring to remember the delayed intention (i.e., switching back to the monitoring task).

Generally, despite several limitations in this study, our results will be helpful in understanding individuals' attention regulation in a monitoring task while performing a secondary task. The results may potentially help locate high-risk drivers in the age of automated driving. Future experimental investigations are needed to further support these results when considering the influence of secondary task characteristics (e.g., being interrupted more often, more attractive distractions, or tasks consistent with real-life conditions), and monitoring conditions' ecological efficiency (e.g., using an automated driving or aviation system). The potential countermeasures against the shorter percentage of time looking front road and the longer glance duration to the secondary task were suggested, respectively.

Acknowledgments. This work was supported by the National Natural Science Foundation of China [grant numbers 71771132, 71801156 and 71901033].

References

1. Solís-Marcos, I., Ahlström, C., Kircher, K.: Performance of an additional task during level 2 automated driving: an on-road study comparing drivers with and without experience with partial automation. Hum. Factors **60**, 778–792 (2018)
2. Miller, D., Sun, A., Johns, M., Ive, H., Sirkin, D., Aich, S., Ju, W.: Distraction becomes engagement in automated driving. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. pp. 1676–1680 (2015)
3. Hartono, B., Gozali, M.: Secondary task and risky driving behavior: a test of the mediating effect of situation awareness. Int. J. Occup. Saf. Ergon. **21**, 399–404 (2015)
4. Fracker, M.L.: Attention Allocation in Situation Awareness. In: Proceedings of the Human Factors Society Annual Meeting. pp. 1396–1400, Sage, Los Angeles (1989)
5. Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J.: The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. National Highway Traffic Safety Administration, Washington, DC (2006)

6. Horrey, W.J., Wickens, C.D., Consalus, K.P.: Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. J. Exp. Psychol. Appl. **12**, 67–78 (2006)
7. Nobre, A.C.: Attention. In: Wixted, J.T. (ed.) Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience, pp. 1–75. Wiley, Hoboken (2018)
8. Posner, M.I., Petersen, S.E.: The attention system of the human brain. Annu. Rev. Neurosci. **13**, 25–42 (1990)
9. Petersen, S.E., Posner, M.I.: The attention system of the human brain: 20 years after. Annu. Rev. Neurosci. **35**, 73–89 (2012). https://doi.org/10.1146/annurev-neuro-062111-150525
10. Posner, M.I.: Orienting of attention. Q. J. Exp. Psychol. **32**, 3–25 (1980)
11. Fan, J., McCandliss, B.D., Sommer, T., Raz, A., Posner, M.I.: Testing the efficiency and independence of attentional networks. J. Cogn. Neurosci. **14**, 340–347 (2002)
12. Callejas, A., Lupianez, J., Funes, M.J., Tudela, P.: Modulations among the alerting, orienting and executive control networks. Exp. Brain Res. **167**, 27–37 (2005)
13. Lin, R., Liu, N., Ma, L., Zhang, T., Zhang, W.: Exploring the self-regulation of secondary task engagement in the context of partially automated driving: a pilot study. Transp. Res. Part F Traffic Psychol. Behav. **64**, 147–160 (2019)
14. Liu, P., Yang, W., Yuan, X., Bi, C., Chen, A., Huang, X.: Individual alerting efficiency modulates time perception. Front. Psychol. **6**, 386 (2015)
15. Cohen, A.-L., Hicks, J.L.: Selected topics in prospective memory. In: Cohen, A.-L., Hicks, J. L. (eds.) Prospective Memory: Remembering to Remember, Remembering to Forget, pp. 1–19. Springer International Publishing, Cham (2017)

Accidents

Analysis of Human Factors Failures in an Incident of Self-driving Car Accident

Ashraf Labib(✉), Yoskue Nagase, and Sara Hadleigh-Dunn

University of Portsmouth, Portsmouth, UK
ashraf.labib@port.ac.uk

Abstract. There are always submerged risks involved with advanced technology; therefore, it is necessary for policymakers, inventors and technology companies to scrutinise potential risks when they consider implementing new technology. This paper attempts to extract generic lessons from a failure relevant to autonomous transport systems. We use fault tree analysis (FTA), a reliability block diagram (RBD) approach and failure mode and effect analysis (FMEA), for analysing a fatal pedestrian accident caused by a level-3 self-driving car in 2018. The work highlights the importance of prematurity of test driving self-driving cars on public roads and the potential of an insightful analysis method that can capture human factors. In this work we theorise accident reporting systems, and provide a framework for triple loop learning.

Keywords: Learning from failures · Self-driving car · Accident analysis

1 Introduction: Learning from Failures and Advanced Technology

While it is hoped that technology makes our society more comfortable, more convenient and safer, risks exist in new technologies and tragic failures associated with technologies have occurred. However, in the world of engineering, learning from failures is indispensable, and it is known that various failures have helped societies evolve. Leoncini [1] argues that learning-by-failing is a necessity as the very nature of uncertainty that is associated with innovative activities, can lead to failures. Subsequently, learning and failure are closely connected due to the fact that trial-and-error procedures are among the main elements of discovery [2]. Also Leoncini [1] suggests that the learning process is more focused in the case of failure than it would be in the case of success. In this context the learning process is defined as a search process intended to rectify an organisational behaviour that led to a failure.

Furthermore, there is a shift from micro to macro level of emphasis that also includes a shift from technical to human to socio. This led to an interest among researchers in constructing hierarchy of causes of failures, as originally proposed in the Accimap model, proposed by Rasmussen [3], and other variations of this mental models of hierarchies, and networks, of causal factors such as Hierarchy of Failures [4], STAMP [5], FRAM [6]. Such approaches in framing causal factors can support policymakers, inventors or corporations when considering the implementation of a new

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 221–228, 2020.
https://doi.org/10.1007/978-3-030-50943-9_28

technology. Correspondingly, this paper attempts to examine policies and recommendations related to recent technological failures with respect to implementation of advanced technology in an evolving hybrid modeling approach.

Autonomous transport systems (ATSs), such as self-driving cars and autonomous surface ships, attract attention as a solution for achieving sustainable development goals [7]. Moreover, it is hoped that these technologies will contribute to the maintenance of national wealth in developed countries that will soon face an aging and diminishing population [8, 9]. In the US, some states have allowed social implementation tests of independent-type autonomous vehicles on public roads, and consequently, the number of accidents is increasing [10]. Accordingly, the current paper examines the failures in a fatal accident using fault tree analysis (FTA), a reliability block diagram (RBD) approach and failure mode effect analysis (FMEA).

The paper is organised as follows: Sect. 2 reviews the theoretical literature on the subject; Sect. 3 describes the first case study data on self-driving car accident and demonstrates the use of modelling techniques of FTA, RBD, and FMEA.

2 Literature Review and Theorization Related to Triple Loop Learning from Major Incidents

Barriers to Advanced Technology Adoptions. Barriers to advanced technology adoptions have been discussed in terms of its relation to research policy in different countries [11–14]. However none of them applied root cause analysis methods to assess casual factors and analyse vulnerabilities in the system as investigated in this paper.

Accident Analysis Methods. In terms of accident analysis methods and accident causation, there is a systematic review [15] that examined literature related to accident analysis methods related to sociotechnical systems contexts, mainly; AcciMap, the Human Factors Analysis and Classification System (HFACS), the Systems Theoretic Accident Model and Processes (STAMP) method, and the Functional Resonance Analysis Method (FRAM). The authors have identified the need to upgrade incident reporting systems, or recommendation reports, and have encouraged exploring further opportunities around the development of novel accident analysis approaches, which is also in line with the recommendations by Goode et al. [16].

Learning Loops. The grounding of safety and security in the learning loops is under researched in the literature. Hence we are in agreement with the literature review in the field of safety carried out by Drupsteen and Guldenmund [17] where they concluded that how learning occurs had been rarely studied, and suggested that safety research 'would benefit from input from organizational learning theories, such as [18] models of single and double-loop learning' [17, pp. 94). We extend their argument into the benefits of triple loop learning in safety and security arguing that safety and security are similar in their dealing with prevention and management of hazardous incidents or threats, where the main difference is the intent. Hence, we are also in agreement with the observation of Aven [19] related to the growing interest in applying risk analysis and risk management not only to safety but also to security problems.

In order to conceptualize triple loop learning, we need first to understand the literature related to double-loop learning. The concept of organisational learning can be described as a dichotomy [20]. Single loop learning occurs 'whenever an error is detected and corrected without questioning or altering the underlying values of the system', and double loop learning occurs 'when mismatches are corrected by first examining and altering the governing variables and then the actions', as defined by Argyris [21]. Hence, single loop concerns preserving and improving status quo, whereas, second loop learning implies changing the status quo itself (Labib, 2016). Accordingly, by extending this logic, triple loop learning can then be described as a 'deeper', or 'higher', level than, primary and secondary forms of learning, which implicitly means that this level has greater impact. Yet, as noted by [22] in spite of its perceived importance, conceptualisations of this level of learning do not always make clear how it differs from, or relates to, primary or secondary forms.

One of the most comprehensive conceptualizations of the organizational triple loop learning can be found in [22] in their paper titled 'The origins and conceptualisations of 'triple-loop' learning: a critical review'. They distinguish between three conceptualizations of 'triple-loop learning'. They offer three conceptualisations of triple-loop learning, viz; a) a level superior to single and double-loop learning, a form of shift from operations to strategy; b) a level that involves reflexivity on learning how to learn about the previous two levels i.e. learning about the process of learning; c) a level that involves a change of epistemology; a change in the wisdom in the form of knowing and learning. This third conceptualisation is about a complete, or fundamental, change of belief and opinion. The triple loop learning is about a shift towards 'richness'; as Weick puts it succinctly: 'it takes richness to grasp richness' [23].

In this paper we provide tools and case studies as enablers for realizing second and third loops of learning from failures. In doing so, it is hoped to extend such triple loop conceptualization to both the safety and security fields.

3 Case Study: A Self-driving-Car Accident

On 18 March 2018 in Arizona, for the first time, a self-driving car killed a pedestrian [24]. Recently, some authorities in the US have allowed self-driving tests on public roads, and accidents caused by these vehicles have increased, however, the fatalities in these accidents had been restricted to the drivers. Although there was only one fatality, it is possible that fatal accidents during tests will discourage people and policy makers from accepting this technology. Hence, learning from an accident is a meaningful process for progressing science and technology policy.

3.1 Overview of the Accident

The fatal pedestrian accident happened at around 10 p.m. on a road in Tempe, Arizona [25]. The female victim was walking her bicycle across the road when the vehicle in autonomous mode struck her at about 45 miles/hr. The vehicle was operated by Uber

Technology Inc. and equipped with their original light detection and ranging (or LIDAR) system that enabled the vehicle to drive at night. However, according to the released dashboard-camera recording, the system did not work very well and the driver inside the vehicle looked away without taking evasive action. Moreover, the US National Transport Safety Board (NTSB, 2018) reported that although the system detected an object six seconds before the accident, the system needed emergency braking manoeuvre 1.3 s before the accident when the system determined it was a bicycle, and the car was not designed to reduce speed and alert the driver while under computer control.

3.2 The Accident Factors and an Analysis

First, to confirm the accident factors, the current paper used FTA and an RBD approach (Figs. 4 and 5). Using an overview of the accident, the causes were divided into two types: the self-driving car's faults and the pedestrian's faults. The self-driving car's faults were further divided into machine errors and human errors. Finally, the machine errors were classified as being derived from inherent defects (e.g. computer programme, code or mechanical structure) or acquired defects (breakdown), and the human errors were classified as being derived from external factors (invisibility because of light or weather conditions) and internal factors (carelessness or intention). However, a 'basic event 5' in Fig. 4 could be removed as it is difficult to determine the next action in a sudden event, as in this case, and it is also unreasonable in this situation that a sensing system could be expected to work appropriately. Moreover, the pedestrian's faults could include intentional and non-intentional behaviours (Fig. 1). Correspondingly, the RBD, based on the FTA, illustrates a 3-line parallel structure (Fig. 2).

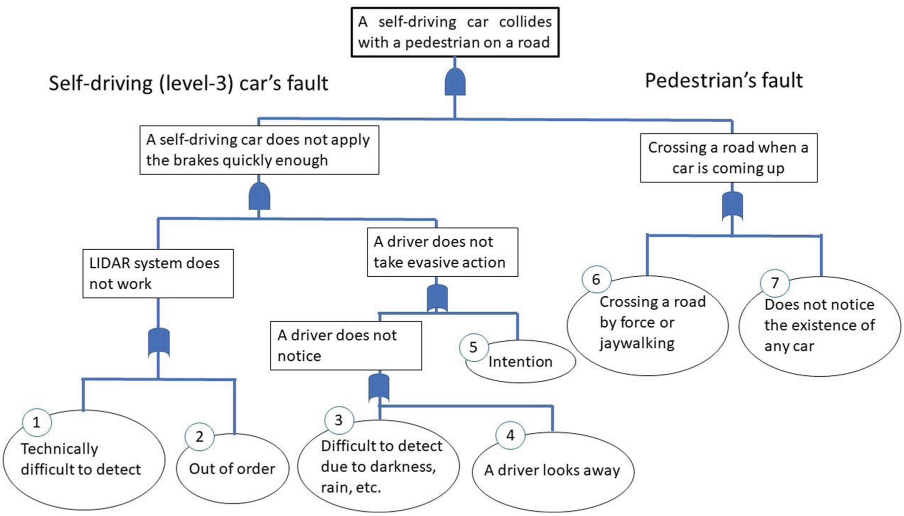


Fig. 1. Fault tree analysis of Case 1.

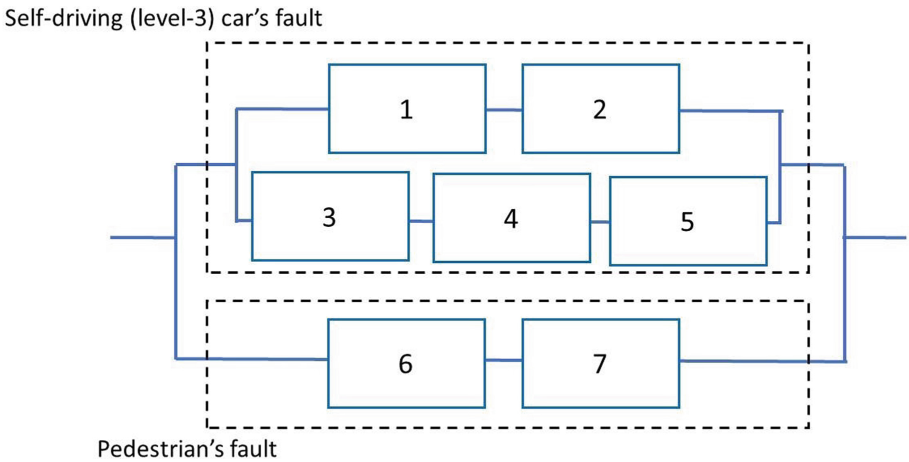


Fig. 2. Reliability block diagram for Case 1.

Second, the current paper conducted FMEA to consider each risk-priority number (RPN) from the results of the FTA and RBD (Table 1). Initially, for occurrence, the likelihood of bad conditions for driving and pedestrian road crossing are highly correlated. Next, for severity, the degrees for invisibility and intentional crossing were lowered because people can take evasive actions. Lastly, for detection, almost all the internal factors, except for carelessness, were relatively undetectable. Consequently, Table 1 highlights the criticality of the machine errors.

3.3 Discussion: Lessons from the Accident and the Limitations of Failures and Advanced Technology

First, a noteworthy point when analysing factors is to take into account the current autonomous level. Because a driver is necessary to take evasive action in the case of an emergency, the current technology level of self-driving cars, including the one which caused the accident, is level-3 (Fig. 3). This directly means that an AND-gate should be added as a fail-safe to the side of the self-driving car's fault compared with a level 4 or higher autonomous level (basic events 3 and 4, shown in Fig. 1) and, accordingly, reliability blocks were added (Fig. 2). As a result, the RBD forms a simple 3line parallel structure; considering each RPN, however, technological difficulties become remarkable (see Table 1). To summarise, currently the status is that a driver of a level-3 self-driving car assumes an important role in avoiding critical dangers.

Table 1. Failure mode effect analysis (existing condition)

Potential failure mode	Potential effects of failure	Potential failure causes	Current controls	O	S	D	RPN
A selfdriving car does not apply the brakes quickly enough (selfdriving car's fault)	LIDAR system does not work	Technically difficult to detect a pedestrian	Debugging	3	5	5	75
		Out of order	Regular maintenance	3	5	4	60
	A driver does not take evasive action because the driver does not notice the existence of a pedestrian or does not take intentionally	Invisibility (darkness, rain, fog, etc.)	Reducing speed	5	3	1	15
		A driver looks away	Guideline, regulations	4	5	1	20
		Intention	Guideline, regulations	1	5	5	25
(Pedestrian's fault)	A pedestrian crosses a road when a car is coming	A pedestrian crosses a road by force or jaywalks	Education	5	4	1	20
		A pedestrian does not notice the existence of a car coming	Education	5	5	1	25

Note: O = Occurrence, S = Severity, D = Detection and RPN = risk-priority number (all shown as a 5-point scale)

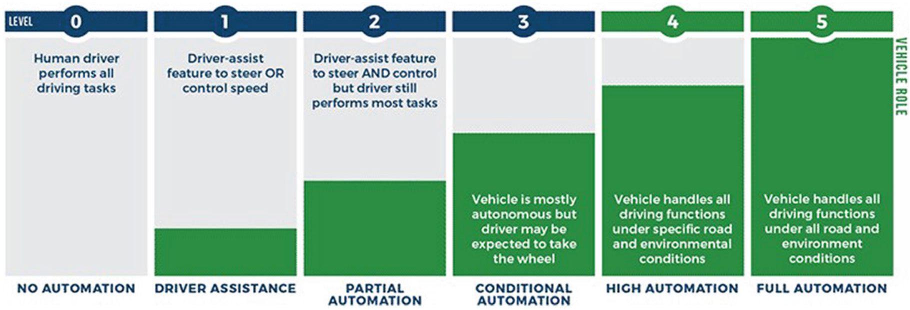


Fig. 3. Self-driving car's autonomous levels (Guerra, 2017).

Second, it is obvious that the technical difficulty of detecting people is the greatest problem for the safeguarding of pedestrians (Table 1). To achieve further high autonomous levels, the development of highly robust systems is indispensable because if the autonomous level reaches level-4 or higher, there is no driver command. Even recently, what some US states permitted is to allow technology companies to test selfdriving cars on roads with a driver onboard [24] (Nakata, 2018); the safety benchmark have not been clarified (Guerra, 2017) [24]. While the NTSB (2018) focused on how the sensing systems worked and indicated the necessity for at least compulsory braking and alert systems, no probable cause for the accident has not been specified. The situation surrounding the detection capability of selfdriving cars illuminates the prematurity of the implementation of level-3 or higher autonomous vehicles.

Third, considering this prematurity, it is possible that assumed variables on public roads are too difficult to deal with. This means that the hybrid environment, where people and cars simultaneously exist, makes the challenge too complicated for the state-of-the-art technology. In addition, there is a wide range of variables, including the shapes of roads, road signs, the weather, brightness and other objects. Therefore, it appears necessary for policy makers to reduce such variables as much as possible, not to suddenly allow tests on public roads. However, reducing failures by changing the prerequisites seems beyond the analysis of FTA, which seems a useful analysis method for analysing risk factors and considering palliative countermeasures.

Finally, in this paper recommendations were derived from a hybrid approach in accident causation analysis. In this paper we used FTA an analysis to study causal factors. The FTA is a logical ‘language’ that consists of two main logical symbols (a language of just two letters). Such constraint can act as both a point pf weakness and a blessing; the former is due to its simplicity, whereas the latter is due to it forcing one to think logically, and this acts as a 1st loop learning. The higher levels, or richness, in learning, occurs when this feeds into RBD (for venerability analysis) FMEA (for occurrence, severity and detection analysis). This subsequently leads to recommendations related to prevention, mitigation and enacting new early warning measures. In order to fully satisfy triple loop learning, we propose future work on benchmarking analysis that is coupled with high reliability organisations (HRO) principles in order to provide a framework that fully satisfies all three loops of learning.

References

1. Leoncini, R.: Learning-by-failing An empirical exercise on CIS data. Res. Policy **45**(2), 376–386 (2016)
2. Chesbrough, H.: Business model innovation: opportunities and barriers. Long Range Plan. **43**, 354–363 (2010)
3. Rasmussen, J.: Risk management in a dynamic society: a modelling problem. Saf. Sci. **27**(2–3), 183–213 (1997)
4. Hatamura Y.: Evacuation and decontamination in response to the Fukushima nuclear power plant accident. In: The 2011 Fukushima Nuclear Power Plant Accident. Woodhead Publishing, Boston (2015)

5. Leveson, N.: A new accident model for engineering safer systems. Saf. Sci. **42**(4), 237–270 (2004)
6. Hollnagel, E.: Barriers and accident prevention. Routledge. (2016)
7. Lim, H.S.M., Taeihagh, A.: Autonomous vehicles for smart and sustainable cities: an in-depth exploration of privacy and cybersecurity implications. Energies **11**(5), 1062 (2018)
8. Department for Transport: The Pathway to Driverless Cars: Summary report and action plan (2015). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/401562/pathway-driverless-cars-ummary.pdf. Accessed Feb 2015
9. Cabinet Secretariat: 日本再生戦略 [Japan revitalization strategy] (2012). https://www.cas.go.jp/jp/seisaku/npu/pdf/20120731/20120731.pdf. Accessed 31 Jul 2012
10. Nakata, A.: 年に急増した自動運転車の事故一覧、米カリフォルニア州で50件超. NIKKEI x TECH (2018). https://tech.nikkeibp.co.jp/atcl/nxt/column/18/00155/110500034. Accessed 7 Nov 2018
11. Baldwin, J., Lin, Z.: Impediments to advanced technology adoption for Canadian manufacturers. Res. Policy **31**(1), 1–18 (2002)
12. D'Este, P., Iammarino, S., Savona, M., von Tunzelmann, N.: What hampers innovation? Revealed barriers versus deterring barriers. Res. Policy **41**(2), 482–488 (2012)
13. Galia, F., Legros, D.: Complementarities between obstacles to innovation: evidence from France. Res. Policy **33**(8), 1185–1199 (2004)
14. Lhuillery, S., Pfister, E.: R&D cooperation and failures in innovation projects: Empirical evidence from French CIS data. Res. Policy **38**(1), 45–57 (2009)
15. Hulme, A., Stanton, N.A., Walker, G.H., Waterson, P., Salmon, P.M.: What do applications of systems thinking accident analysis methods tell us about accident causation? A systematic review of applications between 1990 and 2018. Saf. Sci. **117**, 164–183 (2019)
16. Goode, N., Salmon, P.M., Lenne, M., Finch, C.: Translating Systems Thinking into Practice: A Guide to Developing Incident Reporting Systems. CRC Press, Boca Raton (2018)
17. Drupsteen, L., Guldenmund, F.W.: What is learning? A review of the safety literature to define learning from incidents, accidents and disasters. J. Contingencies Crisis Manage. **22**(2), 81–96 (2014)
18. Argyris, C., Schön, D.: What is an organization that it may learn. In: Argyris, C, Schoen, D. A. (eds.) Organizational Learning: A Theory Perspective, pp. 8–29. Addison-Wesley, Reading (1978)
19. Aven, T.: A unified framework for risk and vulnerability analysis covering both safety and security. Reliab. Eng. Syst. Saf. **92**(6), 745–754 (2007)
20. Cope, J.: Entrepreneurial learning and critical reflection: discontinuous events as triggers for 'higher-level'learning. Manage. Learn. **34**(4), 429–450 (2003)
21. Argyris, C.: On Organizational Learning, 2nd edn. Blackwell, Malden, MA (1999)
22. Tosey, P., Visser, M., Saunders, M.N.: The origins and conceptualizations of 'triple-loop' learning: a critical review. Manage. Learn. **43**(3), 291–307 (2012)
23. Weick, K.E.: The generative properties of richness. Acad. Manage. J. **50**(1), 14–19 (2007)
24. Wakabayashi, D.: Self-driving Uber car kills pedestrian in Arizona, Where Robots Roam. The New York Times. Accessed 19 Mar 2018
25. Griggs, T., Wakabayashi, D.: How a self-driving Uber killed a pedestrian in Arizona. The New York Times. Accessed 21 Mar 2018

Assessing the Effectiveness of Augmented Reality Cues in Preventing Rear-End Collisions: A Driving Simulator Study

Alessandro Calvi(✉), Fabrizio D'Amico, Chiara Ferrante, and Luca Bianchini Ciampoli

Department of Engineering, Roma Tre University, Via Vito Volterra 62, 00146 Rome, Italy
{alessandro.calvi, fabrizio.damico, chiara.ferrante, luca.bianchiniciampoli}@uniroma3.it

Abstract. Rear-end collisions are one of the most common types of road accidents, often caused by the sudden deceleration of the leading vehicle during a car-following situation. In-vehicle applications based on Augmented Reality (AR) technology could optimize the driver's visual attention, providing new and additional visual feedback to improve the driving experience. This study has focused on a new AR application aimed at improving the safety of rear-end driving conditions, by means of different AR video and audio warnings. A driving simulator study was carried out to test the effectiveness of the proposed AR system and assess the ability of drivers to avoid rear-end collisions, with and without the AR warnings, in an urban road scenario. Significant positive effects of the AR warnings on driving speeds and road safety were observed, mainly consisting in lower speeds, decelerations and reaction times and improved surrogate measures of safety.

Keywords: Augmented reality · Driving simulation · Road safety · Driving performance · Rear-end collision

1 Introduction

Improving road safety is certainly one of the most significant challenges worldwide. The number of road accidents, injuries and fatalities is dramatic everywhere [1], especially when considering that several effective safety countermeasures already exist. Drastic action is needed to put these measures in place in order to meet any future global targets that might be set, and consequently save lives.

According to the National Highway Traffic Safety Administration [2], rear-end collisions are the most frequent type of crashes (29% of all vehicle accidents). The majority of rear-end collisions occur when the leading vehicle is stopped or moving at a very slow speed. In most collisions, the driver was following too closely to the vehicle in front of it. In nearly half of rear-end collisions, the driver following the vehicle failed to react to the stopped/slowed vehicle due to being distracted or not paying attention.

New collision avoidance technologies, such as forward-collision warning systems and autonomous braking systems, are becoming more widespread and have significant

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 229–236, 2020.
https://doi.org/10.1007/978-3-030-50943-9_29

potential for reducing the number of accidents. These systems, however, do not replace attentive and alert driving. Conversely, in-vehicle applications based on Augmented Reality (AR) technology could optimize the driver's visual attention providing new and additional visual feedback of other automated functions to improve the driving experience. For instance, preventive knowledge of potentially dangerous situations and conflicts through virtual warning messages that provide an augmented version of the roadway ahead, or of moving vehicles and other road users, could increase the driver's awareness and reduce the risk of accidents. It is reasonable to expect that AR technology, integrated with connected vehicles technology, could enable future systems to detect road hazards earlier and provide the driver with additional information with significant improvement in terms of road safety and operations [3]. Specific warnings about these hazards could improve the driver's reactions by directing the driver's attention to critical events or supporting them during difficult and critical maneuvers [4, 5]. Few studies in literature have evaluated AR in-vehicle applications for preventing rear-end collisions. As an example, Park et al. [6] investigated the effectiveness of an AR in-vehicle system aimed at providing the drivers with collision warning and path guides. The proposed system established the information to be given through a forward situation awareness service and lane-level navigator service and provided the driver with AR information on the windshield using a Head-Up Display (HUD) in the vehicle. The authors confirmed that the system could improve the intuitive cognition of a driver and reduce the driver's distraction.

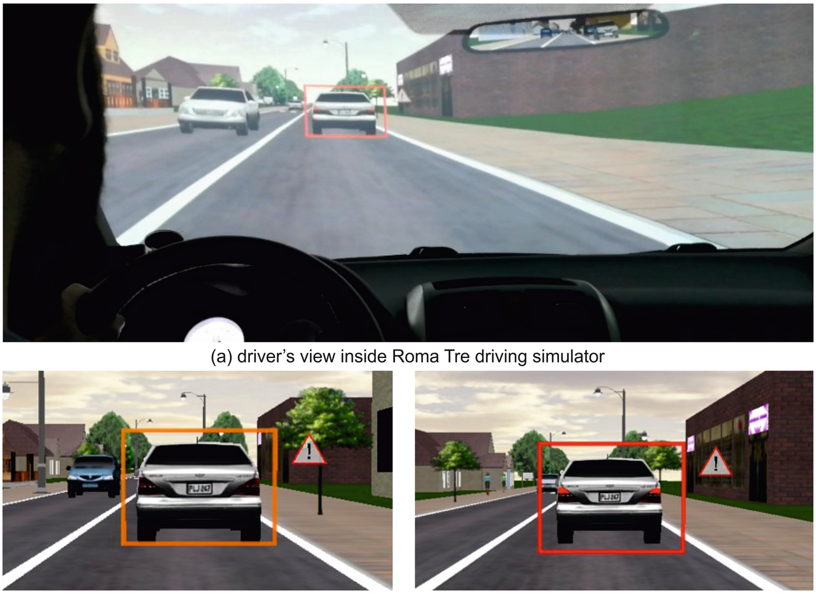
2 Objective

The study presented in this paper is about a new AR application for connected vehicles aimed at improving the safety of driving conditions, by means of different AR video and audio warnings. Specifically, the ability of AR technology to improve the safety of car-following driving condition is tested by providing the driver with visual virtual information and improving driver's risk perception and reaction in the case of a sudden braking maneuver of the leading vehicle. To this aim, a driving simulator study was carried out to test the effectiveness of the proposed AR system and assess the ability of drivers to avoid rear-end collisions, with and without AR warnings, in an urban road scenario.

3 Method

3.1 Driving Simulator

The driving tests were performed using the STI fixed-based driving simulator of the Road Safety Laboratory at the Department of Engineering of Roma Tre University (Fig. 1a). The driving simulator consists in a full cab Toyota Auris with the driver feedbacks controlled by a force-feedback steering wheel, brake and accelerator pedals. Three projectors assure a wide image of 180° field of view on a curved screen located in front of the vehicle. The resolution of the visual scene is 1920 × 1200 pixels with a refresh rate of up to 60 Hz. The reliability of the tool has been fully validated by previous studies [7, 8] that allowed using it for the evaluation of the driving performance in terms of speed, acceleration and trajectory under different driving conditions and road environments [9–12].



(a) driver's view inside Roma Tre driving simulator

(b) AR pre-warning
before the leading vehicle's braking

(c) AR warning
when the leading vehicle starts braking

Fig. 1. (a) Roma Tre driving simulator and AR warnings ((b) and (c)).

3.2 Scenario and Braking Event

A scenario consisting of a two-lane urban road was implemented in the driving simulator and all the braking events of the leading vehicles were randomly simulated in four drives (tests) of the same scenario to investigate the effects of the AR warnings on driving performance and safety, by comparing the results with those obtained in a baseline configuration (namely BC, with no AR warning activated). The urban road (50 km/h speed limit) was characterized by 3.50 m lanes with 4.00 m sidewalks on both sides. Several car-following conditions were simulated with the leading vehicles travelling at a constant speed of 40 km/h. In some of the car-following conditions, after 200 m, the leading vehicle suddenly braked with a deceleration rate of 5.5 m/s^2 and stopped in two seconds. Some vehicles were added randomly in the opposite direction in order to induce the driver to avoid overtaking the leading vehicle.

As illustrated in Fig. 1, three different virtual warning configurations, aimed at improving the drivers awareness of a dangerous situation and reducing the drivers' reaction times by adding an additional information for alerting them of an imminent braking of the leading vehicle, were tested: an AR warning, consisting in a flashing red rectangle placed in the rear of the leading vehicle that appeared when it started braking, namely AR_1 (Fig. 1c); an additional orange rectangle that appeared few seconds before the start of the leading vehicle's braking maneuver (and therefore, before the red rectangle), providing a pre-warning of the sudden braking, namely AR_2 (Fig. 1b); an additional audible warning, namely AR_3, that was activated simultaneously with the virtual coloured warnings: the sound, a "beep", was slow during the pre-warning and became quicker during the braking maneuver. Both the rectangles appeared along with a general warning sign comprising an exclamation mark on a standard triangular sign.

3.3 Sample of Drivers

Forty-six participants (13 women and 33 men, age range between 20 and 35 years) took part in the experiments. The sample, deliberately selected, consisted mostly of young drivers as numerous studies have shown that risk-taking and aggressive behavior are generally associated with males, in particular young males [1]. Three participants were excluded from the analysis as they experienced symptoms of simulation sickness and were unable to complete the driving tests. Moreover, based on a statistical analysis on the speed values recorded 200 m before the beginning of the braking maneuver of the leading vehicle, two drivers were considered outliers and were excluded from the analysis since they adopted a mean speed that was three standard deviations higher than the average speed of the sample of drivers. Therefore, the final sample consisted of forty-one drivers (11 women and 30 men) with a mean age of 26.6 years (SD = 3.0 years), age range of between 21 and 35 years.

3.4 Procedure

Each participant drove the scenario four times on two different days. After a preliminary drive of a training scenario to help the participant become familiarized with the tool, and the filling of a questionnaire with general driver information, the participant has been made aware of the implemented visual and audio AR systems to recognize the warnings in the simulated scenario. Following each driving test, an after-driving questionnaire was filled out by the participant about the type and characteristics of any possible discomforts perceived during driving (nausea, giddiness, fatigue, etc.). The sequence of routes was randomly selected in order to avoid any repetitive influences.

3.5 Data Collection

Several driving performances and surrogate measures of safety were collected for all the drives and the three AR warning configurations were then compared to the baseline configuration in order to determine whether differences in driving performance exist among the different AR warnings and evaluate the effectiveness of the AR systems in avoiding rear-end collisions due to a sudden braking of the leading vehicle during car-following condition. Specifically, the driver's speed where the leading vehicle started braking (S_1) and stopped (S_3) and where the driver started braking (S_2) were collected, as well as the driver's reaction time (RT) and the driver's average deceleration (d_{av}) during the braking maneuver. Moreover, the Time-to-Collision recorded when the driver started braking (TTC_1), when the leading vehicle stopped braking (TTC_2) and the minimum Time-to-Collision (TTC_{min}) during the braking maneuver were collected. TTC, internationally acknowledged as a surrogate safety measure, is defined as the time that remains until a collision between two vehicles would have occurred if they continue along their predicted path with the same speeds and trajectories [13]. RT is computed as the time interval between the first braking of the leading vehicle's and the time when the driver starts braking.

4 Results and Discussion

The Analysis of Variance (one-way ANOVA 4 × 1) was then carried out: the different conditions of the AR warnings (BC, AR_1, AR_2 and AR_3) were considered as independent variables and their effects on each dependent variable described above were evaluated. Before performing ANOVA, all the data collected were subjected to the Kolmogorov-Smirnov test to verify the normality assumption required by ANOVA. Moreover, a sphericity test was carried out in order to test whether sphericity has been violated. In fact, when the sphericity assumption was violated (in this study, when the Mauchly test was found to be significant for a given independent variable) adjustments were then made using the Geisser-Greenhouse epsilon that provides a much stricter F test. For multiple comparisons, the Bonferroni correction was used. Finally, a significance level of 0.05 was adopted for the significance test.

Table 1 summarizes the results of the analysis performed on the variables selected for evaluating the effectiveness of the AR warnings on driving performances and safety. The average values of each variable, along with the standard deviations (in parenthesis), are given in the table, as well as the results of the ANOVA tests.

Table 1. Investigated variables (average and standard deviation) and ANOVA results.

Variable	BC	AR_1	AR_2	AR_3	ANOVA results	
					F	p
S_1 [km/h]	43.36 (4.15)	45.30 (3.89)	38.72 (3.26)	38.33 (4.34)	31.518	<0.001
S_2 [km/h]	41.38 (4.28)	43.14 (4.03)	37.52 (4.21)	37.10 (4.02)	20.825	<0.001
S_3 [km/h]	36.52 (9.37)	28.45 (9.82)	26.09 (9.42)	25.91 (8.80)	10.805	<0.001
RT [s]	3.04 (1.69)	1.21 (0.45)	1.21 (0.85)	1.11 (0.81)	28.013	<0.001
TTC_1 [s]	3.60 (1.55)	4.30 (1.93)	4.67 (1.28)	5.24 (2.17)	5.874	0.001
TTC_2 [s]	3.59 (2.07)	3.63 (1.86)	4.38 (1.71)	4.27 (2.25)	2.746	0.045
TTC_{min} [s]	3.04 (1.69)	3.27 (1.59)	4.07 (1.61)	3.95 (1.78)	3.713	0.013
d_{av} [m/s^2]	3.98 (1.79)	3.91 (1.50)	3.21 (1.42)	3.19 (1.43)	2.974	0.034

ANOVA shows a statistically significant main effect of the AR warnings on all the investigated variables. Specifically, all the AR systems have been quite effective in informing the driver, well ahead of time, about the sudden braking of the leading vehicle, thus prompting the driver react earlier and safely to the braking maneuver.

In fact, the post-hoc comparisons performed on speed variables revealed significant drivers' speed reductions compared to BC configuration when the AR pre-warning was activated (AR_2 and AR_3), even at the beginning of the braking maneuver (S_1 and S_2). As expected, no significant differences in speeds were recorded between BC and AR_1 for S_1 and S_2. At the end of the braking of the leading vehicle (S_3), all the AR configurations revealed significant drivers' speed reductions, up to 30%. For S_3, post-hoc comparisons did not reveal any significant differences in S_3 among the AR configurations.

One of the main positive effects of the AR warnings has been found for the driver's reaction time. In fact, RT significantly decreased in all the AR configurations, up to 63% in AR_3, but almost the same reduction was recorded in the other AR configurations (60% in AR_1 and AR_2). It means that drivers reacted much earlier to the sudden braking of the leading vehicle with important improvements in terms of collision's risk reduction. Post-hoc comparisons did not reveal any significant differences in RT values between the AR configurations.

The same safety improvements have been revealed with Time-to-Collision analysis. In fact, TTC_1, TTC_2 and TTC_{min} significantly increased in the AR configurations, especially when pre-warning was activated (up to 45%). Post-hoc comparisons did not reveal any significant differences in TTC values between AR_2 and AR_3 that provided greater improvements than AR_1.

The driver's braking maneuver was also less aggressive with the AR warning activated and significant reductions in the average deceleration rate (d_{av}) were shown in AR_2 and AR_3, when the pre-warning was activated. Post-hoc comparisons did not reveal any significant differences in d_{av} values between AR_2 and AR_3 that, again, provided greater improvements than AR_1.

5 Conclusions

The great benefits that Augmented Reality cues and connected vehicle technologies could bring to the general safety conditions on the road network have been demonstrated in this study. Significant positive effects of the tested AR warnings on driving performance and road safety during car-following condition and sudden braking of the leading vehicle were observed: reduced drivers' speeds and deceleration rates during braking maneuvers (−30%), reduced drivers' reaction times (−60%), increased Time-to-Collision (+40%). Such differences were statistically confirmed between the tested AR warnings, demonstrating the highest effectiveness of the AR systems with pre-warning activated; no significant additional effects of the audible warning were noted.

Moreover, this study demonstrated the high potentialities and effectiveness of driving simulation in the analysis and evaluation of the effects of the AR technologies on driving performance and road safety. The results should be considered quite important when developing driver warning systems; they could provide the automotive industry with important information on the design of more efficient in-vehicle accident prevention systems.

The results are surely of great interest and quite promising and should be enlarged considering a wider sample of drivers. In fact, the effectiveness of the AR warnings also needs to be examined using different samples of drivers, with particular attention to age, gender, and AR experience. To this end, a study is being conducted to test two different samples of drivers: the first, composed of drivers who are quite familiar with the AR warnings (the same sample used in this study) with the overall aim of testing whether the positive results presented in this paper may persist over time; the second, composed of older drivers, was chosen to determine whether the effectiveness of the AR warnings are comparable between old and young drivers.

Finally, in order to investigate possible side effects of the warnings, such as those related to distraction, a driving simulator study is ongoing to monitor the drivers' eye movements using an eye tracking system in order to examine additional variables to improve the design of the AR warnings.

References

1. World Health Organization: Global status report on road safety 2018 (2018). www.who.int/violence_injury_prevention/road_safety_status/2018/en. Accessed 7 Feb 2020
2. National Highway Traffic Safety Administration (NHTSA). Traffic Safety Facts 2015: 2015 Motor Vehicle Crashes: Overview. Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration (2016). https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812318. Accessed 3 Feb 2020
3. Naujoks, F., Grattenthaler, H., Neukum, A., Weidl, G., Petrich, D.: Effectiveness of advisory warnings based on cooperative perception. IET Intell. Transp. Syst. **9**(6), 606–617 (2015)
4. Moussa, G., Radwan, E., Hussain, K.: Augmented reality vehicle system: left-turn maneuver study. Transp. Res. C-Emerg. **21**, 1–16 (2012)
5. Rusch, M.L., Schall Jr., M.C., Lee, J.D., Dawson, J.D., Rizzo, M.: Augmented reality cues to assist older drivers with gap estimation for left-turns. Accident Anal. Prev. **71**, 210–221 (2014)
6. Park, B.J., Lee, J.W., Yoon, C., Kim, K.O.: A study on augmented reality of a vehicle information using head-up display. In: 6th International Conference on IT Convergence and Security (ICITCS), Prague, Czech Republic, 26–29 September 2016
7. Bella, F.: Driving simulator for speed research on two-lane rural roads. Accident Anal. Prev. **40**(3), 1078–1087 (2008)
8. Calvi, A.: Investigating the effectiveness of perceptual treatments on a crest vertical curve: a driving simulator study. Transp. Res. F-Traffic **58**, 1074–1086 (2018)
9. Calvi, A., D'Amico, F.: Quality control of road project: identification and validation of a safety indicator. Adv. Transp. Stud. **9**, 47–66 (2006)

10. Calvi, A., D'Amico, F., Ferrante, C., Bianchini Ciampoli, L.: Applying perceptual treatments for reducing operating speeds on curves: a driving simulator study for investigating driver's speed behavior. In: Advances in Intelligent Systems and Computing, vol. 964, pp. 330–340 (2020)
11. Calvi, A.: Does roadside vegetation affect driving performance? Driving simulator study on the effects of trees on drivers' speed and lateral position. Transp. Res. Rec. **2518**, 1–8 (2015)
12. Calvi, A., Bella, F., D'Amico, F.: Diverging driver performance along deceleration lanes: driving simulator study. Transp. Res. Rec. **2518**, 95–103 (2015)
13. Hayward, J.: Near misses as a measure of safety at urban intersections. PhD thesis, Department of Civil Engineering, Pennsylvania State University, University Park, PA, USA (1971)

Operational and Geometrical Conditions of Accident Occurrence and Severity at Signalized Intersections

Abdulla Alghafli[1] and Mohamed Shawky[2(✉)]

[1] Unversiti Teknikal, Durian Tunggal, Melaka, Malaysia
[2] Faculty of Engineering, Ain Shams University, Cairo, Egypt
m_shawky@eng.asu.edu.eg

Abstract. This research investigated the effect of road geometric features and operational conditions on the occurrence and severity of traffic accidents at signalized intersections in Abu Dhabi city, UAE. Speed, number of lanes, lane configuration, traffic signal sequence (lead/lag or split phasing), average hourly traffic volume per lane were used as independent variables. The accident occurrence was tested by using Poisson's regression modeling and the accident severity was examined by using multinomial logit modeling approaches. The Poisson model showed that at 4-leg intersections, one of the major causes of the accident, is passing of a street (either minor or major street) through the intersection. It was also found that at 3-leg intersection, the main cause of the accident is minor street passing through the intersection. The research also found that the higher the traffic volume the higher the chance of occurrence of traffic accidents. The multinomial logit model showed that five significant variables affect the severity of traffic accidents occurs at signalized intersections. the significant variables are the speed of the main road, traffic signal sequences, number of through lanes of minor road number of left lanes of main and minor roads.

Keywords: Accident occurrence at intersections · 3-leg intersection · 4-leg intersection · Traffic volume · Road geometric feature

1 Introduction

The main objective of traffic safety is to adopt programs, plans, traffic regulations as well as to implement measures to prevent or limit traffic accidents. Then the ultimate goal is to guarantee the safety of the people and property and to safeguard the nation's safety and its human and economic components. It is because of this that United Arab Emirates (UAE) Government puts lots of effort in to keep the roadway networks are safer and enhance the rank of UAE in the international ranking of the road safety indicators.

Therefore, factors affecting the occurrence of traffic accidents and its severity need to be more understood and more investigation To this end, a number of theoretical models have been developed to help explain safety issues in road systems and to help in the identification of safety performance indicators. These models have guided traffic

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 237–246, 2020.
https://doi.org/10.1007/978-3-030-50943-9_30

infrastructure planning in the reduction of accidents and incidents occurrence on the roads. Examples of such models are Multiple Linear Regression, Nested Logit Regression, Negative Binomial regression and Poisson's regression. However, the use of Multiple Linear Regression and Nested Logit Regression are discouraged for a number of reasons. This research applies Poisson's regression to study how the various geometric configurations of road intersections contribute to traffic accidents. The paper also investigates the contribution of traffic volume to the country's road accidents. For accident severity, the models that describe the categorical ordered parameters such as multinomial logistic regression model and ordered logit model is the common approach to analyze the different levels of accident severity.

2 Literature Review

2.1 Models of Evaluating Road Traffic Accident Occurrences

In order to study factors causing road accidents, researchers have applied a number of crash-based models. The idea behind these crash-based models is to determine the probability of occurrence traffic accidents in the area under study. These crush-based models according to Ouni [1] are: Multiple Linear Regression, Nested Logit Regression, Negative Binomial regression and Poisson's regression. Haleem [2] used nested logit to investigate relationship between geometric properties of road and traffic occurrence. The results indicated that indicated that several factors affect crash severity. Such factors include: the number of left and right lanes on the primary approach, right and left shoulder width, downstream and upstream distance to the nearest signalized intersection.

A number of researchers such as Aty [3] have extensively used Negative Binomial (NB) to establish the relationship between traffic crashes and relevant contributing factors. Poisson model has also been shown to help road safety specialist to predict road crashes and also determine the association between road crashes and the contributing factors [4]. For example, Cai [5] used Poisson lognormal model to conduct a comparative study and analysis of several geographic units for model analysis of crash-based issues to provide safety guidance in planning in Florida.

It is, however, important to note that multiple linear regression is not effective in modeling factors affecting traffic occurrences. This is because of the occurrence of non-normal error term distributions; the existence of relationships between the mean and variance of accident frequency; and the nonnegativity of the dependent variable. Because of this, Poisson's regression is a better approach to modeling such traffic accidents.

In order to demonstrate the accuracy associated with linear regression, [6] used both Poisson's regression and linear regression to study the relationship between accident occurrences and road geometric factors. This research found that linear regression does not accurately present the relationship between some dependent. Poisson's regression, on the other hand, produced relatively accurate results.

However, Negative Binomial regression is equally a good approach. In fact, according to Aty [3] Negative Binomial regression is even better than Poisson regression when modeling traffic accidents especially when factors affecting traffic accidents are dependent. This is because in such scenarios, dispersion is expected.

2.2 How Road Geometry and Traffic Volume Affect Road Traffic Accidents Occurrence

Research has shown that there exist a number of factors that contribute to affect traffic accidents. One such factor is the geometric design of roads and intersections and it has been studied by a number of researchers. Zewde [7] applied Negative Binomial regression to investigate how the geometric design of highways and traffic volume affects traffic accident occurrence. This research showed that traffic accidents tend to increase with road length. This study also found that the increase in the sharpness of the road curvature increases the occurrence of a traffic accident and vice-versa. Zewde [7] also found Average Annual Daily Traffic volume has a positive effect on the traffic occurrence. This means that the higher the traffic volume the higher the number of traffic accidents.

Agent [8] studied how road geometry (the number of lanes) affects traffic accident occurrence in Kentucky, United States of America. They found that the four-lane undivided highway is associated with the highest number of injuries and fatal accidents. They also found that two-lane highways are associated with high traffic accident occurrences especially at sharp road curvatures.

Milton [9] studied how road geometry and traffic volume affect traffic occurrence in Washington State. This study showed that sharp horizontal curvature, high traffic volume, and narrow shoulder width tend to increase the frequency of occurrence of traffic accidents. The researchers also found that the more the number of lanes in a highway the higher the frequency of traffic accidents.

Regarding the accident severity modeling, several researches used multinomial logit models [10–13]. Other researchers used the ordered probit model to investigate the contributing factors affecting accident severity [14, 15]. These studies investigated several variables that affect the accident severity including road feature, at-fault drivers' characteristic weather and lighting factors and vehicle types with different results and explanations.

2.3 Accident Frequency and Severity Modeling

Poisson's Model. Poisson's regression models tend to assume that the dependent follow Poisson distribution. Another assumption is that observations are independent. It is defined as shown below.

$$\lambda_i = \beta x_i + \varepsilon \, [16] \tag{1}$$

Where: λi = frequency of occurrence of traffic accident at a given time interval; β = vector parameter; xi = independent variables; ε is the error term. The exponential of the error term (ε) has a gamma distribution having a variance of $\alpha 2$ and a mean of 1 [11].

It is also important to note that the probability distribution of the k number of accident occurrences at a given t time interval is represented by the equation:

$$P_i(k) = [(e^{(\lambda i)t}) \times (\lambda_i)^k] \div k! [17] \tag{2}$$

Where: e = exponential; k = the number of accident occurrence at t interval of time.

Testing for the Goodness Fit. To test for the goodness fit of the model, the deviance value 2 (LL(b) − LL(0)) was used. It (deviance value) follows X^2 distribution [18, 19]. The log – likelihood ratio $\rho^2 = 1 - LL(\beta)/LL(0)$ was also used to test the goodness fit of the model [20].

Multinomial Logit Model. Amultinomial logistic regression model or called Multinomial Logit Model (MNL) is used to express the accident severity which is considered as ordered dependent variable. In this model, the accident severity is categorized into the ordered level from the slight injury to fatal. Then the multinomial data approach is to nominate one category of the independent variable as a baseline or reference category and calculate log-odds for the other category levels relative to the reference. In the case of accident data in AD, the injury-severity is categorized into four categories: 1) slight injury, 2) medium injury, 3) severe injury and 4) fatality.

Hence, if the first category is the reference, then, for m = 2, …, N,

$$\ln\frac{P(Y_i = m)}{P(Y_i = 1)} = \alpha_m + \sum_{k=1}^{k} \beta_{mk} X_{ik} = Z_{mi} \tag{3}$$

For each case, there will be M−1 predicted log odds, one for each category relative to the reference category. (Note that when m = 1 you get ln (1) = 0 = Z_{11}, and exp (0) = 1).

In case of more than 2 groups, computing probabilities is a little more complicated than it was in logistic regression. For m = 2, …, N,

$$p(Y_i = m) = \frac{\exp(Z_{mi})}{1 + \sum_{k=2}^{N} \exp(Z_{hi})} \tag{4}$$

For each category, the N − 1 log odds are computed and exponentiate it,

For the reference category (category = 1),

$$p(Y_i = 1) = \frac{1}{1 + \sum_{h=2}^{N} \exp(Z_{hi})} \tag{5}$$

3 Data Description

3.1 Data Collection

The used data in this research were obtained from the accident database of Abu Dhabi (AD) Traffic Police. The data of the traffic volumes and control type were obtained from AD municipality database. Intersection geometry data were obtained from a field survey. The data were collected at 76 different signalized intersections in AD city.

The data covered all traffic accidents that were occurred at the signalized intersections for five years from 2013 to 2017. This data includes full information about each accident such as accident location, type, injury and fatality information, weather condition, etc. the intersections were categorized into two classes; 3 leg and 4-leg intersections. And the traffic signal control sequence was classified into two classes split phasing and lead/lag phasing systems.

3.2 Methodology and Data Analysis

STATA statistical software was used to perform the developed regression models in this research s. Two different types of models were used to investigate the impact of the road feature traffic control factors on the occurrence and severity of the accidents that were occurred at the 76 intersections.

About 593 severe accidents (i.e., any accident resulted in at least one injury or fatality) from the year 2013 to the year 2017 were analyzed. The number of accidents per intersection rages from 1 to 35 accidents. The hourly traffic volumes pass the intersections from all approaches ranges from 806 veh/h to 9,205 veh/hr.

The registered accidents resulted in about 1,246 casualties: 93% are minor injuries, 6% severe injuries and 1.0% fatalities. About 65% of the investigated accidents are caused due to red-light running and 64% are right-angle collusion.

4 Accident Occurrence Modelling Results and Discussion

4.1 Relationship Between Road Geometry and Traffic Occurrence

At 4-Leg Intersection

Table 1. Poisson regression result of the relationship between traffic accident occurrences and road geometry at 4-leg intersection.

	Direction 1		Direction 2	
Variables	Coefficient	p-value	Coefficient	p-value
Constant	−0.136	0.663	−0.081	0.803
Main through	0.366	0.000	0.358	0.000
Main left	0.031	0.702	0.012	0.904
Minor through	0.216	0.000	0.266	0.000
Minor Left	0.224	0.011	0.117	0.279
Deviance goodness fit [2(LL(b) − LL(0))]	227		224	
Poisson goodness fit [(r2 = 1 − LL(b)/LL(0)]	258		262	

Table 1 above shows the results of Poisson regression of the relationship between traffic accident occurrences and road geometry at 4-leg intersection. From the table, it is observed that the geometric configuration of the road has a great influence on the occurrence of traffic accidents. At direction 1 approaches, for instance, it is observed that there is a positive relationship between the number of traffic accident occurrences and when the main street passes through the intersection. A coefficient is 0.366 and p-value is 0.000. Also, the number of accidents tend to be high when minor street passes through a 4-leg intersection. A coefficient is 0.216 and p-value is 0.000. A similar result was observed in direction 2 approaches (see Table 1 for more details). These results indicate at 4-leg intersections, one of the major causes of the accident, is passing of a street (either minor or major street) through the intersection.

At 3-Leg Intersection

Table 2. Poisson regression result of the relationship between traffic accident occurrences and road geometry at 3-leg intersection.

	Direction 1		Direction 2	
Variables	Coefficient	p-value	Coefficient	p-value
Constant	−2.208	0.397	−0.954	0.485
Main through	1.011	0.063	0.571	0.206
Main left	0.560	0.372	0.000	0.000
Minor through	0.578	0.050	0.451	0.047
Minor left	−0.462	0.215	0.418	0.142
Deviance goodness fit [2(LL(b) − LL(0))]	0		2.66	
Poisson goodness fit [(r2 = 1 − LL(b)/LL(0)]	0		2.62	

Table 2 above is the results of Poisson regression analysis of the effect of road geometry on traffic accident occurrences at 3-leg intersection. From the table it is observed in direction 1 approaches, the main cause of the accident is minor street passing through the intersection. The coefficient is 0.578 and p-value is 0.050. A similar result is observed in direction 2 approaches. A coefficient is 0.451 and p-value is 0.047. This probably as a result of an increase in traffic volume in the outgoing minor street and the small number of lanes in this street. This result is similar to the results of Agent [8] who studied how road geometry (the number of lanes) affects traffic accident occurrence in Kentucky, United States of America. They found that two-lane highways are associated with high traffic accident occurrences. Other streets (main left, main through, and minor left) do not affect traffic occurrence.

4.2 Traffic Volume and Traffic Occurrence

4-Leg Vs 3-Leg Intersections. Tables 3 shows the results of Poisson regression of the relationship between traffic accident occurrences and traffic volume at 4-leg and 3-leg intersections respectively. It is observed that traffic volume at any given hour influence on the occurrence of a traffic accident. This is such that the higher the traffic volume the higher the chance of occurrence of traffic accidents. In both the intersections, the coefficient for this relationship is positive and statistically significant. Similar results were obtained by Zewde [7] who also found a relationship between traffic volume and the occurrence of accidents.

Table 3. Poisson regression result of the relationship between traffic accident occurrences and traffic volume at 3-leg and 4-leg intersection.

	3-leg intersections		4-leg intersections	
Variables	Coefficient	p-value	Coefficient	p-value
Constant	0.362	0.646	1.6409	0.000
Traffic volume	0.0046	0.034	.0000851	0.000
Deviance goodness fit [2(LL(b) − LL(0))]	9.06			274
Poisson goodness fit [(r2 = 1 − LL(b)/LL(0)]	8.3			330

5 Accident Severity Modelling Results and Discussion

5.1 Modeling Approach

The Multinominal Logit Model (MNL) is used to examine the contributing factor affecting the accident severity at the 76 signalized intersections. The MNL model is the preferable modeling approach for categorically ordered parameters such as accident severity levels. In the analysis, the severity is classified into four levels from category 1 that represent minor injury to level 4 that represent fatality, passing by level 2 (medium injury) and level 3 (severe injury). The reference level in our case was used level 1 (i.e., minor injury). In case of different severity levels for one accident, the accident was classified based on the worst level. Then, the independent variable is the accident severity and the dependent variables were the total number of lanes for all approaches, average hourly traffic volumes per lane, umber of lanes for though and left movements for main and minor road for both approaches, traffic signal phasing type (split/lead-lag), intersection type (3-leg/4-leg), and speed of the main road.

5.2 MNL Model Results and Discussion

The model goodness of fil of the developed model which shows the significant value is 0.029 that means the developed model is significant at level 95%. Table 4 shows the results of the estimated parameters of the model. the results show five significant variables affect the severity of the traffic accidents occurs at signalized intersections. The significant variables are the speed of the main road, traffic signal sequences, number of through lanes of minor road number of left lanes of main and minor roads. In addition, by using the minor injury level as the reference category and refer the three other accident-severity levels to the minor-injury level. The results show that using lead/lag signal phasing sequence increase the probability of severe injuries refer to a minor-injury. In addition, the geometric design of minor roads in terms of the number of through and left lane movements affect significantly the accident-severity levels.

Table 4. The MNL model parameters estimation results

Effect	Model fitting criteria	Likelihood ratio tests		
	−2 log likelihood of reduced model	Chi-square	df	Sig.
Intercept	142.495	6.246	3	.100
Main_Speed	145.148	8.900	3	.031*
Traffic volume/lane	137.455	1.207	3	.751
4-leg/3-leg	136.701	.453	3	.929
Lead-lag/split	148.421	12.173	3	.007*
Total_Thr.Main road	139.328	3.080	3	.379
Total_Thr.Minor road	143.791	7.542	3	.056**
Total_left.Main road	143.114	6.866	3	.076**
Total_left.Minor road	147.966	11.718	3	.008*

* significant at level 95%
** significant at level 90%

6 Conclusion

The research studied how road geometric and traffic volume configuration may affect the occurrence of traffic accidents in UAE.

Regarding the accident occurrence, it was found that:

i. These results indicate at 4-leg intersections, one of the major causes of the accident, is passing of a street (either minor or major street) through the intersection. Therefore, it is recommended that 4-leg intersections be replaced with different kinds of intersection such as roundabouts and flyovers.
ii. At 3-leg intersection, the main cause of the accident is minor street passing through the intersection. This probably as a result of an increase in traffic volume in the outgoing minor street and the small number of lanes in this street. Therefore, it is suggested at 3-leg intersections, only the main street should be allowed to pass through the intersection.
iii. The higher the traffic volume the higher the chance of occurrence of traffic accidents. Therefore, it is recommended that policies should be adopted by the government of UAE to reduce traffic volume on the road at any given time.

Regarding the accident severity, it was found that:

iv. Five significant variables affect the severity of traffic accidents occurs at signalized intersections. The significant variables are the speed of the main road, traffic signal sequences, number of through lanes of minor road number of left lanes of main and minor roads.

v. By using the minor injury level as the reference category and refer the three other accident-severity levels to the minor-injury level. The results show that using lead/lag signal phasing sequence increase the probability of severe injuries refer to a minor-injury. In addition, the geometric design of minor roads in terms of the number of through and left lane movements affect significantly the accident-severity levels.

References

1. Ouni, F., Belloumi, M.: Modeling traffic accident occurrence at hazardous road locations: a case study in Tunisia (2018)
2. Haleem, K., Abdel-Aty, M.: Examining traffic crash injury severity at unsignalized intersections. J. Saf. Res. **41**(4), 347–357 (2010)
3. Abdel-Aty, M.A., Radwan, A.E.: Prevention, modeling traffic accident occurrence and involvement. Accident Anal. **32**(5), 633–642 (2000)
4. Miaou, S.-P.: The relationship between truck accidents and geometric design of road sections: poisson versus negative binomial regressions. Oak Ridge National Lab, TN, United States (1993)
5. Cai, Q., Abdel-Aty, M., Lee, J., Eluru, N.: Comparative analysis of zonal systems for macro-level crash modeling. J. Saf. Res. **61**, 157–166 (2017)
6. Joshua, S.C., Garber, N.J.: Estimating truck accident rate and involvements using linear and Poisson regression models. Transp. Plann. Technol. **15**(1), 41–58 (1990)
7. Zewde, T.: Determinants that lead drivers into traffic accidents: a case of Arba Minch city, South Ethiopia. Sci. J. Appl. Math. Stat. **5**(6), 205–210 (2017)
8. Agent, K.R., Deen, R.C.: Relationships between roadway geometrics and accidents (1974)
9. Milton, J.C., Mannering, F.L.: The relationship between highway geometrics, traffic-related elements, and motor vehicle accidents (1996)
10. Xie, Y., Zhao, K., Huynh, N.: Analysis of driver injury severity in rural single-vehicle crashes. Accident Anal. Prevent. **47**, 36–44 (2012)
11. Ye, F., Lord, D.: Comparing three commonly used crash severity models on sample size requirements: multinomial logit, ordered probit and mixed logit models. Anal. Methods Accident Res. **1**, 72–85 (2014)
12. Dong, C., Richards, S.H., Huang, B., Jiang, X.: Identifying the factors contributing to the severity of truck-involved crashes. Int. J. Inj. Control Saf. Promot. **22**(2), 116–126 (2015)
13. Penmetsa, P., Pulugurtha, S.S.: Modeling crash injury by road feature to improve safety (2017)
14. Abdel-Aty, M.: Analysis of driver injury severity levels at multiple locations using ordered probit models. J. Saf. Res. **34**(5), 597–603 (2003). https://doi.org/10.1016/j.jsr.2003. PMID: 14733994
15. Shawky, M., Hasssan, H., Garib, A., Al-Harthei, H.: Examining the factors affecting the severity of run-off-road crashes in Abu Dhabi. Can. J. Civil Eng. **43**, 132–138 (2016)
16. Basu, S., Saha, P.: Regression models of highway traffic crashes: a review of recent research and future research needs. Procedia Eng. **187**(Supplement C), 59–66 (2017)
17. Imran, M., Nasir, J.A.: Road traffic accidents. Prof. Med. J. **22**(06), 705–709 (2015)

18. Razali, M.S.M.A., Zamzuri, Z.: Modeling motorcycle road accidents with traffic offenses at several potential locations using negative binomial regression model in Malaysia (2016)
19. Mohanty, M., Gupta, A.: Factors affecting road crash modeling. J. Transp. Lit. **9**(2), 15–19 (2015)
20. Yang, H., Ozbay, K., Ozturk, O., Xie, K.: Work zone safety analysis and modeling: a state-of-the-art review. Traffic Inj. Prev. **16**(4), 387–396 (2015)

Effects of a Background Arrangement on Collision-Prediction Accuracy for Approaching Objects

Yohsuke Yoshioka(✉) and Hinako Tanaka

Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba City, Japan
yoshioka.yohsuke@faculty.chiba-u.jp

Abstract. This study implemented virtual environment technology to examine how the spacing and height characteristics of a column-type background arrangement affected collision-prediction accuracy for an approaching object set in vast virtual space. Participants were divided into two groups for testing, including the "safety-oriented type" and "accurate judgment type." Results showed that participants in the safety-oriented type group tended to make collision/non-collision judgments at earlier stages, when columns were higher and more narrow. On the other hand, subjects in the "accurate judgment type" group tended to make accurate collision judgements at later stages (i.e., as column spacing increased).

Keywords: Human factors · Collision prediction · Visual information · Virtual environment technology

1 Introduction

In some marine accidents, it is difficult to accurately determine the size of the other boat or its distance from one's own boat, thus leading to collision [1, 2]. This is because there are few objects that can be used for visual comparison while on the sea, thereby making it difficult to gain a sense of scale and distance [3].

Accidents that occur at intersections with good prospects are referred to as "collision course accidents," many of which may occur due to the characteristics of human peripheral vision, which make it difficult for operators to perceive non-moving objects [4, 5]. For road driving conditions, poles are often intermittently installed on roadsides to create a contrast between approaching vehicles and their changing backgrounds. Previous research has confirmed that these measures effectively reduce the probability of collision by increasing each driver's awareness of other approaching vehicles [6]. The practice of arranging specific background objects to create contrasts between moving objects from behind not only aids in visualizing the moving object, but also provides a reference for the distance between oneself and the moving object, especially while on the sea or in countryside areas where there are no other points of comparison. As such, an examination of the relationship between method of arranging the background objects and resulting accuracy of estimating distance can provide insight into how vehicle operators determine distance between themselves and other moving objects in vast spaces.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 247–252, 2020.
https://doi.org/10.1007/978-3-030-50943-9_31

2 Purpose

In this study, we conducted experiments using virtual environment technology to examine the effects of a background arrangement on collision-prediction accuracy for approaching objects while operating in vast space. Based on the specific effects of the background environment, the main purpose was to obtain useful information for creating spatial designs aimed at reducing collision probability.

3 Method

We conducted two experiments using virtual environment technology. For both, participants entered a vast virtual space that was constructed through the Vizard 5.0 virtual reality (VR) software via the HTC Vive head-mounted display. Once operating in the virtual space, participants perceived that they were traveling forward at a speed of 4 m/s while a sphere measuring 1 m in diameter approached from the right side at an angle that would orthogonally cross their paths. Participants were asked to brake if they believed they would collide with the object within a few seconds.

Further, different types of rectangular columns were placed in lines behind the approaching objects in each experiment. This was done to test for the influences of column width, height, and spacing, which were set as variables to determine their respective effects on braking accuracy when attempting to avoid colliding with the approaching objects.

3.1 Experiment 1

Experiment 1 was conducted using the following 3×4 conditions (i.e., participants randomly experienced 12 conditions based on random matches between I–III and A–D):

The distance between the spherical object and operator viewpoint at the intersection:

I. The viewpoint and sphere collide
II. The viewpoint passes in front of the sphere
III. The viewpoint passes behind the sphere

Column shapes and arrangements:

A: Column widths and spacings are smaller than the diameter of the spherical object
B: Column widths and spacings are equal to the diameter of the spherical object
C: Column widths and spacings are larger than the diameter of the spherical object
D: No columns are presented

Both the viewpoint and spherical object were halted 7.5 m before the collision point. Immediately after stopping, participants answered whether they would have collided with the object if they had not stopped by answering yes or no. The effects of column arrangement could be determined based on whether these answers were correct or incorrect (Fig. 1).

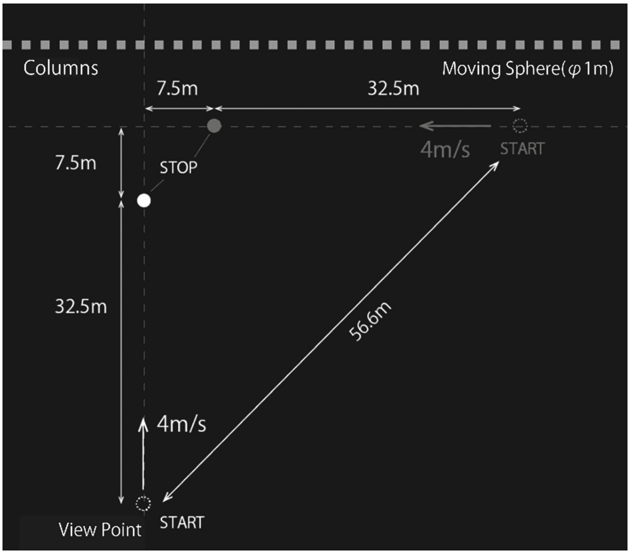


Fig. 1. Participant viewpoint and trajectory of the moving sphere in the collision condition

3.2 Experiment 2

In Experiment 2, participants were asked to apply the brake based on the movement seen from their viewpoint by pressing a button when a collision was predicted. The effects of column arrangement were thus verified on collision prediction by comparing respective braking times.

Along with items A–D from Experiment 1, the following conditions were added to the column shapes in Experiment 2:

E: Column height is greater than the diameter of the spherical object
F: Column height is lesser than the diameter of the spherical object
G: Column depth is greater than the diameter of the spherical object
H: Column depth is lesser than the diameter of the spherical object

4 Results and Discussion

In each experiment, the distance from the participant's viewpoint to the trajectory of the spherical object was initially set at 40 m, while the velocities of both were set at 4 m/s. As such, participants needed to predict or judge a possible collision within 10 s if the viewpoint and sphere would actually collide. In this context, collision-prediction accuracy was thus determined based on column arrangement by quantifying the amount of judgement time needed within the 10-s period.

4.1 Results of Experiment 1

Experiment 1 was conducted among eight college students. Table 1 shows the correctness rate for determining collisions under each experimental condition. In the post-experimental questionnaire, seven participants answered that the speed of viewpoint movement seemed to change depending on column shape (i.e., slower for thicker columns and faster for thinner columns).

As shown in Table 1, the highest correctness rate was found for the condition with no columns. This is because the spherical object appeared to move straight ahead in comparison to the viewpoint when no columns were presented. For this condition, the experimental task may have simplified the problem of avoiding the moving sphere to the right or left based on the experimental setting.

We expected that cube-shaped columns would facilitate the perception of distance and position, thereby resulting in a higher correctness rate. However, results showed that this condition produced the lowest correctness rate. Participants who misjudged the collision tended to think that the spherical object would pass behind them.

Some participants commented that it was difficult to determine their own size in the virtual environment while using the head-mounted display. As such, the boundary between collision and non-collision was ambiguous. In particular, collision detection could vary based on certain preconditions (e.g., the extent to which the user estimated their own size within the virtual environment) when operating under conditions in which the sphere passed through backwards. These types of situations likely affected participant judgment.

Table 1. Rate of correctness for collision detection in Experiment 1 (%)

	Thin columns	Cube columns	Thick columns	None
Passed in front of viewpoint	75.0	62.5	62.5	100.0
Collided with viewpoint	75.0	87.5	87.5	75.0
Passed behind of viewpoint	75.0	50.0	62.5	87.5
Average	75.0	66.5	70.8	87.5

4.2 Results of Experiment 2

Experiment 2 was conducted among 10 college students. Participants were divided into two groups. This included the “safety-oriented type,” which applied brakes even when they would not collide with the spherical object, and the “accurate judgment type,” which never applied brakes if no collision would occur.

Multiple comparisons were then made using the data produced between groups. Here, significant intergroup differences were found under the conditions in which column depth was “small” and when column height was “large.” Further, the “accurate judgment type” group revealed a significant difference between conditions involving “small” and “large” column widths. On the other hand, the “safety-oriented type” group revealed significant differences between the conditions in which column heights were “large and medium” and “large and small” (Fig. 2).

When depth was "small," the accurate judgment type waited to make a judgment just prior to the collision, while the safety-oriented type tended to apply the brake at a relatively early stage. When column depth is small, the gaps between columns become large. Because of the higher contrast between the columns and gaps, participants began to pay more attention to the column than the spherical object, thus affecting their ability to detect collisions with the spherical object. Such difficulty likely influenced the "accurate judgment type" to take more time when making judgements, but influenced the "safety-oriented type" to judge more impatiently.

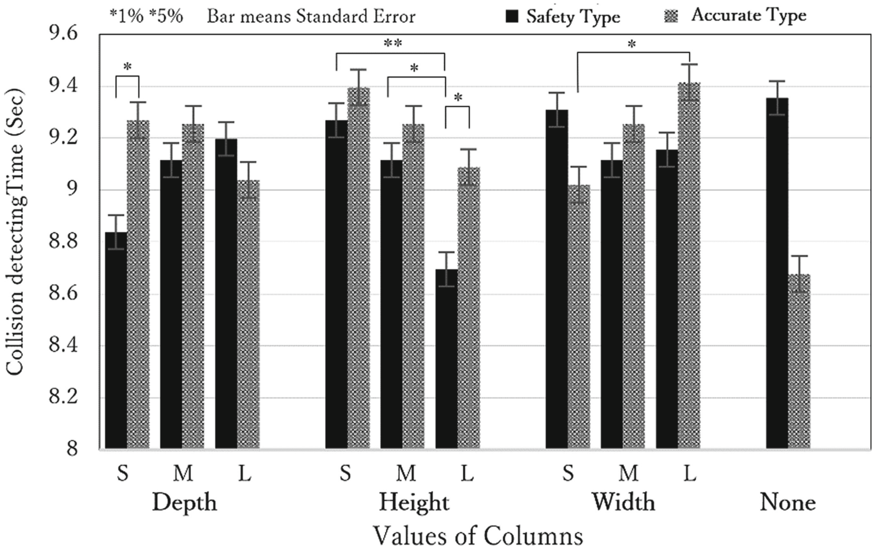


Fig. 2. Estimated time required for detecting collisions in each condition during Experiment 2

The greater the column height, the earlier the safety-oriented type predicted collisions. On the other hand, prediction times increased when column heights were lesser. This is because greater column heights increased their areas within the field of view, thus giving the impression that the columns were approaching. Conversely, reduced column heights gave the impression that there was more approaching time, thereby possibly slowing reaction times. Further, narrow columns appeared to move faster than thicker columns. It can thus be considered that thinner columns facilitate earlier collision prediction, while thick columns delay collision prediction.

5 Conclusion

This study's experiments revealed that participants in the "safety-oriented type" group tended to make collision/non-collision judgments at earlier stages (i.e., when columns were higher and more narrow), while participants in the "accurate judgment type" group tended to make accurate collision judgements at later stages (i.e., when column spacing increased). Results further confirmed that collision-prediction accuracy for the moving sphere-shaped object was affected by both the size and spacing of background columns. Notably, collision-judgment accuracy also differed based on each participant's personality.

Many factors are involved in collision incidents. However, it may be possible to statistically clarify whether the safety-oriented or accurate judgment types are more likely to cause actual collisions. In this context, our experimental findings can be used to determine the exact value settings needed for background objects when attempting to reduce the probability of collision accidents. In addition, we will use the same findings to examine the detailed settings of the background columns to improve collision-prediction accuracy through verification experiments involving the factors that affected the results for both the safety-oriented and accurate judgment types.

Acknowledgments. This research was supported by JSPS KAKENHI Grant Number JP17H03359.

References

1. Nishimura, T.: A study on effects of the horizontal distance between two masthead lights on lookout by other vessels - a case of collision between guided missile destroyer "Atago" and fishing boat "Seitokumaru" –. J. Jpn. Inst. Navig. **131**, 33–39 (2014)
2. Inoue, K.: Safety criteria for the passing distance to achieve between high speed and conventional craft. J. Jpn. Inst. Navig. **86**, 1–10 (1992)
3. Toyoda, K., Iwai, N., Oyama, S.: A study of peripheral visual field characteristics at good prospect intersection. IEEJ Trans. Electron. Inf. Syst. **131**(12), 2148–2153 (2011)
4. Saunier, N., Sayed, T., Lim, C.: Probabilistic collision prediction for vision-based automated road safety analysis. In: 2007 IEEE Intelligent Transportation Systems Conference, Seattle, WA, pp. 872–878 (2007)
5. Berthelon, C., Mestre, D.: Curvilinear approach to an intersection and visual detection of a collision. Hum. Factors **35**, 521–534 (1993)
6. Charlton, S.G.: Restricting intersection visibility to reduce approach speeds. Accid. Anal. Prev. **35**, 817–823 (2003)

Influence of Passive Fatigue and Take-Over Request Lead Time on Drivers' Take-Over Performance

Ali Muhammad Hadi, Qingkun Li, Wenjun Wang(✉), Quan Yuan, and Bo Cheng

School of Vehicle and Mobility, Tsinghua University, Beijing, China
wangxiaowenjun@tsinghua.edu.cn

Abstract. At SAE Level 3 of driving automation, human drivers are not expected to continuously control the vehicle or monitor the driving environment. However, due to system limitations, drivers would be prompted to intervene and "take-over" the vehicle in certain instances. This brings into question the capability of human driver to take control after being out of the loop for prolonged periods. Distractedness has been identified and modelled in previous studies however, paucity exists with regards influence of drowsiness in the absence of a secondary task on take-over performance. This study discusses the effect of parameters like take-over request lead time and the degree of driver's drowsiness on take-over performance. Experiments were conducted in a high fidelity driving simulator on 12 licensed drivers. Results reveal significant influence of lead time and degree of drowsiness on different performance parameters. Subjective scores of drivers also correlated with objective parameters of drowsiness.

Keywords: Human factors · Automated driving · Drowsiness · Fatigue · Take-over performance

1 Introduction

At SAE level 3 of automated driving (Conditionally automated driving), drivers act as fallback and their vigilance is not consistently required. In instances like exiting of operational design domain (ODD) or automation failure, drivers are expected to intervene and take-over vehicle control [1] upon system generated take-over request (TOR). The necessity of driver intervention brings into question the ability of a human driver to take over control after being out of the loop during prolonged periods of automation [2]. The state of driver holds important value in a conditionally automated vehicle [3]. The two primary states that have been researched for drivers in automated driving include distractedness (caused by secondary task engagement) and drowsiness (also defined as fatigue or passive fatigue) [4]. Several studies in the past have thoroughly investigated the influence of distracted driving with different non-driving related tasks (NDRT). From critical to non-critical take-over request lead times (TORlt), a review paper has explored drivers' reaction time in different task modalities. The findings reveal average lead time to be 6.37 ± 5.36 s and average reaction time to be 2.96 ± 1.96 s for the

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 253–259, 2020.
https://doi.org/10.1007/978-3-030-50943-9_32

referred studies [5]. Due to criticality and complexity of experiment design, most experimental studies have been conducted in driving simulator with different lead times [6], traffic densities [7] and task modalities [8, 9]. One study compares driving simulator studies with on-road experiments involving take-over requests concluding that greater risk perception decreased reaction time in on-road experiments but strong positive correlation exists between simulator and on-road experiment results [10]. Both driving simulator [11] and on-road [12] experiments have also been conducted to investigate the effect of fatigue/drowsiness on take-over performance. A between subject study on 47 participants revealed no significant difference between take-over time however for fatigued drivers, certain indicators like higher longitudinal acceleration (due to braking force) and lower situational awareness (due to participants not checking mirrors or environment) suggested a startle reaction due to fatigue [13]. In most studies requiring fatigue assessment, observer rated scales like Karolinska Sleepiness Scale (KSS) [14] are used for which independent observers are needed to trigger TOR at different levels [11]. However, as discussed in [13], an objective fatigue assessment system (FAS) was also employed in parallel for the research. This system as described by authors in another paper [15] uses eye tracking instrument to attain values of percentage of eye closures (PERCLOS), number of microsleep events (NMS) and head movement behavior. Using a decision tree classifier, high accuracy of system was attained using only PERCLOS and NMS. With a sensitivity of 90% and specificity of 99.2%, when compared with observer rating, it can be established that integration of such simple quantified metrics in driver monitoring systems (DMS) can be a step towards accurate driver state classification for future CAD applications.

2 Method

2.1 Experiment Design and Scenario

A total of 12 participants with normal vision and no previous experience of on-road automated driving were recruited for this driving simulator based study with an average age of 32.16 years (SD = 9.03, Min = 22, Max = 49) and a driving experience of 6 years (SD = 2.95, Min = 3, Max = 14). A repeated measures design with 12 sub-experiments on each participant was used for this study. The driving environment consisted of a three lane highway with ego vehicle positioned in the middle lane. The traffic in left lane was absent whereas a consistent traffic density of 20 vehicles per kilometer (50 m distance between vehicles) was maintained in the right lane for each sub-experiment. Traffic speed in the right lane was 80 km/h whereas the automated ego vehicle maintained a continuous speed of 100 km/h until take-over request. Upon TOR, middle and left lane were blocked by the presence of traffic cones and participants had to monitor their position, surrounding traffic and distance to boundary within the time budget to effectively take-over by switching to right lane. Due to difference in cruise speed with respect to surrounding traffic, the positions of front and rear vehicles of right lane traffic in all experiments were random variables. Figure 1 shows the driving environment scenario at TOR. A total of Four lead times (4, 6, 8 and 10 s) were used in order to study the effect of lead time on take-over performance.

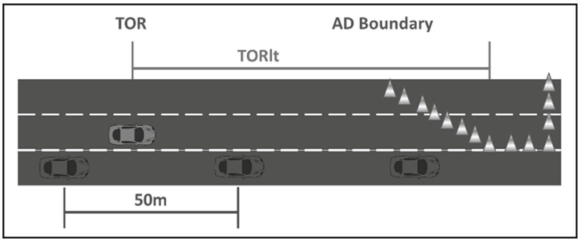


Fig. 1. An illustration of TOR scenario.

Since, this study's focus is to evaluate the effect of fatigue on take-over performance, the take-over request was triggered if drivers exhibited signs of drowsiness. The subjective evaluation of fatigue during the experiment was carried out by two observers in the control room and they were also responsible to trigger a TOR. For the convenience of observers, the camera module displayed the facial characteristics of participants including the eye state (1 = close and 0 = open) along with a blink rate indicating average eye state over the last minute (from 0 to 1). In the first 8 experiments (phase 1), participants had the freedom to succumb to their drowsiness and sleep whereas in the last 4 experiments (phase 2), they were required to resist their urge to sleep. In both cases however, monitoring of driving environment was not a requirement. Furthermore, they were told not force themselves to sleep by keeping their eyes closed for prolonged periods. A take-over would take place upon ±2° change in steering wheel or 5% change in acceleration or brake pedal position (0.5 MPa for brake pressure). Table 1 shows TOR lead time of each experiment.

Table 1. Take-over lead times of all experiments.

Experiment number	1	2	3	4	5	6	7	8	9	10	11	12
TORlt in seconds	6	8	10	4	8	4	10	6	6	10	4	8

2.2 Procedure

Participants were informed in advance to reduce two hours of their usual sleep night before the experiment and avoid any caffeinated beverages before experiment. Upon arrival to the driving simulator facility, participants were told to set aside all distractible articles (watches, phones, books, wallets, etc.) before sitting in the driving simulator vehicle. After they were positioned inside the vehicle, eye tracker was mounted on their heads and gaze calibration was performed. After calibration, an acclimation phase began in which drivers were instructed to familiarize themselves with the simulated driving environment and navigate through the three lane highway with traffic in the right lane. They were instructed to perform different maneuvers like maintaining a certain speed and headway distance and changing lanes without collisions. After assuring that drivers were familiar with the setup and could comfortably perform the driving task, two training experiments similar to the actual testing experiments were performed on drivers in awake state with the purpose of assuring the take-over capability of drivers in simulated environment and removing of practice effects normally associated with repeated measures study. Participants were told in advance about the capabilities of the autonomous system to maintain longitudinal and lateral control. They were further told that whenever their intervention would be necessary, a take-over request in the form of an audio-visual alarm will prompt them to take over the vehicle control. Before TOR, participants were supposed to keep their hands off the steering wheel and feet off the pedals to ensure no difference in performance parameters. After training experiments, both phases of experiments were carried out. After each sub-experiment, participants were asked to evaluate their degree of drowsiness, the difficulty of task and their take-over performance on a 10-point scale.

3 Results and Discussion

3.1 Frequency of Collision

Collision in this study refers to any physical contact of ego vehicle with surrounding objects including cones, traffic vehicles, road shoulders etc. that should have been avoided for a successful maneuver in the simulated environment. For the total dataset of 144 sub-experiments, the collision frequency was 30/144 (20.8%) with 9 out of 36 cases (25%) occurring in 4 s lead time compared to 7/36 (19.4%) in 6, 8 and 10 s TORlt. This shows diminishing effect of increasing lead time beyond 4 s on collision prevention. In this paper, collision is used to check validity of a case for inclusion in performance analysis. The total number of valid cases therefore becomes 114.

3.2 Degree of Drowsiness and Take-Over Performance

In this study, degree of drowsiness is assessed using both objective and subjective methods. Objective criteria includes "Percentage of eye closure" or PEC measured in the duration from 60 s and 120 s before TOR to the beginning of the TOR termed as 60sPEC and 120sPEC respectively. PEC is similar to PERCLOS but also includes blinks. For take-over performance evaluation, parameters like reaction time (RT), time to collision (TTC), distance to collision (DTC), maximum brake pressure (MBP) and maximum steering wheel angle (MSWA) were used. Furthermore, parameters like eyes on road time (EoRT) also known as gaze reaction time and eyes on mirror time (EoMT) which is the amount of time after TOR until first fixation falls on right rear view mirror aided in understanding the effect of fatigue on situational awareness and responsiveness.

The results of regression analysis in Table 2 between 60sPEC (independent variable) and take-over performance metrics and in Table 3 between 120sPEC (independent variable) and take-over performance metrics reveal negative influence of degree of drowsiness on take-over performance. The positive correlation of reaction time and negative correlation of TTC and DTC with PEC indicated the negative influence of drowsiness on take-over quality and performance. Furthermore, the analysis reveals that higher PEC resulted in longer gaze reaction time and delay in first glance on right rear-view mirror. This indicates a negative effect of degree of drowsiness on situational awareness required for safe vehicle control transition from automation.

Table 2. Regression analysis of 60sPEC with different take-over performance metrics with $p < 0.05$ criteria used.

Variable	R^2	Unstandardized beta (β)	Standardized beta (Sβ)	Probability value (p)
RT (seconds)	0.174	0.017	0.418	0.000
TTC (seconds)	0.059	−0.014	−0.243	0.009
DTC (meters)	0.055	−0.325	−0.234	0.012
MBP (MPa)	0.003	−0.003	−0.058	0.538
MSWA (degrees)	0.010	0.019	0.098	0.301
EoRT (milliseconds)	0.195	8.592	0.441	0.000
EoMT (milliseconds)	0.155	23.676	0.394	0.000

Table 3. Regression analysis of 120sPEC with different take-over performance metrics with $p < 0.05$ criteria used.

Variable	R^2	Unstandardized beta (β)	Standardized beta (Sβ)	Probability value (p)
RT (seconds)	0.161	0.018	0.401	0.000
TTC (seconds)	0.055	−0.015	−0.235	0.013
DTC (meters)	0.048	−0.334	−0.219	0.021
MBP (MPa)	0.009	−0.006	−0.095	0.321
MSWA (degrees)	0.008	0.019	0.088	0.385
EoRT (milliseconds)	0.186	9.394	0.431	0.000
EoMT (milliseconds)	0.153	26.237	0.392	0.000

3.3 Influence of Lead Time

Another purpose of this study was to evaluate the effect of take-over lead time (TORlt) on different take-over performance parameters. From the data in table, the average value of parameters like reaction time change significantly with increase in lead time. Similarly, a greater lead time led to increase in values of distance and time to collision indicating that a safer maneuver can result from increment in lead time (Table 4).

Table 4. Average and standard deviation value of take-over performance parameters for different lead times.

Variable	4 s	6 s	8 s	10 s
RT (seconds)	2.60 (1.28)	2.91 (0.95)	3.33 (1.26)	3.67 (1.96)
TTC (seconds)	1.98 (1.21)	2.90 (1.56)	4.50 (1.86)	5.10 (2.00)
DTC (meters)	43.7 (26.6)	62.4 (36.0)	101.6 (41.75)	120.5 (48.9)

3.4 Comparison with Subjective Scores

After each sub-experiment of this study, feedback was collected from participants about their degree of drowsiness, the difficulty of task and their take-over performance on a 10-point scale. A strong positive correlation was attained between 60sPEC and sleepiness score of participants ($R^2 = 0.423$, $p < 0.05$, $\beta = 0.036$) showing validity of PEC in evaluating degree of drowsiness. Furthermore, difficulty of scenario score positively correlated with 60sPEC ($R^2 = 0.096$, $p < 0.05$, $\beta = 0.016$) indicating that higher degree of drowsiness would result in greater perceived difficulty. In addition, 60sPEC negatively correlated with quality of take-over score ($R^2 = 0.097$, $p < 0.05$, $\beta = -0.021$) signifying the negative influence of drowsiness on take-over performance and drivers' understanding of their performance. Table 5 shows average subjective rating of scenario difficulty and take-over quality for each lead time. It can be seen that with increasing lead time, average rating for difficulty of scenario dropped whereas average rating for quality of take-over improved.

Table 5. Average and standard deviation value of subjective score of performance parameters for different lead times.

Variable	4 s	6 s	8 s	10 s
Difficulty of scenario	7.26 (1.023)	5.90 (1.345)	5.66 (1.421)	5.28 (1.830)
Take-over quality	5.74 (1.831)	5.93 (2.034)	6.52 (1.299)	6.69 (1.929)

4 Conclusion

The results of this study reveal that the higher the degree of drowsiness is, the worse the take-over performance becomes and the greater the difficulty of task is perceived by the driver. In addition, the longer the lead time is, the better the take-over performance becomes and the lesser the difficulty of task is perceived by the driver. Furthermore, the significance eye state is established for both 1 and 2 min before TOR indicating that the evaluation of eye state by future driver monitoring systems should not be confined to the last minute or 30 s since inducing drowsiness takes longer compared to inducing distraction.

Acknowledgements. The authors of this paper gratefully acknowledge the support provided by TOYOTA Motor Company.

References

1. Wintersberger, P., Green, P., Riener, A.: Am I driving or are you or are we both? A taxonomy for handover and handback in automated driving. In: 2017 Driving Assessment Conference, Vermont (2017)
2. Louw, T., Merat, N., Jamson, H.: Engaging with highly automated driving: to be or not to be in the loop? In: 2015 Driving Assessment Conference, Salt Lake City, Utah (2015)
3. Rauch, N., Kaussner, A., Kruger, H.-P., Boverie, S., Flemisch, F.: The importance of driver state assessment within highly automated vehicles. In: 16th ITS World Congress and Exhibition on Intelligent Transport Systems and Services, Stockholm, Sweden (2009)

4. Naujoks, F., et al.: From partial and high automation to manual driving: relationship between non-driving related tasks, drowsiness and take-over performance. Accid. Anal. Prev. **121**, 28–42 (2018)
5. Eriksson, A., Stanton, N.A.: Takeover time in highly automated vehicles: noncritical transitions to and from manual control. Hum. Factors J. Hum. Factors Ergon. Soc. **59**(4), 689–705 (2017)
6. Ito, T., Takata, A., Oosawa, K.: Time required for take-over from automated to manual driving. In: SAE 2016 World Congress and Exhibition, Detroit (2016)
7. Gold, C., et al.: Taking over control from highly automated vehicles in complex traffic situations the role of traffic density. Hum. Factors J. Hum. Factors Ergon. Soc. **58**(4), 642–652 (2016)
8. Yoon, S.H., Ji, Y.G.: Non-driving-related tasks, workload, and takeover performance in highly automated driving contexts. Transp. Res. Part F Traffic Psychol. Behav. **60**, 620–631 (2019)
9. Wandtner, B., et al.: Non-driving related tasks in highly automated driving - effects of task modalities and cognitive workload on take-over performance. In: AmE 2018 - Automotive Meets Electronics; 9th GMM-Symposium, Dortmund, Germany (2018)
10. Eriksson, A., Banks, V.A., Stanton, N.A.: Transition to manual: comparing simulator with on-road control transitions. Accid. Anal. Prev. **102**, 227–234 (2017)
11. Kreuzmair, C., Gold, C.M., Meyer, M.J.: The influence of driver fatigue on take-over performance in highly automated vehicles. In: 25th International Technical Conference on the Enhanced Safety of Vehicles (ESV), Detroit Michigan, United States (2017)
12. Weinbeer, V., Muhr, T., Bengler, K., Baur, C., Radlmayr, J., Bill, J.: Highly automated driving: how to get the driver drowsy and how does drowsiness influence various take-over aspects? (2017)
13. Bengler, A., Kroll, D., Bengler, K.: Wake up and take over! The effect of fatigue on the take-over performance in conditionally automated driving. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA. IEEE (2018)
14. Miley, A.Å., Kecklund, G., Åkerstedt, T.: Comparing two versions of the Karolinska Sleepiness Scale (KSS). Sleep Biol. Rhythms **14**(3), 257–260 (2016)
15. Feldhütter, A., et al.: A new approach for a real-time non-invasive fatigue assessment system for automated driving. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 62, no. 1 (2018)

Modeling of a Vehicle Accident Prediction System Based on a Correlation of Heterogeneous Sources

Pablo Marcillo(✉), Lorena Isabel Barona López,
Ángel Leonardo Valdivieso Caraguay,
and Myriam Hernández-Álvarez

Departamento de Informática y Ciencias de la Computación,
Escuela Politécnica Nacional, Ladrón de Guevara E11-25 y Andalucía,
Edificio de Sistemas, 170525 Quito, Ecuador
{pablo.marcillo,lorena.barona,angel.valdivieso,
myriam.hernandez}@epn.edu.ec

Abstract. Statistics affirm that traffic accidents are the main cause of death in developing countries. The indicators are alarming, so governments, manufacturers, and researchers have been looking for solutions to mitigate them. Despite all efforts to face this problem, the number of victims remains high. A significant percentage of traffic accidents are caused by external factors, so the search for solutions that use information from multiple sources is crucial. This article presents a traffic accident prediction system based on heterogeneous sources using data mining techniques and machine learning algorithms. The development of this system includes the following tasks: collecting information from different sources, performing cluster analyses and feature selection, generating new datasets, performing machine learning algorithms to define accident rates, and sending traffic rate levels to the vehicles. For this article, we focused on performing cluster analyses to determine high-risk clusters that identify drivers with risky driving patterns.

Keywords: Traffic accident prediction · Heterogeneous sources · High-risk clusters · Machine learning · Data mining

1 Introduction

International financial agencies have established that traffic accidents slow down the economic growth in developing countries because most of the victims are in working age [1]. According to local statistics, the primary cause of death in men and women aged 20 to 30 years is for traffic accidents [2]. Because the indicators about traffic accidents are alarming, states and local governments have implemented policies and technology infrastructure to mitigate them. Furthermore, vehicle manufacturers have incorporated driver assistance systems such as Vehicle Collision Avoidance System (VCAS) [3] or Advanced Driver Assistance System (ADAS) [4] on their vehicles. Although these systems have reduced the accident rate, the number of casualties remains high.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 260–266, 2020.
https://doi.org/10.1007/978-3-030-50943-9_33

Much research in recent years has focused on the design and the development of traffic accident prediction systems based on machine learning, data mining, cloud computing, and others. For instance, proposals about frameworks for prevention and prediction of accidents based on machine learning algorithms and recognition techniques, deep learning models for traffic accident prediction based on neural networks and heterogeneous sources, or the optimization of the existing machine and deep learning algorithms have been presented.

The limitation with current prediction systems is about making decisions based only on one or very few data sources. A significant percentage of accidents are caused by external factors such as traffic and weather conditions and medical data, physical, and physiological conditions of drivers [5]. Thus, it is essential having a prediction system that permits to collect, merge, and process multiple data sources and distribute processed information to the users. In this study, a traffic accident prediction system that correlates multiple data sources is introduced. Additionally, some high-risk driving patterns that identify drivers with a high risk of suffering an accident are presented.

This article is organized as follows. Section 2 introduces some proposals related to the subject of this study, followed by Sect. 3 that presents our solution. Then, Sect. 4 describes the methodology. The results and the conclusions of this study are shown in Sect. 5 and 6, respectively.

2 Related Work

Some proposals for traffic accident prediction have been presented. For instance, Xiong et al. [6] proposed a framework for the prevention and prediction of traffic accidents based on machine learning algorithms and recognition techniques. It introduces the concept of Chain of Road Traffic Accident, which establishes that a traffic accident is preceded by a series of incidents. Therefore, it aims to avoid traffic accidents by identifying them at an early stage.

Meanwhile, Ren et al. [7] proposed a deep learning model for traffic accident prediction based on a recurrent neural network (RNN) and heterogeneous sources. It established the following traffic-related factors: traffic accident, traffic flow, geographical position, weather, air quality, time period, and holiday. The Granger analysis was used to rank the predictive power of factors and features related to traffic accidents.

Also, Park et al. [8] proposed a traffic accident prediction model that uses a framework based on MapReduce algorithm which accomplishes the following steps: processing of data sets, combination and creation of training data sets, performing of over-sampling technique to manage imbalanced training data sets, conducting the classification process and checking the accuracy of the results.

Finally, Yuan et al. [9] proposed a new version of ConvLSTM named Hetero-ConvLSTM which is a deep learning framework with spatial graph features. In this approach, the study area is divided into grid cells and all the features are matched with each grid cell. Considering that traffic accidents may be caused by one factor or another, it considered the spatial heterogeneity of the environment. A common feature of these proposals is the use of one or more types of data sources, however, this feature has been considered as a limitation for our proposal because other types of data sources must also be included.

3 Traffic Accident Prediction System

Our solution proposes the development of a traffic accident prediction system based on diverse data sources. Figure 1 illustrates the schema of our solution. Firstly, information from different data sources is collected. Secondly, cluster analysis to determine high-risk clusters is carried out. Thirdly, a feature selection is performed, and new datasets are generated. Fourthly, machine learning algorithms are performed to define accident rates. Finally, traffic rate levels are sent to the vehicles. This article focuses on the second stage of the schema.

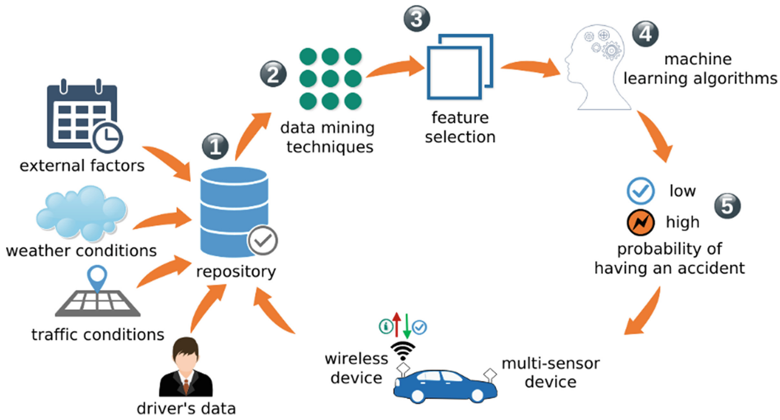


Fig. 1. Schema of our traffic accident prediction system

The data sources defined for this system are vehicle data, driver's data, weather and traffic conditions, and external factors or events. Every data source has many variables that could be included in the system. Figure 2 presents some variables for every data source. One part of them can be collected through On Board Diagnostic (OBD) readers, GPS devices, and multi-sensors, meanwhile another part via web services provided by Internet services such as Open Weather Map, Google Maps, Waze, and others.

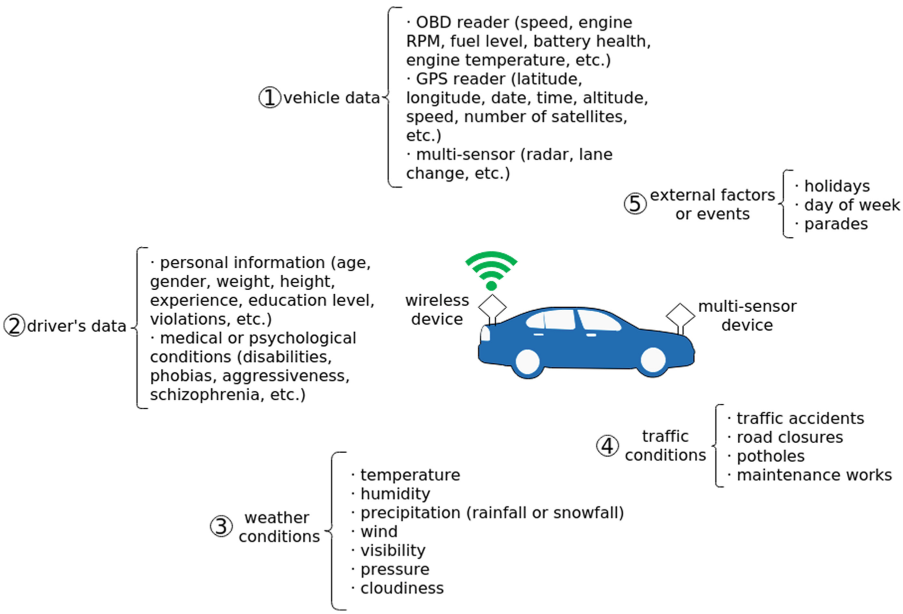


Fig. 2. Heterogeneous data sources

4 Methodology and Materials

The dataset proposed by Santana et al. [10, 11] was used for this analysis. It contains around 7 h of driving data for three different drivers and has a sampling frequency of 100 Hz. Driving data includes measures for variables such as timestamp, blinker, gear choice, gas, speed, brake, revolutions per minute (rpm), steering angle, latitude, longitude, and others. One sample of the driver named '*dog*' was used to perform this analysis. We used *Python* as language programming, and *h5py*, *sklearn*, *numpy*, and *matplotlib* libraries. Besides, the analysis is a four-task process. At first, the Hierarchical Data Format (HDF5) file format is transformed into CVS. Later, variables are chosen and related them to each other. Then, the optimal number of clusters (k) for these relations is calculated [12]. In the end, the K-Means clustering algorithm is performed to identify high-risk clusters.

5 Results

After performing the Elbow method and the K-Means clustering, six high-risk clusters were identified. Figure 3a and Fig. 4a present the analysis for the relation 'speed vs. brake'. The first cluster identifies the drivers who speed and break abruptly. This combination could cause brake pad over-heating and a brake system malfunction. In this case, a value of $k = 16$ was used for the clustering algorithm. Figure 3b and Fig. 4b present the analysis for the relation 'speed vs. rpm'. The second one identifies the drivers who drive at high speed and high-rev limit. This combination could cause severe damage to vehicles and unsuccessful and risky overruns. The optimal number of clusters for this case is $k = 10$.

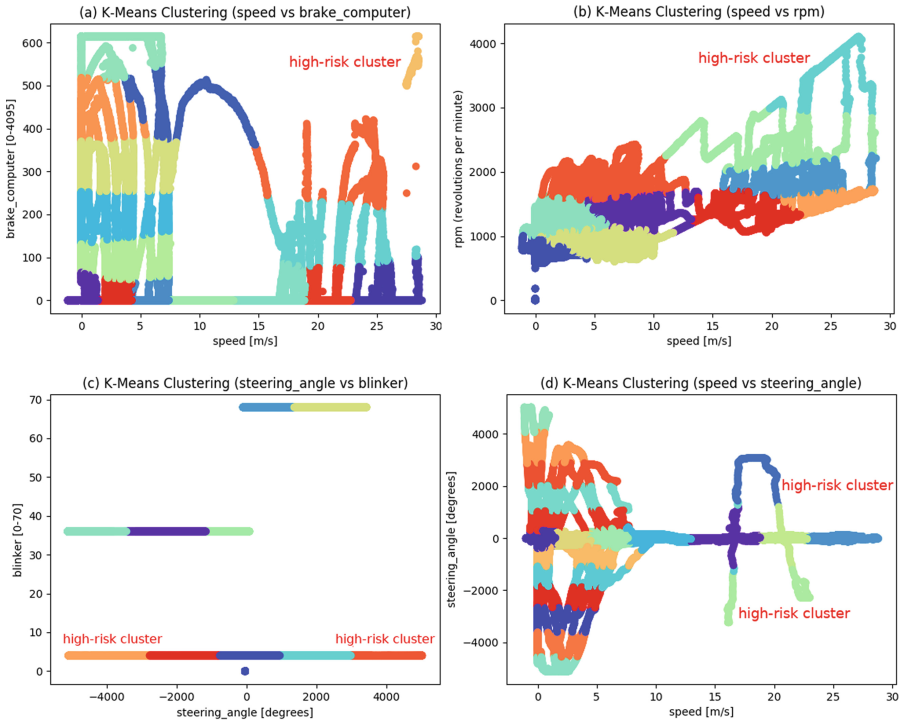


Fig. 3. K-Means clustering analyses

Figure 3c and Fig. 4c present the analysis for the relation 'steering angle vs. blinker'. The third and fourth clusters identify the drivers who make turns to the left or the right and not use blinkers. This combination could cause crashes at the moment of performing unannounced turning maneuvers. In this case, $k = 10$ was the optimal value. Figure 3d and Fig. 4d present the analysis for the relation 'speed vs. steering angle'. The two last clusters identify the drivers who take turns at high speed. This combination could cause skids, loss of vehicle control, or loss of traction. In this case, it was used a value of $k = 20$.

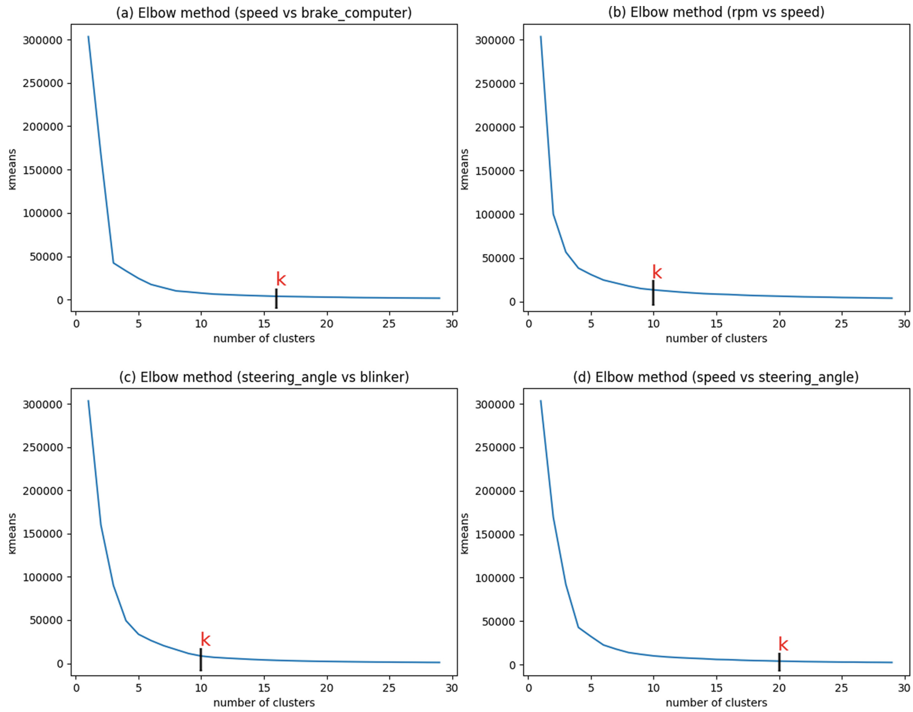


Fig. 4. Elbow method analyses

6 Conclusions

Considering that a significant percentage of traffic accidents are caused by factors beyond vehicle data or traffic conditions, it is crucial to have a system that correlates many more data sources. The high-risk clusters found after the analysis identifies drivers with risky driving patterns. To accomplish this aim, it was necessary to probe with a different number of clusters (k) until to identify clearly the high-risk clusters. In that way, it is possible to identify drivers who break abruptly, take turns at high-speed, drive at the high-rev limit, and not use blinkers properly. Even, it is possible to predict faulty functioning or abusive use of vehicles components in the future.

References

1. The World Bank (2018). https://www.worldbank.org/en/news/press-release/2018/01/09/road-deaths-and-injuries-hold-back-economic-growth-in-developing-countries
2. INEC (2017). https://www.ecuadorencifras.gob.ec/documentos/webinec/Estadisticas_Economicas/Estadistica%20de%20Transporte/2017/2017_TRANSPORTE_PRESENTACI-ON.pdf
3. Wong, C., Qidwai, U.: Vehicle collision avoidance system. In: Proceedings of IEEE Sensors, pp. 31–319. IEEE Press, Vienna (2004)

4. Ziebinski, A., Cupek, R., Grzechca, D., Chruszczyk, L.: Review of advanced driver assistance systems (ADAS). In: AIP Conference Proceedings, vol. 1906, p. 120002. AIP Publishing LLC (2017)
5. Zhang, X., Huang, F., Zheng, C.: Causes analysis of the serious road traffic accident cases. In: 2016 5th International Conference on Energy and Environmental Protection (ICEEP 2016). Atlantis Press, Beijing (2016)
6. Xiong, X., Chen, L., Liang, J.: A new framework of vehicle collision prediction by combining SVM and HMM. IEEE Trans. Intell. Transp. Syst. **19**, 699–710 (2018)
7. Ren, H., Song, Y., Wang, J., Hu, Y., Lei, J.: A deep learning approach to the citywide traffic accident risk prediction. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), pp. 3346–3351. IEEE Press (2018)
8. Park, S., Kim, S., Ha, Y.: Highway traffic accident prediction using VDS big data analysis. J. Supercomput. **72**, 2815–2831 (2016)
9. Yuan, Z., Zhou, X., Yang, T.: Hetero-ConvLSTM. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining – KDD 2018, New York, pp. 98–992 (2018)
10. Santana, E., Hotz, G.: Learning a driving simulator. arXiv preprint (2016)
11. Biasini, R., Hotz, G., Khalandovsky, S., Santana, E., Van der Westhuizen, N.: The people’s comma (2016). https://github.com/commaai/research/blob/master/Logs.md
12. Pham, D., Dimov, S., Nguyen, C.: Selection of K in K-means clustering. Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci. **219**, 103–119 (2005)

Comfort and Posture

A Theoretical Framework for Occupant Comfort in Future Shared Autonomous Vehicles

James Jackson(✉) and Davide Salanitri

IDIADA Automotive S.A., L'Albornar, 43710 Santa Oliva, Spain
james.jackson@idiada.com

Abstract. People want to be comfortable, and perception of what makes each user comfortable can vary greatly. This makes it important that vehicle products meet the different expectations of users through flexibility in their interior environment design. One of the most prominent predictions in personal mobility progression is a move towards shared autonomous vehicles. SUaaVE is a European project aiming at improving acceptance, trust and comfort of future shared automated vehicle users, through development of a system concept known as ALFRED. This paper presents a theoretical framework for an adaptive model, with comfort componentised into attributes. Application of the model aims to drive a response to vehicle user's individual characteristics and preferences - such that their comfort can be optimised intelligently by manipulation of vehicle features and functions as part of the ALFRED concept.

Keywords: Human factors · Automotive · Automated vehicles · Comfort

1 Introduction

Asking a group of people what makes them comfortable is likely to yield responses that could, at very least, be varied. There may be some consensus in influencing factors, but even then it is likely there will be variety in what each perceives as best.

As technology of vehicles continues to progress into an automated and connected world, it is likely that we will see significant changes to the way vehicles are designed and used [1]. As models for use and ownership change it could well be possible that the needs of users change, with a consequential shift in what is perceived as 'just right' across a number of vehicle attributes – none more so than user comfort [1]. Equally, as technology moves forward, there is likely to be utilisation of advancement to increase the capability of how vehicles respond to influence levels of user comfort.

This move towards autonomous and connected vehicles, and integrating technological advancement is the primary focus of the SUaaVE (SUpporting acceptance of automated VEhicles) project. The aim is to ensure that vehicles are able to meet the expectations and characteristics of their users. The project includes a number of different factors, with this paper reviewing a framework for occupant environment comfort.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 269–275, 2020.
https://doi.org/10.1007/978-3-030-50943-9_34

1.1 The SUaaVE Project

SUaaVE [2] is European Union funded project under the Horizon 2020 research framework. The project seeks to close gaps between advancement of shared, connected, and autonomous vehicles and user acceptance, ethics, emotion, and comfort.

The project sets out to develop system able to control intelligent responses of vehicle systems in accordance with human emotions. This comprises of (a) evaluation of a passenger state, and (b) an adaptive interface and vehicle functionality. This is with the aim of manipulating aspects of interior environment, and vehicle dynamic behaviour. Also a key output of the project is guidelines to support public authorities, aiming to represent a breakthrough in public acceptance and ethics of use surrounding future CAVs. Development of components will follow an initial investigative stage of establishing models for each component. This is followed by multiple iterative testing stages with representative naïve users across multiple simulation platforms.

1.2 The ALFRED Concept

The systems and their associated functions developed as part of the project will be integrated into a concept known as ALFRED (Automation Level Four+ Reliable Empathic Driver). Evaluation and response, considering changes and opportunities of automation and connectivity, is central to the concept. It intends to be agnostic of vehicle application, and instead define a theoretical amalgamation of technology and functions for integration into vehicles of the future. This is to be achieved through a series of models covering user acceptance, ethics, emotions, and of particular concern to this paper; comfort. These models unobtrusively evaluate the user and scenario, adapting the system so as to improve user experience of the vehicle.

2 Definition of Comfort

Within literature there is no clear consensus regarding a definition, nor is there an agreed theory regarding 'comfort'. Shen and Vértiz [3] summarised the principal theories by which comfort is defined under two categories; comfort as a state of ease, and comfort as an absence of discomfort. The first of these is defined [4] as the ease or well-being surrounding physical, psychological, and physiological harmony between a person and their surrounding environment. A neutral state is defined by a lack of disturbance in a person's environment, and general state with no sensation of 'heightened' well-being, satisfaction, or perception of being at ease [5]. This is visualised in Fig. 1.

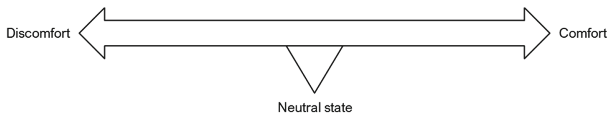


Fig. 1. Multidirectional continuum of comfort

The second explains comfort in terms of an absence of discomfort [6, 7]. Under this theory a state of 'comfort' is regarded as the optimum and is subsequently a starting point of a continuum of discomfort. Here comfort is quantified by the degree to which a person identifies discomfort. This theory is shown in Fig. 2.

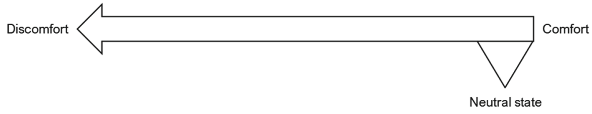


Fig. 2. Unidirectional continuum of comfort

Following on from these models, more recent definitions seek to separate the concepts of comfort and discomfort. Vink and Hellback [R] describes it being a reaction to the environment with comfort as a pleasant state or relaxed feeling and with discomfort as an unpleasant sate. These definitions both find origins in literature describing factors influencing comfort and discomfort. One of these developed as part of a model by De Looze et al. [8] separated the concepts of discomfort and comfort under distinct characteristics. Here, both are considered as influenced by human, product, and environmental characteristics. Subsequently Vink and Hellback [9] developed a model taking into account the earlier models by De Looze et al. [8] and others such as Moes [10].

Considering the basis from existing theory and research, it can be said that there is no absolute definition of comfort. That said, there is some consensus in its components. The quantification of comfort in terms of positive and negative sensation governing overall state is widely used as a reference point of various theory. Generally physical, physiological, psychological, and emotional components are considered as influences on human perception of comfort. Likewise, environment, user traits, and product attributes are seen as factors that can affect and manipulate how comfort is perceived.

2.1 Vehicle Comfort

The comfort of vehicles has been widely studied throughout the history of automotive research and development, dating right back to the early 20th century [11] and continuing up until to present day. As available technology has evolved the means of enhancement to the comfort of vehicle occupants has become progressively more advanced. This is evidenced by increases in available features intended to impact perception of vehicle comfort, and development in associated capability and performance. Further to these developments, have been the implementation of sensors and the increased intelligence of adaptive systems, which have increased the possibilities regarding how a vehicle occupant's experience of comfort can be influenced [12].

Vehicle comfort closely connects with general comfort definition, with several identifiable product attributes regarding vehicle design [12–15]. User characteristics depend on the physiological, psychological, and physical state of the occupants. These are associated with personal traits (e.g. personal preferences, physical characteristics) and their momentary state (e.g. state of stress, presence of pain, tiredness). User characteristics might also take into account the context and the scenario (e.g. trip urgency occupant activity, trip purpose). Environmental factors take into account elements within which the vehicle operates (e.g. weather condition, road surface, vehicle).

3 Solutions to Be Investigated

3.1 Development of Ambient Comfort Model

As described, and central to the SUaaVE project, implementation of connected and autonomous vehicles has the potential to greatly change in-car comfort perception. This is in part attributable to the potential change in characteristics that the vehicle has to respect in order to optimise comfort, and also attributed to technological advancement.

Based upon the principals behind both, there are two aspects that must be developed – first an approach by which different factors influencing comfort can be assessed and measured so as to quantify the level of comfort for users. Second will be the definition of a means to manipulate elements influencing the degree of comfort for users. This will function such that vehicle systems are able to work to improve their level of comfort. This represents a key innovation on the part SUaaVE, with this 'intelligent' response based upon understanding of the user and scenario of use central to the project.

Vehicle comfort can broadly be divided into two categories according to the engineering challenges posed – dynamic comfort, and interior (ambient) comfort. The influence of both is considered by the SUaaVE project and, whilst separate, dynamic comfort is set to follow a framework close to that defined in this paper. During later ALFRED concept integration there will be an assimilation of the two. Following definition, the comfort model will be evaluated by means of iterative testing across multiple phases. This is intended to be carried out within a simulated scenario and environment.

3.2 Components of Ambient Comfort

In response to evaluated theory and models within existing literature, a process of comfort definition took place as part of the SUaaVE project. A parallel framework for dynamic and ambient comfort was used, separated into a series of components representing overall comfort when comprised. At the component level this will allow for assessment of state, whilst forming the basis of combined assessment of comfort controlling the ALFRED response. Seven components were identified from a larger pool:

- *Thermal Environment* covers the thermal environment of the vehicle cabin, following the individual characteristics and state of each user.
- *Acoustic Environment* refers the vehicle cabin sound and vibration
- *Visual Environment* refers to occupant environment visible components
- *Postural position* refers to components of the occupants physical position when inside the vehicle cabin, following characteristics and state of users

- *Environmental Hygiene* is comprised of a variety of factors governing user sensations of cleanliness and hygiene whilst traveling within the vehicle.
- *Spatial Environment* defines level of perceived space within the occupant environment with relation to the demands of the user
- *Tactile Interaction* concerns perception whilst interacting tactile surfaces

3.3 ALFRED Integration

The foundation for inclusion the SUaaVE model, is the basis for comfort components to be influenceable by the generic vehicle features and functions included in the scope of control the ALFRED concept. Individual comfort components have been linked to those features which have a corresponding impact (Fig. 3.)

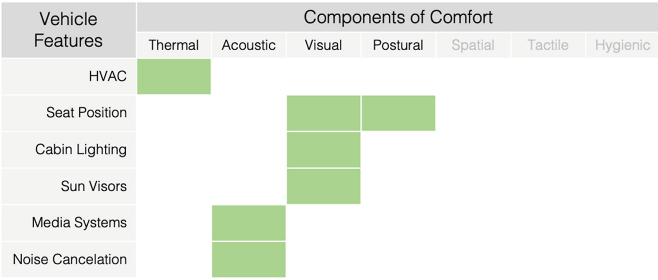


Fig. 3. Components of Comfort in ALFRED

Control of features used to manipulate the comfort of occupants will be presented through the aforementioned interface developed in conjunction with the ALFRED concept. This interface will have comfort functions driving adjustments to corresponding vehicle feature settings. The interaction strategy itself will present recommendations based upon the output of the model, requesting confirmation from the user.

3.4 Ambient Comfort Concept

It will be the role of the ALFRED concept sensor technology to determine the comfort status of occupants. The underlying assessment methodology at the components level determines this comfort state by data regarding occupant state of, their traits and preferences, situational factors, and environmental conditions. This will quantify sources of discomfort regarding the user or scenario and measure discomfort at the component level. Output will then dictate functional recommendations to the ALFRED user interface.

Referring to the aforementioned concepts, assessment is to follow a unidirectional continuum of Fig. 2. This is measured for each data source and provides an overall output. Also included in the output is determination of the degree of discomfort for each component and its theoretical route cause. With this functionality causes behind states of discomfort can be identified and counteracted through ALFRED – improving overall sensation of comfort. This through shaping conditions controllable by the vehicle, and closer to what might be regarded as 'just right' for each specific user and scenario (Fig. 4).

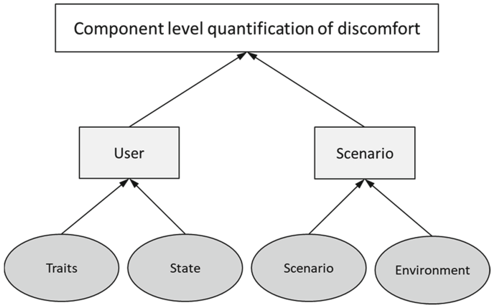


Fig. 4. Model for (dis)comfort assessment

4 Continued Development and Evaluation

An initial version of ambient comfort profiles must be built up in the next phase of the SUaaVE project. This leading phase must include investigation around each component of ambient and posture settings with relation to identified vehicle features. The aim is to identify specific feature functionality in terms of component level comfort impact. This is to be based upon existing literature and in conjunction with expert assessment – replicating individual influencing factors associated with each assessment.

The outcomes from this initial investigative phase are then planned to be tested with a sample of participants. A first stage of evaluation will follow directly from the initial phase, with separate components assessed independently. This subsequent data set will be applied to developing the ALFRED functions responding intelligently to manipulate comfort state. Subsequent testing stages will combine comfort components within progressively diverse simulation platforms and, through iterative testing phases, refine combined functions able to impact vehicle occupant comfort perception. During this phase the ALFRED ambient comfort functions will be joined with dynamic comfort modelling and assessment of emotions investigated elsewhere in the SUaaVE project.

5 Conclusion

Comfort is evidently defined under may different models and structures. There is, however, some consensus within the previous investigation and definition with theories finding their basis in the separation between states of comfort and discomfort. Regarding vehicle comfort, there is also a wide variety in approaches surrounding its investigation and development. Most apparent from comfort properties of vehicles is the continuous trajectory of development, with its scope only set to be amplified by the technology enabled as vehicles move towards autonomy and connectivity. This is also set to run in parallel with moving user expectation as models of use change.

SUaaVE sets out to develop the ALFRED concept and provide a basis for implementation of future technology, with the consideration of users at its heart. The first phases of development defining ambient comfort assessment and response in conjunction with this intelligent system are defined within this paper. Subsequent development phases objectively assess the model, with progressive stages of iterative investigation, development, and evaluation. The final aim is to present a framework assessing user comfort, with output influencing intelligent vehicle functionality.

Acknowledgments. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 814999. The content of this publication is the sole responsibility of the authors, and in no way represents the view of INEA or European Commission.

References

1. Diels, C., Erol, T., Kukova, M., Wasser, J., Cieslak, M., Payre, W., Miglani, A., Mansfield, N.J., Hodder, S.G., Bos, J.: Designing for comfort in shared and automated vehicles (SAV): a conceptual framework. Presented at the 1st International Comfort Congress (ICC2017), Salerno, Italy (2017)
2. SUaaVE European project. http://www.suaave.eu
3. Shen, W., Vértiz, A.M.: Redefining seat comfort. SAE Trans. **106**, 1066–1073 (1997)
4. Osbourne, D.J., Clarke, M.J.: The development of questionnaire surveys for the investigation of passenger comfort. Ergonomics **16**(6), 855–869 (1973)
5. Slater, K.: Human Comfort. C.C. Thomas, Springfield (1985)
6. Hertzberg, H.T.E.: Seat comfort. Annotated Bibliography of Applied Physical Anthropology in Human Engineering, WADC Technical report 56-30 (1958)
7. Branton, P.: Behaviour, body mechanics and discomfort. Ergonomics **12**(2), 316–327 (1969)
8. De Looze, M.P., Kuijt-Evers, L.F., Van Dieen, J.A.A.P.: Sitting comfort and discomfort and the relationships with objective measures. Ergonomics **46**(10), 985–997 (2003)
9. Vink, P., Hallbeck, S.: Editorial: comfort and discomfort studies demonstrate the need for a new model. Appl. Ergon. **43**, 271–276 (2012)
10. Moes, N.C.C.M.: Analysis of sitting discomfort, a review. Contemp. Ergon. **2005**, 200–204 (2005)
11. Warner, J.A.C.: Riding-qualities research. SAE Trans. 193–226 (1924)
12. Cieslak, M., Kanarachos, S., Blundell, M., Diels, C., Burnett, M., Baxendale, A.: Accurate ride comfort estimation combining accelerometer measurements, anthropometric data and neural networks. Neural Comput. Appl. 1–16 (2019)
13. Elbanhawi, M., Simic, M., Jazar, R.: In the passenger seat: investigating ride comfort measures in autonomous cars. IEEE Intell. Transp. Syst. Mag. **7**(3), 4–17 (2015)
14. Zhang, E., Zhang, Q., Xiao, J., Hou, L., Guo, T.: Acoustic comfort evaluation modeling and improvement test of a forklift based on rank score comparison and multiple linear regression. Appl. Acoust. **135**, 29–36 (2018)
15. Danca, P., Vartires, A., Dogeanu, A.: An overview of current methods for thermal comfort assessment in vehicle cabin. Energy Procedia **85**(Supplement C), 162–169 (2016)

Detection and Classification of Unconscious Movements with Body Pressure Distribution Measurement for Ride Comfort Evaluation in Vehicle Seat

Junya Tatsuno[1(✉)], Koki Suyama[2], Hitomi Nakamura[1], and Setsuo Maeda[3]

[1] Faculty of Engineering, Kindai University, 1 Takaya-Umenobe, Higashi-Hiroshima City, Hiroshima 739-2116, Japan
tatsuno@hiro.kindai.ac.jp

[2] Graduate School of Systems Engineering, Kindai University, 1 Takaya-Umenobe, Higashihiroshima, Hiroshima 739-2116, Japan

[3] School of Science and Technology, Nottingham Trent University, College Dr, Clifton, Nottingham NG11 8NS, UK

Abstract. We have focused on the unconscious movements of participants. Previously, we checked all video off-line to obtain the frequencies of participants' movement. Although we could discuss the relationship between the subjective ratings of participants and the frequencies of the unconscious movements, we recognized that the visual judgment was too heavy for us to carry out. Thus, in this paper, we installed the body pressure measurement system. First, we analyzed time-series data of participant's body pressures on the seat during some typical movements. As a result, we found the relationship between the movement types and the characteristics of the COP trajectories. Next, we carried out the driving simulator experiments. When we made a comparison between the frequency with the visual inspection of video images and those counted with the body pressure data, we found that there were some false detections with the analysis of the body pressure data.

Keywords: Ride comfort · Objective evaluation · Whole-body vibration · Driving simulator · Fidget

1 Introduction

It is important for the automotive manufacturer and supplier to evaluate human discomfort during traveling by any method. Previously, many subjective and objective methods have been utilized for ride comfort evaluation. Recently, we have focused on an objective measure of discomfort via the analysis of unconscious movements in vehicle seats [1–3].

In the previous studies, it was shown that the frequency of participants' seat fidgets and movements (SFMs) increased with the duration of driving, and the rate of participants' SFMs had a high correlation with the subjective discomfort rating [4]. On the

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 276–282, 2020.
https://doi.org/10.1007/978-3-030-50943-9_35

other hand, through these researches, we found that there were several problems to detect participants' SFMs because we video-recorded the participants' movements during all experiments and checked all video off-line. The most serious problem was that the visual judgment method is a heavy workload. Besides, when we use a single video camera in the experiments, there is a possibility that we might overlook participants' movements that occur opposite side to the camera.

The study of the SFMs is inspired by the In Chair Movement (ICM) study of Fenety et al. [5]. They utilized an interface pressure mat and analyzed participant's center of pressure (COP). Similarly, pressure distribution measurement has been applied to automotive seat evaluation in several reports [6]. Thus, in this paper, we investigated the possibility of SFMs detection with a body pressure measurement system (BPMS).

2 Detection of SFMs Using Body Pressure Measurement

Figure 1 shows the experimental environment where we placed a pressure mat on the seat of the driving simulator. In the previous research on the vehicle seat evaluation, they set two mats on the seat-surface and the seat back. We also utilized sensor sheets (BPMS, Nitta Co., Japan) on both surface and back. By using this measurement, we can obtain the body pressure distribution of 44 columns and 48 rows on each mat.

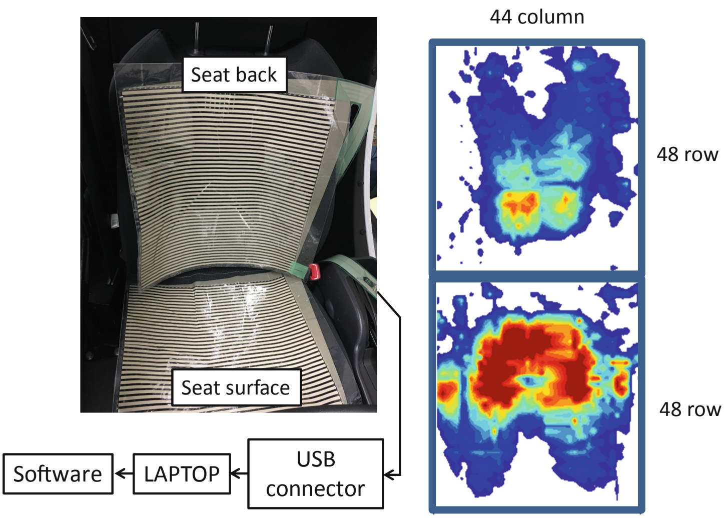
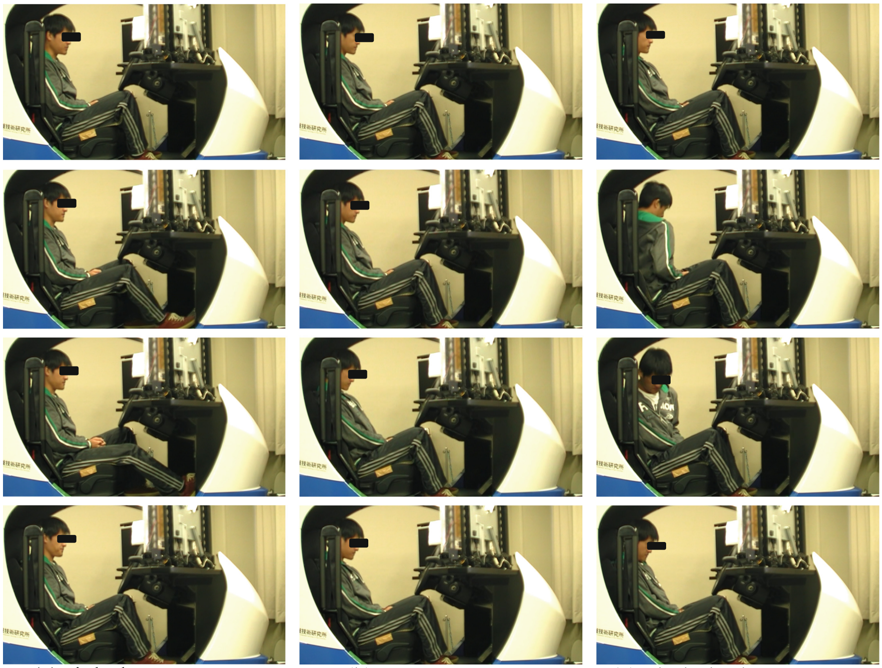


Fig. 1. Body pressure mats fixed on the seat-surface and the seat back.

In order to investigate the relationship between the participant movement types and the output of BPMS, we asked a participant to perform three typical types of movement as shown in Fig. 2. In the previous research, we focused on ride comfort evaluation and improvement of passenger or autonomous vehicle driver [7]. Since our interesting is discomfort when occupants do not drive, we gave participants reading task instead of driving task. In Fig. 2(a) and (b), the participant moved his right leg or his torso, respectively. Figure 2(c) shows the whole-body movement of the participant. When the participant performed each movement, we recorded the body pressure distribution at the sampling rate of 2 Hz.

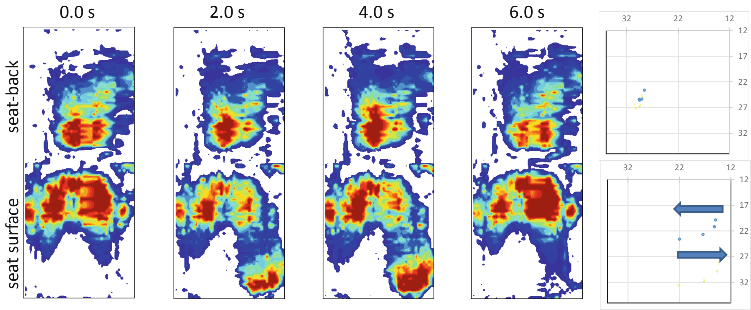


(a) right leg movement (b) torso movement (c)whole-body movement

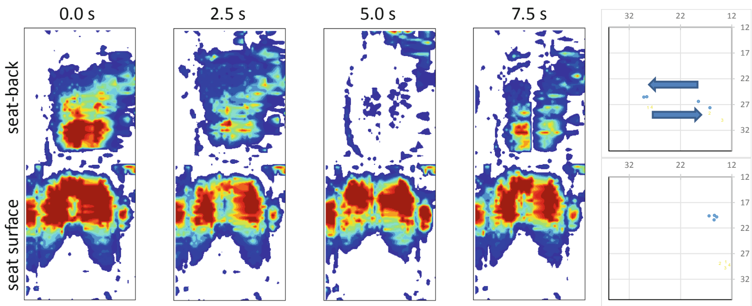
Fig. 2. Frame shots when participant moved his body on the driving simulator.

Figure 3 shows the body pressure distribution changes during each movement. From the body pressure distribution change shown in Fig. 3(a), we could understand the load acted on the left leg when he moved his right leg. When the participant stretched his back, as shown in Fig. 3(b), the body pressure distribution on the seat back changed much while the data on the surface did not vary. It was found that his back was away from the seat at five seconds. In Fig. 3(c), the participant twisted his whole-body. In this case, the body pressure distribution on both seat-surface and seat back changed.

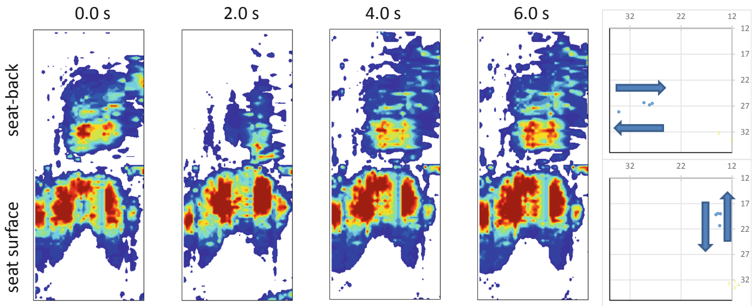
Next, to find any signature from the distribution change, we calculated the trajectories of the center of pressure (COP) from the time-series distribution data. From the trajectories of COP, as shown in Fig. 3(a), the COP on the surface moved to the lateral direction when the participant moved his leg. Meanwhile, when the participant moved his torso, the COP on the seat back moved longitudinally, as shown in Fig. 3(b). Besides, when the participant moved his whole-body, the COPs on both surface and back were shifted synchronously. Thus, when we obtain the trajectories of the COPs through analyzing the time-series body pressure distribution, we can not only detect participants' unconscious movement but also classify the movements to each type.



(a) right leg movement



(b) torso movement



(c) whole-body movement

Fig. 3. Body pressure distribution change and trajectory of COP (center of pressure).

3 Driving Simulator Experiments

3.1 Experimental Outline

The objective of the experiment is to investigate the detection and the classification of participants' unconscious movements using body pressure measurement. Figure 4 shows the configuration of the experimental apparatus. For whole-body vibration exposure to participants, we used a six DOF driving simulator (Fuji Heavy Industries, Japan). Since the participants were exposed to whole-body vibration as passengers in this experiment, the experimenter operated the driving simulator at the speed of 60 km/h with the external controller for an hour. During traveling, a reading task instead of a driving task was given to the participants. We created the test course with some bumpy sections to expose a certain magnitude of whole-body vibration. By connecting the endpoint and the initial point virtually, we could provide the semi-permanent traveling. In each driving, we recorded the body pressure distribution data and video-recorded for their behavior observation. The trial participants included five male students with driving licenses. Before the experiment, permission was obtained from the bio-ethics committee of the Faculty of Engineering, Kindai University. After each test, we noted occurrence time and movement type through the visual judgment of recorded movies and calculated the COP trajectory around the occurrence time.

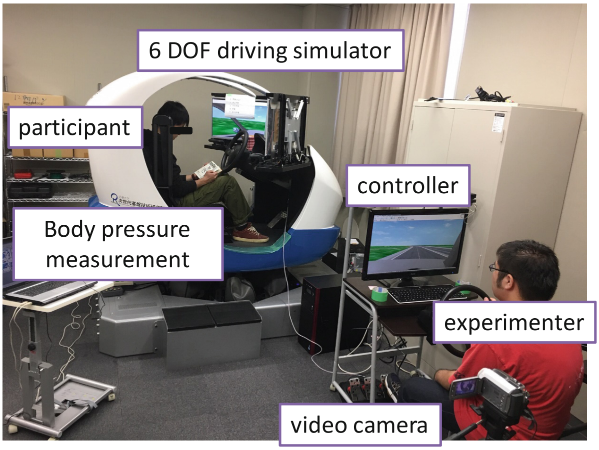


Fig. 4. Photos of experimental apparatus.

3.2 Result and Discussion

We summarized the number of the participants' unconscious movements that could be extracted with both methods. From this result, the counting numbers with the body pressure distribution were lower than the visual judgement. We could consider several causes. First, it was difficult to detect movement of lower arms or uncles with body pressure data. Although we could enumerate such movements when we observed the movies, the COP trajectories did not vary so much. We consider that we should confirm whether such small movements are caused by increasing discomfort. On the other hand, under the analysis with the COP trajectory, some movements were classified into the whole-body while they were classified into upper body or lower body with the visual judgement. We will solve this problem by adjusting threshold for analyzing the COP trajectories (Table 1).

Table 1. Number of the participants' unconscious movements with both visual judgement and the COP data.

	Upper body		Lower body		Whole-body		Total	
	Visual	Pressure	Visual	Pressure	Visual	Pressure	Visual	Pressure
P01	4	2	5	1	0	2	9	5
P02	11	9	4	3	1	1	16	13
P03	5	4	5	3	1	1	11	8
P04	0	0	9	6	0	0	9	6
P05	3	2	11	1	2	12	16	15

4 Conclusion

We have focused on the Seat Fidgets and Movement (SFMs) to evaluate human discomfort in the vehicle seat. In the previous research, we video-recorded the participant's movements and checked all video off-line to obtain the frequencies of participants' movement. We found that it was too heavy for us to detect the SFMs of participants, and there is a possibility that we overlook participant's movement in case of the visual judgment of video-recorded movies. Therefore, in this paper, we installed the BPMS. Since we found the relationship between the movement types and the characteristics of the COP trajectories, we carried out the driving simulator experiments. As a result, we found that there were some false detections with the analysis of the body pressure data. To solve this problem, we have to investigate whether small movement depends on the increasing discomfort. Besides, we need to develop a software program that can analyze the body pressure data automatically, to improve the classification ability and decrease workload for analyzing data.

References

1. Tatsuno, J., Maeda, S., Sammonds, G., Mansfield, N.: Possibility of automobile seat evaluation with seat fidgets and movements. In: Proceedings of the AHFE 2017 International Conference on Human Factors in Transportation, California, USA, pp. 474–483 (2017)
2. Tatsuno, J., Maeda, S.: Development of simple detection method of involuntary movements to evaluate human discomfort in vehicle seat. In: Proceedings of the AHFE 2018 International Conference on Human Factors in Transportation, Florida, USA, pp. 857–865 (2018)

3. Tatsuno, J., Suyama, K., Mitani, H., Nakamura, H., Maeda, S.: Experimental investigation of the relationship between human discomfort and involuntary movements in vehicle seat. In: Proceedings of the AHFE 2019 International Conference on Human Factors in Transportation, Washington D.C., USA, pp. 400–411 (2019)
4. Sammonds, G.M., Fray, M., Mansfield, N.J.: Effect of long term driving on driver dis-comfort and its relationship with seat fidgets and movements (SFMs). Appl. Ergon. **58**, 119–127 (2017)
5. Fenety, P.A., Putnam, C., Walker, J.M.: In-chair movement: validity, reliability and implications for measuring sitting discomfort. Appl. Ergon. **31**(4), 383–393 (2000)
6. Inagaki, H., Taguchi, T., Yasuda, E., Iizuka, Y.: Evaluation of riding comfort: from the viewpoint of interaction of human body and seat for static, dynamic, long time driving. SAE Trans. **109**(Section 6), 960–964 (2000)
7. Tatsuno, J., Maeda, S.: Driving simulator experiment on ride comfort improvement and low back pain prevention of autonomous car occupants. In: Proceedings of the AHFE 2016 International Conference on Human Factors in Transportation, pp. 511–523 (2016)

An Ergonomic Analysis on the New E-Traysikel

Nouriet Rocel San Juan, Kristiana Louise Abaa(✉), Daniel Jairoh Alto, Enrico Fernando Jr., and Benette Custodio

Department of Industrial Engineering and Operations Research, College of Engineering, University of the Philippines, 1101 Quezon City, Philippines
nourietsanjuan@gmail.com,
{kmabaa,bpcustodio}@up.edu.ph, dan.alto31@gmail.com,
erfernandojr@yahoo.com

Abstract. Tricycles, consisting of a motorcycle for the driver and a sidecar for passengers, are a staple of Philippine transportation. With the amount of CO2 produced by tricycles, the Department of Energy and the Asian Development Bank spearheaded the creation of the "E-Traysikels". While being environmentally friendly, it is also important for the design to be ergonomic, thus, the study aimed to assess the fit and comfort to Filipinos of the "E-Traysikel". Dimensions of the e-tricycle were compared to corresponding Filipino anthropometric dimensions. Postures of selected passengers and drivers were taken and assessed using Rapid Entire Body Assessment and Rapid Upper Limb Assessment. Perceived comfort when riding the e-tricycle were also gathered. Results showed multiple parts of the "E-Traysikel" are smaller than corresponding anthropometric measurements. REBA and RULA scores range from acceptable to medium risk. Recommendations were provided to increase the comfort of the "E-Traysikels" for the passengers and drivers.

Keywords: E-trikes · Anthropometry · Postural assessment

1 Background of the Study

One of the most reliable modes of public transportation in the Philippines is the tricycle. It consists of a motorcycle and a sidecar attached to it where passengers are seated. According to data released by the Philippine Statistics Authority in 2018, 75% of the registered motorcycles in the country were tricycles.

A study done by the Department of Energy (DOE) and the Asian Development Bank (ADB) [1] reported that tricycles produce 10 million tons of CO2 annually. With this, ADB built and tested their own version of an e-tricycle which proved successful in being eco-friendly. The success pushed the DOE to hold a contest to design the new e-trikes while incorporating certain aspects from the tested e-tricycles.

Currently, these "E-traysikels" have been more visible around Metro Manila with the DOE making bigger plans to make it a more stable unit starting 2020 [2]. Since the current model of the "E-traysikels" is still relatively new, no ergonomic analysis has been done to evaluate the vehicle.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 283–290, 2020.
https://doi.org/10.1007/978-3-030-50943-9_36

2 Objectives of the Study

With the increasing number of e-trikes in the country, and in the pursuit of new eco-friendly transportation alternatives, there is a lack of ergonomic analysis done on the latest version of the e-trikes, that would assess its fit and comfort to passengers and drivers. The study aims to provide an anthropometric, postural, and subjective analyses of the "E-Traysikels" and recommend E-trike's dimensions catered specifically to Filipinos.

3 Methodology

3.1 Anthropometric Analysis

An anthropometric analysis was done to provide insight on how the model fits the general users. The interior dimensions of the "E-Traysikel" were measured using a flexible tape measure. These measurements were compared to corresponding anthropometric data of Filipinos [3] associated with certain percentiles based on design of extremes principle. A 15% allowance [4] was added to the anthropometric data to give space for ample mobility and comfort within the vehicle (Table 1).

Table 1. E-trike dimensions with anthropometric data and assigned percentile and sex [5–9].

E-trike interior	Anthropometric data	Percentile
Passenger		
Passenger entrance height	Standing height + 15% allowance	95th male
Passenger entrance width	Shoulder width (Bideltoid) + 15% allowance	95th, male
Passenger seat height	Popliteal height	5th, female
Passenger seat 1 width (right)	Shoulder width (Bideltoid) + 15% allowance * 2	95th, male
Passenger seat 2 width (left)	Shoulder width (Bideltoid) + 15% allowance * 3	95th, male
Passenger seat depth	Buttock popliteal length	5th, female
Passenger seat backrest height	Shoulder height (sitting) + 15% allowance	95th, female
Passenger compartment height	Sitting height + 15% allowance	95th, male
Passenger compartment width	Buttock knee length + 15% allowance * 2	95th, male
Driver		
Driver seat height	Popliteal height	5th, female
Driver seat width	Sitting hip breadth + 15% allowance	95th, male
Driver seat depth	Buttock popliteal depth	5th, female
Driver seat backrest height	Shoulder height (sitting) + 15% allowance	95th, female

3.2 Postural Analysis

Postural analysis was done to determine the risk of discomfort or pain to passengers and drivers. Six passengers falling under the 5th, 50th and 95th percentile of both sexes for standing height [8] were asked to ride the e-trikes. Rapid Upper Limb Assessment (RULA) was conducted on postures of passengers while riding the vehicle, and Rapid Entire Body Assessment (REBA) was conducted on postures of passengers while entering and leaving the vehicle. The same was done for postures of drivers with the addition of using RULA for postures while driving, turning and using the handbrake. Scores and their meanings are as follows (Table 2):

Table 2. Level of musculoskeletal disorders risks [10, 11]

Score	Level of MSD Risk
Rapid Upper Limb Assessment (RULA)	
1-2	Acceptable Posture. No action required.
3-4	Further Investigation. Change may be needed.
5-6	Further Investigate. Change Soon.
7	Investigate and Implement Change.

Score	Level of MSD Risk
Rapid Entire Body Assessment (REBA)	
1	Negligible Risk. No action required.
2-3	Low Risk. Change may be needed.
4-7	Medium Risk. Further Investigate. Change Soon.
8-10	High Risk. Investigate and Implement Change.
11+	Very High Risk. Implement Change.

3.3 Subjective Analysis

Six passengers falling under the 5th, 50th and 95th percentiles of both sexes for shoulder width as well as the aforementioned 6 passengers for standing height were asked to ride the vehicle. A survey based on the Cornell Musculoskeletal Discomfort Questionnaire (CMDQ) [8] was given to both passengers and drivers before entering, and after riding or driving the vehicle to determine their perceived state of comfortability and satisfaction of the fit of features of the "E-Traysikel".

4 Results and Discussion

4.1 Results of Passenger Compartment Analysis

Anthropometric Analysis. Result of comparison of E-Trike passenger interior measurements and corresponding Filipino anthropometric measurements showed that only seat height and seat depth provide a good fit (see Table 3).

Proper headroom clearance and extra width allowance for the passenger entrance must be applied to ensure comfortable ingress and egress of passengers. Although the seat width measurements were almost exact with the anthropometric measurements of Filipinos, an allowance must be given to secure the comfortability and aid the mobility of passengers. In addition, the backrest height is too low and nonconforming to the sitting shoulder height of the 95th Filipino female. Thus, discomfort may arise.

Table 3. Comparison of e-trike interior measurements and Filipino anthropometric measurements for passenger entrance.

E-trike interior	Actual measurement (cm)	Ideal measurement based on Anthro (cm)	Conformance
Passenger entrance height	140	204.7	x
Passenger entrance width	50	56.81	x
Seat 1 width (right)	98.7	113.62	x
Seat 2 width (left)	150.5	173.08	x
Seat height	35	36	/
Seat depth	38.5	40	/
Seat backrest height	33.5	69.17	x
Passenger compartment height	132	159.85	x
Passenger compartment width	136	142.37	x

Postural Analysis. REBA and RULA scores of the passengers' postures range from acceptable to medium risk. High REBA scores of the 50th and 95th percentile participants support the earlier conclusions regarding the nonconforming entrance height and width. On the other hand, RULA scores were acceptable (see Table 4).

Table 4. RULA/REBA scores of participants

Percentile	Ingress (REBA)	Ride (RULA)	Egress (REBA)
5th male	3	2	4
50th male	7	2	7
95th male	7	2	7
5th female	3	1	3
50th female	6	2	7
95th female	7	2	6

Subjective Analysis. Some passengers experienced pain in their shoulders, upper and lower back, hips, and buttocks. The first three could be attributed to the position of the backrest as it is positioned too low while the latter can be attributed to the hardness of the seat cushion. Some passengers also felt more comfortable after riding the e-trikes. This may be due to external factors such as being seated after walking with backpacks.

Survey results showed that more than half of the passengers felt that the width of the center aisle is too small which supports the earlier findings that it doesn't conform to Filipino anthropometric data. Moreover, half of the passengers rated the height and amount of backrest, along with the height and width of the passenger entrance as too small. This supports the earlier findings regarding their inadequacy to conform to Filipino anthropometric measurements (Table 5).

Table 5. Perceived comfort of standing height and shoulder width (Bideltoid) representatives.

	Standing height representatives			Shoulder width (Bideltoid) representatives		
Percentiles	Ingress	Ride	Egress	Ingress	Ride	Egress
5th male	2	3	3	3	3	3
50th male	4	5	2	2	3	1
95th male	3	4	2	2	3	2
5th female	4	4	4	5	3	5
50th female	2	4	3	2	4	1
95th female	2	2	1	5	4	4

The perceived comfort of passengers during the ride was adequate. On the other hand, perceived comfort during ingress is relatively low and lowest during egress. This supports earlier conclusions that the passenger entrance height is too small, affecting how passengers embark and disembark.

4.2 Results of Driver Compartment Analysis

Anthropometric Analysis. Comparison of E-Trike driver interior measurements and corresponding Filipino anthropometric measurements showed that only the seat depth provides a good fit (see Table 6).

Table 6. Anthropometric conformance of E-trike interior dimensions for driver seat area.

E-trike interior	Actual measurement (cm)	Ideal measurement based on anthropometry (cm)	Conformance
Seat width	40	47.15	x
Seat height	42.5	36	x
Seat depth	40	40	/
Seat backrest height	38.5	69.17	x

Postural Analysis. Despite the low RULA scores of drivers, the drivers had to fully extend their arms while making a turn. The relatively high REBA scores arose from the analysis of the neck, trunk, and legs since the drivers had to bend all 3 body parts when alighting the e-trikes due to the curved design of the entrance (Table 7).

Table 7. RULA/REBA scores of drivers.

	Driver 1	Driver 2
	Score	Score
Ingress (REBA)	6	4
Driving (RULA)	3	3
Turning (RULA)	3	3
Handbrake (RULA)	2	2
Egress (REBA)	4	7

Subjective Analysis. During the interview with the drivers, several concerns came up regarding the e-trikes (see Table 8).

Table 8. Driver concerns regarding certain E-trike interiors.

E-trike interior	Drivers' concerns
Backrest	• Supports only a small portion of their backs; leaves other parts of the back in awkward positions • Foam is too hard; uncomfortable during long rides
Seatbelt	• Too short and tight; usually unused • Seatbelt design in cars was preferred
Handlebars	• Too far; has minimal adjustable range • Drivers have to fully extend arms to reach handlebars
Handbrake	• Too far; fully extended arms are necessary for use • Positioned on the left side; should be positioned on the right–the standard in Philippine vehicles
Foot pedal	• Drivers have to move forward to reach foot pedal

5 Recommendations

Many parts of the e-trike do not conform to Filipino anthropometric measurements and are uncomfortable for the use of most Filipinos, especially for drivers. The researchers recommend that nonconforming E-trike dimensions should be adjusted accordingly with the recommended measurements shown in Table 9. In addition, a full backrest for the passenger and driver, as well as a headrest for drivers, should be provided to give better support, address discomfort and provide safety. Lastly, it is recommended that the driver seat be adjustable.

Table 9. Recommended measurements of the E-trike interior features.

E-Trike Interior	Recommended Measurement (cm)
Passenger Area	
Entrance height	204.7
Entrance width	56.81
Seat 1 width (right)	113.62
Seat 2 width (left)	173.08
Seat backrest height	69.17
Compartment height	159.85
Compartment width	142.37

E-Trike Interior	Recommended Measurement (cm)
Driver Area	
Seat width	47.15
Seat height	36
Seat backrest height	69.17

References

1. Hasnie, S.: Republic of the Philippines: Market Transformation through Introduction of Energy Efficient Electric Tricycles Project. Asian Development Bank, pp. 1–31 (2012). https://www.adb.org/sites/default/files/project-document/74849/43207-013-phi-pam.pdf. Accessed 13 Nov 2019
2. "Electric Tricycles" Department of Energy. https://www.doe.gov.ph/alternative-fuels/electric-tricycles-0. Accessed 13 Nov 2019
3. Del Prado-Lu, J.: Anthropometric measurement of Filipino manufacturing workers. Int. J. Ind. Ergon. **37**, 497–503 (2007). https://doi.org/10.1016/j.ergon.2007.02.004
4. Ismaila, O.S., Musa, A.I., Adejuyigbe, S.B., Akinyemi, O.D.: Anthropometric design of furniture for USE in tertiary institutions. Eng. Rev. **33**(3), 179–192 (2013). https://www.researchgate.net/publication/286217730_Anthropometric_design_of_furniture_for_USE_in_tertiary_institutions_in_Abeokuta_South-Western_Nigeria. Accessed 30 Nov 2019
5. Dawal, S., Ismail, Z., Yusuf, K.: Determination of the significant anthropometry dimensions for user-friendly designs of domestic furniture and appliances - experience from a study in Malaysia. J. Int. Meas. Confed. (IMEKO) **59**, 205–215 (2015). https://www.sciencedirect.com.ezproxy.upd.edu.ph/science/article/pii/S0263224114004102#t0015. Accessed 30 Nov 2019
6. Chuan, K., Hartono, M., Kumar, N., Chuan, T.: Anthropometry of the Singaporean and Indonesian population. Int. J. Ind. Ergon. (2010). https://www.academia.edu/21009989/Anthropometry_of_the_Singaporean_and_Indonesian_populations
7. Gutierrez, J., Perez, E., Quilantang, L., Custodio, B.: Ergonomic design and assessment of footbridges for Filipino children and elderly. Jpn. J. Ergon. **53**(Supplement) (2017). https://www.jstage.jst.go.jp/article/jje/53/Supplement2/53_S636/_pdf/-char/ja. Accessed 1 Dec 2019
8. Godilano, E.C., Villanueva, R.C., Ang, W.Y., Anonuevo, A.B., Insorio, B.C., Morita, M.V., Padios, M.A.C., Realce, H.R.G., Diego, J.C.S.: Risk assessment towards innovative ergonomic design of a tricycle. In: 2015 International Conference on Industrial Engineering and Operations Management (IEOM) (2015). https://ieeexplore.ieee.org/document/7093704. Accessed 13 Nov 2019
9. Molenbroek, J.F.M., Albin, T.J., Vink, P.: Thirty years of anthropometric changes relevant to the width and depth of transportation seating spaces, present and future. Appl. Ergon. **65**, 130–138 (2017). https://www.sciencedirect.com/science/article/abs/pii/S0003687017301308. Accessed 30 Nov 2019

10. Hignett, S., Mcatamney, L.: Rapid entire body assessment (REBA). Appl. Ergon. **31**(2), 201–205 (2000)
11. Mcatamney, L., Corlett, E.N.: RULA: a survey method for the investigation of work-related upper limb disorders. Appl. Ergon. **24**(2), 91–99 (1993)

Investigation on Driving Posture and Operating Habits of Heavy Truck Drivers

Junmin Du[1], Haoshu Gu[1], Weiyu Sun[2], Xin Zhang[3](✉), Huimin Hu[3], and Yang Liu[4]

[1] School of Transportation Science and Engineering, Beihang University, Beijing 102206, China
{dujm, haoshu_gu}@buaa.edu.cn
[2] Hefei Hualing Co., Ltd., Hefei 230001, China
wysun@buaa.edu.cn
[3] SAMR Key Laboratory of Human Factor and Ergonomics, China National Institute of Standardization, Beijing 100191, China
{zhangx, huhm}@cnis.ac.cn
[4] Beijing Foton Daimler Automotive Co., Ltd., Beijing 101400, China

Abstract. Ergonomic design of heavy truck cabs has a significant impact on road safety and driver health. In the design phase of the cab, a full understanding of the driver's driving posture and operating habits is the necessary foundation for the human-machine interface design and the requisite guarantee of a good driving experience. However, at present, in the design of heavy truck cabs in China, insufficient consideration is given to the driving habits of drivers, resulting in poor ergonomics and user experience complaint. Through interviews and questionnaire surveys with heavy truck drivers with the static method of significance testing, this article investigated and analyzed the characteristics of professional truck drivers' driving postures and the frequency of operating actions which provides improvements suggestion to the ergonomic design and the driving experience of the truck.

Keywords: Heavy truck driver · Driving habit · Investigation · Cab ergonomics design · Human-machine interface

1 Introduction

The ergonomics of truck cab has an important impact on road safety and driver's health, especially for heavy truck, which is the main tool for long distance road transportation and usually result in more serious consequences than other types of vehicle if an accident occurs. Ergonomic design of the cab is the basis for ensuring the driving experience, driver health, and road safety. Excellent ergonomic design can reduce driver fatigue and maloperation rate, improve drivers' attention and his driving comfort. Based on this, the ergonomic design of truck cabs is receiving increasing attention. Due to the differences in the human body size and driving habits of people from different countries, there are different requirements for the size, layout, and operating power of the various components inside the cab. However, in the current

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 291–298, 2020.
https://doi.org/10.1007/978-3-030-50943-9_37

domestic (China) heavy truck cab design, these differences have not been fully recognized or understood, leading to health problems and hidden dangers to road safety. For example, Due to the lack of data on the driver's driving posture, truck seat designs are generally poor, which makes truck drivers quickly tired and uncomfortable on the road. Spinal diseases such as neck and back pain have a high incidence among professional drivers [1–3]. Fatigue on the feet or other parts of the body is also common [4]. After a long period of accumulation of pain or fatigue, it may turn into chronic occupational musculoskeletal injury. These problems that cause health damage may even damage the driver's control of the vehicle when driving long distances, and bring serious and irreparable consequences. The death rate and injury rate per vehicle for heavy goods vehicles are much higher than for all types of vehicles [5]. Therefore, the understanding of the characteristics of the heavy truck driver group and the design of the human-machine interface of the cab should attract more attention from designers. Full consideration of driver characteristics and driving habits in the cab design process, as the basis of the cab man-machine interface design, will help improve the ergonomics of the cab man-machine interface, and enhance driving safety and driving experience.

Based on the above considerations, through interviews and questionnaires, this article obtained the characteristics of the user group and driving habits of Chinese heavy truck drivers. The research results are helpful to understand the characteristics and the needs of Chinese heavy truck driver groups, and provide a reference for the design of the human-machine interface of the heavy truck cab.

2 Method

This study first interviewed 9 heavy truck drivers, and installed a driving recorder on the trucks of 6 of them to record their driving behavior for one day. Based on interviews and driving records, a questionnaire design was developed. Afterwards, in a logistics company in Beijing Daxing District and a heavy truck service area in Changping District, a quantitative questionnaire survey was conducted using on-site questionnaires, and a total of 100 valid quantitative questionnaires were recovered. Drivers participating in interviews and questionnaires must be professional heavy truck drivers with at least one year of driving experience in heavy trucks, and the vehicle weight must be at least 8 tons. Before interviews and questionnaire, the purpose of the study was explained to the drivers to obtain their consent. The data analysis in this article used data from 100 questionnaires.

The main questions in the questionnaire were as follows.

1. Basic information of truck driver: including gender, age, height, weight, and weight of truck.
2. Driving habits: including driving sitting habits, steering wheel gripping posture, in-vehicle devices (seat, steering wheel, rear-view mirror, switches/buttons, foot pedals) adjustment frequency or using frequency.

3 Results

3.1 The Basic Information of Respondents

The respondents of 100 valid questionnaires were all males. It also showed that among the participants of truck drivers, males had an absolute advantage in number. The age of the interviewed drivers was 34.6 ± 7.6 years, with young adults around 30 years old the most. Their response speed, strength, and bearing capacity were at the best level. As people age or decrease, the number of people involved in the cargo driver profession decreased.

The driver interviewed was 172.14 ± 4.86 cm tall. According to GB10000-88 "Chinese Adult Human Body Size", adult males aged 18–60 have a P50 height of 167.8 cm and a P90 of 175.4 cm [6]. Considering that the size of Chinese adults has increased in recent years, it can be judged that the height level of the respondents was medium to high. The driver weighed 77.32 ± 9.22 kg. According to GB10000-88 "Chinese Adult Human Body Size", adult males aged 18–60 have a P95 weight of 75 kg and a P99 of 83 kg [6]. Considering that the body weight of Chinese adults has increased in recent years, it can be roughly judged that the respondent's weight level was medium to heavy.

Of the 100 respondents, those who drove during the day accounted for 34% of the total, 23% drove at night, and 43% drove during the day and night. The majority (88%) drove daily, 6% drove several times a week, 4% drove several times a month, and 2% drove several times a year. The mileage of more than 45,000 km in the past year accounted for nearly one-third, and the vast majority (99%) of the respondents had driven more than 5000 km in the past year, and only one person had less than 5000 km. It can be seen that most of the interviewees had high driving frequency and long mileage, and their views would be able to reflect the behavior of the truck driver and the problems of the truck itself.

3.2 Sitting Posture

The sitting posture during driving is shown in Table 1. It can be seen that the relative positions of the respondent's body and the seat during driving were not the same. 43% of the respondents had their backs completely resting on the back of the seat (A) while driving, 9% had the exact opposite, the back of the body was almost completely separated from the seat (D); another 28% of the respondents said The back rested only partially on the back of the chair (B), and 20% of the respondents only leaned their waist on the back of the chair (C). Significant tests of other parameters (age, vehicle load, and length of a single drive) of the four postures showed no significant differences ($p > 0.05$).

Table 1. Relative position of body and seat

Relative position of body and seat	A. The back rested completely on the seat back	B. The back rested partly on the seat back	C. Only the waist leaned on the back	D. The back of the body was almost completely separated from the seat
Schematic diagram				
Respondent number	43	28	20	9

3.3 Steering Wheel Gripping Mode

See Fig. 1 for common steering wheel gripping posture mode. There were more than one gripping mode commonly used by individual respondents. It can be seen that the most commonly used gripping posture was to hold with both hands, with both hands at the same height (selected by 78 respondents); followed by left-handed single-handed holding, with the right hand suspended or placed on the gear lever (22 (interviewees choose); followed by holding with both hands, the left hand was in a higher position (12 respondents choose).

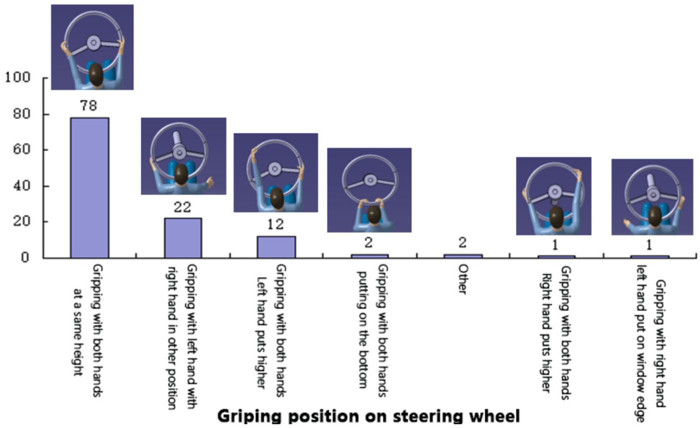


Fig. 1. Steering wheel gripping posture mode

Significant tests of other parameters (age, vehicle load, and length of a single drive) of the steering wheel gripping mode showed no significant differences ($p > 0.05$).

It can be seen that age, vehicle type, and driving time had no significant effect on the sitting posture and steering wheel gripping style, and the driver's performance was more a habit.

3.4 Adjustment or Using Frequency of In-Cab Device

Seat, Steering Wheel and Mirror Adjustment. When asked whether the respondent would adjust the seat position according to different road conditions (such as highways, urban and rural roads), 73% answered no and 27% answered yes. Respondents who adjusted the seat position according to road conditions indicated that in terms of height adjustment, they were accustomed to raising the seat during high-speed road conditions and lowering the seat during low-speed road conditions. In terms of front-to-back adjustment, they were accustomed to tilting the backward the seat back properly at high speeds, which made them feel more comfortable; when the roads were complex in urban areas, the seats were fine-tuned forward to make them fit and keep tension.

Regarding the adjustment of the steering wheel (Table 2), more than 70% of the respondents indicated that they would hardly adjust the steering wheel. Nearly 20% said that they would adjust it several times a year, and a few said that they would adjust it several times a month or adjust once every 1–2 days.

Table 2. Adjustment frequency of steering wheel

Adjustment frequency	Rarely	Several times a year	Several times a month	Once every 1–2 days
Respondent number	72	18	6	2

Regarding the adjustment of the rear-view mirror (Table 3), nearly one-third of the respondents said that they often adjusted the rear-view mirror outside or inside the truck. The rest of the respondents said that they adjusted the rear-view mirror rarely.

Table 3. Adjustment frequency of the rear-view mirror

Rear-view mirror	Rarely adjust	Always adjust
In-car mirror	67	13
Out-car mirror	–	18

Note: The number in the table is the respondent number.

The Using Frequency of In-Cab Switches/Buttons. The frequency of using the system switch/button in the driving state and the non-driving state is shown in Table 4.

It can be seen that in the driving state, the switches/adjustment buttons of the entertainment system, air conditioning system, and lighting were used frequently. Among them, the proportion of respondents who often used air conditioning had reached nearly 80%. The proportion of respondents who often used entertainment components and lighting components also reached about 50% each. Except for the respondents who often used these three systems, most of the others released they would occasionally use them, and only few respondents said that they rarely used these systems while driving.

Table 4. The using frequency of in-cab switches/buttons

System switch/button in the cab		Often used	Seldom used	Rarely used
In driving	Entertainment system	50	39	11
	Air conditioner	76	21	2
	Lights	52	41	7
Non-driving	Entertainment system	25	54	19
	Air conditioner	23	18	55
	Lights	34	35	28

Note: The number in the table is the respondent number.

Compared with the driving state, in the non-driving state, the proportion of the respondents who frequently used these three systems decreased significantly, to about 30%. The proportion of respondents who occasionally use these three systems varies from system to system. The proportion of people who occasionally used entertainment was the highest (55%), followed by occasional used of lighting (35%), and finally the occasional used of air conditioning (20%). The percentage of respondents who rarely used air condition in non-driving state was almost 60%. 30% and 20% of the respondents rarely used lighting and entertainment system.

By t test, the frequency of use of these three systems had no significant relationship with the driving time period (day, night) and age of the respondents.

The Using Frequency of Foot Pedals. A total of 80 respondents ranked the using frequency of the three pedals: the accelerator, brake, and clutch (the other 20 respondents skipped this question). The results are shown in Table 5. Among them, 61 respondents ranked the accelerator pedal first (the most frequently used), 1 ranked the brake pedal first, and 18 ranked the clutch pedal first. 12 respondents had the accelerator pedal ranked second, 31 had the brake pedal ranked second, and 37 had the clutch pedal ranked second. The ranking results were converted into 2 points in the first place, 1 point in the second place, and 0 points in the third place. After calculation, the accelerator pedal had 134 points, the brake pedal had 33 points, and the clutch pedal had 73 points.

It can be seen that compared with the three pedals, most of the respondents (strong 3/4) used the accelerator pedal most frequently, and the remaining respondents used the clutch pedal most frequently. In terms of the pedals ranked third, more than half of the respondents (48 people) used the brakes least frequently, and nearly half of the respondents (35 people) used the clutches least frequently.

Table 5. The using frequency of foot pedals

	Ranking (according to the using frequency)			Points
	First place (used most frequently)	Second place	Third place	
Accelerator pedal	61	12	7	134
Brake pedal	1	31	48	33
Clutch pedal	18	37	35	73

Table 6 shows the foot feeling evaluation after a day of driving. It can be seen that of the 98 respondents who made the evaluation, only 17 said they had no discomfort in their feet after a day of driving, and the remaining 81 respondents all expressed discomfort. This discomfort was manifested on the right foot (66 people) for the vast majority of the respondents. If add sore aches on, the number would be higher.

Table 6. Foot feelings after a day of driving

	Both feet were tired						
	Left foot was tired	Right foot was tired	Left foot was more tired	Right foot was more tired	Feel sore aches	No discomfort	Total
Respondents number	5	57	2	9	8	17	98

4 Conclusion

According to the survey results, for siting posture, 43% of the participants rested their backs fully to the seat back, 28% of them leaned their backs partially on the seat back, 20% of them leaned only their waist on the seat back, 9% of them separated their bodies almost completely from the seat back. For the way to grip the steering wheel, most of them were holding with two hands and the two hands were at the same height, followed by left hand holding, right hand hanging or resting on the gear lever. Total 73% of the participants would not adjust their seats on different road, while the rest of them (27%) would adjust the seat position according to the road conditions. Most of the participants did not adjust the steering wheel. Nearly one third of them often adjusted the rear-view mirror outside or inside the truck. In driving status, the entertainment system, air conditioning system and lighting light were used more frequently than in non-driving status. For foot pedal, the order of using frequency from high to low was accelerator pedal, clutch pedal, and brake pedal.

When designing the cab of a heavy truck, considering these driving habits of the drivers, combined with the individual characteristics and needs of the driver, will help to improve the ergonomics of the cab, as well as benefit the driving experience and road safety.

Acknowledgements. This research was supported by 2017NQI project (2017YFF0206601), Project of the President Foundation of China National Institute of Standardization (project number 522018Y-5984), and National Natural Science Foundation of China (project number 71601007).

References

1. Massaccesi, M., Pagnotta, A., Soccetti, A., Masali, M., Masiero, C., Greco, F.: Investigation of work-related disorders in truck drivers using RULA method. Appl. Ergon. **34**(4), 303–307 (2003)
2. Okunribido, O.O., Magnusson, M., Pope, M.: Delivery drivers and low-back pain: a study of the exposures to posture demands, manual materials handling and whole-body vibration. Int. J. Ind. Ergon. **36**(3), 265–273 (2006)
3. Okunribido, O.O.: An investigation of posture and manual materials handling as risk factors for low back pain in delivery drivers. Ergon. Soc. S. Afr. **28**(2), 19–27 (2016)
4. Du, J., Lu, H., Sun, W., Zhang, X., Hu, H., Liu, Y.: Investigation on driving habits of Chinese truck driver. HCI 2017, vol. 713, pp. 526–531 (2017)
5. Mayhew, D.R., Simpson, H.M., Beirness, D.J.: Report: Heavy Trucks and Road Crashes. Traffic Injury Research Foundation, Ottawa, Ontario, Canada (2004)
6. GB10000-88. Chinese Adults Human Body Dimensions, State bureau of technical supervision, China (1988)

Research on the Comfortable Joint Angle for Chinese Automobile Drivers

Linghua Ran[1(✉)], Yang Gao[2], Weinan Ju[3], Chaoyi Zhao[1], and He Zhao[1]

[1] China National Institute of Standardization, Beijing, China
ranlh@cnis.ac.cn
[2] Capital University of Economics and Business, Beijing, China
[3] CATARC Software Testing (Tianjin) Co., Ltd., CASTC, Tianjin, China

Abstract. This study used the Vicon motion capture system to record the spatial positions of the points and to obtain the range of five joint angles in comfortable driving positions for Chinese men and women. The data could provide basic reference data for the establishment of the human body template to do the ergonomic verification of Chinese automobile cab.

Keywords: Automobile · Driving · Joint angle

1 Introduction

Driving comfort is an important consideration for consumers to choose and buy automobiles. Good man - machine design of the cab will greatly improve the comfort and safety of the vehicle. The sitting comfort of the driver is an important factor in the process of driving.

For the comfortable driving posture, some scholars abroad have conducted in-depth research on it. For the comfortable driving posture, some scholars abroad have conducted in-depth research on it. After in-depth analysis of the driver's operating characteristics, Rebiffe [1] used simulation methods to establish a biomechanical human body model to simulate the actual position and driving posture of the driver while driving, and calculated the relevant driving joint angle through theoretical calculations. On the basis of Rebiffe, Grandjean [2] kept the head and feet in a similar position, and calculated the comfortable angle in the driving posture of each joint. Different from the theoretical calculations of the above scholars, Porter and Gyi [3] conducted the actual test to acquire the most comfortable driving posture and compare with the previous data. After the research and analysis of the above scholars, Se Jin Park et al. [4] Measured the comfortable driving postures of 43 healthy humans on the driving platform of the driving posture monitoring system developed by them, obtained relevant measurement characteristics, and obtained the data with those of previous scholars. The data of driving comfort angle are comparatively analyzed. There are not many studies on joint angles in comfortable driving posture in China. Ma Jia [5] measured 20 subjects on two models of PASSAT and POLO, and took photos and records. The photos were imported into the CAD system for analysis. The driving

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 299–303, 2020.
https://doi.org/10.1007/978-3-030-50943-9_38

comfort angle range and related relationship of the experimental models were obtained, which provided relevant information for driving comfort design in accordance with. However, due to the specific model, universality is lacking.

This study obtained the joint Angle range of Chinese men and women in the comfortable driving posture through experiments on the comfortable driving posture with people in different height percentiles, and provided basic reference data for the establishment of the human body template to do the ergonomic verification of Chinese automobile cab and the formulation of relevant standards of Chinese human body template.

2 Test Device

This experiment was conducted on a driving simulator. Six typical automobile parameters were selected: the angle of steering wheel α; the pedal Angle drive β; H30, representing the height from human body point H to the vertical direction of the plane at the bottom of the bridge; L11 represents the distance from the center point of the steering wheel to the horizontal direction of the heel point, and L6 represents the horizontal distance from the center point of the steering wheel to the tread point, H17 represents the height from the center point of the steering wheel to the bottom plane of the bridge. According to the combination of the six parameters, a total of 8 set of experiments were set up, as shown in Table 1.

Table 1. Eight sets of experiment

H30 (cm)	L11 (cm)	L6 (cm)	H17 (cm)	Angle α (°)	Angle β (°)
180	534	600	578	23	65
	484	550			
270	507	600	646	25	6–2
	457	550			
	407	500			
360	425	550	715	27	51
	375	500			
	325	450			

The vertical and horizontal directions of the simulated cab seats, as well as the tilt angle of the seat surface and back can be adjusted according to the driver's subjective feelings until the driver reaches the most comfortable driving state.

3 Experimental Methods

VICON motion capture system is used to collect 3D coordinate data of human body surface and calculate the required joint angle.

Each subjects were labeled with seven markers. Three were labeled at the upper extremity: acromion point, external superior ankle of humerus, ulnar cauline process, and the other four were labeled at the lower extremity: great trochanter, lateral ankle of femur, lateral malleolus, and distal end of little finger. According to these seven markers, five human joint angles under driving condition can be measured, as shown in Table 2.

Table 2. Five human joint angles

Number	Angle name	Angel define
1	The elbow joint angle	The angle between the upper arm and the forearm
2	The trunk angle	The angle between trunk and forearm
3	The angle between trunk and thigh	The angle between upper limb and thigh
4	The knee joint angle	The angle between the thigh and the lower leg
5	The ankle angle	The included angle between the lower leg and the foot

4 Subject Recruitment

In this experiment, a total of 30 subjects were recruited, including 15 male and 15 female subjects. The subjects were screened according to the percentiles of height. According to the percentiles of height given in GB/T 10000-1988 “Chinese adult body size”, 2 male and female subjects of P5, P50 and P95 were selected, and 4–5 male and female subjects of p5–p50 and p50–p95 were selected. The height of the subjects was covered from p5 to p95. And at the same time, the recruited participants are required to have a driving age of more than 1 year, and often travel by car in daily life.

5 Experimental Process

First, the subjects filled in the basic personal information and signed the informed consent. The basic size of the body was measured, including height, weight, sitting height, thigh length, leg length, upper arm length, and forearm length and leg height with foot. VICON motion capture system was calibrated. 7 reflective balls were labeled as bone markers on each subject.

Then the first set of experiments began: the positions of all control parts of the simulated bridge were adjusted to the initial state; the subjects entered the simulated bridge and adjust the simulated bridge continuously until it reached the most comfortable driving state. VICON motion capture system was used to collect the marker balls fixed on the subjects’ bodies to obtain the three-dimensional coordinates.

At the end of the experiment, the subjects rested for 5 min and moved on to the next experiment. A total of 8 experiments were conducted for each subject.

6 Experimental Results and Analysis

6.1 Data Statistics and Analysis

VICON's processing software was used to derive the 3D coordinates of each marker point and calculate the joint angle value of each subject. Ideally, the comfort values of each parameter have the characteristics of normal distribution, that is, the population tending to the mean distribution is dense, while the samples far away from the mean distribution are relatively few. Therefore, in this study, the p5 and p95 percentiles of experimental results were taken as the upper and lower limits to give the recommended range to meet the needs of the majority of the population. The range of joint comfort angles for male subjects, female subjects and the general population is as shown in Table 3.

Table 3. The range of joint comfort angles

Gender	Angle of elbow (°)	Angle of trunk (°)	Angle of trunk and thigh (°)	Angle of knee (°)	Angle of ankle (°)
Male	86–132	12–43	82–99	103–125	94–123
Female	72–112	10–31	86–109	107–132	93–117
All	77–126	11–42	83–104	104–131	93–122

Through the non-parametric test ($P < 0.05$), there were significant differences between men and women in the five parameters. For men, except for the trunk angle, the other four parameters varied among people of different heights. For women, except for the difference in knee angle and ankle angle between different height groups, there was no significant difference in the other three parameters between different height groups.

6.2 Comparison with SAE Data

According to the SAE standard [6], the comfort range between the trunk and the thigh is 95°–115°. The test results in this study are about 10° smaller than the range in SAE, which indicates that Chinese people drive with a slightly straight back. The knee Angle of SAE is 100°–145°, the upper limit is 14° higher than the angle in this experiment. The Angle of the SAE ankle is 87°–110°, and the Angle of the upper and lower limits of the range is 6° and 12° less than that of this study, indicating that the Chinese have a larger angle of ankle bending when driving.

By comparing with the comfortable joint movement angle in the SAE standard, we can see that when the Chinese are in a comfortable driving position, the back is relatively straight, the knee angle is relatively small, and the ankle angle is relatively large. When carrying out the internal space layout of domestic cars, the driving comfort angle of the Chinese should be considered.

7 Conclusions

In this study, the driving comfort measurement experiment was used to measure the joint angles of subjects of different heights on the simulated driving platform. Using pre-labeled points, the Vicon motion capture system was used to record the spatial positions of the points and obtain the range of joint angles in comfortable driving positions for Chinese men and women. The research results could provide basic reference data for ergonomic design and verification of Chinese car cabs.

Acknowledgments. This work is supported by China National Institute of Standardization through the "special funds for the basic R&D undertakings by welfare research institutions" (project number 522018Y-5984) and the National Key Technology R&D Program (2017YFF0206602).

References

1. Rebiffe, R.: An ergonomic study of the arrangement of the driving position in motor cars. In: Journal of the Institute of Mechanical Engineers Symposium (1966)
2. Grandjean, E.: Sitting posture of car drivers form the point of view of ergonomics. In: Oborne, D., Levis, J.A. (eds.) Human Factors in Transport Research, pp. 205–213. Taylor & Francis, Milton Park (1980)
3. Porter, J.M., Gyi, D.E.: Exploring the optimum posture for driver comfort. Int. J. Veh. Des. **19**(1), 255–266 (1998)
4. Park, S.J., Kim, C.-B., Kim, C.J., Lee, J.W.: Comfortable driving postures for Koreans. Int. J. Ind. Ergon. **26**, 489–497 (2000)
5. Ma, M., Fan, R., Ying, R., Song, L.P., Feifei, L., Qiang, S.: The test of comfortable driving posture and the fuzzy evaluation of the comfort. Ind. Eng. Manag. (4), 121–125 (2008)
6. SAE J826-2008 Devices for Use in Defining and Measuring Vehicle Seating Accommodation

Vulnerable Road Users

Effects on Driver's Yielding Behavior of a Pedestrian Collision Warning System in Different Road Environments

Francesco Bella(✉), Chiara Ferrante, Manuel Silvestri, and Maria Rosaria De Blasiis

Department of Engineering, University of Roma TRE, via Vito Volterra n. 62, 00146 Rome, Italy
francesco.bella@uniroma3.it

Abstract. Three different road environments (urban, suburban and rural) were implemented in a fixed – base driving simulator. Forty – five participants drove the three road environments in which two zebra crossing with a pedestrian that crosses the road were implemented: in one case, the driver was helped by the Pedestrian Collision Warning System (PCWS) which provided a combined visual-auditive warning, while, in the other pedestrian crossing, the warning system was absent. The recorded interaction conditions between driver and pedestrian during the tests were classified in three groups according to Time To Zebra arrive $\left(\mathrm{TTZ}^{*}_{\mathrm{arr}}\right)$. Results showed positive effects of PCWS that induced the driver to reach lower speeds in presence of that system, improving the general risk conditions at which the pedestrian is exposed during the interaction with drivers at crossroads.

Keywords: Pedestrians · Driver's behavior · Pedestrian Collision Warning System · Driving simulator

1 Introduction

Walking, as alternative transport mode, has great potential benefits in terms of physiological and psychological health benefits and its simplicity, associated with little cost, makes it economically accessible and, thus, one of the best ways to improve the life quality [1]. However, all – around the world, the Vulnerable Road Users (VRUs) are not always protected at all. They often have to face the lack of pedestrian paths suitable for safe transit. The global statistical data show that every year more than 270,000 pedestrians lose their life in road related crashes, which represents the 22% of all the road deaths [2]. In addition, in the decade 2006–2015, the accident data provided by the European Road Safety Observatory (ERSO) show an increasing trend of the accident rate of pedestrian fatalities on the total road deaths [3].

Several studies have been carried out with the aim of increasing pedestrians' safety, analyzing the effects of pedestrian collision warning system (PCWS) to alert the driver of the pedestrian presence. Such systems have an enormous potential in prevent and reducing vehicle – pedestrian crashes because of the timely warning about possible

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 307–314, 2020.
https://doi.org/10.1007/978-3-030-50943-9_39

collisions, which can also increases the yielding compliance towards the pedestrians and decreasing the probability of collision.

Almost all the literature researches are referred to the urban areas, where the interactions between driver and pedestrian are recurring and the likelihood of crashes are higher. However, the problem of crashes involving pedestrians on rural areas should be considered too, due to the higher speed collision between vehicle and pedestrian and, thus, to the higher probability of fatality for the pedestrians [4, 5]. In these areas, in fact, the fatality rates (number of pedestrian fatalities every 100 crashes involving pedestrians) are almost six times higher than the rates for urban roads [6].

The aim of this study is evaluating the effectiveness of a pedestrian collision warning system on the driver's yielding behavior under different road environments and different conditions of driver – pedestrian interaction. With this aim, three road environments (urban, sub-urban and rural) were implemented in a driving simulator. The present study is part of a wider research program whose some preliminary results are reported in [7], while, in the following, the outcomes of the statistical analysis on the variables which describe the driving behavior are analyzed and discussed.

2 Methodology

2.1 Equipment

The present study was conducted by using the fixed-base driving simulator located in the simulation laboratory of the Department of Engineering in the University "Roma Tre". The driving simulator has been validated [8] and successfully used in past studies [9–12], which demonstrated its reliability for the study of the driver's behavior. The main parts of the simulation system consist in a) a real vehicle, on which the hardware interfaces (pedals, wheels, etc.) are installed and b) on three projection screens which form a 135° field of view and on which the road scenarios are projected. During the simulation, the system records many dynamic parameters (such as the vehicle's speed, the acceleration values, etc.) which describe the driver's behavior.

2.2 Road Simulated Scenarios

Three road scenarios were designed and implemented in the driving simulator to simulate the driving condition in the urban, sub-urban and rural environments (Fig. 1. a–c). All the cross-sections, typical of Italian road, were designed according to Italian Highway Code [13].

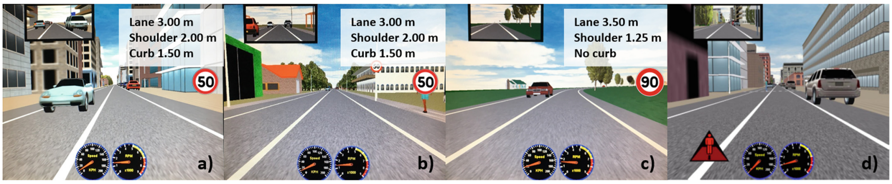


Fig. 1. Road scenarios: a) urban; b) suburban; c) rural d) visual warning of PCWS

In urban and suburban scenarios, the stripes of the zebra crossings are 2.5 m long and 0.5 width, in the rural scenario they are 4.0 m long and 0.5 m width. In urban and suburban scenarios, in advance of the pedestrian crosswalk, a parking restriction 13.2 m long was located by horizontal marking. The triggering point of the pedestrian movement was based on different vehicle distances from the zebra crossing, depending on the road type: 55.6 m for urban scenario, 66.6 m for sub-urban scenario and 88.8 m for rural scenario. Considering a driver's speed of 50 km/h on urban scenario, of 60 km/h on sub-urban scenario and 80 km/h on rural road, these distances represent the theoretical value of Time To Zebra arrive (TTZ_{arr}, the time left for the vehicle to arrive at the zebra crossing) equal to 4 s.

2.3 Warning System Design

A visual-auditive pedestrian collision warning system was adopted; this system provides a visual warning (Fig. 1d) displayed on the central projection screen (simulating a HUD display) for a duration of 3.5 s, and provides an auditive non-directional message (a "beep" sound, provided by the audio system of the vehicle) which alerts the driver of a pedestrian who is crossing,

The warning was activated when the driver reached a point that was 49 m, 59 m and 78 m in advance of the crosswalk for the urban, suburban and the rural road, respectively. Considering the same speeds used for the estimation of the theoretical TTZ_{arr} on the three road environments (50 km/h on urban scenario, 60 km/h on sub-urban scenario and 80 km/h on rural road), these values represent a time to collision (TTC) equal to 3.5 s [14, 15].

2.4 Procedure

First of all the experimenter explained to the participants the duration of trials and the use of the automatic gear, steering wheel and pedals. After that a training drive for approximately 10 min was submitted to the participants to familiarise with the tool. After the filling of a personal data questionnaire, the drivers were requested to drove the three scenarios. The sequence of the three scenarios was counterbalanced to avoid influences due to the repetition of the same order in the experimental conditions. Finally, a questionnaire with the simulation sickness or discomfort perceived was submitted to the drivers.

2.5 Participants

The initial sample of participants used in this experiment included forty-five volunteers, but three of them did not end the driving due to sickness and one participant experienced a high level of discomfort, thus, he was excluded from the sample. Therefore, the sample was composed by 41 participants, balanced in gender (21 males and 20 females), and aged from 20 to 74 (average 33).

3 Data Processing

The driver's speed behaviors when approaching a zebra crossing where a pedestrian started the crossing were analyzed. The dependent variables were: i) V_i: the driver's initial speed, i.e. the speed in the moment in which the driver perceives the zebra crossing and decreases his speed, releasing the acceleration pedal; ii) L_i: the distance from the zebra crossing at which V_i is registered; iii) V_{min}: the minimum speed registered at the end of the deceleration phase; iv) L_{min}: the distance from the zebra crossing at which V_{min} is registered; v) d_{av}: the average deceleration value during the deceleration phase.

The speed profiles were classified on the basis of TTZ^*_{arr} value, calculated as the ratio between L_i and V_i. This parameter gives an important information about the driving behavior, because low values of TTZ^*_{arr} indicate low willingness to yield, so characterizing aggressive drivers, while high values of TTZ^*_{arr} indicate more cautious driving behaviors.

The actual driver-pedestrian interactions were classified in the following three groups of TTZ^*_{arr}: i) $TTZ^*_{arr} \leq 4$ s; ii) $4\,s < TTZ^*_{arr} \leq 6\,s$; iii) $TTZ^*_{arr} > 6\,s$. The numerousness of each group is reported in Table 1.

Table 1. Recorded driver-pedestrian interactions

	PCWS			No PCWS		
	Urban	Suburban	Rural	Urban	Suburban	Rural
$TTZ^*_{arr} \leq 4$ s	16	10	10	15	13	17
$4\,s < TTZ^*_{arr} \leq 6\,s$	11	18	20	12	11	18
$TTZ^*_{arr} > 6\,s$	11	10	8	11	12	5

4 Results

A set of analyses of variance (ANOVA) was performed. The aim of the analysis was to assess how the driver's yielding behavior was affected by the pedestrian collision warning system and how this influence occurred for the three road environments and the three conditions of vehicle-pedestrian interactions (and implicitly by the driver's characteristics). More specifically, the three within – subject factors (road environment, PCWS and TTZ^*_{arr} group) were manipulated as in the following: i) Road environment:

urban, suburban and rural (3 levels); ii) PCWS: presence and absence (2 levels); iii) TTZ^*_{arr} group: $TTZ^*_{arr} \leq 4$ s, $4\,s < TTZ^*_{arr} \leq 6\,s$, $TTZ^*_{arr} > 6\,s$ (3 levels).

The analysis was carried out on the dependent variables V_{min}, L_{min}, d_{av}, descriptive of the driver's yielding behavior. The variables V_i and L_i were excluded from the statistical analysis, because they were used to determine, through the parameter TTZ^*_{arr}, the different drivers attitudes to yield to the pedestrian. In other words, the dependent variables V_i and L_i were excluded from the statistical analysis due to the mathematical correlation with the factor TTZ^*_{arr}. Statistical analysis was carried out through SPSS (Statistical Package for Social Science). For the multiple comparisons, the Bonferroni correction was used.

4.1 Driver's Minimum Speed

Results showed a significant effect of the road environment ($F_{(2,291)} = 4.021$, $P = 0.019$); the minimum speed for the rural road (32.35 km/h) was significantly higher than that for the urban road (20.45 km/h, mean difference = 11.90 km/h; P = 0.000) and for the sub-urban road (23.61 km/h, mean difference = 8.74 km/h; P = 0.000). No other difference was statistically significant. Also the effect of PCWS condition was statistically significant ($F_{(1,291)} = 8.210$, $P = 0.004$). More specifically, the recorded minimum speed for the condition of PCWS absence (30.07 km/h) was significantly higher than that for PCWS presence condition (20.86 km/h, mean difference = 9.20 km/h; P = 0.004). Results showed a significant effect of the pedestrian interaction condition ($F_{(2,291)} = 4.098$, $P = 0.018$); the mean value for $TTZ^*_{arr} \leq 4$ s was 30.47 km/h, which was significantly higher than that for $4\ s < TTZ^*_{arr} \leq 6$ s (25.66 km/h, mean difference = 4.81 km/h; P = 0.011) and for $TTZ^*_{arr} > 6$ s (20.28 km/h, mean difference = 10.19 km/h; P = 0.002). No other difference was statistically significant.

4.2 Distance from the Zebra Crossing Where the Braking Maneuver Ends

The effect of the road environment on the distance from the zebra crossing where the braking maneuver ends was statistically significant ($F_{(2,291)} = 7.470$, $P = 0.001$). The mean value of L_{min} for the urban road (25.56 m) was significantly lower than that for the rural road (38.86 m, mean difference = −13.30 m; P = 0.017) and not significantly different compared to that of the sub-urban road (30.77 m, mean difference = −5.20 m; P = 0.163). No other difference was statistically significant. Results showed that the effect of PCWS condition was not statistically significant ($F_{(1,291)} = 0.976$, $P = 0.324$) (although when PCWS was present, L_{min} was higher of about 3 m than the value for PCWS absence condition). However, the interaction effect road environment by PCWS condition was statistically significant ($F_{(2,291)} = 3.315$, $P = 0.038$). The results showed that the benefits of PCWS were remarkable in the rural road environment, where the driver significantly increased the distance from the zebra crossing in which he ended the yielding maneuver (45.70 m) compared to the condition of PCWS absence (32.03 m) (Fig. 2).

Fig. 2. Distance from pedestrian crossing where yielding maneuver ends: interaction road environment by PCWS condition

ANOVA showed also a significant effect of the pedestrian interaction condition ($F_{(2,291)} = 29.006$, P = 0.000). The mean value for $TTZ^*_{arr} \leq 4$ s was 21.76 m, which was significantly lower than that for 4 s $< TTZ^*_{arr} \leq 6$ s (35.17 m, mean difference = −13.41 m; P = 0.000) and for $TTZ^*_{arr} > 6$ s (38.28 m, mean difference = −16.52 m; P = 0.000). No other difference was statistically significant.

4.3 Average Deceleration

Results showed that the effects of the road environment and PCWS condition were not statistically significant ($F_{(2,291)} = 0.197$, P = 0.821; $F_{(1,291)} = 1.689$, P = 0.195, respectively), while the effect of the pedestrian interaction condition was statistically significant ($F_{(2,291)} = 33.112$, P = 0.000). The average deceleration for $TTZ^*_{arr} \leq 4$ s was 2.88 m/s^2, which was significantly higher than that for 4 s $< TTZ^*_{arr} \leq 6$ s (1.56 m/s^2, mean difference = 1.32 m/s^2; P = 0.000) and for $TTZ^*_{arr} > 6$ s (0.73 m/s^2, mean difference = 2.15 m/s^2; P = 0.000). No other difference was statistically significant.

5 Discussion and Conclusion

The present study aimed at evaluating the effectiveness of a combined visual-auditive pedestrian collision warning system in different road environments and driver – pedestrian interactions. The experiment was carried out using the driving simulator of Roma Tre University Civil Engineering Department.

Results showed that the effect of PCWS was significant on the variable V_{min}. The lower value of V_{min} was recorded in presence of PCWS. The timely warning of PCWS helped the driver to adopt a more correct yielding behavior towards the pedestrian. Thus, the driver was aware about the presence of the pedestrian, adopting an earlier reaction and an advanced braking response. When PCWS was present, in fact, the lower minimum speed was recorded (V_{min} = 20.87 km/h) compared to the PCWS absence condition (V_{min} = 30.07 km/h). It should be noted that the importance in reducing vehicle speed during interactions with pedestrian is consistently with the aim of PCWS; the speed of vehicle, in fact, is one of the most important parameter which is directly correlated to the fatality risk for pedestrians (an extended literature review is reported in [16]). Thus, considering the effect of speed on vulnerable road users probability of fatality, the presence of PCWS produces an important decreasing of the risk for the pedestrian in case of impact.

Furthermore, the outcomes showed that the road environment affected the drivers' behavior. On overall, the driving behavior was significantly different in the rural road compared to the urban and sub-urban one for which, instead, the driving behavior was comparable. In relation to the effect of the pedestrian interaction, the results confirmed that aggressive drivers adopted delayed yielding maneuver with higher minimum speeds and deceleration rates.

An interesting interaction effect of the road environment by PCWS condition was found for the variable L_{min}. Results showed that the effectiveness of PCWS was recorded for the rural road environment, in which the difference between L_{min} values for PCWS presence (45.70 m) and PCWS absence condition (32.03 m) was higher (13.67 m). This outcome highlights that in the road environment with higher speed, PCWS contributed to help the driver to end the braking maneuver farther from the zebra crossing. Considering the higher probability of fatal accident in case of collision at higher speeds, the contribute on the pedestrian safety highlighted in this case is remarkable.

The result showed that for the rural road the drivers, when aided by the driving assistance system, were able to end the yielding maneuver farther from the pedestrian crosswalk. This finding highlights the great potential of the warning systems in those road environments where, due to the high speeds, the consequences of impact are the worst for the vulnerable road users.

Acknowledgments. The authors acknowledge funding from the Italian Ministry of Education, University and Research (MIUR), in the frame of the Departments of Excellence Initiative 2018–2022, attributed to the Department of Engineering of Roma Tre University.

References

1. ACSM: Walking for health: easy, inexpensive and effective. In: 55th Annual Meeting of ACSM (2011). http://www.acsm.org/search-results?q=walking
2. World Health Organization, Global status report on road safety (2013)
3. ERSO: Traffic basic facts: pedestrians. Technical report, European Commission (2017)
4. Pasanen, E.: Driving speeds and pedestrian safety; a mathematical model. Transport Engineering Publication 77 (1992)
5. Varhelyi, A.: Drivers' speed behaviour at a zebra crossing: a case study. Accid. Anal. Prev. **30**(6), 731–743 (1998). https://doi.org/10.1016/S0001-4575(98)00026-8
6. ACI-ISTAT 2017. Rapporto ACI - ISTAT sugli incidenti stradali (2016)

7. Bella, F., Natale, V., Silvestri, M.: Driver-pedestrian interaction under different road environments. Transp. Res. Procedia **27**, 148–155 (2017). https://doi.org/10.1016/j.trpro.2017.12.093
8. Bella, F.: Driving simulator for speed research on two-lane rural roads. Accid. Anal. Prev. **40**, 1078–1087 (2008). https://doi.org/10.1016/j.aap.2007.10.015
9. Bella, F., Silvestri, M.: Survival model of drivers' speed reduction time at bicycle crossroads: a driving simulator study. J. Adv. Transp. **2018** (2018). https://doi.org/10.1155/2018/4738457
10. Bella, F., Silvestri, M.: Driver's braking behavior approaching pedestrian crossings: a parametric duration model of the speed reduction times. J. Adv. Transp. **50**, 630–646 (2016). https://doi.org/10.1002/atr.1366
11. Bella, F., Silvestri, M.: Effects of directional auditory and visual warnings at intersections on reaction times and speed reduction times. Transp. Res. Part F **51**, 88–102 (2017). https://doi.org/10.1016/j.trf.2017.09.006
12. Bella, F., Silvestri, M.: Interaction driver–bicyclist on rural roads: effects of cross-sections and road geometric elements. Accid. Anal. Prev. **102**, 191–201 (2017). https://doi.org/10.1016/j.aap.2017.03.008
13. Ministry of Infrastructures and Transports: Nuovo codice della strada, D.L. 30 aprile 1992 n.285 e successive modificazioni. Gazzetta Ufficiale della Repubblica Italiana. 114 (1992)
14. Chang, S.H., Lin, C.Y., Fung, C.P., Hwang, J.R., Doong, J.L.: Driving performance assessment: effects of traffic accident location and alarm content. Accid. Anal. Prev. **40**, 1637–1643 (2008). https://doi.org/10.1016/j.aap.2008.05.003
15. Chang, S.-H., Lin, C.-Y., Hsu, C.-C., Fung, C.-P., Hwang, J.-R.: The effect of a collision warning system on the driving performance of young drivers at intersections. Transp. Res. Part F Traffic Psychol. Behav. **12**, 371–380 (2009). https://doi.org/10.1016/j.trf.2009.05.001
16. Rosén, E., Stigson, H., Sander, U.: Literature review of pedestrian fatality risk as a function of car impact speed. Accid. Anal. Prev. **43**, 25–33 (2011)

Identified Risk Factors Among Truck Drivers Circulating in France

Anabela Simoes[1](✉), Patricia Delhomme[2], Blazej Palat[2], Alexandra Gheorghiu[2], Jean-Pascal Assailly[2], Teodora Stefanova[2], Giulio Bianchi Piccinini[3], Loic Josseran[4], Gilles Vallet[5], and Juan Pérez[6]

[1] LUSOFONA University, Lisbon, Portugal
anabela.simoes@ulusofona.pt
[2] UGE, LaPEA, Versailles, France
{patricia.delhomme, jean-pascal.assailly}@univ-eiffel.fr, blazej.palat@yahoo.fr, gheorghiu.anda@yahoo.com, teostanstef@gmail.com,
[3] Chalmers University of Technology, Gothenburg, Sweden
giulio.piccinini@chalmers.se
[4] Equipe de Recherche HandiResp– EA4047, Université de Versailles, St Quentin, France
loic.josseran@aphp.fr
[5] UGE, UMRESTTE, Bron, France
gilles.vallet@univ-eiffel.fr
[6] Facultad de Psicología. Departamento de Psicología Social, University of Valencia, Valencia, Spain
juan.a.perez@uv.es

Abstract. The growth of the European market for road-freight transport has recently led to important changes. Moreover, due to the geographical context of France, truck drivers from different countries circulate on French roads to deliver their goods. Having road safety concerns in mind and based on a literature review, as well as interviews, a French questionnaire has been developed. Aiming at collecting data among truck drivers from different European countries, the questionnaire has been translated into seven languages and has been administered at four highway rest areas in France. The collected data were analyzed by means of multiple correspondence analysis, which pointed out new links between working conditions, driving behaviors, perceptions of the road environment, job skills, and risk factors. The practical implications of the results for improving the safety of the infrastructure used by truck drivers and their well-being are discussed.

Keywords: Truck drivers · Road-Freight transport · Working conditions · Crash risk · Health risk

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 315–321, 2020.
https://doi.org/10.1007/978-3-030-50943-9_40

1 Introduction

According to a review of social issues in road freight transport [1] in 2015, the estimated activity of European road-freight transport was 1,768 billion tons/kilometer. This activity grew by 2.4% compared to 2014. However, the increase differed across countries, having progressed 5.3% in the new European Union members compared to 1.2% in old ones. In 2015, European road freight transport was carried out by 549,000 companies with more than two thirds of them from the old member countries. While the number of road-freight companies decreased in the old member countries, it increased in the new member states, especially in Romania (10.3%) and Poland (2.5%).

Despite a decline in mortality from road crashes involving heavy trucks by 48% in France between 2000 and 2010 and maintained a 2% decrease every year thereafter, an increase by 4.2% between 2015 and 2016 has been noticed. New policies pertaining to legal daily working hours and the means used to check whether truck drivers abide by those policies, coupled with improved technical conditions of trucks, improved infrastructure and road signage, have helped increase in truck drivers' safety, notably in France [2–4]. However, crashes involving heavy trucks in France still account for 4.9% of all road crashes, while their victims account for 14.2% of all road deaths. There are 176 persons killed in every 1000 crashes involving heavy trucks, almost three times as many as in road crashes in general, where the count is 60 deaths in every 1000 crashes [1]. Fatal crashes involving heavy trucks occur more often on divided highways (31%) compared to fatal road crashes in general (10%) since heavy trucks cover a greater proportion of their itineraries on divided highways as compared to other types of vehicles [1]. As a consequence, 18% of persons killed and 27% of persons wounded in crashes involving heavy trucks in 2016 are victims of crashes that occurred on divided highways [5].

Concerning sociodemographic characteristics, truck drivers under 27 and over 63 years old appear to be at a greater risk of being involved in a road crash than truck drivers in other age categories [6]. However, other studies have shown that older drivers (60 and more) are more attentive to safety aspects, such as seat belt use or alcohol consumption [7]. Regarding gender, the field of professional truck driving is less feminized [8], even if some fluctuations have been noticed [9]. In 2015, women filled less than 6% of the truck driving positions in France according to the "Bilan Social Annuel du transport routier de marchandises", 2018. Men, more often than women, tend to engage in risky driving behaviors, such as driving under influence, speeding, or failure to use the seat belt [1].

Considering many risk factors of traffic safety and truck drivers' health discussed herein, the influence of those factors in the population of truck drivers circulating in France was studied. Given the central location of France, road-freight transport employment and work conditions depend on a broader economic context, and in particular, that of the European Union. In this changing context and in the absence of recent studies on risk factors involving that population, a questionnaire-based survey was carried out.

2 Method

The survey addressing truck drivers' characteristics and behaviors was carried out using a questionnaire that was administered in highway rest areas in France (East, Southeast, Southwest, and Ile de France). The rest areas were chosen based on two criteria: their use by different nationalities of truck drivers and the easiness in obtaining the permission for data collection.

2.1 Participants

The sample was composed of 515 truck drivers (11 women) with an average of 20.5 years of work experience (σ = 11.42, range: 0–48), who volunteered to take part in the survey. They covered an average of 588 km (σ = 162.34, range: 100–900) per day with their truck. The respondents were distributed by three regions of Europe (east, south, or west), according to their own language (Table 1) and nationality. Their nationality was distributed as following: 19% of eastern Europeans, 19.4% of southern Europeans, and 61.6% of western Europeans.

Table 1. Proportion of each of the eight languages in which the questionnaire was administrated

Region of Europe	Language	Respondent Nationality	N	% of the region	% of the sample
East	Bulgarian	Bulgarian	25	25.5	4.9
	Polish	Belarusian	32	32.7	6.2
		Polish, Czech			
		Ukrainian			
	Romanian	Moldavian	41	41.8	8
		Romanian			
West	English	Belgian, British	18	5.7	3.5
		Dutch, German			
	French	Belgian, French	299	94.3	58.1
		French Belgian			
		Dutch, German			
		Luxembourgian			
		Swiss			
South	Italian	Italian	29	29	5.6
	Portuguese	Portuguese	36	36	7
	Spanish	Spanish, Portuguese	35	35	6.8

2.2 Questionnaire

The content of the questionnaire was based on a literature review and two series of preliminary, semi-directive interviews. The aim of the interviews was to obtain up-to-date knowledge of truck drivers' experiences on the road and during rest periods, their opinions and impressions about their occupation, and their driving behaviors, difficulties, and work conditions in general.

The questionnaire developed in French was translated into Bulgarian, English, Italian, Polish, Portuguese, Romanian, and Spanish. Most of the questions required answers on five-point Likert scales indicating frequency (from 1 = never to 5 = very often), and intensity or agreement (from 1 = not at all to 5 = absolutely). Several questions required a simple yes or no answer.

Measures. Different types of data were collected: continuous, ordinal, and categorical. The exploratory method used for data analysis could handle all of these types of information, aiming at identifying relationships between the following five variable groups:

a) Driving context with Perceived accident or crash risk, Expectations regarding training of truck drivers, Near crashes, Past occurrence of incidents related to fatigue, Perceived difficulty of being behind the wheel of a truck, Perceived frequency of roadside checks, Perceived frequency of dangerous behaviors by car drivers and motorcycle/scooter riders, Perceived truck-driving skill, Risk-taking behind the wheel of the truck, Vulnerability to distraction behind the wheel.

b) Employment context with the following variables. Attitude toward foreign truck drivers, Average daily time spent behind the wheel, Average daily working time, Company size, Turnover, Factors of the effort-reward imbalance (ERI) model, Frequency of transporting various types of goods, Inconveniences of working as a truck driver, Job experience, Monthly income, Perceived annoyances at rest areas in France, Perceived safety of rest areas in France, Perceived competition from foreign road-transport companies, Zone of job activity.

c) Health and health-related behaviors with Blood-pressure problems, Body mass index (BMI), Daily alcohol consumption, Illegal drug intake, Number of cigarettes smoked daily, Number of energy drinks consumed daily, Self-reported state of health, Soft-drink consumption, Sports.

d) Psychological characteristics such as Burnout, Daytime sleepiness, Insomnia severity, Mind wandering, Perceived stress, Wellbeing.

e) Demographic characteristics such as Gender, Age, Driving experience, Daily kilometrage, Demerit points, Involvement in road crashes as a truck driver, Marital status.

2.3 Procedure

The questionnaire was administered at four highway rest areas in France where different nationalities truck drivers frequently pass by professional interviewers in mid-March 2018. Truck drivers were contacted in the parking zones and/or in the commercial facilities of the rest areas. The questionnaire was displayed on an iPad. The median filling in time was 44 min. A scarf and a coffee were offered as gifts to thank the participants.

3 Results

Our approach consisted of taking into account numerous aspects of a truck driver's job in order to identify the risk factors in the sample participating in the survey. Given that the data collected data were categorical and quantitative, our approach, which is exploratory, had to include both types of variables. The results of a series of comparisons between the drivers originating from eastern and western Europe are in line with those of the previous exploratory approach. However, even though the low level of perceived inconvenience of the truck driver's job was negatively associated with the axis obtained from the MCA, as were also the drivers originating from the east Europe; this second analysis revealed that they actually saw more inconveniences than their western-European counterparts. We obtained an analogous result from the test of differences in job experience. Longer work experience was negatively associated with this axis as were the eastern-European truck drivers, whereas they are actually less experienced than western-European drivers according to the second tests.

4 Discussion

First of all, we found some of the important risk factors already documented in the literature. Firstly, a poor quality of sleep is a cause of daytime sleepiness, which in turn increases the probability of occurrence of near crashes related to fatigue, near crashes in general, and/or crashes. Also involved are various sources of distraction, both intrinsic (mind wandering) and extrinsic (communication tools, elements of the road environment, driving aids, etc.). Furthermore, we found that a poor state of health, along with several behaviors deemed to deteriorate it (smoking, imbalanced diet), are also associated with road risks, even though we are unable to identify the mechanisms underpinning this relationship on the basis of our results.

In addition to the above results that are in line with prior knowledge in the domain, we noted that truck drivers' risk taking behind the wheel, road risks, and probably also sleep quality can probably be linked to distress at work, the latter being especially triggered by certain work conditions (for example, rarely driving on divided highways, "problematic" truckloads, long working hours, annoyances during rest periods, etc.). We also observed that risk exposure is associated with poor skills at manoeuvring, anticipation of danger, keeping safe distance with the vehicle in front, driving downhill, coping with fatigue, and with dangers on the road, with underestimation of the difficulties linked to blind spots in the mirrors, braking while driving downhill, handling sharp bends, driving in the rain, snow, or strong winds coupled with a lack of a desire to develop one's skills during mandatory training.

Moreover, we found a positive relationship between risk perception and risk taking, and a negative one between these variables and the perceived frequency of roadside checks. Hence, we can assume that risky truck drivers are well aware of their risk exposure and the perception of low frequency of roadside checks for sobriety, the truck's technical condition, mandatory safety equipment, the driver's adherence to speed limits and the legal daily driving-time limits; mobile phone use, and seat belt use can facilitate their risk-taking behaviors such as speeding, drunk driving, driving under the influence of marijuana, exceeding daily driving-time limits, and/or cheating with the tachograph (stopping the clock during loading/unloading).

From a practical standpoint, it seems that all means aimed at reducing truck drivers' work inconveniences and stress can be indirectly beneficial for road safety, insofar as they also improve the attractiveness of the profession and motivation to develop one's truck driving skills. More specifically, guaranteeing the conditions for a proper rest (securing truckload, limiting annoyances at rest areas) is desirable. Roadside checks can be seen as deterring truck driven from violating traffic laws, so they should be more frequent and more effective at detecting risky behaviors of drivers behind the wheel of a truck. Finally, promoting a healthy lifestyle among truck drivers' creating conditions for participating in sports during rest periods, and developing healthy eating habits could also improve the well-being of these workers, and indirectly, their road safety.

Our study has several limitations that must be addressed. First, we could not thoroughly examine the causal relationships between the variables, and their possible mediation and/or moderation effects, using the adopted exploratory approach. Our work should be seen as a first step toward gaining further knowledge of European truck drivers' occupational issues and the risk factors specific to this population. Second, using a more diversified sample, also recruited in other countries of Europe, could enable the collection of information likely to account for local specificities (road infrastructure, rest areas, safety levels, etc.) that would allow for between-country comparisons of the influence of local characteristics on the work conditions of truck drivers in different European countries. Third, the majority of the measures used in the survey were developed ad hoc. Hence, their quality and validity were not tested as thoroughly as those in well-established self-report measurement tools. In addition, our ad hoc measures may have been altered in the translation process.

Acknowledgements. This work was supported by the Fondation VINCI-Autoroutes, Paris, France [grant number: RP1-E17088].

References

1. References Commissariat général au développement durable (CGDD-SDES): Service de la donnée et des études statistiques: Bilan social annuel du transport routier de marchandises (2018)
2. Medina-Flintsch, A., Hickman, J.S., Guo, F., Camden, M.C., Hanowski, R.J., Kwan, Q.: Benefit–cost analysis of lane departure warning and roll stability control in commercial vehicles. J. Safety Res. **62**, 73–80 (2017)
3. Naderi, H., Nassiri, H., Sahebi, S.: Assessing the relationship between heavy vehicle driver sleep problems and confirmed driver behavior measurement tools in Iran. Transp. Res. Part F **59**, 57–66 (2018)
4. Pereira de Oliveira, L., Morais Lemos, B., Vieira da Silva, M.A., Jiménez Alonso, F., da Silva Guabiroba, R.C.: Analysis of the event data recorder system regarding criteria of safety, operation and consumption in a Brazilian trucking company. Transp. Res. Part F (2018, in Press)

5. Observatoire national interministériel de la sécurité routière (ONISR): La sécurité routière en France. Bilan de l'accidentalité de l'année 2016. Paris. Direction de l'information légale et administrative (2017)
6. Duke, J., Guest, M.: Age-related safety in professional heavy vehicle drivers: a literature review. Accid. Anal. Prev. **42**, 364–371 (2010)
7. Newnam, S., Blower, D., Molnar, L., Eby, D., Koppel, S.: Exploring crash characteristics and injury outcomes among older truck drivers: an analysis of truck-involved crash data in the United States. Saf. Sci. **106**, 140–145 (2018)
8. Douglas, M.A., Swartz, S.M., Richey, R.G., Roberts, M.D.: Risky business: investigating influences on large truck drivers' safety attitudes and intentions. J. Saf. Res. **70**, 1–11 (2019)
9. Cantor, D.E., Corsi, T.M.: A driver focused truck crash prediction model. Transp. Res. Part E **46**, 683–692 (2010)

Applying the Systems Theoretic Accident Model and Process to Analyze a Downgrade-Truck Collision Caused by a Brake Failure in Vietnam

Do Duy Dinh[1(✉)], Nam Hoai Vu[1], Rich C. McIlroy[2], Katherine L. Plant[2], and Neville A. Stanton[2]

[1] Department of Highway and Traffic Engineering, National University of Civil Engineering (NUCE), Hanoi, Vietnam
dinhdd@nuce.edu.vn
[2] Transportation Research Group, University of Southampton, Southampton, UK

Abstract. This present paper is to apply the Systems Theoretic Accident Model and Process (STAMP) and its corresponding the Causal Analysis using Systems Theory (CAST) as an innovative method to analyze a road traffic collision involving a goods vehicle travelling downhill ostensibly caused by a brake failure on a mountain road in Vietnam. The developed STAMP-CAST model of the collision under study showed that although driver's inexperience, together with the truck's low quality and severe road conditions were found as potential factors directly leading to the collision, the inadequate control actions of various actors residing at higher levels of the Vietnamese road transport system also contributed to the crash.

Keywords: Downgrade truck collision · Sociotechnical systems · STAMP

1 Introduction

Downgrade truck collisions, where a goods vehicle travelling downhill is involved in a collision with infrastructure or other vehicles, is often caused by a brake failure, and commonly results in extensive damage to lives and property. In Vietnam, with its extensive network of mountainous roads, this type of road accident has become a serious traffic safety problem which has been strongly highlighted in the Vietnamese media recently. However, up to date, there has been almost no published research dedicated to identifying the influential factors of this type of accident in Vietnam, while results from police investigations on most downgrade truck accidents in Vietnam often attributed causes to drivers or vehicles' characteristics.

Previous studies have found loss of control as a significant cause of downgrade truck crashes. For example, LTCCS [1] reported that loss of control accounted for 28.6% of truck collisions. Ahmed et al. (2011) [2] indicated that the risks of these types of collisions are generally higher on steeper downgrades as compared to other typical road sections. Although various factors have been found to have effects on downgrade

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 322–329, 2020.
https://doi.org/10.1007/978-3-030-50943-9_41

truck collisions, the causes of these accidents have been mainly attributed to driver's inexperience on downgrade driving, defective brakes, and inadequate signing [3] as well as factors of environment and road geometry [2]. Very little, if any attention has been paid to the higher sociotechnical factors that shape the system from which these types of accidents emerge. Previous studies have only focused on the immediate, proximal causes of collisions, at the expense of an examination of the potentially influential factors arising from other actors within the road transport system, as well as the interactions amongst them. It is argued that it is necessary to seek a different approach to address what remains a serious problem in Vietnam, as well as in other low- and middle-income countries with similar topographies.

One potential alternative is the sociotechnical systems-based approach, a philosophy that has gained increasing attention in the road safety research domain in recent years. The guiding principle of the approach is that safety is an emergent property arising from non-linear interactions between a system's components [4]. To facilitate analyses of accidents in complex sociotechnical systems, a number of systemic safety models have been developed such as FRAM [5], AcciMaps [6], Systems Theoretic Accident Model and Process (STAMP) [7], and EAST broken-links approach [8]. Amongst these models, STAMP approach has been widely used in academic literature and it has been used to analyze accidents in different domains [9]. As highlighted by Stanton [10], STAMP can be used for collision analysis post-event and/or risk analysis pre-event, and provides an overview whereby systems are seen as having hierarchical levels, each with a control structure. Controls enforce constraints, resulting in safe behaviour. Controls and constraints operate bottom-up and top-down between levels. According to STAMP [7], rather than incidents being viewed as the consequence of events, they are seen as resulting from control failures (i.e. inadequate enforcement of constraints or lacking/inaccurate constraints). Accidents occur when component failures, external disturbances, and/or inappropriate interactions between systems components are not controlled [9].

Road transport has been considered as a complex sociotechnical system, comprising many inter-related components [4], therefore it is potential for using STAMP to analyze road crashes in order to find out the role and interaction of factors outside of road users, vehicles and the road environment. Up to date, STAMP approach has been employed to model the safety of the overall road transport systems [11], however there has been no studies using the approach to analyze the causes of a specific road accident while this type of analyses have been successfully applied in other safety domains [9].

The aim of this paper is to analyze the causes of a typical downgrade truck collision involving a brake failure on a mountain road in Vietnam considering the whole factors of road systems. To avoid the limitations of the current method of investigating such collisions, this present study applied Systems Theoretic Accident Model and Process (STAMP) and its corresponding the Causal Analysis using Systems Theory (CAST) [11] as an alternative method to analyze the road accident, and then make recommendations for safety interventions.

2 Method

Causal Analysis using Systems Theory (CAST) is the accident analysis method within the STAMP framework, which is often known as STAMP-CAST. This method proposes a taxonomy of control failures using observations and scenarios as data collection methods. It accounts for cognitive factors by considering the context for decisions and including a 'mental model flaws' category. The first step in carrying out a STAMP–CAST analysis is to model levels of the Sociotechnical Systems (see [6]). Next, the potential or actual collisions to be considered, and the control flaws and hazards are identified (see [10] for more details). Then a model of the functional control structure is created, specifying potential unsafe control actions and how they could occur. Lastly, remedies are suggested. The output consists of two stages: the control structure and then a more detailed analysis of key personnel selected from the control structure.

The present study analyzed a downgrade truck collision that occurred on March 22nd, 2018 on National Highway No. 6 (NH6) in the Hoa Binh province of Northern Vietnam, in which a truck had a brake failure on a mountain road and then crashed into a passenger car and a rock cut-slope. All information and data described the nature of the accident were thoroughly collected from different sources including the on-site traffic accident investigation reports from the traffic police office and reports on road accident black-spots from Directorate for Road of Vietnam (DRV). A number of site visits and field surveys were made to the study road sections. The first two authors of the current article also conducted a number of interviews relating the accident and traffic safety issues of the road sections with a variety of individuals; the drivers and passengers involved in the accident, and their family members; local traffic policeman, and local traffic authorities, and etc. The data collection process took approximately 8 months, from February 2018 to September 2018.

Based on the collected data, at first an Accimap model of the accident was developed as showed details in [13]. In this phase, all aspects of the accident were thoroughly examined to identify all the actors that could have been involved in the accident as well as all decisions/actions/failures associated with the accident. The current study used the actor map for Vietnam developed by McIlroy et al. (2019) [12] as a starting point for identifying the factors that may have contributed to the occurrence and severity of the incident under investigation. The Accimap model was then turn into STAM-CAST model which represents control actions and feedbacks between different levels of actors in the Sociotechnical Systems following Leveson's method [9].

3 The Accident

The incident under analysis happened at the end part of a continuous, long downgrade hill road section with many sharp horizontal curves, as presented in Fig. 1. This downgrade section starts from the Thung Khe mountain pass, approximately 11.6 km from the accident location. A goods vehicle was travelling down from the pass at high speed, having, at some point, lost use of the brakes. After negotiating a number of sharp curves, the vehicle crashed into a passenger car in front on the last curve (see Fig. 1).

During the crash, the truck pushed the car into the roadside safety barriers on the right-hand side, which are made of deposed rubber wheels and a steel frame. The truck then jumped over the car and a part of the truck slid on the top of the barriers until it turned over on the roadway. Damage to the passenger car was especially severe. Incredibly, all the truck driver, the car driver and their passengers survived, though injuries were serious. The location where his truck had brake malfunction was approximately 3.2 km from the accident site.

The truck driver was born in 1989 and received his driving license in 2010. Recently, he used the truck as a commercial transport service for local people. His transport routes were not fixed and depended on the demand of his customers. He used to drive about four or five times per month, with each job lasting one or two days. Before the accident, most of his driving experience had been in the low-lying, plain regions; he rarely drove through mountain regions. The route to Son La province on the day of the accident was the first time the driver travelled on this particular route, and was the first time he had experienced driving on such a difficult mountain road as NH6.

At the time the truck driver realized that brake functionality was lost, his truck was behind a car and another truck, with little separation between vehicles. He continued to drive with using a high gear and, along with the car, passed the truck travelling ahead of him. The truck driver knew that he should switch to a low gear in order to decelerate the truck by the vehicle engine, but he was not able to do this. The failure of the pneumatic system that had rendered the brakes inoperable had also affected use of the gears, as well as making the air horn unusable. From this time until the collision, he did nothing other than try to avoid collision with other vehicles on the road, and alert other vehicles by shouting, hand waving, and clapping on the vehicle door. When the truck reached the curve (the accident site as shown in Fig. 1), the truck was at its high speed and was rapidly approaching a passenger car. The truck driver intended to pass the car but there was a truck on the opposite direction, therefore was unable to avoid crashing into the passenger car in front. The passenger car driver was a local taxi driver with considerable experience, having achieved his driving license (more than 15 years ago). He had driven on that particular road section many times before. At the accident time, he was talking with his passengers and had not realized that the truck behind him was having issues.

Fig. 1. The road section from Thung Khe pass to the accident site

It should be noted that, within the last ten years (from 2008 to 2018), since the road underwent improvement works, according to traffic police, more than 10 similar accidents have occurred along the same stretch of road, causing more than 10 fatalities and a number of injuries. In 2017 alone, three accidents occurred, leading to one fatality. All the three accidents occurred in 2017 were related to semi-trailers losing their control while going downhill and crashing into the roadside barriers or rock cut-slope at almost the same location as the accident under the study. As a response to this serious traffic safety issue, at the end of 2017 tire walls (i.e., roadside safety barriers made of deposed wheels and steel frame) were implemented at two horizontal curves on the section, including the curve where the accident under analysis occurred.

4 Results and Discussion

Figure 2 showed the STAMP-CAST model for the accident under analysis. In line with previous studies [2, 3], the model clearly demonstrated that the truck driver's inexperience and poor vehicle quality, in combination with severe road conditions were identified as factors directly leading to the truck driver's inadequate actuations that finally caused the collision. Apart from these factors, the model also pointed out that the inadequate control actions of various actors residing at higher levels of the Vietnamese road transport system significantly contributed to the crash.

First, the current driving training curricular which is stipulated in the Circular 12/2017/TT-BGTVT issued by Ministry of Transport (MOT) on regulating driver training, driving tests, and issuance of driving licenses for road motor vehicles does not require learner drivers to practice driving on mountain roads, and does not require candidates to drive in these types of environments during testing. These deficiencies along with the loose monitoring the training activities of driver training centers and instructors contributed to the truck driver's inexperience and lack of driving skill on such driving circumstance. Secondly, the unsafe driving environment in this present study may result from inadequate design standards, inadequate communication and missing feedback between different levels of road authorities, and a lack of power to enforce relevant bodies and personnel when building safe roads. More specifically, the technical standard of "Emergency escape ramp – Specification for design" (TCVN 8810:2011) does not specify clear criteria for a requirement for building emergency escape ramps. This limitation combined with budgetary constraints, the lack of a prioritization system on traffic accident black-spot improvement as well as lack of legal power to enforce DRV quickly solving road safety problems raised by its local branches and the voice from local governments are all the primary reasons why the building of an escape ramp had been postponed by the DRV. Thirdly, insufficient policy on controlling the build-quality of trucks resulted in the presence of vehicles with low-safety standards. As presented in Fig. 2, the national technical regulation on safety and environmental protection for automobiles (QCVN09:2015/BGTVT) issued by MOT does not include the requirement of, and testing methods for ensuring vehicle brakes work well after traveling down steep downgrade sections such as that under the current investigation. If this limitation is eliminated, vehicles with safety standards similar to the truck in the accident would not be allowed to be used. Lastly, a lack of campaigns to raise public awareness of the issue of downgrade truck collisions also contributed to this safety problem. Up to date, Vietnam's National Traffic Safety Committee (NTSC) has not issued any policy that requires local governments and other relevant

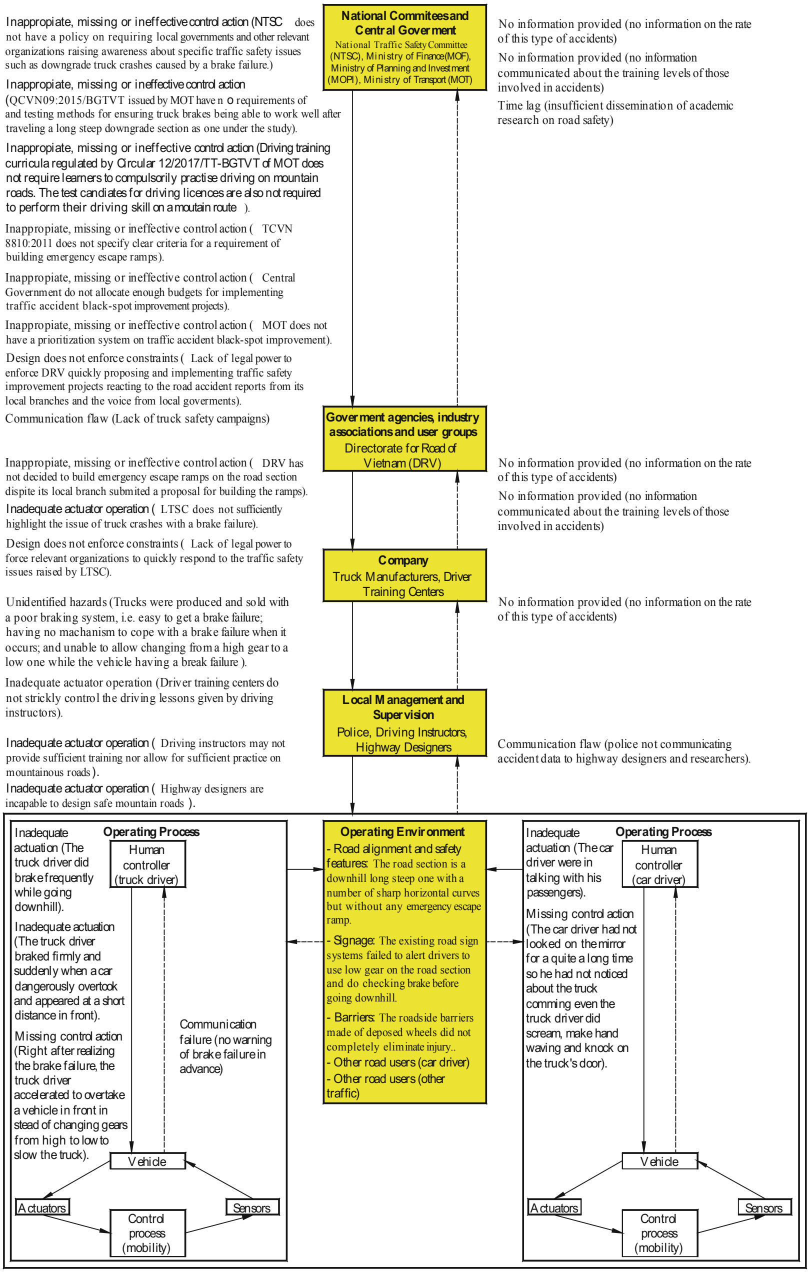


Fig. 2. The STAMP-CAST model for the accident under study

organizations to work on improving traffic safety awareness to public. Figure 2 also shows some flaws in the feedbacks from lower level to higher one on the system especially there is no accident data provided for higher-level organizations making proper decisions on addressing the crash issue. The aforementioned results obviously provided helpful suggestions for reducing downgrade truck collisions in Vietnam and other countries with similar settings.

5 Conclusions

The STAMP method enables analysts to identify contributing factors out-side of road users, vehicles, and immediate road environment. This method should be used as alternative method for investigating such types of collisions in Vietnam and other low and middle-income countries in order to produce appropriate safety countermeasures that go beyond traditional approaches to road safety.

Acknowledgements. This research was commissioned by the National Institute for Health Research using Official Development Assistance (ODA) funding. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

References

1. FMCSA: Report to Congress on the Large Truck Crash Causation Study. Federal Motor Carrier Safety Administration, Washington DC (2006)
2. Ahmed, M., Huang, H., Abdel-Aty, M., Guevara, B.: Exploring a Bayesian hierarchical approach for developing safety performance functions for a mountainous freeway. Accid. Anal. Prev. **43**(4), 1581–1589 (2011)
3. Myers, T.T., Ashkenas, I.L., Johnson, W.A.: Feasibility of Grade Severity Rating System. Federal Highway Administration, Washington DC (1981)
4. Salmon, P.M., McClure, R., Stanton, N.A.: Road transport in drift? Applying contemporary systems thinking to road safety. Saf. Sci. **50**(9), 1829–1838 (2012)
5. Hollnagel, E.: FRAM, the Functional Resonance Analysis Method: Modeling Complex Socio-Technical Systems. Ashgate Publishing, Ltd. (2012)
6. Rasmussen, J.: Risk management in a dynamic society: a modelling problem. Saf. Sci. **27** (2/3), 183–213 (1997)
7. Leveson, N.: A new accident model for engineering safer systems. Saf. Sci. **42**(4), 237–270 (2004)
8. Stanton, N.A., Harvey, C.: Beyond human error taxonomies in assessment of risk in sociotechnical systems: a new paradigm with the EAST 'broken-links' approach. Ergonomics **60**(2), 221–233 (2017)
9. Leveson, N.: Engineering a Safer World: Systems Thinking Applied to Safety. MIT Press, New York (2011)
10. Stanton, N.A.: Models and Methods for Collision Analysis: A Guide for Policymakers and Practitioners. RAC Foundation, London (2019)

11. Salmon, P.M., Read, G.J.M., Stevens, N.: Who is in control of road safety? A STAMP control structure analysis of the road transport system in Queensland, Australia. Accid. Anal. Prev. **96**, 140–151 (2016)
12. McIlroy, R.C., Plant, K.A., Hoque, M.S., Wu, J., Kokwaro, G.O., Vũ, N.H., Stanton, N.A.: Who is responsible for global road safety? a cross-cultural comparison of actor maps. Accid. Anal. Prev. **122**, 8–18 (2019)
13. Dinh, D.D., Vũ, N.H., McIlroy, R.C., Plant, K.A., Stanton, N.A.: A deep examination on a downgrade-truck collision caused by a brake failure in Vietnam: policy implications through a systems-based approach. In preparation (2020)

Transport Planning and Infrastructure Design

Effect of Perceived Contrast Enhancing Lens Technology on Traffic Signal Detection for Color-Deficient Individuals

Cameron Lopez(✉), Jeremy Swan, and Jonas Schmidtler

Oakley Inc., 1 Icon, Foothill Ranch 92610, USA
cslopezl@berkeley.edu, {JSwan,JSchmidtler}@Oakley.com

Abstract. This study examined the effect of Perceived Contrast Enhancing (PCE) lens technology on traffic signal detection and recognition for color-normal and color-deficient observers compared to a neutral density lens. Eighteen color-normal and eighteen color-deficient participants performed a visual-motor task while wearing two different PCE lenses with specific spectral transmissions as well as a neutral-density lens. At random intervals, simulated traffic light signals were presented 5° to the right and left of the participant's focal point, to which participants identified signal color using a three-button input device. Response time and error rate were recorded. We found that lens tint did not have a significant main effect on response time and error rate. The data collected in this study lends considerable evidence to the assumption that PCE lenses will not impair driving.

Keywords: Human factors · Color deficiency · Visual-motor detection and recognition

1 Introduction

Approximately 8% of males and 0.4% of females today are color-deficient, and a large majority of those are color-deficient in the red-green spectrum [1]. A majority (~75%) of the color-deficient population is made up of anomalous trichromats, meaning one of their color receptors is altered as compared to normal. Nearly all of the rest of the color-deficient world are dichromats, meaning they completely lack one of the three color receptors and have two-dimensional color discrimination as opposed to three-dimensional. As a result of these deficiencies, the ability of the color-deficient population to discriminate between red, yellow, and green traffic signals is reduced or, in the case of dichromats, completely absent [2]. Participating in traffic depends on the detection of colors and colored signals. Because tinted sunglass lenses can negatively affect reaction time and recognition rates to those signals for color-deficient users [3, 4], and [5], it is important to determine which sunglass lenses are acceptable to use for the color-normal and color-deficient population. Due to these findings, sunglass standards around the world enforce coloration requirements on sunglasses with the goal of limiting color distortion. These properties are established to define the amount of chromatic influence of red, yellow, and green traffic signals as viewed with a background of average daylight. The American National

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 333–339, 2020.
https://doi.org/10.1007/978-3-030-50943-9_42

Standards Institute (ANSI Z80.3-2018) [6] regulates these chromatic influences with three criteria: 1) *Color limits*: Each respective chromatic coordinate must reside within a specified region of the Commission Internationale de l'Eclairage (CIE) 1931 chromaticity diagram [7], 2) *Transmittance properties*: Each filter must allow a minimum amount of total light and respective colored light through, and 3) *Spectral Transmittance*: All attenuation must not fall below a set threshold.

The objective of our experiment was to assess the extent to which reaction time and error rate may be affected by PCE lens technology for color-deficient subjects when compared to a neutral density lens, using traffic signals that comply with US national standards laid out by the Institute of Traffic Engineers (ITE), Australian national standards, and ANSI guidelines [6–9], and [10]. Our hypotheses are as follows:

H1: *Color-deficient subjects' reaction times and error rates are higher than their color-normal counterparts'.*
H2: *There is no effect of PCE lenses that fail and other lenses that currently pass ANSI standards on reaction time and error rate.*

2 Method

2.1 Study Design

Sport performance frames (Oakley® Radar EV) were used to cover a wide field of view and mitigate backside glare effects. Figure 1 shows the three lenses used which consisted of a neutral density pair (16% visible light transmission VLT, shown in *green*) and two pairs of PRIZM™ lenses (18% VLT, shown in *red*, passes ANSI 4.10.2.3, and 34% VLT, shown in *blue*, fails ANSI 4.10.2.3).

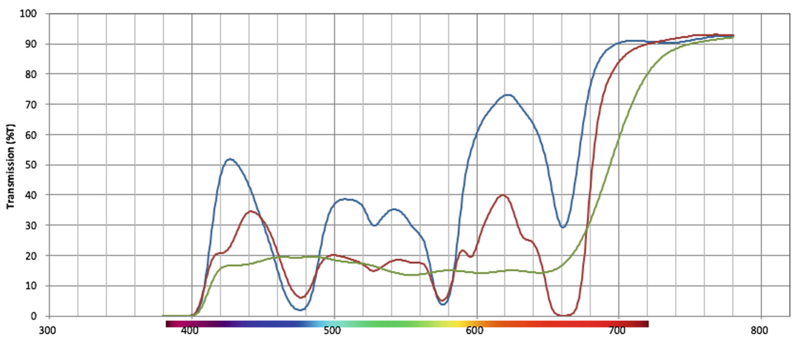


Fig. 1. Spectral transmissions of PRIZM™ Golf (*blue*, 34% VLT, fail ANSI 4.10.2.3), PRIZM™ Road (*red*, 18% VLT, pass ANSI 4.10.2.3) and Dark Violet (*green*, 16% VLT, pass ANSI 4.10.2.3).

2.2 Sample

Our sample utilized thirty-six healthy males between the ages of 14 and 54 (M = 36.4 yrs, SD = 9.0 yrs), consisting of 18 color-normal males and 18 color-deficient males. Pre-study all participants conducted a color discrimination task using a light box and Farnsworth-Munsell 100 Hue Test to ensure their grouping in color-normal or color-deficient. Distinctions were not made between the different types of color deficiencies. All participants had a visual acuity of at least 20/25, and participants wore their non-tinted correctional lenses behind the experiment's tinted lenses if necessary, to achieve this visual acuity.

2.3 Apparatus

The general study apparatus was replicating parts of [5] with minor alterations, discussed below (see Fig. 2). Participants viewed a target in the center of a computer monitor at a 1.5 m working distance to be able to view the secondary task easily on a conventional computer monitor. Three separate bulbs were laid horizontally (spanning 1 cm) on either side in the same space on either side so intensity could be individually regulated. At the working distance of 1.5 m and angle 5º size constancy mandated the use of LED bulbs with 3 mm diameter. Following that, there were no significant positional cues for each of the bulbs that could influence the experiment. Additionally, we positioned our participants' eye level at 7 cm beneath the signal level rather than at the level of the bulbs. This corresponds to a 200 mm traffic light at 100 m being 4.572 m off the ground, which is the minimum height a traffic signal can be in the U.S. according to the ITE [10]. Signals were created using three differently colored, high-intensity LED bulbs (red, yellow, and green). We referenced standards outlined by ITE to find the minimum intensity of a traffic signal from the vantage point of −2.5° vertically and 5° horizontally, to mimic our study design. We converted this to lux at 100 m and used different resistors for our LED bulbs until we were able to match that lux from 1.5 m. For red, a 220 Ω resistor was used to achieve the luminous intensity of 0.033 cd. For yellow, a 47 Ω resistor was used to achieve 0.081 cd, and for green a 1.5 kΩ resistor was used to achieve 0.021 cd. Each of these luminous intensities are the 1.5 m equivalent of the minimum allowed intensity of a 200 mm traffic signal at 100 m, which is the standard Australian practice utilized by Dain [4, 8], and [9]. Bulb rise time to full intensity was negligible. Table 1 references the respective ITE and CIE specifications.

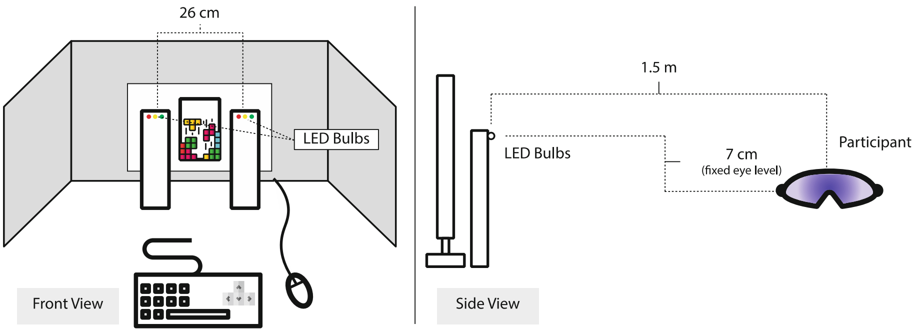


Fig. 2. Apparatus with computer monitor, keyboard, mouse, led bulbs, and measurements. Front view (left), side view (right).

Table 1. Minimum allowed luminous intensity of a traffic light as outlined by ITE and Chromaticity coordinates of each of bulb, from spectral radiance measurements made with an AsenseTek Essence telespectroradiometer, all fall within the required CIE 1931 specifications as laid out by ANSI.

Traffic signals	Red	Yellow	Green
Minimum luminous intensity [cd]	150	373	196
Luminous emittance at 100 m [lx]	.015	.0373	.0196
Right LEDs			
Luminous intensity [cd]	.035	.081	.047
Luminous emittance at 1.5 m [lx]	.0135	.036	.021
Left LEDs			
Luminous intensity [cd]	.033	.08	.051
Luminous emittance at 1.5 m [lx]	.0146	.0355	.018
Bulb color			
x – coordinate	0.583	0.213	0.688
y – coordinate	0.413	0.688	0.309

2.4 Experiment

The experiment was divided into three distinct sections. The first section consisted of measuring reaction times for each finger on a three-click mouse. The reaction times from the first section were used to normalize our response data for the different dexterities among participants and among fingers of each participant. The participant was instructed to place the middle finger of their dominant hand on the center button of the mouse, and their other two fingers on the surrounding buttons. The left button corresponded to red, middle to yellow, and right to green. The lights flashed exclusively on the right-hand side for this section of the experiment, with the participant's gaze fixated directly upon them. The participant was instructed to look at the lights on the right-hand side only and respond as quickly and accurately as possible to the color of these flashing lights. Each light would stay on for 3 s, with a 1 s break in between each individual flash. To ensure we had a good measure of each participants' true reaction time, this task would not stop until the previous 9 responses were all correct and within ~150 ms of each other. The participant carried out this task twice to ensure good data was collected. For data analysis, response times for each button were normalized by subtracting the average reaction time of the last 9 responses for that button.

The second section consisted of a single run of the same procedure as section one, but while wearing a lens. This section served as an adaptation period for the lens in question before the participant moved onto section three. Lens order was randomized using a python code, and the participant would perform sections two and three in succession before switching to the next lens.

The third section consisted of the experiment proper, in which the secondary task was utilized. We used Tetris as our secondary task, as it is a well-recognized visual-motor secondary task. The Tetris was displayed on a monitor behind the LED lights. The center of the game was aligned in the middle of the two sets of LED lights, and was located 7 cm below the level of the lights so as to be even with the participant's eye level. To control the Tetris, the participant used their non-dominant hand on the arrow keys of a standard QWERTY keyboard. In cases where the participant was left-handed but used a mouse with their right hand, they used their left hand on the keyboard. The participant was instructed to play Tetris and keep their gaze fixated straight ahead on the task. As they played, an LED bulb would flash on either the left or the right side, and the participant would respond to the color of the flash with the input device, and their reaction times were recorded. The lights flashed in random intervals of between 6 and 12 s and stayed on for 3 s. Failure to respond to the light in the 3 s that it stayed on was recorded as an error. Each run consisted of 24 randomized presentations of the LED lights, so that each side and color combination was presented 4 times. This was done 3 times for each of the three lenses, for a total of 72 data points for each lens per participant. The participant was given the opportunity to take a short rest in between each trial, but most opted to not take the break. The participants informed the experimenter immediately if they had made a mistake in responding to a light and these data points were not used in analysis. Observers were not given feedback about which lights were correctly or incorrectly identified.

3 Results

We conducted a mixed 2 (*color vision*) × 3 (*lens color*) ANOVA to analyze the interaction of reaction time and error rate. The normality assumption was checked graphically using histograms and Q-Q plots as well as by applying the Shapiro-Wilk test. Effect sizes are classified using Cohen's benchmark: small (η^2, ω^2 = .01), medium (η^2, ω^2 = .06), and large (η^2, ω^2 = .14). Table 2 summarizes the descriptive results for *color vision* and *lens color*.

Table 2. Minimum allowed luminous intensity of a traffic light as outlined by ITE.

Lens	Color-normal				Color-deficient			
	M_{RT} [s]	SD_{RT} [s]	M_{error} [-]	SD_{error} [-]	M_{RT} [s]	SD_{RT} [s]	M_{error} [-]	SD_{error} [-]
Dark Violet	0.695	0.368	0.031	0.054	0.729	0.432	0.195	0.233
PRIZM™ Road	0.695	0.394	0.030	0.042	0.761	0.452	0.211	0.215
PRIZM™ Golf	0.695	0.374	0.040	0.040	0.723	0.453	0.216	0.224

There were significant main effects of color vision on response times ($F_{1,\ 102} = 11.93$, $p < .001$, $\eta = .105$, $\omega = .095$) and on error rate ($F_{1,\ 102} = 39.53$, $p < .001$, $\eta = .279$, $\omega = .27$), but this was expected. There were no significant main effects of lens on response times ($F_{2,\ 102} = 0.12$, $p = .89$, $\eta = .002$, $\omega < .001$) or on error rates ($F_{2,\ 102} = .099$, $p = .91$, $\eta = .002$, $\omega < .001$), which you can see in Fig. 3. There were no significant two-way interactions between lens tint and color vision for reaction time ($F_{2,\ 102} = .11$, $p = .892$, $\eta^2 = .002$, $\omega^2 < .001$) or for error rate ($F_{2,\ 102} = 0.03$, $p = .968$, $\eta^2 = .002$, $\omega^2 < .001$). Additional Bonferroni post-hoc tests revealed that there were no significant differences in response times or error rate between individual lenses.

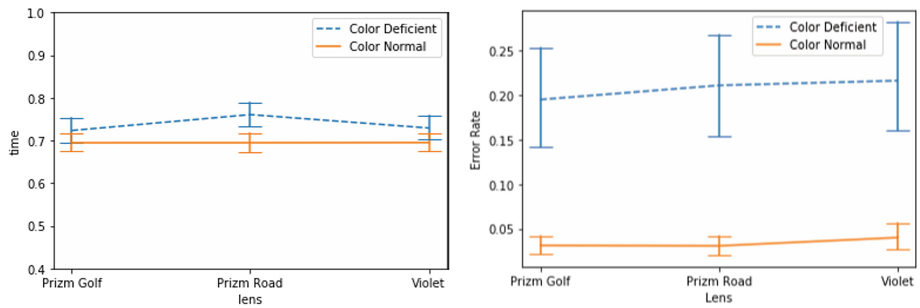


Fig. 3. Reaction times (left), error rates (right).

4 Discussion

There was no significant difference in mean reaction time or error rate across all lenses for color-normal participants. For color-deficient participants, although the difference was not statistically significant, the lens with the slowest mean adjusted reaction rates was the PRIZM™ Road lens, which currently passes the mentioned ANSI standards. PRIZM™ Golf (does not pass ANSI) yielded the fastest reaction times for color-deficient participants. Although the differences in these results were not statistically significant, the question remains as to whether they show any practical significance.

The fact that PRIZM™ Golf yielded the lowest mean adjusted reaction times raises questions about the validity of current standards regarding transmission rates.

The practical significance of our results is best illuminated by a simple rates problem. At 96.56 km/h (60 mph), it takes an average of 4.5 s and around 25 m to stop a car completely. According to our data, and by looking at reaction times (Fig. 3), the highest difference between lenses for color-deficient participants is $\sim$30 ms or 0.81 m travelled distance. Therefore, if a color-deficient person was driving a car while wearing a pair of PCE lenses, they might expect to stop for a sudden red light change 0.81 m later than they would have if wearing a neutral density lens. This distance, when taking into account all of the other external factors that can influence the braking distance of a car, would have minimal effects on braking in time for an intersection as opposed to ending up in the middle of it.

In conclusion, according to the data we gathered, we believe that Oakley PCE lenses offer solely a subjective change in our perception, as they don't appear to cause any significant decrease in reaction time or error rate. Especially in the light of rapidly changing transportation paradigms (Level 5 automation and human-robot interaction) and novel communication between traffic participants (e.g., via external HMIs), additional studies have to be conducted around lens technology and signal perception.

References

1. Hunt, R.W.G., Pointer, M.R.: Measuring Colour. Wiley, Hoboken (2011)
2. Dain, S.J., King-Smith, P.E.: Visual thresholds in dichromats and normals: the importance of post-receptoral processes. Vision. Res. **21**, 573–580 (1981)
3. Atchison, D.A., Pedersen, C.A., Dain, S.J., Wood, J.M.: Traffic signal color recognition is a problem for both protan and deutan color-vision deficients. Hum. Factors **45**(3), 495 (2003)
4. Dain, S.J., Wood, J.M., Atchison, D.A.: Sunglasses, traffic signals, and color vision deficiencies. Optom. Vis. Sci. **86**(4), 296–305 (2009)
5. Australian Standards Association: Australian Standard AS 1067-1990. Sunglasses and Fashion Spectacles. Part 1: Safety Requirements. Australian Standards Association, Sydney (1990)
6. American National Standards Institute: Ophthalmics: Nonprescription Sunglasses and Fashion Eyewear: Requirements: Z80.3-2018. American National Standards Institute, New York (2018)
7. Commission Internationale de l'Eclairage: Commission Internationale de l'Eclairage proceedings, 1931. Cambridge University Press, Cambridge (1932)
8. Standards Australia/Standards New Zealand: Australian/New Zealand Standard Traffic Signal Lanterns AS/NZS 2144:2002. Standards Australia Ltd., Sydney (2002)
9. Fisher, A.J., Cole, B.L.: The photometric requirements of vehicular traffic signal lanterns. In: Proceedings of the Australian Road Research Board, vol. 7, pp. 246–265. Australian Road Research Board, Melbourne (1974)
10. Institute of Transportation Engineers: Vehicle Traffic Signal Control Heads: Light Emitting Diode (LED) Circular Signal Supplement. Prepared by the Joint Industry and Traffic Engineering Council Committee (2005)

Semi-charmed Life - Willingness to Use and Related Contributing Factors Regarding Semi-public Charging Infrastructure for Electric Cars

Ralf Philipsen(✉), Imke Haverkämper, Hannah Biermann, Teresa Brell, and Martina Ziefle

Human-Computer Interaction Center (HCIC), RWTH Aachen University, Campus-Boulevard 57, 52074 Aachen, Germany
{philipsen, haverkaemper, biermann, brell, ziefle}@comm.rwth-aachen.de

Abstract. Particulate matter and greenhouse gas emissions are two of the main challenges of current mobility. The switch to electric drives, especially in motorized individual transport, could be part of the solution. However, the number of registrations of private battery electric vehicles (BEV) in Germany is currently increasing only slowly. Among other things, some user groups are currently difficult to address with electro-mobile offers, especially car drivers without their own parking space who would be dependent on public charging facilities. Semi-public charging stations, which are used by companies or authorities in certain periods of the day and are made available to the public outside these times, could be an additional offer here. Therefore, the aim of this paper was to better understand the users' requirements for semi-public charging stations and to identify possible trade-offs. The results revealed that the station's distance to the destination was the most important attribute followed by the temporal availability, the obligation to re-park, the costs, the surroundings, and finally the charging strategy. In contrast, user factors were of minor importance for the usage decision.

Keywords: Electromobility · Charging strategy · Semi-public charging · Conjoint analysis · Usage decision

1 Introduction

Although e-vehicles are one way to reduce greenhouse gas emissions (especially CO_2) from the transport sector [1], the registration figures in most European countries are so far negligible. In Germany, for example, only 0.3% of all registered cars are battery electric vehicles (BEV) [2]. In particular, private registrations are increasing only slowly, despite government subsidies. In addition to the technical parameters of the vehicle (range, charging time, costs), the availability of charging infrastructure is a key factor in a purchase or usage decision [3]. Therefore, it is not remarkable that private e-vehicle users are currently predominantly real estate owners who are able to establish

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 340–353, 2020.
https://doi.org/10.1007/978-3-030-50943-9_43

private charging facilities on their own property [4] and therefore use public charging infrastructure less often than private charging facilities [5]. In order to help electromobility achieve a broad market penetration, however, people in multi-family houses who may not have their own parking space should also be able to charge.

In addition to pure public charging stations, so-called semi-public stations, which are set up by companies on their own grounds and reserved for their employees or customers during business hours but made available to everyone outside those hours, could also increase the availability of charging options for the aforementioned target group. This raises many questions from the perspective of user acceptance: What effect does temporary availability have, especially if availability and own working time do not match perfectly? Are people willing to have the vehicle charged, e.g., overnight, in a more distant car park? What influence do possible intelligent charging strategies (throttling, scheduling) have? Can the establishment of (semi-)public charging infrastructure motivate a buying decision? And so on.

2 State of Research and Resulting Research Questions

Current research on electromobility from the user's perspective mainly focuses on two topics: On the one hand, the focus is on the general willingness to use or purchase e-vehicles, and on the other hand, all aspects of the charging process are studied in order to better plan the charging infrastructure and roll it out according to demand.

With regard to the willingness to use and purchase, numerous studies mention technical parameters, such as a long charging time or a short range compared to vehicles with combustion engines [6, 7], economic reasons in the sense of high acquisition costs [7–9], lack of model availability [6], and an inadequate charging infrastructure [9, 10] as impediments for adoption by potential users. In order to address the latter barrier, a widespread charging infrastructure network is required that is tailored to the needs of users, which requires a comprehensive knowledge of user requirements – and behavior.

In order to understand the charging process, it is first necessary to look at behavior patterns during charging and underlying triggers. The decision to charge is triggered in particular by the subjective assessment of the battery charge level and the remaining range of the vehicle, as well as by habits and routine with regard to places and times [11, 12]. Habits are more important for e-vehicle users than for users of vehicles with an internal combustion engine, while the reverse is the case for financial reasons triggering a refilling decision [12]. In terms of frequency, it can be seen that e-vehicle users recharge more frequently and tend to recharge smaller amounts, even if the battery is not yet empty, whereas in the fuel context, the majority of users fill up and run empty again [12].

Regarding the requirements for and the use of specific charging points, Krause et al. showed that in terms of distances travelled by car in Germany, public (fast-)charging infrastructure for trips undertaken for leisure and shopping purposes has the greatest potential, followed by business trips [13]. This coincides with surveys from the user's point of view, which, in addition to classic motorway service areas for long-distance journeys, considered workplaces and locations close to leisure and shopping facilities to be particularly important [14, 15]. The recommended charging speed to be offered depends on the duration of the stay in those areas, so slower speeds are possible at work or at home than, for example, along motorways [15]. About the evaluation of public charging stations, users particularly consider the (technical) reliability, the temporal and spatial availability [14], as well as the possibility to carry out other activities parallel to the charging process as relevant [14, 16]. Concerning the temporal and spatial availability, especially the necessary detours [5], the opening hours [14], and the waiting times before a charging station becomes available [16] have a decisive influence on the decision for or against the use of a charging location. Furthermore, the temporal availability must also be compatible with drivers' habits: at present, for example, fast charging points are used more frequently in the morning and afternoon [17], while slow charging speeds are mainly preferred during sleep and work [15]. Disadvantages of the above-mentioned charging station characteristics can be partially compensated for by financially attractive charging offers [16].

What previous research has in common is that it is addressing public charging infrastructure – usually with 24/7 availability – and does neither explore possible effects of limited temporal and spatial availability, which is inherent in the semi-public context, nor address different smart charging strategies like throttled or scheduled charging to relieve the power grid [18].

Therefore, based on the identified gaps in knowledge, the following research questions have emerged for the present paper:

1. *What influence do certain characteristics of a semi-public charging station, such as its position, availability, costs, or charging strategy, have on the user's decision for or against usage and what are the conflicting trade-offs?*
2. *What influence do user characteristics have on usage decisions and which profiles can be derived in terms of users and station ratings?*

3 Methodology

In order to address the research questions, an online questionnaire featuring a conjoint analysis was created. In the following, the structure of the questionnaire and the conjoint analysis' design will be described in detail. Subsequently, the data collection, processing, and analysis will be presented.

3.1 Questionnaire Design

The questionnaire consisted of three parts, the first gathering personal information like age, gender, education, household income, current status of employment, and working hours. Furthermore, questions concerning the personal need for control, the personal need for structure, and separation anxiety regarding one's car were asked. The second part dealt with mobility patterns in which the participants were asked if they had a driver's license, if they were dependent on a car in their everyday life, how high their annually driven mileage was, and how often they themselves refueled or charged their vehicle. The third part consisted of questions regarding the overall attitude towards (e-)

mobility, which included questions regarding one's interest in e-mobility, the intention to purchase an e-vehicle, the current usage of e-vehicles, and the possibility to charge a vehicle at one's home.

3.2 Conjoint Study

In order to study the participants selection process of semi-public charging stations, a choice-based conjoint analysis (CBC) was performed. With this method it is possible to investigate discrete decision making through choice tasks opposed to directly rating single features as it would be done in a conventional questionnaire design.

A two-scenario approach was chosen to consider possible differences in the selection process that might be related to the participants' schedules. The first scenario took place after a regular workday where the respondents had to charge their e-vehicle on their way home. The second scenario described them having a day off and looking for a charging station on their way to a leisure activity of their liking. To be able to better compare the two scenarios to each other, a repeated measures design was chosen. This way every participant had to perform several choice tasks for both scenarios. Furthermore, a forced-choice approach has been used to replicate the necessity to find a suitable charging station in real life.

The first attribute that was picked for the CBC analysis was the distance (for an overview of all attributes see Table 1). Note that the distance is defined indirectly by a given amount of time between 5 and 15 min it would take the participants to walk to the charging station. That approach was chosen because some people might interpret set amounts of kilometers as a longer or shorter detour than others. In addition to the walking distance the option to continue their travels with public transport has been inserted. The temporal availability attribute was divided into three shifts that resemble regular working hours and a control level with 24/7 availability was added. So far, most e-vehicle users have an option to charge their vehicle at home where they can leave it parked for as long as they like. On semi-public charging sites however, the drivers might have to re-park in order to allow someone else to use the charging device. Hence, an attribute was added to vary the obligation to re-park after a full charge or leave the vehicle at the station without any restrictions. As costs usually play an important role in purchase decisions, they were added as another attribute which included a discount option for slower charging, a surcharge in order not to have to re-park, and a basic price that was said to be cheaper than conventional fuels. Previous research has already focused on where to place charging stations. As petrol stations, leisure facilities, public buildings, and shopping facilities were proven to be suitable locations for charging stations [14], those options have been used as possible surrounding areas for semi-public charging stations. Since semi-public charging infrastructure is meant to support everyone's charging needs, it is expected to be rather busy. To avoid an overuse of the power grid, it might be necessary for the vehicles to be charged slower than technically possible. That's why the last attribute of the CBC was the charging strategy. The participants could choose between being notified after their car was fully charged within no set amount of time, having their car be fully charged within a specific period of time, having it be charged at a specific time, and the charging to be done as fast as possible.

Table 1. Attributes and levels used in the CBC

Attributes	*Distance*	*Temporal availability*	*Obligation to re-park*
Levels	Continue with public transport	24/7	Yes
	15-minute walk	10 pm to 6 am	No
	10-minute walk	6 am to 2 pm	
	5-minute walk	2 pm to 10 pm	

Attributes	*Costs*	*Surrounding*	*Charging strategy*
Levels	Discount for slower charging	Petrol station	Notification after variable end
	Surcharge for no re-parking obligation	Leisure facility	End in specified period
	Basic price (cheaper than fuel)	Public building	End at specified time
		Shopping facility	End as soon as possible

3.3 Data Collection, Processing and Analysis

The survey was conducted online. To address users of both conventional and electric vehicles, the survey was distributed via link within the social environment of the authors, on social media platforms, and on forums dealing with e-mobility. No prior knowledge on the technology was mandatory for participation and no financial incentives were given.

Responses that were incomplete, contained contradictory information, or were deemed unreliable due to a very short handling time or unbelievable claims about their age were removed. Furthermore, those who stated they did not have a driver's license were removed in order to focus on current and potential users of electric vehicles.

To analyze the conjoint data, the attributes' relative importance scores and their corresponding part-worth utilities were computed. For the classification of different profiles in the charging station evaluation two-step-clustering was used and optimized according to BIC to determine the number of clusters. The level of significance was set to $\alpha = .05$.

4 Sample

In total, $N = 147$ responses were included in the analysis. *88* (*59.9%*) of the participants were male and *59* (*40.1%*) were female. With a mean of *33.1* years ($SD = 14.0$), the respondents' age ranged from *17* to *71*. The sample consisted of rather highly educated people as *73* (*49.7%*) stated they had a university degree, another *59* (*40.1%*) had graduated high school, and *15* (*10.2%*) had completed secondary school. More than half of the participants were still in education or training ($n = 80$, *54.5%*), while *58* (*39.5%*) were employed, and *7* (*4.8%*) had already retired.

All participants had a driver's license, *78* (*53.1%*) of them declared that they were dependent on a car in their everyday life, while *69* (*46.9%*) were not. The sample expressed a very high opinion on e-vehicles in general ($M = 4.1$, $SD = 1.1$, *scale min* = *0*–*max* = *5*). Furthermore, the majority of respondents stated to have an interest in e-mobility: *23* (*15.6%*) said that they were already using an e-vehicle, *21* (*14.3%*) were interested in e-mobility and had the intention to purchase an e-vehicle, and *90* (*61.2%*) stated they were interested in e-mobility but did not currently intend on buying an e-vehicle. *13* (*8.8%*) people responded that they were not interested in e-mobility. As for charging, more than half of the participants (*n = 92, 62.6%*) expressed that they did not have the means to charge an e-vehicle at their home, while *55* (*37.4%*) stated that they had access to a parking spot with electricity.

5 Results

In the following, the relative importance of the charging station attributes for the usage decision in both scenarios is presented first. Subsequently, the individual levels of the attributes are considered in detail before the focus is on the influence of user factors on attribute perception. Finally, a first approach to clustering decision behavior is presented.

5.1 Relative Importances of Charging Station Attributes

Considering the relative importance of the charging station attributes for the usage decision, it was found that the two scenarios "working day" and "day off" do not differ in the order of the attributes (see Fig. 1). For both scenarios, the *distance* that still has to be travelled from the charging location to the respective destination is the most important attribute and accounts for almost *30%* of the user's decision, followed by the *temporal availability*, the *obligation to re-park*, the *costs*, the *surroundings*, and finally the *charging strategy*.

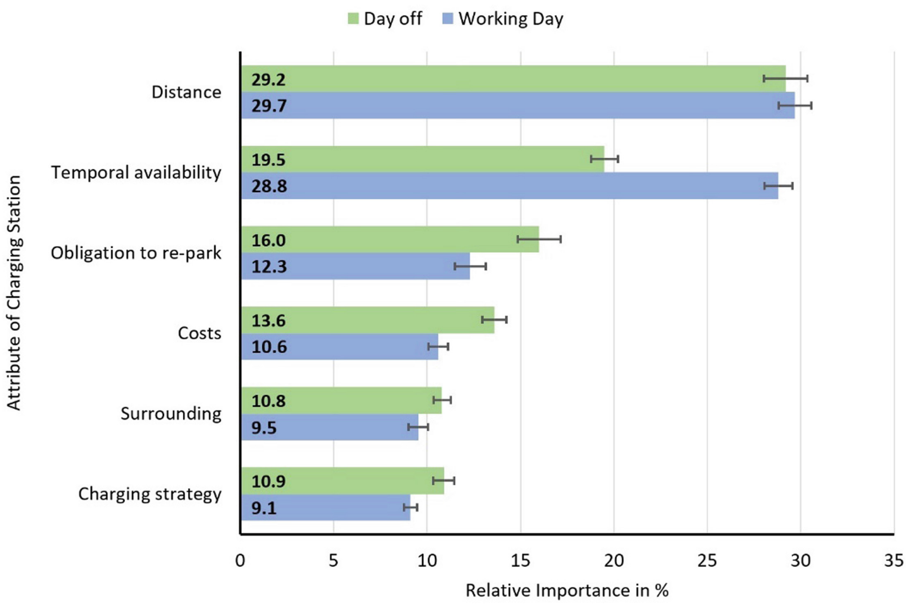


Fig. 1. Relative importances of charging station attributes in percent (means and standard errors) differentiated between a working day and a day off

Basically, the attributes differ from one another regarding their importance. Thus, significant main effects could be found for both the working day scenario ($F(5,730) = 180.0,\ p < .001,\ \eta_p^2 = .552$) and the day off scenario ($F(5,730) = 59.2,\ p < .001,\ \eta_p^2 = .289$). However, a closer look through pairwise post hoc tests showed that for the working day only two groups of attributes can be distinguished with regard to their importance: On the one hand, the *distance* and *temporal availability* with about *30%* importance each for the usage decision and, on the other hand, all other surveyed attributes with approx. $11 \pm 2\%$ share in the decision. With regard to the "day off", however, three attribute groups must be distinguished, which each differ significantly in their importance for the user: The group of highest importance is formed only by the attribute of *distance* (almost *30%* share), while the second group of medium importance consists of the attributes *temporal availability* and *obligation to re-park* (approx. $18 \pm 2\%$ share). The last group with the least importance for the decision-making process is formed by the attributes *costs*, *surroundings*, and *charging strategy* (approx. $12 \pm 2\%$ share). By comparing the two scenarios, it becomes apparent that the difference between them is mainly due to the attribute of temporal availability, which is significantly more important for the working day ($t(146) = 12.0,\ p < .001,\ d = 0.991,\ MD = 9.36\%$).

5.2 The Levels of Charging Stations' Attributes

The contribution of individual levels of charging station attributes to the usage decision can be found in Fig. 2. It was to be expected that levels that were themselves ordered hierarchically would also show values ordered hierarchically in terms of their positive contribution to decision-making. Accordingly, it was found, for example, that the shorter the walking *distance* from the charging station to the destination, the more positive the contribution to the usage decision was, while continuing the journey by public transport had the lowest value at all.

Furthermore, for attributes that had levels with the least possible restrictions for the user, these also had the most positive influence on the choice of charging station. This concerns, e.g., the *temporal availability* of the charging station, where opening around the clock was the preferred level, the *obligation to re-park*, where the lack of it was preferred, or the *charging strategy*, where the option to charge as fast as possible was the most likely to lead to a decision on use. Regarding the *surrounding* of the charging station, shopping facilities were favored by the users, while leisure facilities or public buildings had a less positive influence on the decision to use a location. Conventional petrol stations made the least contribution to the decision-making process. In terms of *costs*, the basic price and discounts for slower charging had roughly the same positive impact on the decision to use the facility, while surcharges to avoid re-parking were not favored. Concerning the *charging strategy*, it is apparent that, in addition to the above-mentioned preference for charging as fast as possible, for the "working day" the completion of charging at a specific point in time was more likely to contribute to the decision to use the charging station than an ending within a time period, whereas the opposite was true for the "day off". Analogous to the importance of the attributes, the greatest differences between working days and days off were to be found for the

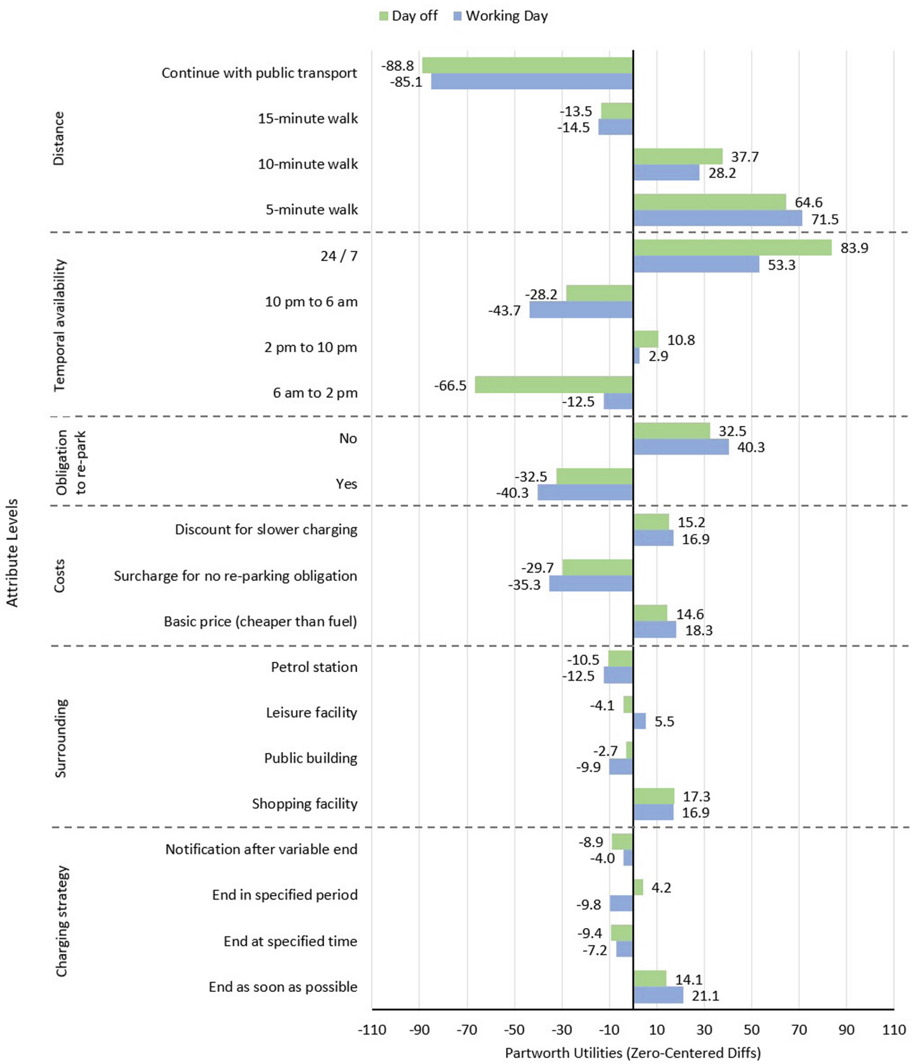


Fig. 2. Partworth utilities of charging stations' attribute levels (zero-centered) differentiated between a working day and a day off

temporal availability of the stations. While on working days especially the time slot between 10 pm and 6 am was unattractive for the user, on days off this was the case for an availability between 6 am and 2 pm.

5.3 The Influence of User Factors

Regarding the possible influence of user factors on the perceived importance of charging station attributes only few significant effects could be found, each affecting only one of the scenarios. For example, a missing possibility to charge at home had a positive influence on the perceived importance of the charging station *surrounding* in the working day scenario ($t(1,145) = 2.093,\ p = .038,\ d = 0.357,\ MD = 2.3\%$). The same applies to the type of drive currently in use. For e-vehicle users, the *surrounding* area of the station was also more important in the working day context than for users of internal combustion engine vehicles ($t(1,145) = 2.147,\ p = .033,\ d = 0.487,\ MD = 3.1\%$). Further on in the working context, it became apparent that participants who worked in shifts ascribed greater importance to the *charging strategy* ($t(1,7.31) = 2.515,\ p = .039,\ d = 1.407,\ MD = 5.7\%$). Also only in the working day scenario, a small but significant correlation between the age of the participants and the importance of the *obligation to re-park* existed: the older the participants, the less important the *obligation to re-park* for the usage decision ($r_s = -.226,\ p = .006$). With regard to gender, it was evident that for men in a day off context the *distance* of the station from the destination was more important than for women ($t(1,145) = -2.520,\ p = .013,\ d = -0.424,\ MD = -5.8\%$).

In contrast, no relationships could be found between the perceived importances of charging station attributes on the one hand and, on the other, the personal need for control, the personal need for structure, the attitude towards electric cars, the separation anxiety regarding the car, the annual mileage driven, the frequency of refueling or charging, the household income, or the users levels of education.

5.4 Patterns in the Stations' Attributes Influencing the Usage Decision

Based on a two-step clustering of the relative importance values of both scenarios, a distinction between two potential clusters was identified (*silhouette coefficient* $SC = 0.3$). The user groups identified in this way mainly differ in the perceived importance of the *distance* of the station to the destination and the *obligation to re-park* the car. While for one group short walking distances are decisive for the selection of a charging station, for the other group the decisive factor in the decision-making process is that they do not have to re-park their vehicles after the charging process (see Fig. 3). Although the evaluation patterns were clearly distinguishable, the use of binomial logistic regression to derive a predictive model for cluster membership of the participants based on the collected user characteristics could not identify a significant model.

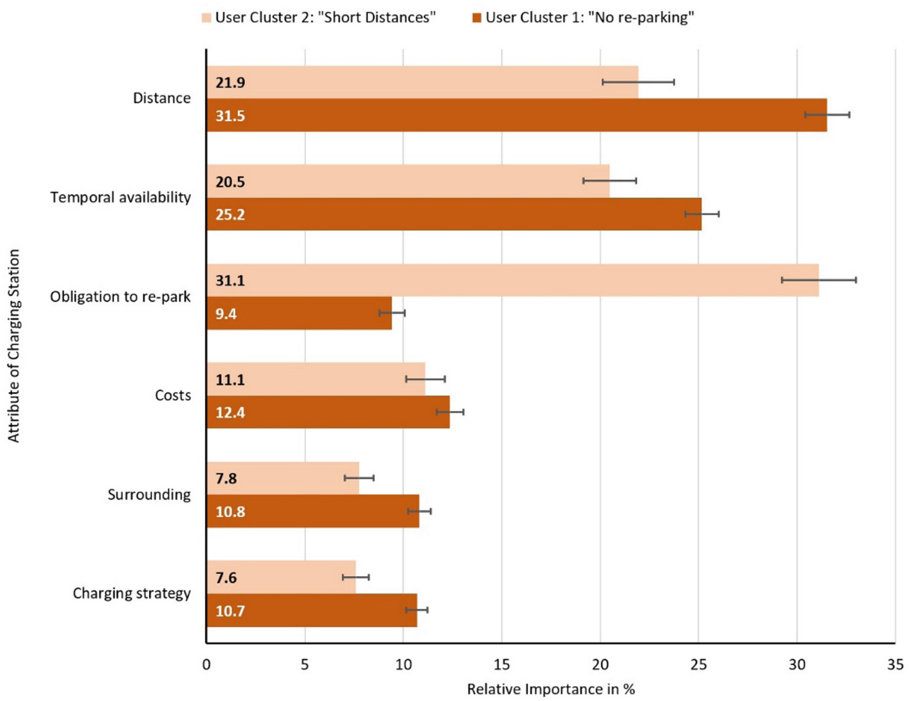


Fig. 3. Relative importances of charging station attributes in percent (means and standard errors) differentiated between identified clusters

6 Discussion

The present study carried out a composite measurement of the decision-making process in the use of semi-public charging infrastructure and was thus able to measure the importance of individual attributes of a charging station as perceived by the user in different scenarios.

First, it is noticeable that the perceived importance of the attributes for the two scenarios "working day" and "day off" hardly differ. In both cases, proximity to the destination is the decisive criterion for the usage decision. Disadvantages of other station characteristics can be compensated by a good position, while the opposite is only partially the case. The importance of distance fits into existing research results. Although the necessary walking distance to and from the charging station was not considered in detail before, distances between charging stations and optimal routes – in terms of necessary detours – had a high influence on the usage intention [5, 12].

Surprisingly, and in contrast to previous studies [12], costs were not the all-decisive criterion, but were even placed in the lower field of attributes in terms of their importance. One possible reason could be that the present study did not focus on absolute price levels, but rather on concrete discounts or surcharges to compensate for features of the charging service, whereas the basic price was predetermined to be a competitive offer. In order to derive business and price models, the surcharges or discounts could be explicitly monetized in future surveys to allow for an absolute classification of different tariffs. This would allow to determine whether there is a tipping point where financial aspects become more important. For example, free or very cheap offers were not considered in the present study, which, although not necessarily suitable as a business model, could make sense as an advertising measure for potential customers of a company. At present, however, based on the results presented, it can be stated that surcharges for additional services are rather not desired, while discounts for disadvantages have a more positive influence on the decision to use a charging offer.

Less surprising was the fact that temporal availability is also an important criterion, which is in line with previous research [14]. If the temporal availability does not match the daily routine of the user, it influences the willingness to use a charging station significantly, especially in the working day scenario. One must consider the temporal fit here in combination with the willingness to re-park. A lack of the latter has a more positive contribution to the decision on use and the users' willingness to re-park is generally to be regarded as low [14]. The present study showed that users are generally not willing to pay extra for the convenience of not having to re-park their vehicle, however, previous studies have shown that time-based fees amongst other factors can influence the willingness to re-park [19]. A low willingness to recharge can become a problem with semi-public charging infrastructure, as these stations may also be used by the operating company itself during business hours. Office hours or shop opening hours are not necessarily compatible with times when residents return from work. Although the present paper only considered subsequent re-parking after the charging process, re-parking for charging, e.g., if a charging station is only authorized for use at a later time, should also be subject to a comparable willingness. In this respect, further research is needed to better assess the impact of semi-public charging infrastructure, e.g., if it is available too late for the typical late afternoon commute home from work. According to the results of this study, compensation through favorable tariffs seems to be possible only to a limited extent.

In the context of temporal availability and willingness to re-park, the two identified evaluation patterns also become relevant. While one group rather refrains from long walking distances, for the other group the later re-parking is more important as a reason against the use of a charging station. The latter will be more difficult to address as potential customers due to the problems of temporal fit described above. However, in order to be able to address both groups with suitable offers, there must be a better understanding of what (causally) influences the two different evaluation patterns. The user characteristics examined in the present study had no influence on the affiliation to one or the other group. Therefore, it can be assumed that further, not yet identified user factors must exist as predictors. To this, it could also be meaningful to take a close look at the minority (identified in this study) who have expressed no interest in electro-mobility in order to explore their reasons why next to possible barriers to use with the potential to address them at the right time and thus possibly even establish motivation for use.

7 Limitations and Outlook

The main limitation of the present study and a possible explanation for the missing effects of user factors is certainly the sample size. Here, the required identifiable effect size is at least .287, which means that small effects would not be identified as significant. In particular, this may have contributed to the failed derivation of explanatory models for the decision profiles (see Sect. 5.4). In order to identify those models and to validate the results, the measurements may be repeated with a larger N. However, the present sample size is sufficient to exclude medium or larger effects as sufficiently unlikely. This simplifies business model development and infrastructure planning to the extent that more universal solutions are possible that do not have to differentiate between various user characteristics.

Another limitation concerns the presence of user characteristics in the sample, especially e-vehicle users. Although the establishing of semi-public charging infrastructure addresses in particular potential users who have not yet made a change in drive technology due to the lack of charging possibilities, a detailed analysis of existing e-vehicle users in future studies is necessary to better understand how such charging infrastructure is integrated into everyday mobility and what additional requirements may exist. To this, research could particularly address female e-vehicle users, their needs and demands in this context, especially since men, who were also proportionally more present in the gender distribution of this study, have already been identified as early adopters and thus considered more closely before [20].

The limitation to only two, albeit important, scenarios leaves questions unanswered regarding the generalizability to other contexts, such as specific trip purposes (shopping, bringing/taking people in/out, other errands). This could be particularly relevant regarding charging strategies, as time pressure can also prevail in non-work contexts, for example, which has not been addressed separately here. Therefore, a closer look at the context of the respective charging decision is necessary.

Due to the limitations and new questions that have arisen, this paper can only be a first step towards a more comprehensive understanding of the potential of semi-public charging infrastructure. Although the results already allow a relative evaluation of possible charging locations and charging services, the absolute effect of such infrastructure on the willingness to purchase or switch to electromobility, especially for residents without private parking space, cannot yet be estimated. This means that the present findings can already be used by infrastructure planners or mobility service providers to position and design charging options. For the development of prediction models for the distribution of e-vehicles, however, further research is needed regarding the potential and influence of semi-public charging infrastructure in distinction to private and public charging options.

Acknowledgements. This research was partly funded by the German Federal Ministry for Economic Affairs and Energy (Projects Cities in Charge, reference no. 01 MZ 18005C, and eMIND, reference no. 01 MZ 18015B). The authors thank all participants for sharing opinions on a novel technology. Furthermore, thanks go to those who see despite blindness for inspiring the paper title.

References

1. Chapman, L.: Transport and climate change: a review. J. Transp. Geogr. **15**, 354–367 (2007)
2. Kraftfahrt-Bundesamt: Fahrzeugzulassungen (fz) - Bestand an Kraftfahrzeugen nach Umwelt-Merkmalen - FZ 13. Statistische Mitteilungen des Kraftfahrt-Bundesamtes (2019). https://www.kba.de/SharedDocs/Publikationen/DE/Statistik/Fahrzeuge/FZ/2019/fz13_2019_pdf
3. Biresselioglu, M., Demirbag Kaplan, M., Yilmaz, B.: Electric mobility in Europe: a comprehensive review of motivators and barriers in decision making processes. Transp. Res. Part A Policy Pract. **109**, 1–13 (2018)
4. Trommer, S., Jarass, J., Kolarova, V.: Early adopters of electric vehicles in Germany unveiled. World Electr. Veh. J. **7**, 722–732 (2015)
5. Krause, J., Ladwig, S., Schwalm, M.: Better (not) charge in your garage! Perceived benefits and optimal positioning of public fast charging infrastructure for electrical vehicles from user's perspective. In: Advances in Human Factors, Sustainable Urban Planning and Infrastructure, pp. 261–269 (2017)
6. Haddadian, G., Khodayar, M., Shahidehpour, M.: Accelerating the global adoption of electric vehicles: barriers and drivers. Electr. J. **28**, 53–68 (2015)
7. Egbue, O., Long, S.: Barriers to widespread adoption of electric vehicles: an analysis of consumer attitudes and perceptions. Energy Policy. **48**, 717–729 (2012)
8. Sierzchula, W., Bakker, S., Maat, K., van Wee, B.: The influence of financial incentives and other socio-economic factors on electric vehicle adoption. Energy Policy **68**, 183–194 (2014)
9. Graham-Rowe, E., Gardner, B., Abraham, C., Skippon, S., Dittmar, H., Hutchins, R., Stannard, J.: Mainstream consumers driving plug-in battery-electric and plug-in hybrid electric cars: a qualitative analysis of responses and evaluations. Transp. Res. Part A PolicyPract. **46**, 140–153 (2012)
10. Halbey, J., Kowalewski, S., Ziefle, M.: Going on a road-trip with my electric car: acceptance criteria for long-distance-use of electric vehicles. In: Design, User Experience, and Usability: Interactive Experience Design, pp. 473–484 (2015)
11. Daina, N., Polak, J., Sivakumar, A.: Patent and latent predictors of electric vehicle charging behavior. Transp. Res. Rec. J. Transp. Res. Board **2502**, 116–123 (2015)
12. Philipsen, R., Brell, T., Biermann, H., Eickels, T., Brost, W., Ziefle, M.: Should I stay or should I go? - influencing context factors for users' decisions to charge or refuel their vehicles. In: Advances in Human Factors of Transportation, pp. 573–584 (2019)
13. Krause, J., Ladwig, S., Saupp, L., Horn, D., Schmidt, A., Schwalm, M.: Perceived usage potential of fast-charging locations. World Electr. Veh. J. **9**, 14 (2018)
14. Philipsen, R., Schmidt, T., van Heek, J., Ziefle, M.: Fast-charging station here, please! User criteria for electric vehicle fast-charging locations. Transp. Res. Part F Traffic Psychol. Behav. **40**, 119–129 (2016)
15. Anderson, J., Lehne, M., Hardinghaus, M.: What electric vehicle users want: real-world preferences for public charging infrastructure. Int. J. Sustain. Transp. **12**, 341–352 (2018)
16. Philipsen, R., Schmidt, T., Ziefle, M.: Well worth a detour? - Users' preferences regarding the attributes of fast-charging infrastructure for electromobility. In: Advances in Intelligent Systems and Computing, pp. 937–950 (2016)
17. Morrissey, P., Weldon, P., O'Mahony, M.: Future standard and fast charging infrastructure planning: an analysis of electric vehicle charging behaviour. Energy Policy **89**, 257–270 (2016)

18. Stroehle, P., Becher, S., Lamparter, S., Schuller, A., Weinhardt, C.: The impact of charging strategies for electric vehicles on power distribution networks. In: 2011 8th International Conference on the European Energy Market (EEM) (2011)
19. Wolbertus, R., Gerzon, B.: Improcing electric vehicle charging station efficiency through pricing. J. Adv. Transp. **2018**, 1–11 (2018)
20. Frenzel, I., Jarass, J., Trommer, S., Lenz, B.: Erstnutzer von Elektrofahrzeugen in Deutschland. Nutzerprofile, Anschaffung, Fahrzeugnutzung [First users of electric vehicles in Germany. User profiles, acquisition, vehicle use]. Deutsches Zentrum für Luft- und Raumfahrt e. V. (DLR), Berlin (2015)

An Investigation of Traffic Noise Levels Around a Major Hospital in Qatar

Khaled Shaaban(✉), Abdelrahman Abouzaid, Ahmad Musleh, and M. Fares Hout

Qatar University, Doha, Qatar
{kshaaban,aa1510753,am1510253,mh1508518}@qu.edu.qa

Abstract. Although hospitals are sensitive to noise and are considered silent zones, many are located in urban areas and are subject to high levels of noise. In this study, an assessment of traffic noise, around a major hospital in the city of Doha, Qatar, was conducted during the morning and evening periods. The results indicated that the noise values exceeded the World Health Organization's reference value of 50 dBA for both morning and evening times, and were higher than the 55 dBA limit at which serious annoyance is generated. Additionally, it exceeded the Qatar local standards for both daytime (55 dBA) and nighttime (45 dBA) noise levels. The results indicated that one of the main sources of noise was road traffic. Different solutions were proposed for a possible action plan to reduce the high level of noise.

Keywords: Traffic noise · Silence zones · Noise pollution · Road traffic

1 Introduction

Noise annoyance is a feeling of displeasure, disturbance, or irritation, and has a negative effect on an individual or a group of individuals. Constant exposure to irritating noise can affect people's health [1, 2]. Therefore, concerns regarding noise pollution and its effects have been increasing. As a result, different studies have been conducted on the impact and prevention of noise pollution [3]. Hospitals are considered noise-sensitive buildings due to the presence of hospitalized patients [4]. This sensitivity to noise is taken into account by different international organizations that have recommended specific maximum noise values for these types of buildings.

The World Health Organization (WHO) proposed a daytime and nighttime threshold of 30 dBA for forward rooms in hospitals due to the health effects that can result from sleep disturbances, which can affect patients' rest and recovery [1]. Furthermore, the WHO established that 50 dBA is considered a moderate daytime and evening annoyance level in general outdoor living areas; values exceeding 55 dBA are regarded as a serious annoyance. Some studies that measured noise around hospitals in different countries found that noise exceeded the WHO recommendations as well as local regulations [5, 6].

In the case of hospitals, noise pollution can cause several problems, including extended hospital stays, increased dosages of pain medication, and irritability [7, 8]. Moreover, high noise levels have been found to have an impact on the stress level of

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 354–360, 2020.
https://doi.org/10.1007/978-3-030-50943-9_44

workers, which affects their job performance, and can even lead to workers changing jobs in some cases [9, 10]. To prevent noise annoyance, a good understanding of the noise levels in an affected area is needed. The purpose of this study is to evaluate the noise levels around Al-Ahli Hospital, a major hospital in Qatar, compare the noise levels to traffic volumes, and compare the measured sound levels with international and national regulations.

2 Data Collection

This study was conducted in the State of Qatar. The local standards in Qatar established a daytime threshold of 55 dBA and a nighttime threshold of 45 dBA for an allowable limit of noise throughout the day. Al-Ahli Hospital, located in the city of Doha, Qatar, is one of the major hospitals in Qatar (see Fig. 1). The city is known for car dependency, heavy traffic, and limited use of public transportation [11–14]. The hospital is located on a major road and is surrounded by commercial businesses.

Fig. 1. The layout of Al-Ahli hospital showing the data collection locations.

To investigate the noise around the hospital, noise measurements were conducted at six locations, including the main road, main entrance, outpatient clinic, emergency entrance, parking garage entrance, and the inpatient entrance, as shown in Fig. 1. These locations were selected to have coverage of all the main locations around the hospital. At the first location, the measurements were taken in front of the main road of the hospital. The main road is a major arterial that consists of three lanes of traffic and a raised median. A service road exists between the main entrance and the main road. At the second location, the measurements were taken at the main entrance. At the third location, the measurements were taken in front of the outpatient clinic entrance. This entrance is located in the corner of two one-way one-lane roads. There is a passenger loading area directly in front of the outpatient clinic, and a parking area exists in front of the outpatient clinic.

At the fourth location, the measurements were taken in front of the emergency entrance. There are a two-way two-lane road and a bus station in front of the entrance. There is also a drop-off/pick-up area in front of the emergency entrance for temporary parking. At the fifth location, the measurements were taken in front of the underground parking garage. There is a one-way one-lane road leading to the parking garage, and there is also a drop-off/pick-up area in front of another entrance to the hospital. For the sixth location, the measurements were taken in front of the inpatient entrance. There is a roundabout and a drop-off/pick-up area at this location as well.

The noise measurements were conducted twice at each location during the morning and evening peak periods for one hour using an Optimus sound level meter (CR 1720). The calibration was done for the Optimus sound level meter before and after conducting the measurement. The microphone was placed 1.5 m above the ground as per the standard, and the noise intensity was measured in decibels (dBA). An anemometer was also used to measure the temperature, wind speed, humidity, etc. A video camera was used to record the traffic at each location in order to measure traffic volume count and the percentage of heavy vehicles.

3 Analysis

All recorded values were higher than the recommended WHO values and the local standards, as indicated in Table 1. The first location, the main road, recorded a noise level of 71.9 dBA in the morning and 69.6 dBA in the evening. The second location, the main entrance, recorded a noise level of 65.1 dBA in the morning and 66.2 dBA in the evening. It should be noted that an electronic gate for the parking area entrance right after the main entrance makes noise when vehicles enter the parking area. The third location, the outpatient clinic, recorded a noise level of 61.7 dBA in the morning and 58.6 dBA in the evening. The fourth location, the emergency entrance, recorded a noise level of 62.6 dBA in the morning, and 65.2 dBA in the evening. The fifth location, the parking garage entrance, recorded a noise level of 60.0 dBA in the morning and 59.8 dBA in the evening. The sixth and final location, the inpatients' entrance, recorded a noise level of 65.4 dBA in the morning and 61.7 dBA in the evening.

As shown in Fig. 2, the hospital is exposed to high noise levels that exceed both the recommended values of the WHO standards and the Qatar local standards. The highest noise level was observed at the main road, which had a value of 71.9 dBA followed by the emergency entrance (69.6 dBA).

Table 1. Daytime and nighttime noise measurement.

Intensity (dBA)	Daytime						Nighttime					
	1	2	3	4	5	6	1	2	3	4	5	6
LAeq	71.9	65.1	61.7	62.6	60.0	65.4	69.6	66.2	58.6	65.2	59.8	61.7
LAFmax	109.6	89.4	84.3	85.7	75.0	97.1	88.9	97.1	75.5	98.1	87.9	93.2
LAF1	77.6	74	67	73.1	65.9	71.5	75.5	78.0	65.4	77.3	68.6	67.4
LAF5	74.4	67.5	64.1	67.4	62.9	66.9	72.8	66.0	61.5	69.8	64.1	64.1
LAF10	73.2	66.1	63.4	64.8	61.5	65.6	71.9	62.9	60.2	65.3	61.8	63.0
LAF50	69.5	63	60.9	58.2	59.1	62.5	68.4	58.6	57.3	58.0	56.9	60.0
LAF90	63.9	59.7	58.5	55.8	58.2	60.1	63.2	56.0	55.4	55.0	55.2	57.0
LAF95	62.4	58.2	57.9	55.4	57.8	59.5	61.4	55.4	54.7	54.5	54.9	56.2
LAF99	59.6	54.4	56.6	54.8	56.3	57.9	58.0	54.0	53.7	53.8	54.3	54.5

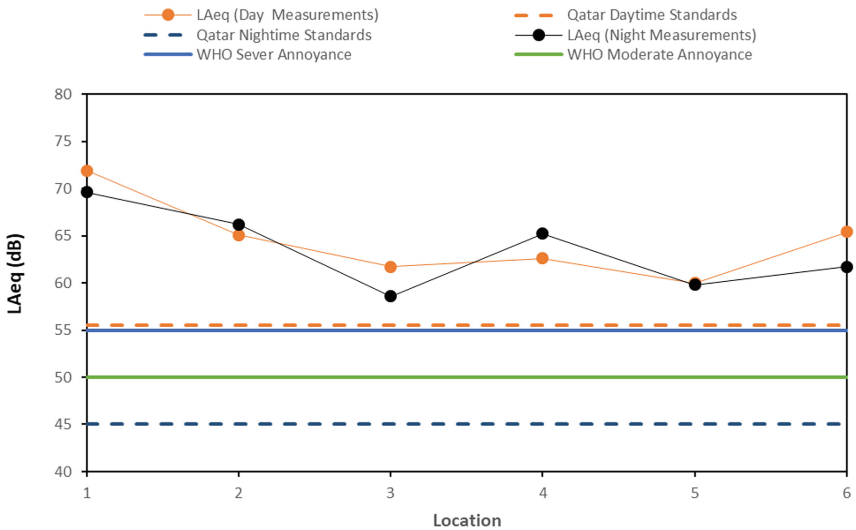


Fig. 2. LAeq for all locations, Qatar local standards, and WHO standards

3.1 Traffic Volumes

The vehicular traffic was recorded using a video camera at all six locations, and the traffic volume was extracted, as shown in Table 2. The highest traffic volumes were observed at the main road followed by the emergency entrance. A scatter plot of LAeq versus the traffic volume was plotted to investigate the relationship between the two. As shown in Fig. 3, the noise level increases with the increase of traffic volume, indicating a strong relationship between the two.

Table 2. Daytime and nighttime traffic volumes.

Vehicle type	Daytime						Nighttime					
	1	2	3	4	5	6	1	2	3	4	5	6
Motorcycles	12	1	0	0	0	0	12	1	0	2	1	1
Passenger cars	733	108	92	196	62	371	639	108	61	231	46	267
Pickup/van	63	6	4	8	0	10	60	1	4	7	2	2
Buses	20	0	0	6	0	0	18	0	0	10	0	0
Truck	43	3	0	12	1		17	1	0	6	3	0
Total	871	118	96	222	63	381	746	111	65	256	52	270

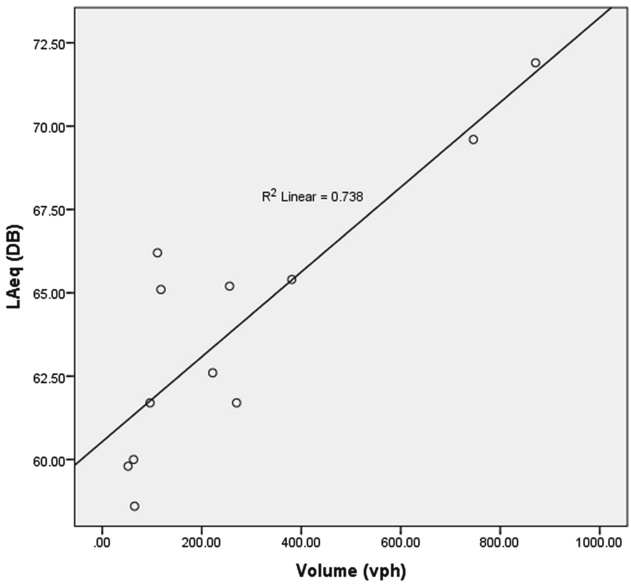


Fig. 3. LAeq for all locations, local standards, and WHO standards

4 Conclusion

This study was conducted to investigate the noise levels around a major hospital in Doha, Qatar. The results indicated that the hospital was exposed to high noise levels. The obtained values confirm that the noise levels exceeded the recommended values of both the WHO standards and the Qatar local standards. The noise levels exceeded the threshold of 55 dBA established by the WHO standards for severe annoyance during the daytime and evening. In addition, it also exceeded the threshold of 55 dBA established by the Qatar local standards for the daytime, and 45 dBA established by the Qatar local standards for the nighttime. It is clear that the road traffic was the main cause of the high noise levels, and the location with the highest noise level detected in this study was the main road. Hence, actions need to be taken in order to reduce the impact of irritating noise levels on hospitals, since they are considered silent zones.

Further, many possible solutions can be implemented to minimize the noise issue in hospitals located in urban areas. During the planning and design phases, major roads should be avoided, and neighborhoods with high potential for walking and cycling should be considered [15–18]. For existing hospitals, soundproof walls [19, 20] and acoustic barriers are the most common solution to reduce irritating noises, such as road traffic [21, 22]. Finally, using systems that monitor, measure, and evaluate noise levels around hospitals such as wireless sensors are needed for continuous noise monitoring and to identify problems on a regular basis and develop solutions [23–25].

Acknowledgements. This publication was made possible by a UREP award [UREP 24-011-2-005] from the Qatar Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the author.

References

1. World Health Organization: Burden of disease from environmental noise: quantification of healthy life years lost in Europe (2017)
2. Beutel, M.E., et al.: Noise annoyance is associated with depression and anxiety in the general population-the contribution of aircraft noise. Plos ONE **11**(5) (2016)
3. Anees, M.M., Qasim, M., Bashir, A.: Physiological and physical impact of noise pollution on environment. Earth Sci. Pak. **1**(1), 08–11 (2017)
4. Giv, M.D., et al.: Evaluation of noise pollution level in the operating rooms of hospitals: a study in Iran. Interv. Med. Appl. Sci. **9**(2), 61 (2017)
5. Khaiwal, R., et al.: Assessment of noise pollution in and around a sensitive zone in North India and its non-auditory impacts. Sci. Total Environ. **566**, 981–987 (2016)
6. Loupa, G., et al.: Mapping the noise in a Greek general hospital. Sci. Total Environ. **646**, 923–929 (2019)
7. Montes-González, D., et al.: Environmental noise around hospital areas: a case study. Environments **6**(4), 41 (2019)
8. Schueler, K., et al.: Detecting parallel moving vehicles with monocular omnidirectional side cameras. In: 2013 IEEE Intelligent Vehicles Symposium (IV). IEEE (2013)
9. Kramer, B., Joshi, P., Heard, C.: Noise pollution levels in the pediatric intensive care unit. J. Crit. Care **36**, 111–115 (2016)
10. Applebaum, D., et al.: The impact of environmental factors on nursing stress, job satisfaction, and turnover intention. J. Nurs. Admin. **40**, 323 (2010)
11. Shaaban, K.: Why don't people ride bicycles in high-income developing countries, and can bike-sharing be the solution? The case of Qatar. Sustainability **12**(4), 1693 (2020)
12. Shaaban, K.: Assessing sidewalk and corridor walkability in developing countries. Sustainability **11**(14), 3865 (2019)

13. Shaaban, K., et al.: Assessing the impact of converting roundabouts to traffic signals on vehicle emissions along an urban arterial corridor in Qatar. J. Air Waste Manag. Assoc. **69** (2), 178–191 (2019)
14. Shaaban, K., Maher, A.: Using the theory of planned behavior to predict the use of an upcoming public transportation service in Qatar. Case Stud. Transp. Policy **8**(2), 484–491 (2020)
15. Shaaban, K., Khalil, R., Muley, D.: An assessment tool to test the complete streets policy in Qatar. Int. J. Sustain. Soc. (2018)
16. Shaaban, K., et al.: A simple method to assess walkability around metro stations. Int. J. Sustain. Soc. **10**(1), 1–19 (2018)
17. Shaaban, K., Muley, D., Elnashar, D.: Evaluating the effect of seasonal variations on walking behaviour in a hot weather country using logistic regression. Int. J. Urban Sci. **22**(3), 382–391 (2018)
18. Shaaban, K., Muley, D.: Investigation of weather impacts on pedestrian volumes. Transp. Res. Procedia **14**, 115–122 (2016)
19. Park, M., Cheon, S., Choi, H.: The study of effectiveness of the soundproof wall to reducing road noise in Pusan City. Audiology **8**(2), 219–223 (2012)
20. Hongisto, V., Oliva, D., Rekola, L.: Subjective and objective rating of the sound insulation of residential building façades against road traffic noise. J. Acoust. Soc. Am. **144**(2), 1100–1112 (2018)
21. González, D.M., et al.: Acoustic screening effect on building façades due to parking lines in urban environments. Effects in noise mapping. Appl. Acoust. **130**, 1–14 (2018)
22. Ow, L.F., Ghosh, S.: Urban cities and road traffic noise: reduction through vegetation. Appl. Acoust. **120**, 15–20 (2017)
23. Zambon, G., et al.: Monitoring and prediction of traffic noise in large urban areas. Appl. Sci. **8**(2), 251 (2018)
24. Fredianelli, L., Del Pizzo, A., Licitra, G.: Recent developments in sonic crystals as barriers for road traffic noise mitigation. Environments **6**(2), 14 (2019)
25. Bello, J.P., et al.: Sonyc: a system for monitoring, analyzing, and mitigating urban noise pollution. Commun. ACM **62**(2), 68–77 (2019)

Analysis of Driving Performance Data Considering the Characteristics of Railway Stations

Daisuke Suzuki[1(✉)], Ayako Suzuki[1], Keiko Shimano[2], Kazuki Kiyota[2], and Yutaka Kakizaki[2]

[1] Ergonomics Laboratory, Railway Technical Research Institute, 2-8-38 Hikari-cho, Kokubunji-shi, Tokyo, Japan
suzuki.daisuke.55@rtri.or.jp
[2] Transportation and Marketing Department, Central Japan Railway Company, 1-3-4, Meieki, Nakamura-ku, Nagoya, Japan

Abstract. Here, we aim to investigate the relationship between characteristics of railway stations and errors in train stop positions. Hence, two kinds of logistic regression analysis were conducted with two different objective variables: train stations with or without the occurrence of delays in braking manipulations and stations with or without the occurrence of misrecognition of stop positions. The explanatory variables included velocities near stations, braking manipulations, and features of the stations. Logistic regression analysis revealed that the delay in braking manipulations was significantly associated with the ratio of the maximum brake notch and the mean of velocities at 200 m before train stops. The delay in braking manipulations occurred frequently at stations where the train velocities when approaching the stations were high and the maximum brake notch was frequently used. Logistic regression analysis further revealed that the misrecognition of stop positions was significantly associated with the existence of a stop sign for four or six vehicles and where there were many stopping velocity patterns. A stop sign before a stop position and decreasing train velocity for a caution signal caused the misrecognition of stop positions.

Keywords: Driving performance data · Characteristics of railway stations · Delay in braking manipulations · Misrecognition of stop positions · Logistic regression analysis

1 Introduction

Driving performance data from railway vehicles are often saved in data-recording devices. These data can be used to analyze accidents, prevent human error, and improve crew skills. In addition, several studies have been conducted with data describing the driving performance of railway drivers.

Sakashita et al. [1] analyzed more than 20 elements of driving performance data from approximately 100 drivers recorded over six months. Their study used the following evaluation indices related to braking operations for stopping a train at a station: (1) duration of strong braking (seven notches) used when stopping, (2) number of

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 361–366, 2020.
https://doi.org/10.1007/978-3-030-50943-9_45

additional brakes stronger than three notches used when stopping, (3) amount of brake notch changes used when stopping, (4) number of times the coasting position was used when stopping, and (5) train speeds at a fixed point (five seconds before the train stops). It was demonstrated that the use of strong braking (1) and many brake notch changes (3) strongly influenced whether the drivers experienced train position errors when stopping at stations. Moreover, it was found that the train speed at a fixed point five seconds before the train stopped varied widely prior to the occurrence of an error. Marumo et al. [2] analyzed braking manipulations of railway drivers in a driving simulator to estimate the mental conditions of the drivers. The results showed that the relationship between the velocity deviation and the braking manipulations is an appropriate indicator of abnormal driving behavior. In another study, Marumo et al. [3] analyzed braking manipulations when stopping at stations and found that simultaneously performing secondary tasks was associated with significantly greater variance in operation of the brake handle. We also previously analyzed driving performance data to investigate how brake manipulations are related to train stop position errors at a station [4]. In our previous logistic regression analysis, the objective variable was whether or not the railway drivers had experienced train stop position errors.

Despite these studies of factors influencing the driving performance of railway drivers, the relationship between characteristics of railway stations and the occurrence of train stop position errors at stations has not been studied in depth. Therefore, the aim of this study is to investigate how railway station characteristics are related to the train stop position errors.

2 Methods

2.1 Analysis Data

Data concerning driving operations (i.e., distance, velocity, and brake notches) when stopping at railway stations were extracted from driving performance data recorded by railway vehicles in Japan over three years. The sampling rate was 1 Hz, and the number of stops at 81 total stations was approximately 1.3 million. A total of 118 railway drivers aged 25 to 61 years (mean age, 33 years; standard deviation (SD), nine years) participated in the current study. The driving experience of the participants ranged from 0 to 29 years (mean, seven years; SD, seven years).

2.2 Classification of Train Stop Position Errors

Eighty-nine train stop position errors at stations were analyzed. The reasons for the errors were divided into two factors: a delay in braking manipulations or a misrecognition of stop positions. The number of delays in braking manipulations was 29, and the number of misrecognitions of stop positions was 60.

2.3 Evaluation Indices and Analysis Method

Evaluation indices for railway station characteristics were the number of stop signs; the existence of stop signs for four vehicles, six vehicles, eight vehicles, or ten vehicles; the number of stopping velocity patterns; and the slopes near stations. In reference to the previous study [1], evaluation indices for driving performance data at each station were as follows: train velocity of 20 m before the train stop, train velocity of 200 m before the train stop, the ratio of using maximum brake notches, the ratio of using additional brake notches, and the number of switching brake notches.

A logistic regression analysis was conducted to identify the influence of the characteristics of railway stations on the occurrence of a train stop position error at a station. Two kinds of objective variables were analyzed: stations with or without the occurrence of delays in braking manipulations and stations with or without the occurrence of misrecognitions of stop positions. The explanatory variables included the station characteristics evaluation indices and the driving performance data at each station.

3 Results

3.1 Delays in Braking Manipulations

Table 1 shows the results of the logistic regression analysis; the objective variable was a station with or without the occurrence of a delay in braking manipulations. Variance-inflation factors (VIFs) were calculated to evaluate the multicollinearity of this study. Generally, a possibility of multicollinearity exists if the maximum VIF is above 10 or the average VIF is considerably more than one [5]. In this case, the maximum VIF among the explanatory variables used in this study was 1.35; thus, it was concluded that multicollinearity was not an issue for this analysis.

A p value < 0.05 indicates that the occurrence of a train stop position error at a railway station is significantly associated with the explanatory variable. The occurrence of delays in braking manipulations was significantly associated with the ratio of using maximum brake notches and mean of velocities at 200 m before the train stop.

Furthermore, evaluating the odds ratios revealed that when other variables do not change, the possibility of delays in braking manipulations increases by 7.62 times when the ratio of using maximum brake notches increases by 1%; similarly, it increases by 1.32 times when the mean of velocities at 200 m before the train stop increases by 1 km/h.

Standardized partial regression coefficients were determined as shown in Table 1. The absolute value of the standardized partial regression coefficient was directly proportional to the influence of the explanatory variable. Thus, the mean of velocities at 200 m before the train stop influenced the possibility of delays in braking manipulations more strongly.

Table 1. Logistic regression analysis with delays in braking manipulations as the objective variable

Explanatory variables	Standardized partial regression coefficient	Odds ratio	*p* value
Ratio of using maximum brake notch	0.78	7.62	$p < 0.05$
Mean of velocities at 200 m before the train stop	1.30	1.32	$p < 0.05$

3.2 Misrecognitions of Stop Positions

Table 2 shows the results of the logistic regression analysis with the objective variable as a station with or without the occurrence of misrecognition of stop positions. In this case, the maximum VIF among the explanatory variables used in this study was 1.97; thus, it was concluded that multicollinearity was not an issue for this analysis.

The *p* values revealed that the occurrence of misrecognitions of stop positions was significantly associated with the existence of a stop sign for four or six vehicles and with the number of stopping velocity patterns.

Furthermore, evaluating the odds ratios revealed that when other variables do not change, the possibility of misrecognitions of stop positions increases by 8.50 times when a stop sign for four vehicles exists and 3.88 times with a stop sign for six vehicles. Moreover, when other variables remain unchanged, the possibility of misrecognitions of stop positions increases by 2.09 times when the number of stopping velocity patterns increases by one pattern.

Standardized partial regression coefficients revealed that among these three variables, the existence of a stop sign for four vehicles had the largest influence on the possibility of misrecognitions of stop positions.

Table 2. Logistic regression analysis with misrecognitions of stop positions as the objective variable

Explanatory variables	Standardized partial regression coefficient	Odds ratio	*p* value
Existence of a stop sign for four vehicles	1.06	8.50	$p < 0.05$
Existence of a stop sign for six vehicles	0.68	3.88	$p < 0.05$
Number of stopping velocity patterns	0.53	2.09	$p < 0.05$

4 Discussion

Logistic regression analysis performed herein revealed that the delays in braking manipulations were significantly associated with the ratio of maximum brake notch and the mean of velocities at 200 m before the train stop. Delays in braking manipulations occurred frequently at stations where the train velocities approaching the stations were high and the maximum brake notch was frequently used. In hearings after delays in braking manipulations occurred, several railway drivers stated that when they braked, the velocity did not decrease as they expected. High velocities at 200 m before the train stop required strong braking to stop the train at the stop position. It is thought that when the velocity at 200 m before the train stop is high and the braking effect is weak, a delay in braking manipulations occurs.

The misrecognitions of stop positions were significantly associated with the existence of a stop sign for four or six vehicles and the number of stopping velocity patterns. In hearings after misrecognitions of stop positions occurred, several railway drivers stated that they mistakenly stopped at the nearest stop sign because passengers lined up at the stop sign for four or six vehicles. If drivers drove railway vehicles more than six, they have to stop trains at more far stop sign. It is therefore possible that a stop sign before a stop position causes the misrecognitions of stop positions. In the same hearings, several railway drivers stated that they misrecognized stop positions because they decreased train velocity for a caution signal aspect and released the brake. Releasing the brake confused drivers to stop trains near or far. Thus, velocity slowing patterns may contribute to the misrecognitions of stop positions.

It is thought that picking up stations that require special attention and sharing them with drivers decreases the possibility of train stop position errors. Further study is required to examine the influence of differences in the features of adjacent stations.

5 Conclusion

In this study, logistic regression analysis on train stop position error was conducted to investigate the relationship between characteristics of railway stations and train stop position errors at stations. Two kinds of objective variables were analyzed: stations with or without the occurrence of delays in braking manipulations and stations with or without the occurrence of misrecognitions of stop positions. The occurrence of delays in braking manipulations was significantly associated with the ratio of using maximum brake notch and mean of velocities at 200 m before the train stop. Delays in braking manipulations occurred frequently at stations where the train velocities approaching the stations were high and the maximum brake notch was frequently used. The occurrence of the misrecognitions of stop positions was significantly associated with the existence of a stop sign for four or six vehicles and the number of stopping velocity patterns. A stop sign before a stop position and decreasing train velocity for a caution signal aspect caused the misrecognitions of stop positions.

References

1. Sakashita, O., Shimano, K.: Statistical analysis of big data for enhancement of train driving skills. Jpn. Railw. Eng. **2018**(201), 5–8 (2018)
2. Marumo, Y., Tsunashima, H., et al.: Evaluation of braking behaviour for train drivers using phase-plane trajectories. Vehicle Syst. Dyn. **46**(Supplement), 729–735 (2008)
3. Marumo, Y., Tsunashima, H., et al.: Analysis of braking behavior of train drivers to detect unusual driving. J. Mech. Syst. Transp. Logist. **3**(1), 338–348 (2010)
4. Suzuki, D., Mizukami, N., et al.: Analysis of driving performance data to evaluate brake manipulation by railway drivers. In: Proceedings of the AHFE 2019 International Conference on Human Factors in Transportation, vol. 964, pp. 282–288 (2019)
5. Netter, J., Wasserman, W., Kutner, M.H.: Multicollinearity, influential observations, and other topics in regression analysis-II. In: Applied Linear Regression Models, pp. 377–416, Richard D Irwin Inc., Homewood (1983)

The Intersection of Spatial Fragmentation and Smart Transport Planning in Gauteng Province, South Africa: Constraints and Opportunities

James Chakwizira(✉)

Department of Urban and Regional Planning, University of Venda,
P.Bag X5050, Thohoyandou 0950, South Africa
James.Chakwizira@univen.ac.za

Abstract. Since the dawn of the new democracy in South Africa, practitioners, policy makers and stakeholders have been resolute on the need to confront spatial fragmentation in order to overcome the spatial inefficiencies that the spatial geography has imprinted on the transportation landscape in the country. Gauteng province is home to the full cycle of impacts and outcomes of apartheid driven spatial and transport inefficiencies. Urban spatial and transportation critics have suggested that adopting a smart transport planning system in the province is one way of seeking to reverse and correct the spatial and transport inefficiencies in Gauteng province. Making use of discourse analysis and systems innovation theory, smart transport and spatial fragmentation integration levers in Gauteng province are unraveled. Transport household databases of the 2003 and 2013 national surveys by the Department of Transport in South Africa as well as Statistics South Africa 2001, 2009, 2011, 2015, and 2018 census and community household surveys complement and provide an interpretative framework for situating spatial fragmentation and smart transport intersections including efficiencies and or inefficiencies in South Africa. To address spatial fragmentation, it is proposed that advanced spatial planning incorporating smart technologies be applied.

Keywords: Spatial fragmentation · Smart transport planning · Spatial and transport inefficiencies · Discourse analysis · Systems innovation theory · Smart transport solutions

1 Introduction

The oxford dictionary defines the word smart as referring to an individual or person being well dressed, fashionable, clever, intelligent, sharp and quick witted. Batty et al., content that the term smart is inherently American as a widely used cliché that refers to ideas and people that provide clever insights and its application has been stretched to city planning through the nomenclature smart growth, smart planning, smart development and smart area or precinct management [1]. Smart city notion is a label that reveals and hides subtle dimensions of urban growth and development making it a

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 367–373, 2020.
https://doi.org/10.1007/978-3-030-50943-9_46

fuzzy concept [2, 3]. Within this paradigm of thinking, smart cities manifest in many dimensions, such as smart people, smart environments, smart economy, smart governance, smart mobility and smart living along a continuum of ever evolving smart cities notions [4]. The notion of smart city and or smart rural areas or regions notational development has undergone a metamorphosis ranging from being conceived as science fiction to being embraced as the new norm and reality of the 21st century new generation urban and rural areas [1]. In any case, urban areas and by implication rural areas have since been conceived as intelligent cities, virtual cities, information cities, networked cities, digital cities, mobile cities, high-technology cities, connected and or wired cities, creative cities, liveable cities, green cities, learning cities, humane cities, knowledge cities, computing cities, software cities, hybrid cities, the utopian sim city, healthy cities, resilient cities, ecological cities, sustainable cities, alternative smart cities, ubiquitous cities and future cities [1, 5–8]. The smart cities cocktail represent modern sites for advancement of big data analytics, information technology analysis and application, expansion and sharpening of human-object interface interactions with the ultimate intention being to address urbanisation and regionalisation development challenges and issues [9–11]. This has implications with respect to how spatial and transportation planning, urban form and systems as well as governance complexities and dynamics across multiple and integrated scales is structured. The use of information communication technologies (ICT) in terms of distributed computation and state of the art human computer interaction (HCI) to valorise integrated spatial and transportation planning is fundamental in this set-up. This study therefore sought to make use of discourse analysis and systems innovation theory, smart transport and spatial fragmentation theoretical frameworks in better understanding and interpreting constraints and opportunities for sustainable smart planning in Gauteng province.

1.1 Research Objectives

The following objectives informed the study, namely:

1. Describing a conceptual framework for marrying spatial fragmentation and smart city solutions;
2. Exploring the implications of advancing intelligent and smart transport solutions in Gauteng province; and
3. Recommending smart transport solutions to counteract spatial fragmentation and transportation inefficiencies in Gauteng province.

2 Research Methods

Two overarching content and thematic lenses in studying smart cities were employed in guiding the study methods, namely the technology driven method (TDM), and the human driven method (HDM). These were chosen solely because TDM is premised on the fact that smart cities are networked places and sites for experimentation with ICTs to achieve the smart city dividend. To complement the ICTs driven smart interpretative framework of analysis, the HDM was incorporated in order to ensure that human capital knowledge, skills and competencies requirements in advancing smart cities were not missed [12, 13]. Additionally, transport and household related databases of the 2003 and 2013 national surveys by the Department of Transport in South Africa as well as Statistics South Africa 2001, 2009, 2011, 2015, and 2018 census and community surveys complemented the generated interpretative analytical framework for situating spatial fragmentation and smart transport intersections including efficiencies and or inefficiencies in South Africa.

3 Seeking a Conceptual Framework for Marrying Spatial Fragmentation and Smart City Solutions

The city can be conceived as a system of systems, and a utopian discourse exposing urban pathologies and their cure [8]. The system of cities is informed by complex adaptive systems, open systems and transactional complexity theory in advancing and building successive waves, micro stages and eco-dynamics with implications for city growth and development cycles [14–18]. Critical discourse theory promotes a conception of urban management in which data and software guide and direct city knowledge, interpretation and specific thematic application. Hollands (2008) and Vanolo (2014) have shown that the idea of the smart city is related to a double lineage in planning literature, i.e. the concept of Smart Growth as theorized by the New Urbanism movement in the USA of the 1980s and, on the other hand, the concept of the technology-based intelligent city [8].

4 Discussion of Results and Findings

The section that follows discusses the major findings in terms of spatial and transport inefficiencies, divided and fragmented transport modes and systems in Gauteng province.

4.1 The Smart City and Spatial Fragmentation Implementation Gaps and Inefficiencies

Since the end of apartheid, planning institutions and departments have embarked on a course aimed at reversing spatial distortion, fragmentation through improving spatial integration efficiencies. Figure 1 presents a spatial fragmentation and inefficiencies storyline in Ekurhuleni municipality that is symptomatic of the spatial deficiencies and inefficiencies existing in Gauteng province.

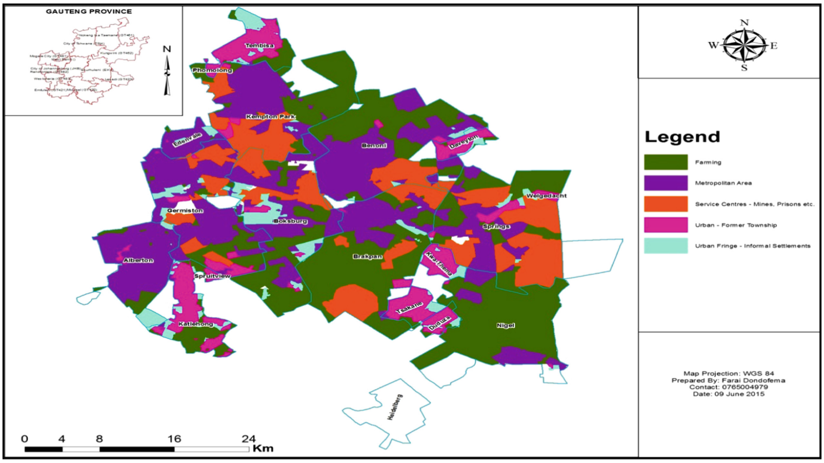


Fig. 1. Fragmentation in Ekurhuleni municipality

Through smart spatial planning in the form of spatial densification, integrated public transport systems and action plans and mixed housing development, efforts aimed at overcoming spatially induced negative externalities have been instituted. Shifts and changes from sprawled suburbanization have been counter-balanced with efforts aimed at spatial consolidation, spatial integration and compaction as well as the promotion of transit-orientated development along major transportation spines such as represented by the Gautrain rapid rail link. All these interventions have led to improved spatial inefficiencies although smart transportation full potential has not yet been achieved. Public transport modes still remain relatively fragmented and disjointed by clusters e.g. Uber, Metrorail, Gautrain, public commuter minibus and subsidised private bus companies, Tshwane bus services, Johannesburg bus service, BRT bus services with a lack of a single integrated ticketing system. While smart systems within modes e.g. BRT Rea Vaya and the application of telematics or Uber and the application of an electronic booking, reservation and payment system are in place, these are however independent and not synchronous and compatible to each other.

4.2 Parallel, Divided, Inadequate Smart Transport Systems and Innovations in Gauteng Province

While different systems such as the rail and road-based systems operate on different platforms, the missing dimension is integrated service provision and interchanges operating from single or common smart based platforms. This is yet to happen with the taxi system not quite ready to embrace technology and cashless payment systems. The development of the Gauteng transport commission is one way through which the smart transport system will be fast-tracked as this will present innumerable opportunities and levers for smart transport governance transition to an integrated smart sustainable public transport system. Figure 2 presents the Gauteng province's road based and railway network systems.

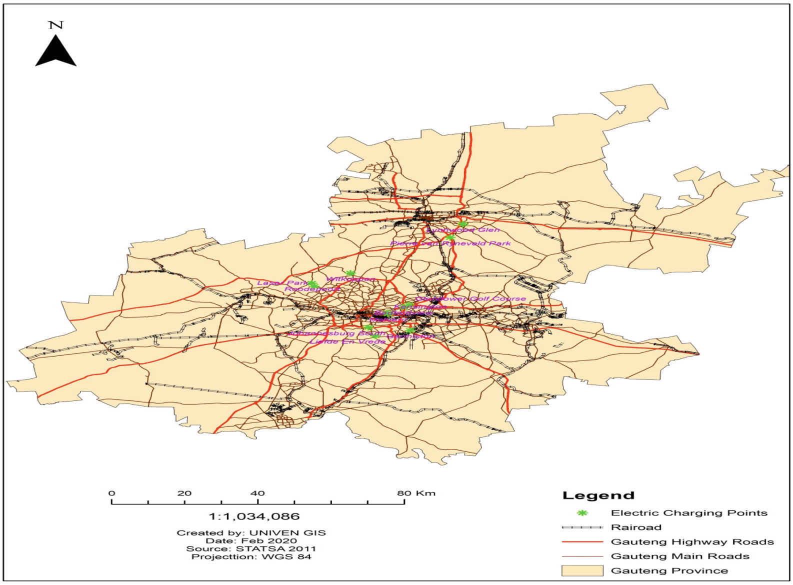


Fig. 2. Electronic vehicle charging stations in Gauteng province

From Fig. 2, we can also deduce that electronic charging stations are available in Gauteng province. Heavy concentration of the electronic charging systems are in Johannesburg especially around the Sandton area while few outlying charging stations are found in Pretoria and Kempton Park. Mass smart public transport systems in Gauteng can be developed taking into account critical value addition dimensions for the road and rail based systems. Such smart city initiatives in support of multi-modal and integrated transport systems will go a long way in embedding smart road and rail systems in Gauteng province.

5 Towards a Futuristic Framework for Anchoring Smart City and Reversing Spatial Fragmentation in Gauteng Province

In thinking smart cities futures, it is critical that existing initiatives are harnessed and directed towards the attainment of a sustainable smart city architectural types that are context driven, timeless, flexible and adaptive to the changing and shifting smart cities contemporary times concept. Resilience, social equity and distributive justice are important pillars of future smart cities [19–26]. Fully exploiting and utilising the capacity and capability of a collaborative and connected society (CCS) in Gauteng province provides opportunities for transport mode and transitions quantum leaps in which mobility as a shared service (MaaS) than a private consumed commodity takes centre stage. While the Gauteng province is yet to reach the smart mobility maturity and spatial fragmentation turn in which the benefits of smart enterprise overrides the negatives, the refreshing finding is that the platform and initiatives necessary to support such a vision and goals are in place and already in motion.

6 Concluding Remarks

The review of Gauteng province smart cities status quo corroborates Hollands recommendation that cities should follow a "fit-for-purpose" pathway to becoming "smarter" cities [27]. Developing a geographically intelligent spatial and transport integrated Gauteng settlement and region requires continuous re-examination of networks, virtual organizations, and managed learning environments within the area with a view to advancing planning, development, and design of planned settlements, neighbourhoods and immediate rural areas within the metropolitan - rural divide in Gauteng province. The discourse and ideology of smart urbanism requires continuous research and development in order to create more sustainable cities and urban areas. One way to achieve this is through developing multiple ways of theorizing, examining alternatives and doing comparative smart cities research [7]. Sensing urban infrastructures and digitizing slums is one practical way of seeking to use the smart economy to bring improved quality of life to residents in Gauteng province. Indeed, there is no straightforward way to implement an integrated concept of smart-sustainable cities that counteract decisively spatial fragmentation and transport inefficiencies. Relief and cues seem to lie in implementing a raft of progressive, context appropriate and driven smart technologies well supported by a lean and smart governance framework that brings value add to the vision and mission of the city/region in question.

References

1. Batty, M., et al.: Smart cities of the future. Eur. Phys. J. Spec. Top. **214**(1), 481–518 (2012)
2. Sun, J., Yan, J., Zhang, K.Z.K.: Blockchain-based sharing services: what blockchain technology can contribute to smart cities. Finan. Innov. **2**(1), 1–9 (2016)
3. Albino, V., Berardi, U., Dangelico, R.M.: Smart cities: definitions, dimensions, performance, and initiatives. J. Urban Technol. **22**(1), 3–21 (2015)
4. Anthopoulos, L.: Smart utopia VS smart reality: learning by experience from 10 smart city cases. Cities **63**, 128–148 (2017)
5. Batty, M.: Big data, smart cities and city planning. Dialogues Hum. Geogr. **3**(3), 274–279 (2013)
6. Artmann, M., et al.: How smart growth and green infrastructure can mutually support each other — a conceptual framework for compact and green cities. Ecol. Ind. **96**, 10–22 (2019)
7. McFarlane, C., Söderström, O.: On alternative smart cities. City **21**(3–4), 312–328 (2017)
8. Söderström, O., Paasche, T., Klauser, F.: Smart cities as corporate storytelling. City **18**(3), 307–320 (2014)
9. D'Acci, L.: A new type of cities for liveable futures. Isobenefit Urbanism morphogenesis. J. Environ. Manage. **246**, 128–140 (2019)

10. Nicolas, C., Kim, J., Chi, S.: Quantifying the dynamic effects of smart city development enablers using structural equation modeling. Sustain. Cities Soc. **53**, 101916 (2020)
11. Vu, K., Hartley, K.: Promoting smart cities in developing countries: policy insights from Vietnam. Telecommun. Policy **42**(10), 845–859 (2018)
12. Neirotti, P., et al.: Current trends in Smart City initiatives: some stylised facts. Cities **38**, 25–36 (2014)
13. Berelson, B.: Content Analysis in Communication Research. Free Press, Glencoe (1952)
14. Samet, R.H.: Complexity, the science of cities and long-range futures. Futures **47**, 49–58 (2013)
15. Cowley, R., Joss, S., Dayot, Y.: The smart city and its publics: insights from across six UK cities. Urban Res. Pract. **11**(1), 53–77 (2017)
16. Gonzalez, R.A., Ferro, R.E., Liberona, D.: Government and governance in intelligent cities, smart transportation study case in Bogotá Colombia. Ain Shams Eng. J. **11**(1), 25–34 (2019)
17. Mora, L., Bolici, R., Deakin, M.: The first two decades of smart-city research: a bibliometric analysis. J. Urban Technol. **24**(1), 3–27 (2017)
18. Ruhlandt, R.W.S.: The governance of smart cities: a systematic literature review. Cities **81**, 1–23 (2018)
19. Chamoso, P., et al.: Smart city as a distributed platform: toward a system for citizen-oriented management. Comput. Commun. **152**, 323–332 (2020)
20. Kumar, H., et al.: Moving towards smart cities: solutions that lead to the Smart City Transformation Framework. Technol. Forecast. Soc. Chang. **153**, 119281 (2018)
21. Yadav, G., et al.: Developing a sustainable smart city framework for developing economies: an Indian context. Sustain. Cities Soc. **47**, 101462 (2019)
22. Chakwizira, J., Bikam, P., Adeboyejo, T.A.: Transport realities and challenges for low income peripheral located settlements in Gauteng Province: are we witnessing the genesis of a new transport order or consolidation of the old transport order? In: International Conference on Applied Human Factors and Ergonomics. Springer (2019)
23. Chakwizira, J., Bikam, P., Adeboyejo, T.A.: Restructuring Gauteng City Region in South Africa: is a transportation solution the answer? In: An Overview of Urban and Regional Planning. IntechOpen (2018)
24. Chakwizira, J.: Searching for sustainable urban transport solutions for Africa: a case study of the Greater Johannesburg region in South Africa. WIT Trans. Built Environ. **130**, 175–186 (2013)
25. Chakwizira, J., Bikam, P., Adeboyejo, T.: Functional and dysfunctional urban mass transportation systems in the greater Gauteng region of South Africa. WIT Trans. Ecol. Environ. **150**, 737–748 (2011)
26. Chakwizira, J., Bikam, P., Adeboyejo, T.A.: Different strokes for different folks: access and transport constraints for public transport commuters in Gauteng Province, South Africa. Int. J. Traffic Transp. Eng. **8**(1), 58–81 (2018)
27. Allwinkle, S., Cruickshank, P.: Creating smarter cities: an overview. J. Urban Technol. **18** (2), 1–16 (2011)

Research on the Bike-Sharing Service from the Users' Perspective and Its Impacts on Their Lifestyles

Tianshi Shen(✉)

School of Architecture and Urban Planning, Tongji University, Shanghai, China
Tianshi_shen@qq.com

Abstract. Shared bicycles have brought new challenges to the urban environment while facilitating the mobility of residents. The main manifestations are the chaos of urban management and the damage to the city's landscape. The literature points out that governments, enterprises, and the urban environment should collaborate to help to improve the bike-sharing system from different aspects such as policy management, operation, maintenance, and public supervision. However, very few pieces of literature investigate the service from users' perspective or what is the impact of the service on their lives. Combined with field surveys and literature studies, I randomly selected 10 users around the Siping campus of Tongji University for questionnaire interviews. The results show that people are satisfied with the bike-sharing service and its corresponding environment. The biggest pain points are inconsistency and unsafe while riding. The bike system and the urban environment match well in public spaces, but the residential areas and the working places need to be improved. Sharing bikes increase people's willingness to travel within short distances and enhance social interaction. Last not least, bike-sharing users are more concerned with the urban environment and are more likely to accept the green traveling lifestyle. Finally, I proposed strategies and recommendations based on environmental policy, planning, design and education aiming to integrate users, service system and urban environments.

Keywords: Bike-sharing service · User experience · Urban environment · Lifestyle

1 Background

Variety bike-sharing services benefit the people that traveling around urban areas, but also bring new challenges to the built environment and city management. For example, the increasing bikes occupy walking paths along the streets, particularly at metro stations. Challenges also come from the service itself. Due to the high cost of maintenance, most of the bike-sharing services need continuous investment [1]. As a new thing, users, operators, and the urban traffic management system are all facing the problems of using, running and managing [2]. The attitude of city authorities, in particular, changes from the acceptance by default to opposition radically. A Beijing

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 374–384, 2020.
https://doi.org/10.1007/978-3-030-50943-9_47

real estate company sued Mobike[1] because the "Mobike bicycles" had been parked randomly that caused chaos to the daily operation of the company. About 4,000 shared bicycles in Huangpu District of Shanghai were confiscated by the regulators because they are parked randomly and some of them occupied the walking streets [3]. However, residents have gradually developed the habit of using bike-sharing services. The growing number of users shows that bike-sharing becomes a big vital part of the future urban mobility system. In this context, how to integrate the bike-sharing system with the existing urban system need in-depth research.

The existing literature on the integration of bike-sharing and urban environments can be roughly divided into two categories. One is top-down macro research. It mainly focuses on how users, enterprises, and urban management systems work together to co-govern the problems caused by the shared bicycles [4, 5]. The other category is bottom-up micro research. User surveys are the main method, specifically the user behaviors and their habits of using shared bikes [6, 7]. In recent years, the user experience has become a popular topic of bike-sharing research, which focuses on the user's experience during cycling, thus proposing a systematic improvement in the process of using shared bikes [8–11]. Yang Liu, Ke Li, etc. used service design methods and tools to study the riding process of bike-sharing services and made proposals to help to improve user experience [11]. Bike-sharing is a system consists of both products and services. From the system perspective, it is not only a subsystem of the urban transportation system but also part of the whole big city system. Therefore, to put the bike-sharing system in a larger urban environment can make up for the shortcomings of simply studying the system itself. This research focuses on people's three main travel activities most related to ridings, such as shopping, commuting, and leisure activities. I also used service design methods and tools to investigate people's bicycle-related travel behaviors and activities. The results are analyzed based on the framework that tells the relationship between people, bike-sharing system and urban environment. Finally, I proposed the most important insights and strategies from different aspects like environmental policies, urban plans, design, and education.

2 Methods

2.1 Questionnaire Design

1. Activity Analysis

   The activity analysis method is based on the assumption that the reason why people travel is to complete specific activities, so travel is a derivative activity of daily activities [12]. Mobility research must start with the research of people's daily activities. Human activities are composed of a series of continuous and related events under a certain space-time environment [12]. Existing studies show that residents mainly use bicycles in commuter, shopping, leisure, and meeting friends

[1] Mobike is one of the largest bike-sharing operators in China.

[6, 13], so this research focuses on people's usage of shared bikes within the three activities.

2. Customer/user flowchart and service blueprints
In service design, experience design, and strategy design, user flowcharts and service blueprints are tools that describe the user's experience when using the service and the user's interaction with the service system. The former focuses on the human experience, while the latter mainly observes the service flow [14, 15]. This research focuses on a typical trip that combines shared cycling with public transport. The research breaks down people's travel activities into a series of touchpoints and actions including starting point, unlocking, riding, parking and locking, public transportation, and again unlocking, riding, parking and locking, and destination. At the same time, to better understand the mechanism behind each touchpoint or behavior, service blueprint is combined with the user flowcharts, see Fig. 1. The questionnaire is designed from the user's point of view and the questionnaire are made of each touchpoint and process in the user flowchart and service blueprint. All the questions are classified into three main parts, they are relationships between users and city, users and bike-sharing systems, bike-sharing systems and urban environment. Finally, we asked the interviewees the open question of how their lifestyles were changed after using a bike-sharing service.

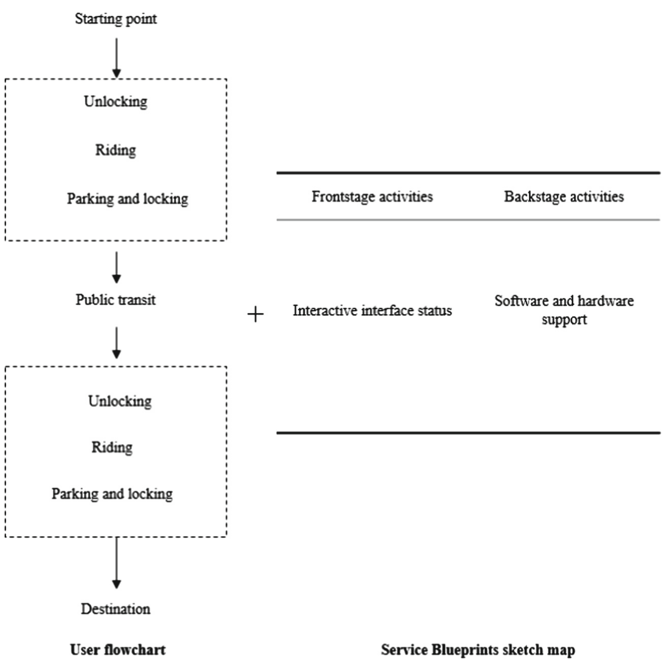


Fig. 1. Combination of user flowchart and service blueprints

2.2 Interview and Investigation

Based on the framework above, we made an open questionnaire we picked up a total of 10 users of bike-sharing services. Previous research indicated that age has a significant impact on the use of shared bikes. The users between the ages of 25–35 are about 50% which accounts for the highest proportion of all ages. The proportion of users under 25 and between 35–45 are not certain, but together they are the second-largest user groups [6, 16]. According to the proportion of age distribution, a stratified sampling approach was used to select 10 typical users as interviewees. All the users were randomly picked up at around the Siping campus of Tongji University. Five of them were between 25 to 35 years old, three of them were between 35–45 years old and the rest two were under 25 years old. The main purpose of the investigation is to collect users' experience when they travel by shared bikes. Respondents were asked to recall at least one bike-sharing experience related to commuting, shopping or leisure. The questions are list in Table 1.

Table 1. The questionnaire on the use of bike-sharing service

Basic information		
Name:		
Gender:		
Age:		
Purpose of travel:		
Private transportation tools		
Categories	Questions	Answers
Uses and urban environment	Where did you use the bike?	
	Do you feel comfortable?	
	Why?	
Uses and bike-sharing system	Where did you do?	
	Do you feel comfortable?	
	Why?	
Bike-sharing and urban environment	How the bike-sharing system affected the urban environment	
	How does the urban environment respond to bike sharing?	
	What do you think is the reason?	
Lifestyles	Has bike sharing changed your lifestyle?	

1. Comfort level is expressed by five levels of −2−2, −2 indicates very bad, −1 indicates bad, 0 indicates usual, 1 indicates good, and 2 indicates very good.
2. Lifestyle manifests as people's interactions, behaviors, and ways of thinking

2.3 Field Research and Extra Data Collection

Respondents were more subjective in their judgment of the relationship between the bike-sharing system and the urban environment. Therefore, we checked out more materials by literature review, looking over the internet and visiting some sites where the conflicts happened.

3 Results

The interviewees ranged from 22 to 44 years old, all had a bachelor's degree or higher. Three of them had no private transport tools, but the rest had a car, e-bike, bicycle, and other transport tools.

3.1 The Relationship Between the Users and the Urban Environment

1. Cycling
   People are satisfied with the urban environment when they use shared bikes. People feel uncomfortable when the route is complicated. The bike lanes are not continuous and mixed with automobiles that make the cyclists feel unsafe. The attitude of the urban authorities to cyclists follows the attitude of traditional bicycle users in the past. The city is more tolerant of violators rather than forbidden. Most of the time they urge the operators to regulate the users' behavior.
2. Parking
   People are satisfied with the parking of shared bikes although parking bikes cause some chaos because of the occupation of the walkways. However, compared with the convenience the service brings, the parking problem is still acceptable. People understand and support the statement "No shared bikes allowed" in residential areas and working places. However, some users living in very large communities expressed fatigue because the bikes were too far to get. The city does not provide particular facilities for bike-sharing, except that certain white lines are marked on the sidewalks or other areas to identify bike parking spaces. In response to the chaos caused by randomly parking to the urban landscape, community managers usually acquiesce and encourage the staff of operators around large parking spots to help reorganize the bikes.

3.2 The Relationship Between the Users and the Service System

The bike-sharing service has become part of people's daily lives including commuting, leisure and shopping activities. Generally, people have good experiences using bike-sharing services. However, there are still some problems with the new system. They are as follows:

1. Bike design and its maintenance
   Since the bike-sharing itself is not designed for a particular type of mobility, so users will have different problems when they are in different situations. For example, commuters with backpacks cannot safely put their bags into small and shallow baskets; bowing down to scan the OR codes is difficult for people with physical disables; high rates of damage have become the biggest obstacle for people to use shared bikes.

2. System optimization
   The e-map that comes with a bike-sharing APP makes it easy for people to see if there are enough vehicles around them. However, after the number of shared bikes has reached a certain number and their locations are stable, the role of the e-map is minimal. People have been used to using Gaode Maps/Baidu Maps to check out the cycling routes in advance. When locked, the APP would not remind the users whether the bike was locked successfully or whether the fee was successfully deducted immediately. The system failure makes people concerned.
3. User education
   People generally use bike-sharing services following the traditional habits of cycling and they do not pay much attention to whether there are settled areas for parking. The bike-sharing service operators do not educate their users on how to use and parking the new shared bikes properly to maintain the urban environment.

3.3 The Relationship Between the Service System and the Environment

1. Traffic facilities
   The emergence of the bike-sharing system has improved the urban transportation environment which originally is designed for vehicle-based mobility. Some part of the roads is allocated into bicycle lanes. However, the mixture with motor vehicle lanes and pedestrian lanes makes the riders feel uncomfortable and unsafe. Except for large road intersections, signal lights specifically set for cyclists are not common.
2. Public spaces
   Urban public spaces get along well with shared bikes, and the city continues to use the original facilities and resources to provide a better parking environment for shared bikes.
3. Residential areas and workplaces
   Although residential areas and workplaces do not positively coordinate with bike-sharing services, their attitudes are acquiescence and compromise. Besides, they allocate a certain area for shared bikes to park. The staff of communities and companies also help the operators with daily work like maintenance or arrangement of the shared bikes.

3.4 Impacts on the Users' Lifestyle

Lifestyle has rich meanings. The Chinese Encyclopedia defines it as "It is a series of systematic features including the form of all activities and behaviors that meet the needs of one's own life under the constraints of certain social conditions and values" [17]. Lifestyle must be expressed in a certain form, and this research uses a simplified definition "people's interactions, behaviors (including clothing, food, housing, rest, entertainment, etc.) and ways of thinking" [18]. The survey asked open questions about what changes and impacts that bike-sharing services have on people's lifestyles. The changes and impacts are as follows:

The new services help to increased users' willingness and frequency of moving. Compared to the time before the emergence of shared bicycles, the respondents were more willing to travel with short distances and the actual number of trips raised, especially to go shopping nearby, meet friends, do business, etc. Another interviewee mentioned that sharing bicycles increased his willingness to go to other areas of Shanghai, while he previously gave up to go out because he was unfamiliar with public transportation or pathways in other districts.

Strengthening social interaction. Bike-sharing has brought new ways of communication, increased social activities, deepened social relations and friendships particularly among young people, making them feel part of the city. Because of bike-sharing service, new forms of riding emerged, for example, a couple riding together as leisure, friends riding at night as a way of a gym or social activity, riding as a new fashion to scene sighting around the city, and so on.

Increasing people's favorability for the city and making them pay more attention to the city. Bike-sharing service has changed the interviewee's thinking of urban mobility. Within a short distance, people become more active to travel by bikes. According to the feedbacks from android APP stores under bike-sharing APPs, not only in Shanghai but also the other cities, people are fond of cities with bike-sharing services. Bike-sharing has also much social influence, at least respondents were more concerned about urban issues than before.

People who use bike-sharing services have a better understanding of ideas such as green mobility and sharing mobility, so they have a preference for these shared mobilities. Since bike-sharing has become a part of their lives, so it is more popular among the users to move in healthier ways such as walking, riding and any other green mobilities. Some interviewees also expressed that they were inspired by sharing bikes so they began to think about sharing other resources like students' textbooks.

4 Discussion

System theory assumes that the relationship between the elements is as important as the elements themselves. By examining the relationship among users, bike-sharing systems, and the urban environment, we can draw preliminary conclusions. As the central element of connecting bike-sharing services with urban environments, users have to maintain and supervise the bike-sharing system and the urban environment system if they want a better riding experience. The bike-sharing system serves not only as a technical supporter but also works as a bridge between users and cities. So that they need to collaborate with different communities. Users need to be educated to formulate good hobbits. Acting as the main educators, bike-sharing service operators play a major role in regulating users' behaviors. Urban environments, which mainly refer to environments where users interact with the service most, need to guide and educate the users of how to ride properly. The government also has to supervise the operation of bicycle-sharing companies, however, more importantly, is that cites have to connect infrastructures or resources with the sharing services and encourage more people to use bicycle-sharing services as green travel ways as shown in Fig. 2.

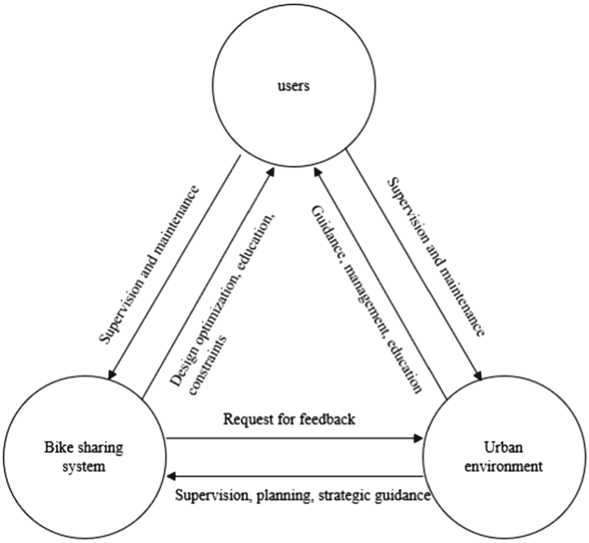


Fig. 2. Relations among users, bike-sharing system, and urban environment

4.1 Environmental Policy and Transportation Planning

The government should unite transportation management and planning authorities, communities and enterprises to work together to protect cyclists' rights on the roads. First of all, design particular bike lanes to enhance cycling consistency and safety. It is shown that the government and users have different attitudes towards bike-sharing. The government prefers a unified plan to address parking issues, while residents expect to focus on addressing the issues of inconsistency and insecurity during cycling. To solve the parking problem, we need to adapt to local conditions, propose reasonable planning especially for residential areas and workplaces. An effective solution is to leave a space empty and then redesign the parking space according to the real demands of bike users. The parking solution in public places is worth learning. For example, when Mobike launched in Shanghai, they discussed with Tongji University to set up a particular parking spot on the campus of the university, see Fig. 3.

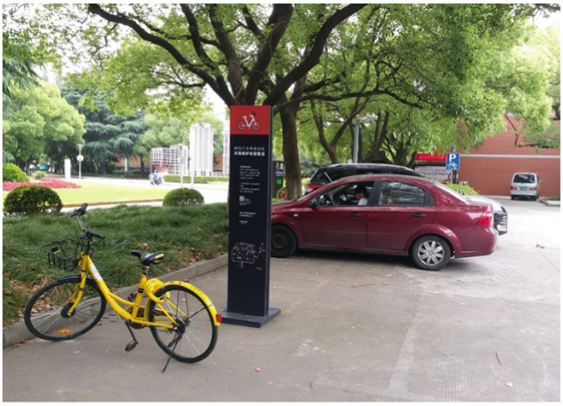


Fig. 3. Parking point of a shared bike at Tongji University

4.2 Design

Design objects are changed from the physical environment to the overall experience. The urban environment has two definitions. In the field of environmental behavior, the urban environment refers to the physical environment system [19], while in the area of environmental design, the urban environment is the "living space ecosystem" around people [20]. Both of them recognize that people and the environment are indivisible, but the design objects are not the same. The former is designed for physical systems, and the latter includes, in addition to physical systems, the design of immaterial systems such as meaning, interaction, service, experience, and lifestyle. The experience must be consistent. The quality of any environment is ultimately reflected in people's experience. The design of the urban environment needs to be changed from the design of individual places to the design of the more systematic experience. The task of design is not simply to create the environment but to undertake the integration of the interests of all stakeholders to achieve common values, in the process of shaping the experience. Experts from service design, experience design, strategy design, etc. have studied people's experiences from different aspects. Experience needs to be examined. For example, a shared bicycle is required to be parked within the white line prescribed by the city. However, some parking spaces in public areas are marked with yellow lines which make the users confused.

4.3 Education

The sharing service operators and governments should collaborate to educate the users. The use of smartphones has made bike-sharing a part of people's daily lives. In 2016, the number of bike-sharing users in China was 4,521,600, and it is expected to reach 10.2615 million in 2019[3]. The education given by the bike-sharing operators to users is effective, and it is better with some fine punishment when users break the rules. The emergence of bike-sharing has exposed many bad habits of people, such as randomly parking, malicious destruction, private occupation and so on. With the collaboration with a national credit system, the above phenomenon has been reduced greatly. Therefore, enterprises should take more responsibilities to guide users to better usage of shared bicycles, help to protect the urban environment and promote green mobility. The role of the government is to support and give help when necessary.

5 Conclusions

The author randomly interviewed 10 typical users of bike-sharing service around Tongji University, investigated their experience with the service and how their life-styles were changed after using shared bikes. The survey found that people are satisfied with the bike-sharing system and its environment including facilitates. The biggest pain point is the inconsistency and safety issues during riding. The bike-sharing system and the urban environment match well in public places, but there is still room that could be improved for communities and working places. From the users' perspective, the parking problem doesn't stop them from using the sharing services. The biggest obstacle that prevents people from using shared bicycles is their interaction with the sharing service and system failures.

Bike-sharing has changed users' lifestyles significantly. The main manifestations are the service to increase people's willingness to travel short distances, improved social interaction. People are more concerned about urban environmental issues, and people are more likely to accept the idea of green and sustainable traveling.

The bike-sharing service cannot be banned in a one-size-fits-all manner due to the issues it brings to society. Environmental policy, urban planning, transportation design, and green mobility education, governments, service operators, communities and civic organizations must cooperate and collaborate, putting human experience at the core to implement different strategies based on different situations. Designers play not only as creators, but they also have the power and capability to integrate the stakeholders and different systems.

In the future, researchers will have several directions. Firstly, people of different ages can be classified for more research. Secondly, more researches can be conducted focusing on user experience at three main scenarios like public spaces, working places, and communities. Thirdly, researchers can focus on the evaluation of the impacts of environmental design on user experience.

References

1. Chen, Y.: Don't forget environmental responsibility when using shared bicycles. Elder. Educ. (Elder. Home) **8**, 28–29 (2017)
2. Liu, Y.: Analysis of the research on shared bicycle development. Times Finan. **8**, 251 (2017)
3. Liu, X.: Prohibition or governance: bike sharing at "crossroads". Decision **6**, 86–88 (2017)
4. Zhang, Z., Wu, W.: The status quo, problems, and suggestions for the development of bicycle sharing. Mod. Bus. **15**, 162–163 (2017)
5. Han, L.: Discussion on the promotion of bicycle transportation under the lead of shared bicycles. Shanxi Archit. **11**, 11–12 (2017)

6. Wang, R.: Analysis of user characteristics of bike-sharing in Shanghai and its relationship with public transport development. Urban Public Util. **004**(006), 46–52 (2017)
7. Li, Y.: A Survey and Analysis of Shared Bike User Experience and Existing Problems-A Case Study of a University Near a University. Economic Research Herald (2018)
8. Xu, Y., Ye, C.: Research on the user experience of shared bikes. Logist. Technol. **8**, 69–71 (2017)
9. Li, Y.: Investigation and analysis of the user experience and existing problems of shared bikes - take the study of a nearby university as an example. Introduction to Economic Research (2018)
10. Li, K.: Study on the improvement of shared bicycle system based on service design (2018)
11. Liu, Y., Li, K., Ren, H.: Analysis of shared bicycle system from the perspective of service design. Packag. Eng. **10**, 23–30 (2017)
12. Chai, Y., Shen, J., Zhao, Y.: The frontiers of research methods on urban traffic travel behavior. China Sci. Technol. Paper Online **05**, 66–73 (2010)
13. Zhu, W., et al.: Research on changes and mechanisms of resident trips under the impact of the public bicycle system: a case study of Minhang District, Shanghai. Urban Plan. J. **05**, 84–89 (2012)
14. Synneva Gjertrudsdotter Storaas: Service blueprints and user journeys. https://www.ntnu.no/documents/10401/1264435841/Synneva+Storaas+PD9+Artikkel.pdf/6fcfdfd2-75d8-4c03-9a4f-396065445b0e
15. Shahrzad Samadzadeh: Customer Journey Map or Service Blueprint? https://www.cooper.com/journal/2015/5/journey-map-or-service-blueprint
16. TalkingData. 2016 bike-sharing users analysis report. http://mi.talkingdata.com/report-detail.html?id=462
17. Wang, Y.: Review of lifestyle research. Sociol. Res. **4**, 41–48 (1995)
18. [Russia] Давиде Гуччи: Dictionary of Applied Sociology: [M]. Heilongjiang People's Publishing House, Harbin (1988)
19. Li, B.: Environmental behavior theory and its extension. Archit. J. **02**, 36–39 (2008)
20. Lou, Y.: New environmental design in the era of global knowledge network. J. Nanjing Univ. Arts (Art Des.) **169**(01), 9–15+238 (2017)

Route Choice, Navigation and Wayfinding

An Evaluation Index System for Wayfinding System and Its Research Applications: The Case of Beijing Subway Line 1

Chuanyu Zou[1(✉)] and Guangxin Wang[2]

[1] AQSIQ Key Laboratory of Human Factors and Ergonomics (CNIS), Beijing, China
zouchy@cnis.ac.cn
[2] Department of Psychology, School of Humanities, Beijing Forestry University, Beijing, China

Abstract. *Object*: To construct a set of comprehensive evaluation index system for wayfinding systems, and conducts trial evaluation using Beijing Metro Line 1 as an example. *Methods:* The research followed the principle of main factors, the principle of dynamics, and used the combination of quantitative and qualitative factors to select the evaluation indexes, which includes functional, normative, safety and coordination indexes. Field tests were performed on Beijing Subway Line 1 using evaluation indexes. *Results:* The overall renovation time in the stations has been half a century, mainly reflected in the absence of platform doors, the dim lighting inside the stations, and the insufficient installation conditions at key intersections. In addition, Line 1 passes through mature urban areas, where the business is prosperous, and the environmental information outside the stations often change. All these also make it difficult for maintain information connecting inter-stations and out surroundings.

Keywords: Wayfinding system · Evaluation index · Beijing Subway

1 Introduction

1.1 Background

Beijing Subway system is the earliest metro system built in China. There are 22 operation lines of Beijing Subway, including the first urban subway line built in China, line 1, which was put into trial operation in 1969, and Beijing Daxing International Airport Express, which was opened and operated simultaneously with Beijing Daxing International Airport in 2019.

Till December 2018, there are 22 subway lines operating in Beijing, covering 11 municipal districts of Beijing, with an operating mileage of 775.6 km [1] and 391 stations, ranking the second in China after Shanghai. The annual passenger traffic volume is nearly 4 billion, and the maximum daily passenger traffic volume is nearly 14 million.

Beijing subway has a large passenger flow and high mobility, so the subway wayfinding system is an important medium for passengers and station space, passengers

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 387–392, 2020.
https://doi.org/10.1007/978-3-030-50943-9_48

and station surrounding environment cognition, and an important infrastructure to instruct passengers to travel safely and orderly and arrive at the destination quickly. The wayfinding system of Beijing Subway includes direction signs, location signs, space schematic diagram of the station, surrounding environment schematic diagram outside the station, barrier free signs, etc.

1.2 Current Standards

At present, three institutions in China have issued national standards related to Subway wayfinding system and its evaluation indexes, namely: GB/T 18574 urban rail transit passenger service sign, issued by the Ministry of housing and urban rural development [2]; GB/T 36953.3 evaluation method of urban public transport passenger satisfaction - Part 3: urban rail Transport, issued by TC 529 (National Technical Committee on urban passenger transport Standardization) [3]; GB/T 15566.4 guidance system for public information - setting principles and requirements - Part 4: public transport station [4], issued by TC59 (National Technical Committee on graphic symbol Standardization).

ISO/TC 145 is responsible for the development of relevant standards of guidance system in the world. At present, it has published ISO 28564 "public information guidance systems" [5], setting design guidelines for wayfinding systems such as guiding signs and location signs, and ISO 7001 "public information graphic symbols" [6].

1.3 Research Status

Tyler Duke et al. [7] conducted a research to assess the usability and effectiveness of the interactive wayfinding system at a large hospital complex in Houston, Texas. Issues that participants consistently faced when given various tasks were organized and later developed into suggestions for system design, such as organizing information effectively, reducing memory overload, and allowing the user to control their pace when given automated information.

CY Zou et al. [8] evaluated the satisfaction of subway wayfinding system using the objective structural equation model to give weights. The study also build an assessment questionnaire with three dimensions: the safety signs, the traffic signs, and the continuity and rationality of the signs.

2 Methods

Based on the relationship between designers, environment, and the users, this research studied and analyzed the impact of typical public place space and process management on the setting and evaluation of the wayfinding system. This research also studied technical requirements, and the formation of corresponding setting specifications and evaluation methods, so as to improve the standardization, systematization and performance of typical wayfinding systems in public place, to ensure the coordination and unification between the wayfinding system and the environment.

2.1 Selection Principle of Evaluation Index

The research followed the principle of main factors, the principle of dynamics, and used the combination of quantitative and qualitative factors to select the evaluation indexes, which includes functional, normative, safety and coordination indexes.

Principle of Main Factors. The indexes shall be able to comprehensively reflect the main aspects of subway station operation and management problems and passenger transfer problems, and ensure the normal operation demand of the station, the flow line inside the subway station and the transfer conditions outside the subway station.

Principle of Dynamics. The subway wayfinding system has different construction requirements in different development periods, different types of subway stations and different regions. The evaluation work is a dynamic optimization process. The evaluation results should be the comprehensive results of each indexes relative to national standards, local standards, urban planning, specific lines and stations, which can not only reflect the degree of established indexes status, but also instruct the optimization of subway wayfinding system.

Combination of Quantitative and Qualitative Factors. Experts and passengers are used to evaluate the subway guidance system qualitatively, and relevant data released by the subway are used for quantitative evaluation.

2.2 Composition of Evaluation Indexes

The evaluation indexes system for wayfinding system in Beijing Subway includes three indexes, which are functional, normative, safety and coordination indexes.

Functional Indexes. a) The wayfinding system can meet the guidance needs of public places where it is located; b) The wayfinding system can meet the needs of people using equipment and services in public places.

Normative Indexes. a) Public information graphical symbols can meet the requirements of national standards; b) The design of guidance elements can meet the requirements of national standards; c) The installation of guidance elements can meet the requirements of national standards; d) Main flows of moving are continuous in public places. e) The wayfinding system in located public place can be connected to other surrounding wayfinding systems in an orderly manner. f) Accessibility signs meet the requirements of national standards.

Safety Indexes. a) The shape and appearance of the guidance elements is not a potential safety hazard; b) The installation of the guidance elements is not a potential safety hazard.

Coordination Indexes. a) The specifications of guidance element are appropriate for the environment; b) The wayfinding system is coordinated with other visual elements in the environment.

2.3 Questionnaire Constitution

The questionnaire consists of the following parts:

Explanatory Text. Such as greetings, investigator's identity, investigation purpose and confidentiality commitment.

Instruction Text. Guide passengers to fill in the questionnaire.

Discrimination Part. Select respondents who meet the requirements and reject the unqualified questionnaires.

Main Part. Questions of the evaluation indexes on Beijing Subway wayfinding system.

Open Part. Collect satisfaction, least satisfaction and most concerned aspects of passengers to the Beijing Subway wayfinding system.

Background Information. Respondents' gender, age, occupation, education level, subway line, time, frequency, purpose, etc.

Investigation Process Record. The investigators fill in the basic information of the evaluation work, including the subway station, the time of the investigation, the name of the investigator, etc.

2.4 Respondents

The respondents were able to represent the common passengers in the subway station, and had recently taken the subway in the station. Passengers who have not taken the subway at the station should not participate in the investigation.

2.5 Subway Station Sampling

Beijing Subway Line 1 was selected for evaluation test. Line 1 is the first subway line constructed in China. It starts from Pingguoyuan station in the West and ends at Sihuidong station in the East. It runs in a straight line with a total length of 31.04 km. There are 23 stations in total. The average daily passenger capacity of Beijing Subway Line 1 is more than 2 million people.

According to the list of Beijing Subway stations of Line 1, the sampling of stations was carried out in a random way. Taking Beijing Subway Line 1 as an example, the sampling starts from Sihui station in a random way, and then every n^{th} ($n > 1$) subway station is selected into the sample.

When sampling subway stations, pay attention to distinguish ordinary subway stations from transfer stations. The transfer stations are sampled separately, and the data are compared with other Lines' transfer stations. In the multi-level overall sampling, if the transfer station is encountered, the next or two subway stations are selected into the sample.

2.6 Field Test Process

In May 2019, in each selected subway station, according to the principle of random sampling, the subjects of the evaluation index of the subway wayfinding system were investigated. The number of respondents selected from each subway station shall not be less than 30.

3 Evaluation Results

Based on field tests performed on Beijing Subway Line 1 using evaluation indexes questionnaire, scores of Beijing Subway Line 1 wayfinding system evaluation were listed as Table 1.

Table 1. Scores of Beijing Subway Line 1 wayfinding system evaluation

Basic index	First level	Second level	Scores
Functional	Information presentation	Accuracy of information	95.3
		Accuracy of key direction and orientation	90.7
	Element configuration	Location signs at key points and service places	98.6
		Direction signs at key intersection points.	97.5
Normative	Design	Graphical symbols and colours	97.3
		Text	99.8
		Guidance element	98.2
		Accessible signs	98.9
	Setting	Guidance elements at key intersection points	93.3
		Visual convenience	92.9
	Systematic	Information consistency	97.8
		Connect orderly with other wayfinding systems	88.2
	Continuity	Direction information	99.1
		Way-in	99.5
		Way-out	87.2
		Interchange	89.7
Safety	Shape	Sharp protrusions or edges	98.8
		Non-slip Floor sign	99.6
	Installation	Installed firmly	99.2
		Not potential safety hazard	83.1
Coordination	Surrounding	Specifications of guidance element are appropriate for the environment	86.2
		Environmental protection and energy saving	83.3
		Coordinated with other visual elements in the environment	81.8
	Maintenance	Maintenance level	80.5

4 Conclusion

Beijing Subway Line 1 is the longest built and used one in Beijing. The environment and supporting facilities in the stations of Line 1 are relatively older compared to other lines. The overall renovation time in the stations has been half a century, mainly reflected in the absence of platform doors, the dim lighting inside the stations, and the insufficient installation conditions at key intersections. In addition, Line 1 passes through mature urban areas, where the business is prosperous, and the environmental information outside the stations often change. All these also make it difficult for maintain information connecting inter-stations and out surroundings.

Acknowledgments. This research was supported by China National Institute of Standardization through the “special funds for the basic R&D undertakings by welfare research institutions” (522019Y-6668).

References

1. China Association of Metros. http://www.camet.org.cn/
2. GB 18574 The passenger service signs for urban rail transit [S]
3. GB/T 36953.3. Evaluation method of urban public transportation passenger satisfaction—Part 3: Urban rail transit [S]
4. GB/T 15566.4 Guidance system for public information - Setting principles and requirements - Part 4: Public transport station [S]
5. ISO 28564 Public information guidance systems [S]. International Organization for Standardization (ISO) (2016)
6. ISO 7001 Public information graphic symbols [S]. International Organization for Standardization (ISO) (2007)
7. Duke, T., et al.: I’m lost: an evaluation of an interactive wayfinding system. In: Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care (2018)
8. Zou, C.Y., Wang, G., Chen, Y.: Study on evaluation of subway passenger wayfinding systems based on SEM. In: Krömker, H. (ed.) HCI in Mobility, Transport, and Automotive Systems, HCII 2019. Lecture Notes in Computer Science, vol. 11596. Springer, Cham (2019)

Empirical Study on Evaluation of Railway Wayfinding System: A Case Study of Shandong Province, China

Chuanyu Zou[1(✉)], Yongquan Chen[1], and Ziding Chen[2]

[1] AQSIQ Key Laboratory of Human Factors and Ergonomics (CNIS), Beijing, China
zouchy@cnis.ac.cn
[2] National Railway Administration, Beijing, China

Abstract. Object: Based on the preliminary research results of the national standards for the evaluation indexes of railway passenger station wayfinding systems, Shandong Province was selected for field-test. The test process and results provided first-hand information for the improvement and optimization of national standards. It can also find problems in the wayfinding system of Shandong Railway Passenger Station and make suggestions for improvement. Methods: First, a questionnaire was prepared based on the preliminary evaluation indexes proposed in the national standard. Second, based on the sampling method in the National Standard, combined with the actual situation of the Shandong Railway Passenger Station, Qingdao Stations, Jinanxi Station, and Weifangbei Station were tested as sampling stations. Results: Among three railway stations tested in Shandong Province, Jinanxi Railway Station had lower scores in terms of basic specifications, safety, and coordination.

Keywords: Wayfinding systems · Railway Station · Evaluation index

1 Introduction

1.1 Research Background

As a node of the railway network in a city, a railway passenger station usually connects with other urban transportation systems to form a comprehensive transportation hub, and is an important public transportation facility for passenger travel. The total number of railway passenger transport in 2019 reached 3.66 billion, an increase of 8.4% over the previous year and accounting for 20.8% of the total; the passenger transport turnover reached 1470.66 billion person-km, accounting for 41.6% of the total [1].

Integrated transportation has been the focus of urban transportation development in recent years. The purpose is to provide cities with convenient and energy efficient transportation services on the basis of fully integrating urban transportation infrastructure. Through the transportation connection system, various transportation modes are internally connected, different transportation modes, private transportation and public transportation, urban transportation and external transportation are effectively connected, and the overall benefits of the transportation system are exerted. The railway

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 393–398, 2020.
https://doi.org/10.1007/978-3-030-50943-9_49

passenger station is connected to the inter-city transportation system and the urban transportation system, and is the transfer node of the city's internal transportation network and external transportation network. At the same time, the railway passenger station is a passenger flow distribution platform that integrates multiple modes of transportation. Therefore, railway passenger stations play a very important role in the process of building integrated transportation in cities.

The quality of public space and use efficiency of railway passenger stations to a certain extent determine the quality and experience of railway travel of passengers, and the construction level of public information guidance system of railway passenger stations is one of the important factors that determine the efficiency of public space usage of railway passenger stations.

1.2 Current Situation of Railway Wayfinding System

During 2016–2017, 57 railway passenger stations' wayfinding systems were field-investigated. More than 290 major problems were found in four categories, mainly focusing on non-standardized use of information element, non-standardized design of wayfinding elements, poor systematized setting of wayfinding elements, and disconnection between wayfinding system planning and architectural design [2].

1.3 Current Situation of Railway Network in Shandong Province

The Shandong Province Railway network is operated by Jinan Bureau. It faces the coast in the east, the Central Plains in the west, and the Beijing Bureau in the north. It is adjacent to the Shanghai Bureau in the south. Throughout the north and south, the Jiaoji Line, Qingrong Line, Hanji Line, Wari Line, Hejiri Line, Delongyan Line and other lines across the east and west, covering the entire territory of Shandong Province, covering 140 counties and cities. The current operating mileage is 5114.6 km, of which 1205.7 km is a high-speed railway. There are 59 railway passenger stations of different levels.

2 Research on Evaluation Index

2.1 Standard Development

Railway passenger station wayfinding system, as an information system directly facing and serving passenger activities, is an integral part of the construction and management of station infrastructure. It not only facilitates passengers to identify public facilities, identify directions, and find roads. To a large extent, it meets and facilitates the needs of passengers for their lives, work and tourism, and at the same time it can effectively play a role in maintaining public order and improving station capacity.

Based on the results of previous investigations, in 2018, the National Technical Committee for the Standardization on Graphical Symbols launched the national standard development of "Public Information Guidance System—Evaluation Requirements—Railway Passenger Stations."

2.2 Evaluation Indexes

The evaluation index of the railway passenger wayfinding systems is composed of 4 dimensions and 10 first-level indicators.

Functionality. The degree to which information is presented to meet the needs of rides within the station; the degree to which the elements are configured to meet the needs of rides within the station;

Normativeness. The degree of standardization of the design of the guiding elements; the degree of standardization of the setting of the guiding elements; the degree of systematization of the guidance system inside and outside the station; the degree of continuity of the guidance system.

Safety. The degree of safety of the appearance of the guiding elements; the degree of security of the installation of the guiding elements;

Coordination. The degree of coordination between the guidance system and the environment; the level of maintenance of the guidance elements.

3 Methods

3.1 Station Sampling Method

According to the requirements of the sampling plan in the standard, and the passenger volume of the hub station and the level of the railway passenger station, Jinanxi Station was selected as a verification sample on the Beijing-Shanghai Railway Line. Qingdao Railway Station of Jiaoji Railway Line was selected as the verification sample. The first-class station Weifangbei Railway Station was selected as the verification sample for the Beijing-Shanghai Railway Line 2 (see Table 1).

Table 1. Samples of Shandong Railway Station for wayfinding system evaluation

Railway line	Railway station	Station level
Beijing-Shanghai Railway Line	Jinanxi Railway Station	Super level
Jiaoji Railway Line	Qingdao Railway Station	Super level
Beijing-Shanghai Railway Line 2	Weifangbei Railway Station	First level

3.2 Subject Sampling Method

It can be seen from Table 2 that among all the valid samples of Qingdao Railway Station, Jinanxi Railway Station, and Weifangbei Railway Station, the number of attending school dominates, accounting for 33 people, including 27 men and 6 women. In addition, travel due to personal affairs occupied the second place in the statistical sample, with a total of 25 people, including 13 men and 12 women.

Table 2. Subjects cross-list

Gender	To school	Goto work	Private affairs	Tourism	On business
Men	27	3	13	5	10
Women	6	1	12	5	10

3.3 Questionnaire

The questionnaire includes:

- *Introduction*: introduction of the background and purpose of the test;
- *Demographic information of the subject*: basic information such as the age and gender of the subject;
- *Select part*: select eligible subjects;
- *Questionnaire text*: questions based on the evaluation indexes of railway wayfinding system;
- *Information related to the evaluation activities of the subjects*: collecting information about the subjects' travel routes, time of travel, frequency of travel, purpose of travel, etc.;
- *Open-ended questions*: the satisfaction of the subjects with the wayfinding system, the least Satisfied areas, areas most concerned by the participants, etc.

Questionnaire Scale: Questions in the questionnaire include two types, e.g., multiple choice questions and single choice questions. Multiple choice questions are clearly marked in the questionnaire, this kind of questions are often used in the select part, and the information related to the evaluation activities of the subjects, etc. While single choice questions are often used in the main body of the questionnaire. The single choice question uses 5-level text scale and a digital scale as options for the question (see Table 3). The text scale selects the appropriate expression according to the question. For example, the question "Is the sign firmly installed?", the text scale can be "very weak, less strong, average, strong, very strong."

Table 3. Text scale and digital scale

Very poor	Poor	Normal	Good	Very good
1	2	3	4	5

3.4 Data Quality

According to the evaluation plan, data collection was performed and the quality of the collected data was evaluated. The data of poor quality are mainly two kinds, e.g., "ceiling" type data: all options are fixed one, which is expressed as fixed high end or low end; "random" type data: not based on your own judgment and decision, instead, answer at will, inconsistently, logically inconsistently.

The methods to avoid bad data from affecting the evaluation result are: cleaning the "ceiling" type data; adding duplicate questions to the questionnaire randomly.

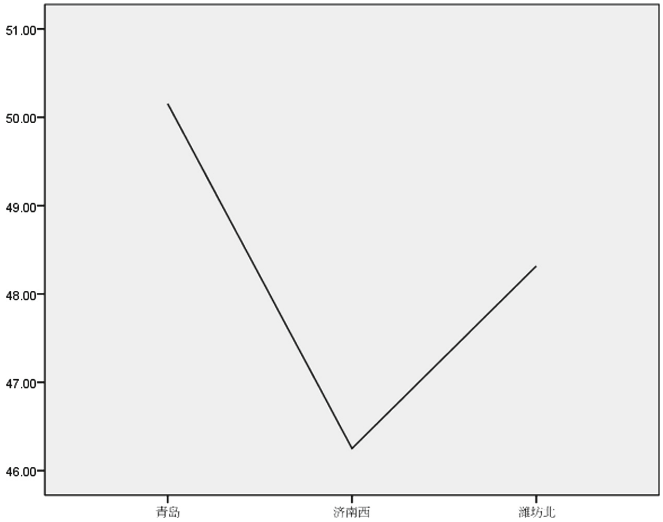


Fig. 1. Normativeness evaluation score of each station

4 Conclusion

In the evaluation of Normativeness index of wayfinding systems, subjects are satisfied with Qingdao Railway Station, so Qingdao station scores higher in this dimension. Jinanxi Railway Station's low score indicates that subjects' satisfaction with the station is low, so the evaluation score is low (Fig. 1).

Using the Normativeness index as the dependent variable and the railway stations as the grouping variable, the results of the single factor analysis of variance (ANOVA) are as follows: On the Normativeness index score, $F(2,89) = 12.07$, $p < 0.05$, displayed on the Normativeness index of each station, there are significant differences. Further multiple comparisons revealed that Normativeness index score of the Jinanxi Railway Station was significantly lower than the Qingdao Railway Station ($p < 0.05$), and also significantly lower than the Weifangbei Railway station score ($p < 0.05$).

The same results are also shown in the Safety index of wayfinding system. Subjects are more satisfied with the safety of Qingdao Railway Station, so Qingdao Railway Station scores higher in this dimension and Jinanxi Railway Station scores lower. The one-way analysis of variance (ANOVA) results are as follows: On the safety score, $F(2,90) = 9.957$, $p < 0.05$, showing that there is a significant difference between stations on this dimension. Further multiple comparisons revealed that the safety index score of the Jinanxi Railway Station was significantly lower than the Qingdao Railway station ($p < 0.05$), and also significantly lower than the Weifangbei Railway station score ($p < 0.05$).

On the Coordination dimension of wayfinding systems, subjects are satisfied with the coordination of Qingdao Railway Station, so the station scored higher in this dimension. Jinanxi Railway Station scored low. The results of the one-way analysis of variance (ANOVA) were as follows: On the coordination index score, $F(2,91) = 7.22$, $p < 0.05$, showing that there were significant differences between the stations on the coordination index score. Further multiple comparisons revealed that the Jinanxi Railway Station's coordination index score was significantly lower than the Qingdao Railway Station ($p = < 0.05$) and also significantly lower than the Weifangbei Railway Station's score ($p < 0.05$).

Acknowledgments. This research was supported by China National Institute of Standardization through the "special funds for the basic R&D undertakings by welfare research institutions" (522019Y-6668, and Project "International Standards Development project: ISO 28564-4 Public Information Guidance Systems Part 4: Guidelines for installation").

References

1. Statistical Communique of the People's Republic of China on 2019 National Economic and Social Development. http://www.stats.gov.cn/
2. Zou, C.Y.: Research on Standardization of Public Information Guidance Systems for Railway Stations. AHFE (2019)

Empirical Study on Evaluation of Subway Wayfinding System: A Case Study of Shandong Province, China

Chuanyu Zou[1](✉), Yongquan Chen[1], and Jindong Gao[2]

[1] AQSIQ Key Laboratory of Human Factors and Ergonomics (CNIS), Beijing, China
zouchy@cnis.ac.cn
[2] Shandong Institute of Standardization (SDIS), Jinan, China

Abstract. *Object:* The proposed evaluation indexes of the subway wayfinding systems was verified in Shandong Province. On the one hand, the evaluation indexes will be further revised and improved, and on the other hand, the subway wayfinding system in Shandong Province will be comprehensively evaluated to provide suggestions for further improvement. *Method:* Questionnaire and multi-level sampling method was used to sample Qingdao City. *Results:* The score of Line 3 is higher than that of Line 2 and Line 11. Using the Normative index score as the dependent variable and the subway line as the independent variable, a one-way analysis of variance was performed. It was found that there were no significant differences in the three subway lines in the Normative index. Line 2 and Line 3 have no significant difference in Normative index.

Keywords: Wayfinding system · Subway · Evaluation index

1 Introduction

In order to better evaluate the construction level of China's subway wayfinding system and provide feasible suggestions to improve service quality, a study on the evaluation criteria of subway wayfinding systems was launched in 2018. At the beginning of 2019, the evaluation indexes for the subway wayfinding systems were basically drafted. Based on this, this study chose Shandong Province to conduct field verification. On the one hand, the results may help further revise and improve the evaluation indexes, and on the other hand, a comprehensive evaluation of the subway wayfinding system in Shandong Province was carried out to provide further suggestions for improvement.

Shandong Province, a coastal province in East China. At present, only Jinan City and Qingdao City have subway lines. As a province with a late start on subway construction, the subway wayfinding systems in Shandong Province are representatives of the level of subway wayfinding system construction in recent years. This empirical study selected Qingdao, the first city in Shandong Province to conduct the research.

Qingdao is located on the southeast coast of the Shandong Peninsula and in the forefront of the China-Japan-Korea Free Trade Area. Qingdao is an important coastal city and an international port city, Northeast Asia international shipping hub and

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 399–404, 2020.
https://doi.org/10.1007/978-3-030-50943-9_50

maritime sports base. As early as 1935, Qingdao proposed to build a subway, which was the earliest city in China to propose subway planning in urban transportation planning. In 1989, Qingdao compiled and completed the “Second Line and One Ring” network plan for the urban area. Qingdao is the first city in Shandong Province which has subway lines. Its first subway line, Line 3, opened for trial operation on December 16, 2015. Till December 2019, Qingdao Subway has opened 4 lines, they are: Line 3, with a total of 22 stations and a total length of 25.2 km; Line 2 Phase 1, with a total of 21 stations, a total length of 25.2 km; Line 11 has 22 stations with a total length of 58.4 km. There are 21 stations in the southern part of Phase 1 and Phase 2 of Line 13, with a total length of 67.0 km. In 2019, Qingdao Subway sent a total of 188 million passengers, with a maximum daily passenger volume of 848,800 [1].

2 Research Background

2.1 Standard Drafting Background

The subway wayfinding system, as an information system directly facing and serving passenger travel activities, is an organic part of the infrastructure construction and management of subway stations. It not only facilitates passengers to identify public facilities, identify directions, and find routes. To a large extent, it meets and facilitates the needs of passengers for their lives, work and tourism, and at the same time can effectively maintain the public order of the station and improve the capacity of the station.

Especially with the rapid development of China’s subway, the subway is changing from a single line operation to network operation. The construction trend of subways has become diversification of investment and operating entities. If there is no standardized wayfinding system that is easily identifiable and versatile within the station, it will greatly affect the function of the subway station. However, how can we set up a high-quality subway wayfinding system or improve the existing wayfinding system at the station? This requires the evaluation standards of the subway wayfinding system to evaluate and test the rationality and scientific of the subway wayfinding system, and to provide a scientific basis for the construction and acceptance of the station construction. However, this area is still blank. For this reason, the establishment of basic standards for evaluation has become an urgent problem that needs to be solved at this stage.

In 2018, the Wayfinding systems Sub-Committee of the National Technical Committee for Standardization of Graphical Symbols initiated the compilation of the national standard “Public Information Guidance System– Evaluation Requirements for Subway Stations”. In 2019, the draft of the standard was completed.

2.2 Evaluation Indexes

The evaluation index of the subway wayfinding system is composed of 4 dimensions and 10 first-level indexes.

2.3 Functional Indexes

Extent to Which Information Presentation Meets the Needs of Wayfinding Inside the Station. a) The key information of the subway station presented on the wayfinding elements is accurate, and the information is expressed in the form of graphic symbols, text, line number and line color; the content of the information includes the name of the station, the entrance/ exit code (or number), and the line (line number), Line color, full station names, ticket (self-service ticketing, manual ticketing), security check, etc.; b) The key directions and orientation information presented on the wayfinding elements are clear, mainly including the way-in direction, train running direction, transfer direction, way-out direction, inter-station orientation, and entrance, etc.

Extent to Which the Element Configuration Meets the Needs of Riding Inside the Station. a) Key equipment facilities and services are provided with location signs, including entrances, security checks, platforms, and bathrooms; b) Key nodes (such as entrances and exits, security checks, stairs, escalators, barrier-free elevators), and intersection points have installed direction signs.

2.4 Normative Indexes

Extent of Standardization of Wayfinding Elements Design. a) The design of wayfinding elements is normative, such as the relative position and relative proportion of information elements [2]; b) The design of barrier-free signs is normative, including the location signs and direction signs of various barrier-free facilities.

Extent of Standardization of Wayfinding Elements Setting. a) the reasonable setting of signs at key nodes and intersection point in the station (including the setting of points, setting methods, etc.) [3]; b) the setting of signs is easy to read.

Extent of Systematization of Wayfinding Elements. a) Same information is presented in the main wayfinding elements; b) Information presented by the wayfinding elements in the station is necessary to connect with the wayfinding system of other transportation facilities outside the station, such as the subway wayfinding system and the public transport wayfinding systems are connected.

Extent of Continuity of Wayfinding System. I.E. the continuous transmission of wayfinding information within the station and the uninterrupted guidance chain include: a) Continuous wayfinding information; b) Continuous inbound wayfinding route; c) Continuous outbound wayfinding route; d) Continuous transfer wayfinding route.

2.5 Safety Indexes

Safety of Shape and Appearance of the Wayfinding Elements. a) No sharp protrusions or sharp edges at the outer surface of wayfinding element; b) No slip threaten for the wayfinding element set on the ground.

Safety Installation of the Wayfinding Elements. a) The wayfinding elements are installed firmly; b) There is no potential danger after the wayfinding elements are installed.

2.6 Coordination Indexes

Coordination between the Wayfinding System and the Environment. a) The specifications of the wayfinding elements are suitable for the space in which they are located; b) Wayfinding elements are environmentally-friendly and energy-saving.

Maintenance of Wayfinding Elements: The wayfinding elements have been maintained normally, such as materials, information content, light sources, etc. that can meet actual needs.

3 Methods

3.1 Station Sampling

The selection of subway stations uses multi-level sampling. First, identify and sample subway lines, as there are fewer subway lines, all subway lines are included in the sampling. Second, register all subway stations in sequence for each subway line. Third, sampling subway stations according to the list of subway stations randomly. Taking Qingdao Subway Line 2 as an example, starting from the transfer station, every N (N > 1) subway stations are selected as samples.

According to the requirements in the standard, in accordance with the multi-level sampling method, 12 ordinary stations and 3 transfer stations were selected as verification samples on 4 subway lines in Qingdao. Among them, Line 2 randomly selects Licun Park as the first station, and then every 6th station is selected as a sample; Line 3 randomly selects Qingdao North Station as the first station, and then every 5 stations are selected as samples; Line 11 randomly selects Shandong University as the first station, and then every 3 stations are selected as samples; Line 13 randomly selects Yinzhu Station as the first station, and then every 2 stations are selected as samples.

In addition, in consideration of the current situation of passenger flow, 3 transfer stations of Qingdao Subway were selected as samples for verification.

3.2 Participant Sampling

In each of the selected subway stations, a systematic evaluation questionnaire was issued in accordance with the principle of random sampling.

A total of 260 questionnaires were distributed and 260 were returned, of which 251 were valid questionnaires, and the effective rate was 96.5% (see Table 1, Table 2).

Table 1. Cross list of participants' gender, age and place of residence.

Place of residence			Age							Subtotal
			<20	20–29	30–39	40–49	50–59	60–69	>70	
Local	Gender	M	16	46	21	12	5	8	1	109
		F	25	29	15	2	1	4	0	76
			41	75	36	14	6	12	1	185
Other provinces	Gender	M	8	22	5	1	0			36
		F	9	16	3	1	1			30
			17	38	8	2	1			66
Total	Gender	M	24	68	26	13	5	8	1	145
		F	34	45	18	3	2	4	0	106
			58	113	44	16	7	12	1	251

Table 2. Cross list of subway line and place of residence.

		Local	Other provinces	Subtotal
Subway line	Line 2	70	24	94
	Line 3	57	36	93
	Line 11	58	6	64
Total		185	66	251

3.3 Questionnaire

The accuracy, rationality and scientific of the survey results depend on the design of the survey questionnaire. In order to truly reflect the problems existing in the proposed indexes and provide a basis for the preparation of the standard, the content of the survey questionnaire, the questionnaire form, and the amount of information have been carefully designed. In accordance with the principles of "conciseness, science, and effectiveness", the questionnaire is divided into two parts, the investigator questionnaire and the participants questionnaire.

The investigator questionnaire includes the survey date, survey time, survey station, survey line, construction period, etc. The feedback from the investigator questionnaire is convenient for us to understand the actual situation of the verification site and provide support for later data analysis.

A questionnaire was developed based on the evaluation indexes in the national standard. The questionnaire was divided into 4 dimensions, 25 items, and 5 points were used, including forward and reverse scoring. The higher the score on the questionnaire, the higher the participants' evaluation of the subway wayfinding system. The questionnaire also designed basic information such as the reason for taking the subway, gender, age, and place of residence, as well as three open-ended questions to understand the participants' satisfaction with the subway wayfinding systems.

4 Conclusion

The data was analyzed in accordance with the mathematical model of evaluation indexes in the national standard "Public information guidance system– Evaluation requirements for subway stations".

In the evaluation of three subway lines in Qingdao, Line 3 got the highest score, Line 11 got the lowest score, and line 2 ranks second. Using the wayfinding system scores as the dependent variable and the subway line as the independent variable, a one-way analysis of variance was performed. It was found that the three subway lines had no significant difference in the wayfinding system evaluation.

The score of Line 3 is higher than that of Line 2 and Line 11. Using the Normative index as the dependent variable and the Line as the independent variable, it was found that there were no significant differences in the three Lines in the basic norm. Line 2 and Line 3 have no significant difference in Normative index.

Acknowledgments. This research was supported by China National Institute of Standardization through the "special funds for the basic R&D undertakings by welfare research institutions" (522019Y-6668, and Project "International Standards Development project: ISO 28564-4 Public Information Guidance Systems Part 4: Guidelines for installation").

References

1. Baidu Baike. https://baike.baidu.com/item/青岛地铁/9670422
2. GB/T 20501 Public information guidance systems—Design principles and requirements for guidance elements
3. GB/T 15566.4 Guidance system for public information - Setting principles and requirements - Part 4: Public transport station

Analysis on the Redesigned Metro Safety Signs Based on Eye Tracking

Guilei Sun[1](✉), Yanhua Meng[1](✉), Qin Li[2], Zijie Wan[1], and Yaqi Wang[1]

[1] Department of Safety Engineering, China University of Labor Relations, Beijing 100048, China
sunguilei@126.com
[2] Beijing Shunjinsheng Construction Engineering Supervision Co., Ltd., Beijing 101300, China

Abstract. To improve the identification efficiency of safety signs in the subway, investigations on the safety signs of the subway were carried on. 158 participants were invited to participate the questionnaire test and the experiment to the redesigned safety signs. Tobii X2-30 was used to collect eye movement data and the area of interest (AOI). Wilcoxon rank sum test and Mann-Whitney U test were used to analyze the difference of the new safety signs and the original signs on the first saccade latency, the first fixation duration, and the total fixation duration in AOI. The results show that the value of P for four signs are less than 0.05. The analysis to eye movement are consistent with the selected subjects for the participants, which indicates that the new designed safety signs are better than the original. And the parameters that mentioned can be used for the test of signs.

Keywords: Safety signs · Eye movement indexes · Tobii eye tracker · Subway · Significance test · Redesigned signs

1 Introduction

Safety signs are designed to enable people to quickly identify potential hazards and effectively avoid dangerous accidents. The subway safety sign plays a vital role for the safe operation. The imperfection of the safety sign will reduce people's vigilance against the danger source of the subway. Research shows that the safety signs in the subway have defects such as unclear content and inconsistent use [1]. These problems seriously affect the efficacy of the security signs. Eye movement response is closely related to psychology, combined with eye movement features to improve design and achieve better design results. It has been used in road traffic, sports psychology, and landscape interface design [2]. At present, some scholars use eye trackers to analyze and study dashboard designing, safety signs, and attention [3–6]. Li et al. [7] proved that the color and shape of safety signs will affect people's attention; Chen et al. [8] studied the cognitive load of safety sign design and proposed that red and graphic content cannot improve the accuracy of cognition, but it can be recognized and understood faster in visual search and scanning activities; He et al. [9] evaluated the

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 405–412, 2020.
https://doi.org/10.1007/978-3-030-50943-9_51

effectiveness of the safety signs in densely populated places, and showed that the significance of safety signs was conducive to the control effect of the crowd. Li et al. [10] suggested that people's understanding to the safety signs would be affected by gender, education and age, and pointed out that the original safety signs were not simple enough to be designed, and the recognition is low, which could not meet the needs of different groups.

It is very important for the validity of safety sign recognition. This study focused on the typical safety signs in the subway, combining subjective. Questionnaires, critical flicker frequency, and eye tracking data analysis are used in the experiment.

2 Experiment Design

2.1 Subjects and Stimulating Material

The safety signs of the subway were taken as the research objects. 500 subway passengers, 274 males and 226 females, were randomly surveyed at the peak rush hour to get the gender ratio and the safety signs that easy misleading. According to the gender ratio, 158 people were selected through external recruitment. After screening, 150 valid samples were obtained, 82 males and 68 females, all of whom have no eye movement test experience. And 9 types of safety signs according to the frequency of occurrence were redesigned to the survey results. Based on the design flaws of the original safety signs, a total of 27 safety signs were redesigned, as shown in Table 1. The proportions, dimensions and colors of the design drawings refer to China Standard GB/T 2893.1-2013 [11]. 9 types of safety signs in the two categories of prohibition and warning.

2.2 Equipment for Experiment

Tobii X2-30 eye tracker was used to collect the eye movement data; Microsoft HD camera was used to record the whole process of the subject. The screen size is 19.7 in. (44 cm $\times$ 24 cm), the screen resolution is 1920 $\times$ 1080.

2.3 Experiment Procedure and Data Processing

The test distance from the subjects to the display screen was about 45 cm–60 cm. Each set of safety signs was interspersed with text explanations so that the participants could know the type of the sign he need to choose. In order to avoid the tendency of selection when the subject selected, the positions of the original images were set randomly in every page. When a set of safety signs appears, the subjects selected the pictures that they considered as the most relevant to the meaning of the signs. After the experiment, the participants were interviewed about the reasons for selecting the safety signs and the subjects' satisfaction to different safety signs of the same type.

Area of interest (AOI) was used to analyze the eye movement data pre-processing. The test stimuli included a total of 36 regions of interest. The AOI was divided into the inner area of each safety signs outline in the appearing page, and the data was sorted and analyzed by SPSS V21.0.

Table 1. 9 Typical metro safety signs & redesign

Safety signals	No Jumping Down	No Touching	No Cart Riding Ladder	Prohibition into Tunnel	No sale	Watch Out Steps	Watch Your Hand	Watch Your Head	Warning! Electric Shock
Original									
1									
2									
3									

Note: 1, 2, and 3 means different redesigned sign, the follows are the same.

3 Data Analysis

Subjective data is mainly based on questionnaire to obtain the most appropriate sign that fitting the given meaning. Eye movement data of the safety signs includes gaze maps, heat maps, first saccade latency in the AOI and the first fixation, the duration of the first gaze point entering the AOI.

The proportions are shown in Table 2. It shows the selection ratio of the redesigned signs is higher than the original safety signs except “No Sale”, which indicates that most of the original safety signs does need to be improved to attract passengers’ attention or avoid ambiguity. The relevant literature indicated that people’s understanding to the signs is open. When choosing a right sign, the participants considered a variety of different factors, such as personal experience, habits, and environment [12]. Therefore, according to the safety signs selected by the participants during the experiment, the factors affecting the subjects’ choice are divided into three categories: personal factors, situational factors, and empirical factors. The redesigned signs with the highest proportion among each type of safety signs are selected and have been compared with the original safety signs. It indicates that the reasons which affect the choice of the subjects are vivid and easy to understand. Most of the signs that chosen as the best signs are the redesigned ones except “No Sale”.

Table 2. Proportions of people choosing different safety signs

Signs name	Number	Proportion (%)	Signs name	Number	Proportion (%)	Signs name	Number	Proportion (%)
No Jumping Down	Original	3.00	Prohibition into Tunnel	Original	15.00	No Touching	Original	18.00
	1	17.00		1	52.00		1	12.00
	2	63.00		2	12.00		2	53.00
	3	17.00		3	21.00		3	17.00
Watch Your Hand	Original	7.00	No Sale	Original	47.00	Watch Your Head	Original	18.00
	1	25.00		1	7.00		1	20.00
	2	63.00		2	15.00		2	45.00
	3	5.00		3	31.00		3	17.00
No Cart Riding Ladder	Original	5.00	Watch Out Steps	Original	18.00	Warning! Electric Shock	Original	23.00
	1	15.00		1	10.00		1	10.00
	2	12.00		2	10.00		2	5.00
	3	68.00		3	62.00		3	62.00

3.1 Subjective Data Discussion

3.2 Eye Movement Data Analysis

1. Descriptive statistical analysis of eye movement indicators

"No Jumping Down" is taken as an example to analyze the original safety sign and the redesigned safety signs to research which is the best (that is, the highest selection ratio) by the eye movement data. It is necessary to study whether the safety signs need to be improved and optimized.

The descriptive statistics of the "No Jumping Down" eye movement data, only the original and the best redesigned, are shown in Table 3.

Table 3. Descriptive Statistics of Gaze Data of "No Jumping Down"

AOI		First saccade latency/s		First fixation duration/s		Total fixation duration/s	
		Mean	SD	Mean	SD	Mean	SD
No Jumping Down	Original	1.432	0.836	0.427	0.275	1.547	1.360
	Redesigned	1.446	0.936	0.441	0.244	1.726	1.319

It can be seen from Table 3 that the first saccade latency, the first fixation duration, and the total fixation duration of the redesigned safety signs are compared, and the gaze data of the redesigned sign corresponding to "No Jumping Down" is better than the original image, which is 1.00%, 3.3%, and 11.6% higher than the original image. Here, the first saccade latency means the time from the presence of the stimulus material to the first entry into the area of interest. The shorter the first saccade latency, the more noticeable the interest zone [13]. The first fixation duration refers to the gaze time of the first gaze point in the area of interest. It reflects the amount of information in the gaze area or the degree of attraction of the area to the subject [14, 15]. The total fixation duration reflecting the difficulty of information extraction, the longer the duration, the more difficult it is for the subject to obtain information from the area or the more interested the area [16]. In this research, the meaning of the sign has been told the subject, therefore, it means the subject is interested in the sign.

By analyzing the eye movement data of the "No Jumping Down" sign, the following results can be obtained: The average time of the first saccade latency of the redesigned signs is higher than the original, which indicating that the original is likely to attract the attention of the subject and is more conspicuous than the redesigned. The average of the first fixation duration is higher than the original, indicating that the participants carefully obtain the information conveyed by the redesigned and the redesigned sign attracts the participants more. The average of total fixation duration is higher than the original, that is, the subject's fixation duration for the redesigned is longer under the condition of knowing the meaning of the signs. Further data analysis needs to proceed to find the difference between the original and the redesigned.

2. Difference analysis

In order to study whether there are significance differences between the original safety signs and the redesigned in the eye movement data. The analysis variables are continuous variables. If the two variables of the index meet the normal distribution at the same time, independent-samples t test is used. If two variables are not satisfied at the same time for the state distribution, Wilcoxon rank sum test is used. The Wilcoxon rank sum test is selected by exploring and testing the data. The results of the difference test between the original and the redesigned at the $\alpha = 0.05$ level are shown in Table 4.

Table 4. Difference analysis between original sign and redesigned sign

Index	Sig.
First saccade latency	0.796
First fixation duration	0.008
Total fixation duration	0.026

Table 4 shows that there is no significant difference between the original "No Jumping Down" sign and the redesigned one on the first saccade latency of the AOI area. The results are not statistically significant. However, there is a significant difference between the original and the redesigned signs in the first fixation duration and the total fixation duration. The eye movement gaze plots and the heat maps of the subject are shown in Fig. 1 and Fig. 2, respectively. Picture 2 in Fig. 1 is more vivid and can attract the attention of the subject, which means it can achieve better expression and is easy to understand. It is consistent with the subjective test results that 63% of the subjects choose picture 2. And it shows that the redesigned signs are more attractive to the participants than the original. Through the experimental data, if the redesigned could attract more attention, it would be better to use the new designed "No Jumping Down" sign than the original.

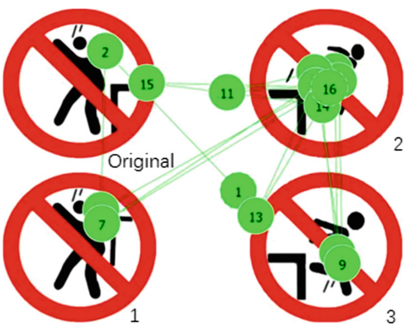


Fig. 1. Gaze plots of "No Jumping Down"

Fig. 2. Heat map of "No Jumping Down"

3. Significant test results of other "safety signs"

According to the analysis of the "No Jumping Down", the remaining eight types of safety signs are analyzed for differences. The results of the significance test at $\alpha = 0.05$ level is shown in Table 5.

Table 5. The result of significance tests for other safety signs

Safety signs	First saccade latency	First fixation duration	Total fixation duration
No Touching	0.000**	0.001**	0.000**
No Cart Riding Ladder	0.000**	0.003**	0.000**
Prohibition into Tunnel	0.000**	0.001**	0.000**
No Sale	0.662	0.387	0.146
Watch Out Steps	0.773	0.116	0.287
Watch Your Hand	0.000**	0.006**	0.106
Watch Your Head	0.000**	0.245	0.466
Warning! Electric Shock	0.039*	0.000**	0.000**

Note: **Extremely significant difference (P < 0.01); *Significant difference (P < 0.05)

According to the results of the significance test in Table 5, the results of the significant analysis to the first saccade latency show that 6 types of safety signs having significant difference to the original, as "No Touching", "No Cart Riding Ladder", "Prohibition into Tunnel", "Watch Your Hand", "Watch Your Head", and "Warning! Electric Shock". It indicates that the redesigned 6 signs are much more noticeable while the original signs of "No Sale" and "Watch Out Steps" are better than the redesigned. Because the meaning of the signs has been given, the first fixation duration, the first fixation duration and the total fixation duration is nearly consistent on the signs except "Watch Your Hand" and "Watch Your Head".

In summary, it shows that the redesigned signs of "No Touching", "No Cart Riding Ladder", "Prohibition into Tunnel", and "Warning! Electric Shock" are more conspicuous, and the information obtained is more comprehensive. These four types of signs can be considered as new signs.

4 Conclusion

1. The first saccade latency and the first fixation duration can be used as indicators of whether the identification of safety sign is eye-catching. The total fixation duration can be used to characterize the passenger's attention to the safety sign.
2. The subjective and objective test and the significant differences shows: the poor identifiability signs such as "No Touching" and "No Cart Riding Ladder" should be

replaced. “No Jumping Down”, “Prohibition into Tunnel”, and “Warning! Electric Shock” should improve the function to obtain more attention to replace the original signs.

3. The design elements of the safety sign should be closer to the reality to make it meet the requirements of vivid and easy to understand rather than oversimplification.

Acknowledgements. The presented work has been supported by General project of China University of Labor Relations (20XYJS020).

References

1. Wang, L.: Safety certification logo in China: situation and tasks. Saf. Coal Mines **07**, 79–80 (2007)
2. Chen, J.J., Yan, H., Xiang, J.: Study of decoding mental state based on eye tracks using SVM. Comput. Eng. Appl. **47**(11), 39–42 (2011)
3. Sun, G., Li, Q., Meng, Y., Ran, L.: Research on optimization of information coding for car dashboard based on eye movement analysis. In: International Conference on Applied Human Factors and Ergonomics, pp. 122–135. Springer, Cham (2019)
4. Sun, G., Li, Q., Fu, P., Ran, L.: Analysis and optimization of information coding for automobile dashboard based on human factors engineering. China Saf. Sci. J. **28**(8), 68–74 (2018)
5. Guilei, S.: Research on location of emergency sign based on virtual reality and eye tracking technology. In: International Conference on Applied Human Factors and Ergonomics, pp. 401–408. Springer, Cham (2019)
6. Sun, G., Li, Q., Meng, Y., Ran, L., Fu, P.: Design of car dashboard based on eye movement analysis. Packing Eng. **41**(2), 148–160 (2020)
7. Li, L.N., Jiang, W., Zhang, Y.N.: Background shape and the background color of safety sign. China Public Secur. **1**, 117–120 (2014)
8. Bian, J.: Research on the Perception Mechanism of Safety Management: A Neuro-Industrial Engineering Perspective. Zhejiang University (2014)
9. He, Q.H., Song, S.X., Gu, Y.B., Xiao, C.Y.: Practical example evaluation on effectiveness of safety signs in assembly occupancies. J. Saf. Sci. Technol. **11**(09), 132–137 (2015)
10. Li, S.B., Jiang, W., Zhang Y.N., Li, L.N.: Analysis of individual influence factors on safety sign understanding in mine workplace. Min. Saf. Environ. Prot. **42**(4), 113–116 + 120 (2015)
11. GB 2893.1: Graphic Symbols Safety Colors and Safety Signs-Part 1: Design Principles for Safety Signs and Safety Marking (2013)
12. Hu, Y.C., Zhou, X.H., Wang, L.: Investigation into recognizability of safety signs: sign features and user factors. China Saf. Sci. J. **23**(3), 16–21 (2013)
13. Cui, C.C.: Study on Safety Sign’s Recognition Based on Eye Tracking Technology. Henan Polytechnic University, pp. 36–45 (2014)
14. Zhang, X.F., Ye, W.L.: Review of oculomotor measures in current reading research. Stud. Psychol. Behav. **03**, 236–240 (2006)
15. Zhang, K., Cui, C.C., Niu, G.Q., Jing, G.X.: Research on the visual attention and identification features of the shapes and colors of the safety signs. J. Saf. Environ. **14**(06), 18–22 (2014)
16. Liu, C.: The Evaluation Method of the Machinery Cab Interior Based on The Eye-moment Research. Shandong University. pp. 81–84 (2010)

TIAMBIENTA Smart Technologies for the Motor-Home Sector

Giuseppe Lotti[1(✉)], Marco Marseglia[1], Elisa Matteucci[1], Margherita Vacca[1], Irene Fiesoli[1], Claudia Morea[1], Alessio Tanzini[1], Francesco Cantini[1], Lu Ji[1], and Eleonora Trivellin[2]

[1] University of Florence, Florence, Italy
{giuseppe.lotti,elisa.matteucci}@unifi.it
[2] University of Ferrara, Ferrara, Italy
eleonora.trivellim@unifi.it

Abstract. The contemporary scenario sees the substantial affirmation of innovative technologies as digital tools necessary for the creation of an increasingly connected society. These technologies have a systemic relevance as they feed the value of the production system chain and have the ability to innovate processes, products and services in all economic sectors of human activity. The TIAMBIENTA project is focused on creating an ecosystem of home automation services through the use and integration of the new tools of digital technology for the design of the interior spaces of the motorhome. The project involves the construction of a cloud home automation sensor system. The sensor network will be interconnected through a control unit capable of collecting data, sending it to the central cloud platform and introducing direct implementation protocols on the internal environment control systems of the motorhome. Among the various enabling technologies that we have taken into consideration, a fundamental role has been given to ICTs, information and communication technologies, fundamental for collecting the data and recording and adjusting the user experience.

Keywords: Motorhome · Design · Smart technologies · ICT · 4.0 industry · KETs · Artificial intelligence · Wellness · Safety maintenance

1 Introduction

Technology - no matter how well designed - is just a magnifying glass of intent and human ability (Toyama, 2010).

In line with the objectives set in the 2030 Agenda, the European Commission is working to ensure that all citizens can fully benefit from the great opportunities offered by innovative information technologies and by a truly connected society. In a contemporary scenario that sees the substantial affirmation of digital tools, these technologies are better known in the economic-productive field with the name of

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 413–418, 2020.
https://doi.org/10.1007/978-3-030-50943-9_52

fundamental enabling technologies (KETs)[1], which - paraphrasing what is reported in the Horizon 2020 program - have a systemic relevance since they feed the value of the production system chain and have the ability to innovate processes, products and services in all economic sectors of human activity [1].

We hear about Scenario 4.0, or better Industry 4.0, for the first time in 2011 at the Hannover Fair in Germany. On that particular occasion, in fact, a group of three engineers - Henning Kagermann, Wolf-Dieter Lukas and Wolfgang Wahlster[2] - presented a programmatic report to the German federal government for the implementation of the Country's Industrial Plan. In Italy, however, the term officially only appeared in 2016, in the 2017–2020 "Industry 4.0" National Plan document, within which a set of measures capable of encouraging investments in the new sector is exposed. The aim of the plan was to integrate some new digital technologies in the industrial sector to improve working conditions, create new business models and increase the productivity and production quality of the plants. This scenario is rapidly leading companies towards an increasingly automated industrial production, or towards the fourth industrial revolution [2].

Industry 4.0 is therefore a direct consequence of the digitization in the production field which has been affirming for a few years now. A process defined by experts as digital transformation[3], which will lead companies to face a double reality for the first time: physical and virtual resources must be managed at the same time, considering them as a single industrial production system. Especially with regard to virtual and digital resources, industries will need to know their potential in order to use them to their advantage [3].

2 Tiambienta Project

Starting from these premises, the Tuscany Region, following the dictates of the aforementioned National Plan "Industry 4.0", promptly aligned itself with the new contemporary scenario, promoting calls for the realization of strategic research and development projects that provide for the application of the new digital technologies in

[1] "Fundamental enabling technologies (KET) are investments and technologies that will allow European industries to maintain competitiveness and capitalize on new markets. The Industrial Technologies (NMP) program focuses on four KETs: nanotechnology, advanced materials and advanced production and treatment (production technologies) and biotechnology". See https://ec.europa.eu/programmes/horizon2020/en/area/key-enabling-technologies.

[2] The German term Industrie 4.0 was first introduced in 2011 at the Hannover Fair in Germany and was invented and promoted by three engineers: Henning Kagermann (physicist and one of the founders of SAP), Wolfgang Wahlster (professor of intelligence artificial), and Wolf-Dieter Lukas (German physicist and senior official from the Federal Ministry of Education and Research) who spoke about it during a press conference.

[3] In this regard, the definition on digital transformation given by The Agile Elephant, one of the first consultancy agencies that studied and theorized the phenomenon, is interesting: "Digital transformation involves a change of leadership, a different way of thinking, new models of business and greater use of technology to improve the experience of employees, customers, suppliers, partners and all interested parties of the organization".

the regional production system. The TIAMBIENTA project - an acronym for the extended title Intelligent Technologies for Living Environments - is the happy result of these avant-garde policies adopted by the Tuscan territory [4].

The aim of the project is to create an ecosystem of home automation services through the use and integration of the new tools of digital technology for the design of living environments belonging to the sectors of the nautical, camper and spaces for education, sectors these of great importance and competitiveness for the productive fabric of the Tuscan territory. Specifically, with regard to technological solutions, the TIAMBIENTA project involves the construction of a system of home automation sensors in the cloud that can be installed in different living spaces. The sensor network will be interconnected through a control unit capable of collecting data, sending it to the central cloud platform and introducing direct implementation protocols on the control systems of the reference environment. Finally, the cloud platform will have to manage the different data structures, collect and process them through data analysis and control modules. Among the various enabling technologies that will be taken into consideration in the development of the research project, a fundamental role will be given to ICT, i.e. the information and communication technologies that will be used to collect the data collected from the range of devices used (sensors and actuators), recording and regulating the user experience [5, 6].

With a view to multidisciplinarity, the project sees the complementary collaboration of productive districts of considerable interest for the competitiveness of the Tuscan regional welfare: the Experimental Center of Furniture and Furnishings (Interior District) and Navigo (Nautical District). Going into more detail, among the industrial partners involved there are Trigano Spa for the camper sector, Spazioarredo for spaces dedicated to children and education, Overmarine Group for the nautical sector. As for the scientific part, the bodies involved are Corobotics, spinoff of the University of Pisa (IoT, robotics and cloud), Studio Sgro (safety in the nautical field) Cubit (IoT) and the Department of Architecture and Design of the University of the Studies of Florence (design, interface, usability) [7].

The objectives of the research project include cross-cutting themes of considerable interest: on the one hand the improvement of well-being, comfort and safety for users, and on the other the control and monitoring of living environments and products with a view their better management and maintenance. The final objective is the competitive growth of the companies involved who see the application of digital technologies as a necessary strategy to increase the value of the offer and therefore the competitiveness of their respective products and services, for the benefit of the end user. The quid necessary to determine the competitive plus may come from an overall and multidisciplinary design that takes account of the important contribution of the digital team, but also of the user experience.

Overall, the proposed technological and digital solutions focus on two specific topics of interest in the field of home automation: services to people and businesses and the competitiveness of companies. With this in mind, the TIAMBIENTA project foresees for the industrial partners involved the creation of prototypes, relating to certain living environments: sleeping space of the living room of the camper, the child's bedroom of the yacht and a collection of furniture for kindergarten. In the design and implementation of the project outputs, an important role is occupied by the contribution made by the scientific research bodies, as well as the contribution and support given by the two districts involved.

The development of the project saw in the preliminary phase the definition of user needs, services, needs and technological solutions for each production sector involved and the selection of the most suitable technologies capable of providing an effective response. Through plenary meetings with all the partners and more specific focus groups among only the scientific partners, it was possible to outline a state of the art of technological and digital solutions on the market, assuming design concepts for the respective prototypes. Of fundamental importance, thanks to the help of the district, was the understanding of the actual needs of the companies involved in terms of market competitiveness and possible development scenarios for the companies themselves.

The general brainstorming phase also saw a visit to the company headquarters and an inspection of the production units for a better understanding of the living environments in question. Indeed, the study and assessment of the positive and negative aspects that can be found in the user experience in reference to the project objectives, namely safety, well-being and maintenance, were fundamental. All the information collected was deepened in the focus group sessions between the research bodies involved. The quantity and range of data collected was very valuable in defining a critical map associated with each of the three industrial partners, in which the different problems encountered were divided into macro-categories, representing the initial objectives of safety, well-being and maintenance of the living environments under consideration.

After the conclusion of the state-of-the-art definition phase regarding the possible technological and digital solutions to be adopted in the design phase, another interesting contribution was made by the University. In fact, as part of the didactic offer of the Master of Science in Design of the University of Florence, a thematic seminar was entitled, entitled Design 4.0, in collaboration with the company Spazioarredo, one of the three industrial partners of TIAMBIENTA. The aims of the seminar - held in the first semester of the 2018/19 academic year and coordinated by the Design Laboratory for the Sustainability of the Design Campus of Calenzano - were to redefine and rethink furnishings and spaces for children, according to a logic more sensory and emotional, in line with the principles of home automation, smart objects and the Internet of Things [8]. The students who participated in the seminar were therefore able to take advantage of the research defined by the scientific part, also studying the critical issues and strengths of the Spazioarredo company and finally developing the design concept not of an entire collection, but of an object or a series of objects that was in line with the initial brief. The projects developed by the Design students showed from the outset how much in the design field it is necessary and saving to change the point of view of approach to the problems. Indeed, precisely from this perspective, the contribution given by the participants to the seminar was a significant source of inspiration in terms of the formal, conceptual and technological solutions proposed. In addition, a curiosity towards digital technology, towards the development of the sensorial and emotional aspect that its use entails, as well as the desire to deepen these issues and develop the projects themselves presented during the seminar as a thesis conclusion of the cycle of studies.

The TIAMBIENTA project started in 2018 and is still ongoing. Currently, the continuation of the work is awaiting the development of the first prototypes of the sensors and actuators by the relevant research bodies (Cubit and CoRobotics). In parallel, the Design Laboratory for Sustainability is carrying out - in close relationship with the three companies involved - the work on the development of the design concepts for the realization, as scheduled, of the final project prototypes. Specifically, among the three prototype projects, the one at a more advanced stage is the collection of smart furnishings for children's spaces, named DOT, in collaboration with the Spazioarredo company. The DOT collection is structured in a system of container furniture equipped with sensors for their sanitation and electric locks, some game tables, a set of shelves with a sensor for safety and a piece of furniture for the teacher equipped with a control panel. After a preliminary phase, fundamental for the definition and approval of the concept, the project proposal moves towards the phase of approval of the executive drawings and production of the prototypes (Figs. 1 and 2).

Fig. 1. Infographic banner of the TIAMBIENTA project

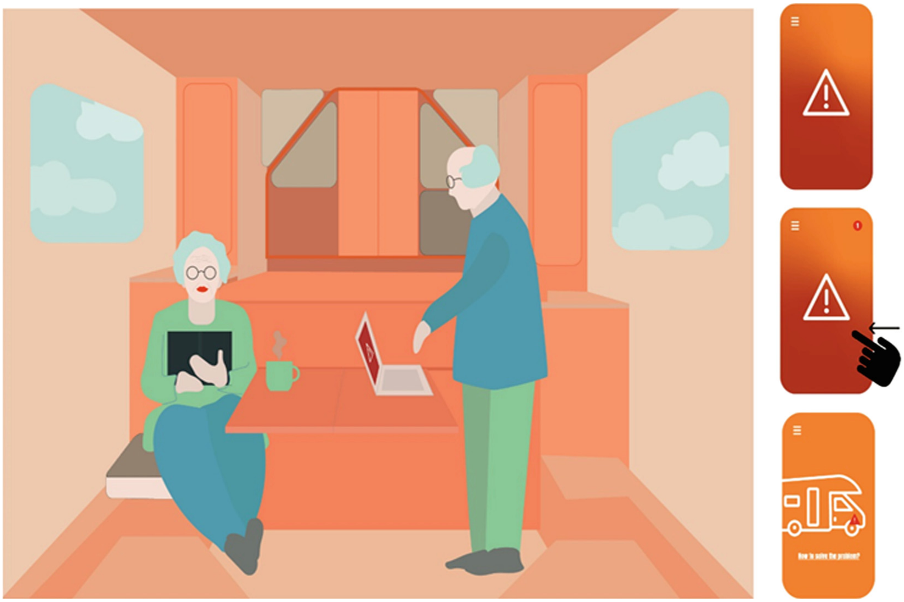


Fig. 2. One of the concept of the TIAMBIENTA project

Acknowledgements. The text was thought of authors all together. In particular Giuseppe Lotti wrote the first and second paragraph, Margherita Vacca wrote the third and fourth paragraph and the references, Elisa Matteucci the fifth and sixth paragraph, Francesco Cantini wrote the seventh paragraph, Alessio Tanzini wrote the eighth paragraph, Irene Fiesoli wrote the ninth paragraph, Claudia Morea wrote the tenth paragraph, Marco Marseglia wrote the eleventh paragraph, Eleonora Trivellin wrote the twelfh paragraph and Abstract, Lu Ji found and edited the images.

References

1. Ryan, J.: Storia di Internet e il Futuro Digitale. Einaudi, Torino (2011)
2. http://www.mise.gov.it › images › stories › documenti › guida_industria_40
3. Caperna, A., Mirabelli, P. (eds.): Partecipazione e ICT per una città vivibile. Gangemi, Roma (2013)
4. Villari, B.: Design per il territorio. Un approccio community centred, Franco Angeli, Milano (2012)
5. Tosi F., Patti I.; Camper. Design, ergonomia, innovazione. Alinea, Firenze (2010)
6. Trivellin, E.: Abitare on the Road. Alinea, Firenze (2006)
7. Lotti, G., Giorgi, D., Marseglia, M.: Prove di design altro. Didapress, Firenze (2017)
8. Za, S.: Internet of things. Persone, organizzazioni e società 4.0. Luiss University Press, Roma (2018)
9. Maldonando, T.: Reale e Virtuale. Feltrinelli Editore, Milano (2015)

Human Factors in Transportation: Maritime

Preventing Unruly Technologies in Maritime Navigation: A Systems Approach

Katie Aylward(✉), Scott N. MacKinnon, and Monica Lundh

Chalmers University of Technology,
Hörselgången 4, 412 96 Gothenburg, Sweden
{katie.aylward, scottm, monica.lundh}@chalmers.se

Abstract. The rapid development of artificial intelligence, big data, and automation have the potential to transform the maritime industry. While change is inevitable, automated solutions do not guarantee navigational safety, efficiency or improved seaway traffic management. This paper describes lessons learned throughout the Sea Traffic Management (STM) Validation project. The STM project tested low-level automation functions intended to enhance navigational safety and efficiency. The results indicate that although the additional information was useful, there may be a disruption in current working practices, and potential gaps in the existing regulatory framework, and training and skill development of future mariners. The aim of this paper is to argue for a systems approach to better understand and prevent technologies from becoming "unruly" in the maritime context. This approach will consider how the technologies impact various system elements to support a more safe and sustainable shipping industry.

Keywords: Human factors · Human-Automation · Navigation · Safety · Sociotechnical system · Maritime

1 Introduction

The shipping industry contributes to approximately 80–90% of the international world trade, employing around 1.6 million seafarers globally [1, 2]. Seafaring is considered to be a high-risk occupation with unique health and safety challenges [1]. For decades, the shipping industry has cited "human factors" or the "human element" as the leading cause (approximately 85%) of maritime accidents [1, 3, 4]. In response, an industry-wide solution has been to reduce human involvement by reducing manning and increasing automation, particularly within navigation activities [1]. While many of these changes have improved maritime safety, they have also altered typical "navigation tasks" and the role of the human operator on board [5–7]. Implementing new technologies in a sociotechnical system requires a careful assessment of the many other system elements that will be impacted through technology integration. These elements include: the technical subsystem (technology, tasks, user interface, equipment), personnel subsystem (operators, teams, training and education), and work design/environment subsystem (processes, regulations, standards). An understanding of interactions between these subsystems is important to determine how technologies have the potential to be

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 421–428, 2020.
https://doi.org/10.1007/978-3-030-50943-9_53

disruptive, or "unruly" [8]. Technologies can become "unruly" as they exist in a social context which is shaped by individuals, organizations, and the system where it is being used [9]. This allows for the technology to be used in different ways than perhaps intended when first developed and creates a conflict between the "work as imagined" and the "work as done" [10].

In order to provide insight over technology development within the domain, the International Maritime Organization (IMO) formally adopted e-Navigation and published their strategy for developing an implementation plan, which was completed between 2015–2019. e-Navigation is defined as "the harmonized collection, integration, exchange, presentation and analysis of marine information on board and ashore by electronic means to enhance berth to berth navigation and related services for safety and security at sea and protection of the marine environment" [11]. One of the key points of this directive is that it must be driven by user needs, including both seafarers and shore side services. The Sea Traffic Management (STM) Project (2014-EU-TM-0206-S) is a European Union (EU) funded project aimed at exploring concepts and applications that are described in the e-Navigation strategy [12]. The STM project examines digitalization and automation through connecting and updating maritime stakeholders (ships, ports, vessel traffic services (VTS), service providers, shipping companies) in real time through information sharing and exchange. The results from the STM Project are selected as the case to discuss how technology can become unruly, leading to potentially disruptive work situations.

This aim of this paper is to argue for a systems approach towards understanding how technologies can become unruly in maritime navigation, based on lessons learned from the STM Project. This paper identifies existing gaps in the maritime sociotechnical system created by a traditional technology-centered development and implementation approach. In order to prevent technologies from becoming unruly, it is important to understand how the technology will be shaped by the technical, personnel, and work environment subsystems.

2 Technology Classification

Automation can be defined as "*the full or partial replacement of a function previously carried out by a human operator. This implies that automation is not all or none, but can vary across a continuum of levels, from the lowest level of fully manual performance to the highest level of full automation"* [13]. Many industry stakeholders have described human-automation interaction with the adoption of a Levels of Automation (LOA) scale. The scales usually range anywhere from 0–10, from 0 = no automation to 10 = fully autonomous, and mixed human-automation task allocations in between [13–16]. The IMO recently published the Strategic Plan for the six-year period 2018–2023 (Resolution A.1110 (30)) which specifically lists a strategic direction (SD2) to "integrate new and advancing technologies in the regulatory framework" [17]. Part of this plan is a regulatory scoping exercise of Maritime Autonomous Surface Ships (MASS). The shipping industry has yet to agree on a unified definition and application of LOA. However, as part of the scoping exercise, the IMO has addressed autonomy of MASS with four degrees, from automated processes/decision support to fully autonomous ships [18]. The STM technology studied aligns with Degree One: “Ship with automated processes and decision support: Seafarers are on board to operate and control shipboard systems and functions. Some operations may be automated and at times be unsupervised but with seafarers on board ready to take control” [18]. The scales and degrees of automation should be seen as a continuum with overlap between the types and degrees of automation, instead of discrete levels. For the purposes of this paper, the technologies discussed (STM Functions) can be classified as “low-level automation”.

3 The Case: STM Validation Project

The goal of the STM Validation project was to develop, validate and verify the STM infrastructure and functions [12]. One of the three testbeds used to validate the concept is the European Maritime Simulator Network (EMSN), a testbed that enables the testing of project navigation functions in complex and large-scale simulation-based traffic situations. The EMSN connected 30 ship simulator bridges across Europe (Sweden, Norway, Finland, Spain, Germany, United Kingdom), tested hundreds of human operators, offering unprecedented opportunities for training and research. The technologies that were developed and tested throughout the STM Validation Project are referred to as the “STM Functions”. They were intended to reduce administrative burden, and increase safety, situational awareness, operational efficiency, and transparency. A description of the functions is provided in Table 1.

Table 1. Description of the STM functions

Ship-to-ship route exchange (S2SREX)	This function provides the navigator with a route segment consisting of the next 7 waypoints of the monitored route of another vessel. Route segments are broadcasted through Automatic Identification System (AIS)
Rendezvous function (RDV)	As an integral part of the S2SREX, this function allows the navigator to view where own ship will meet a target ship if both vessels continue on their monitored broadcasted route with the present speed over ground. This function provides route-based Closest Point of Arrival (CPA) and Time to Closest Point of Arrival (TCPA) based on AIS information
Shore-to-ship route exchange	This function allows the VTS to send a suggested route to the ship, to be reviewed by the bridge team and then either accepted or rejected
Navigational warnings	This function provides a notification which overlays a Navigational Warning Message directly on the Electronic Chart Display and Information System (ECDIS). If the Navigational Warning involves a geographical area to avoid or be aware of, this will be automatically plotted onto the ECDIS

(*continued*)

Table 1. (*continued*)

Chat function	A standalone communication software similar to other programs such as Skype, or WhatsApp which is integrated on the ECDIS station. This function allows chatting between STM enabled ships or with VTS stations
Enhanced monitoring and route cross check	After having received a ship's monitored route and schedule, the VTS Operator will be able to detect if the ship deviates from the monitored route. The VTS has the ability to receive any planned route and cross check the route against any navigational dangers and if necessary and send a route suggestion back to the ship

Two separate data collection efforts were carried out between November 2017 and October 2018 within the EMSN. Detailed information about the respective experiments can be located in the STM Final Report [12], controlled scenarios report [19], and additional project publications [20, 21]. Navigation outcomes were compared between repeated scenarios that either had the STM functions available or not for decision-making by the test participants. Safety and efficiency outcomes were assessed using a mixed-methods approach incorporating both qualitative and quantitative methods. The methodological tools included: post-simulation debriefings, questionnaires, numerical assessment of ship behavior, observations, and subject matter expert (SME) review. The lessons learned from the STM Validation project are summarized in Sect. 4.

4 Lessons Learned

The STM Validation project is the largest e-Navigation simulator testbed, providing an excellent case to study the lessons learned [12]. The STM functions provide relevant information about the surrounding traffic situations and other vessels intentions. It is important to note that these functions do not replace the human action of decision-making, navigation execution and implementation. They are exclusively intended to be used for decision support and strategic navigation, distance and time permitting [13, 22]. However, even with the addition of the low-LOA, it was clear that the STM functions were shaped by the individuals and system in which they are being used, causing small changes in the subsystems. A systems approach is necessary to evaluate how to ensure that the potential for these technologies to become unruly and lead to poor navigation practices is prevented.

4.1 The Technical Subsystem: Human-Automation Interaction and Usability

The additional information and content (i.e. projected route of another vessel) provided by the STM functions was appreciated by the users. The questionnaires and post-simulation debriefings reflected this positive view in terms of function content [12, 20]. However, in addition to the content, the format, interface and usability of technology have been reported to have a powerful impact on trust and use of technologies even if this does not reflect the true capabilities of the system [23, 24]. The maritime industry has traditionally adopted a reductionist approach, implementing individual problem-patching, technology-driven solutions. Human-centered design (HCD) has been advocated for years in the maritime domain, and was specifically identified in the e-Navigation strategy as a priority in the development and implementation of new technology [11, 25]. The results from the STM Validation project indicate that the functions were not adequately user tested prior to implementation. While the content was highly appreciated, the usability issues relating to; the overcrowded navigation screens, the number of "clicks" required to use a function, and the lack of notification of information were frequently cited issues. These usability problems could be easily addressed through a feedback loop from users to manufacturers to produce a better, safer product. Without this feedback and understanding of how the technology is being used, the potential for the propagation of an unruly technology increases.

4.2 The Personnel Subsystem: Training, Skills and Education

Today there are feelings of uncertainty and insecurity for workers in the maritime industry and it is hypothesized that the role of the seafarer (and other actors within the system) will evolve and change. The IMarEST Maritime Autonomous Surface Ship Special Interest Group (MASS SIG) has an ongoing investigation attempting to understand the role, skills, and responsibilities of the "future seafarer" while also identifying the major gaps to better prepare the industry for the future evolution of the industry [26]. The identification of the skills, training and education required to prepare the future generation of seafarers will create a safer match between the operator and the automation/technology.

All the participants involved in the STM testing were provided with the same technology and opportunities for familiarization and were recruited based on similar vocational training and experience. Although the EMSN data collection provided important insight into how the STM functions were used, it was a controlled simulator experiment which has limitations. These limitations primarily include the generalizability of the results and the behavior of the participants given the simulated environment. In reality, the system will be made up of complex combinations of humans, technologies, harsh operating environments, rules and legislation. As new functions are added to the existing tasks and work systems, it is imperative that training and familiarization is considered prior to implementation and use of these functions in real traffic situations. In parallel, there should be training related to the limitations of the functions and how the data are captured and presented. If people understand how automation works, their level of trust and reliance can better match the capabilities of the technologies [24]. This would promote a safer balance of the authority, ability, control and responsibility between humans and automation, and reduce the chances of the technologies embedded in the automation becoming unruly [27].

4.3 The Work Environment Subsystem: Regulatory Considerations

The existing IMO regulatory framework in the maritime industry does not adequately support the inclusion of higher LOA, the reduced human involvement and the complex combinations of these situations. Today regulations are prescriptive in nature, specifying the technical measures needed to achieve compliance. However, in the 1990s the IMO recognized the need for an alternative to prescriptive regulations caused by the rapid development of computing technologies and automation [28]. The alternative is performance or goal-based standards (GBS) which identify the goal that needs to be achieved while leaving the means to achieve the goal open-ended. The GBS approach promotes flexible, novel solutions while still complying within a safely defined limit. The IMO has already implemented GBS for several regulations, including the Polar Code, and ship construction for bulk carriers and oil tankers [28]. While this is a positive step for the maritime industry, there is still much progress to be made. GBS require a drastically different approach to current maritime operations, incorporating a different way of thinking about compliance, liability and navigational safety.

The results from the numerical evaluation based on the ships' physical positions found that the Conventions on the International Regulations for Preventing Collisions at Sea (COLREGs) were breached more frequently when the STM functions were used compared to traditional navigation [19, 21]. Although the COLREG breaches never led an accident or reportable incident this finding requires further investigation. The STM functions are classified as low-LOA, providing users with additional information about meeting points with surrounding traffic. Yet, this additional information managed to complicate the adherence to the formal rule system (COLREGs). Carey (2017) outlined the legal and regulatory barriers to autonomous ships and identified some of the major unresolved issues [29]. One of the barriers the identified was the ability for companies operating autonomous ships to comply with COLREGs as they are written today [29, 30]. Although the barriers described by Carey are related to autonomous ship(s) operations, the results from the STM project indicate that these barriers are possibly relevant for lower LOA. These technologies will likely migrate up the LOA hierarchy, slightly altering the available information for operators which might have an impact upon COLREG compliance. Without proper assessment of the potential impacts of technology upon the entire system, these challenges have the opportunity to contribute to the barriers and the unruliness of technology.

5 Conclusions

The integration of automated functions will transform maritime navigation. This transformation will cause changes in the technical, personnel, and work environment subsystems. Understanding how new technologies will impact the subsystems should help prevent the technologies from becoming unruly and negatively disrupting the maritime sociotechnical system. A systems approach is necessary to better enable industry stakeholders to prepare for a safer and more sustainable shipping industry.

References

1. Hetherington, C., Flin, R., Mearns, K.: Safety in shipping: the human element. J. Saf. Res. **37**, 401–411 (2006)
2. https://www.ics-shipping.org/shipping-facts/shipping-and-world-trade/global-supply-and-demand-for-seafarers
3. Sanquist, T.F.: Human factors in maritime applications: a new opportunity for multi-modal transportation research. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 1123–1127. SAGE Publications Sage, Los Angeles (1992)
4. Han, D.F., Ding, S.: Review of human factors in maritime system. In: Zhao, H., Liu, K., Yu, X. (eds.) Advanced Design and Manufacturing Technology Iii, Pts 1–4, vol. 397–400, pp. 679–682 (2013)
5. Lützhöft, M.H., Dekker, S.W.A.: On your watch: automation on the bridge. J. Navig. **55**, 83–96 (2002)
6. Lützhöft, M.: The technology is great when it works: maritime technology and human integration on the ship's bridge. Linköping Studies in Science and Technology. Dissertations, vol. 907, p. 108. Linköping University Electronic Press, Linköping (2004)
7. Conceição, V., Carmo, M.B., Dahlman, J., Ferreira, A.: Visualization in maritime navigation: a critical review. In: International Conference on Applied Human Factors and Ergonomics, pp. 199–212. Springer, Heidelberg (2017)
8. Man, Y., Lundh, M., MacKinnon, S.N.: Managing unruly technologies in the engine control room: from problem patching to an architectural thinking and standardization. WMU J. Marit. Aff. **17**, 497–519 (2018)
9. Levitt, M.: Assisted reproduction: managing an unruly technology. Health Care Anal. **12**, 41–49 (2004)
10. Hollnagel, E.: Why is work-as-imagined different from work-as-done? In: Resilient Health Care, vol. 2, pp. 279–294. CRC Press, Boca Raton (2017)
11. International Maritime Organization: The e-navigation Strategy Implementation Plan (SIP) (2014)
12. STM: Sea Traffic Management Validation Project: Final Report (2019)
13. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. IEEE Trans. Syst. Man Cybern.-Part A Syst. Hum. **30**, 286–297 (2000)
14. Kaber, D.B.: Issues in human–automation interaction modeling: presumptive aspects of frameworks of types and levels of automation. J. Cogn. Eng. Decis. Making **12**, 7–24 (2018)
15. Vagia, M., Transeth, A.A., Fjerdingen, S.A.: A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? Appl. Ergon. **53**, 190–202 (2016)
16. Endsley, M.R., Kiris, E.O.: The out-of-the-loop performance problem and level of control in automation. Hum. Factors **37**, 381–394 (1995)
17. Imo: Strategic plan for the Organization For THE six-year period 2018 to 2023. In: International Maritime Organisation, I. (ed.) Resolution A., vol. 1110, no. 30, p. 22 (2017)
18. http://www.imo.org/en/MediaCentre/PressBriefings/Pages/08-MSC-99-MASS-scoping.aspx
19. Weber, R., Man, Y., MacKinnon, S., Saunders, J., Mills, T.: Test report: ship to ship route exchange (S2SREX) - controlled simulation trials (2019)
20. Aylward, K., Weber, R., Lundh, M., MacKinnon, S.N.: The implementation of e-navigation services: are we ready? In: Architects, T.R.I.O.N. (ed.) International Conference on Human Factors, London, UK (2018)

21. Aylward, K.: Automated functions: their potential for impact upon maritime sociotechnial systems mechanics and maritime sciences volume, Licentiate Thesis Chalmers University of Technology (2020)
22. Sheridan, T.B., Parasuraman, R.: Human-automation interaction. Rev. Hum. Factors Ergon. **1**, 89–129 (2005)
23. Corritore, C.L., Kracher, B., Wiedenbeck, S.: On-line trust: concepts, evolving themes, a model. Int. J. Hum Comput Stud. **58**, 737–758 (2003)
24. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**, 50–80 (2004)
25. Mallam, S.C., Lundh, M., MacKinnon, S.N.: Integrating Participatory Practices in Ship Design and Construction. Ergon. Des. **25**, 4–11 (2017)
26. Meadow, G., Ridgwell, D., Kelly, D.: Autonomous Shipping: Putting the human back in the headlines Institute of Marine Engineering, Science and Technology (IMarEST) (2018)
27. Flemisch, F., Heesen, M., Hesse, T., Kelsch, J., Schieben, A., Beller, J.: Towards a dynamic balance between humans and automation: authority, ability, responsibility and control in shared and cooperative control situations. Cogn. Technol. Work **14**, 3–18 (2012)
28. http://www.imo.org/en/OurWork/Safety/SafetyTopics/Pages/Goal-BasedStandards.aspx
29. Carey, L.: All hands off deck? The legal barriers to autonomous ships (2017)
30. MacKinnon, S.N., Lundh, M.: Gaps in Regulations, Pedagogic Needs and Human/Automation Interactions in the Shipping Industry Lighthouse (2019)

Artificial Intelligence in Maritime Navigation: A Human Factors Perspective

Scott N. MacKinnon(✉), Reto Weber, Fredrik Olindersson, and Monica Lundh

Division of Maritime Studies, Chalmers University, SE-412 96, Göteborg, Sweden
{scottm, reto.weber, fredrik.olindersson, monica.lundh}@chalmers.se

Abstract. Artificial intelligence (AI) may be the panacea for improving safety and efficiency in shipping. Solutions to navigation problems are often challenged by information uncertainty, complexity and time demands. Tactical decisions to understand traffic patterns and future vessel encounters can be compared to a game of chess where an agent has goals and considers the next several moves in advance. AI approaches to machine learning is a reactive tactic but remains relatively "weak" and relies on computational power and smart algorithms to recreate each decision every time. Ships are required to follow the International Regulations for Preventing Collisions at Sea. While assumed to be the defining rules of the road, these may be "violated" to solve traffic situations in practice without creating increased risk to the situation. In order to create safe and reliable technologies to support autonomous shipping, the system cannot just rely on where it has to go but anticipate the goals of the surrounding vessels. This paper will explore the challenges, knowledge and technology gaps regarding AI in the shipping sector.

Keywords: Ship navigation · Decision-making · Artificial Intelligence

1 AI: A 1960's Vision

2001: A Space Odyssey is a 1968 film (that details a sentient computer named HAL (Heuristically Programmed ALgorithmic Computer) that possessed the capacities of speech and facial recognition, natural language processing, automated reasoning and could interpret emotional behaviours. Not only could HAL pilot the spacecraft upon which it was situated but could play competitive chess with its human shipmates. While HAL personifies an interesting prophecy of future society, such artificial intelligence (AI) is still in its infancy stages and the elusive singularity [1] remains more fiction than fact.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 429–435, 2020.
https://doi.org/10.1007/978-3-030-50943-9_54

2 Decision-Making, Artificial Intelligence and Automation in the Shipping Industry

Like most transportation sectors, shipping is facing a growing exposure to and reliance upon AI and automation to achieve operational goals such as safety and efficiency [2]. Various organisations have defined levels of automation (LOA) [3 (as example)] and researchers have considered the impact of these technologies [4, 5] upon the future of shipping operations and logistics. The challenges described within this sociotechnical evolution [6] in shipping include the potential conflicts between AI, automation technologies, operator capacities and system safety.

The current state of the art of applied AI can be classified broadly as "weak"; implementation which is focused on limited tasks and does not approach "super-human intelligence". Weak AI is reactive and is far removed from a more generalizable theory of mind [7], which is necessary to fulfill the goals and needs of autonomous, unmanned vessels.

There is considerable debate about the heuristics employed when a person makes decisions "in the wild" [8, 9]. It becomes even more difficult to understand how situation awareness, intuition and problem-solving evolve in complex sociotechnical systems containing large number of agents, both human and artefact [10, 11]. Human intelligence requires the capacities to gather and process information (learn), manipulate information (reason) and consider the outcomes of these manipulated data (understand) [12]. When navigating a vessel in a safe and efficient manner a navigator will create a voyage plan (create goals). During the actual voyage, there is a continuous gathering of information (e.g. surrounding vessel traffic, hydrographical and meteorological information, shore side communications) which is needed to validate the navigation goals or modify the plan as required. For example, one might think that Game Theory models how navigators undertake navigation. Broadly described, in Game Theory multiple agents choose from a set of rational choices creating a system where agent choices will effect the choices of others [13]. However, this artificial intelligence AI paradigm is difficult to apply when considering naturalist decision-making processes and largely ignores the challenges of spatial and temporal aspects typical of navigation in complex traffic situations. While criteria for success should be safety or efficiency (or even both) in navigation tasks, these should not be evaluated in the terms of "winning or losing" [14].

3 Rules Governing Collision Avoidance in Navigation

The International Maritime Organizations' (IMO) International Regulations for Preventing Collisions at Sea 1972 (COLREGs) consists of 19 operational "rules of the road" that must be followed by ships and other vessels at sea to improve safety and prevent collisions [15].

The COLREGs would seem to be the obvious departure point for developing AI and automation solutions. However certain decision-making characteristics which have evolved in navigation practices may create difficulties for those on the technical side of

AI to use appropriate machine/deep learning approaches to develop situation-appropriate AI. While there are “rules of the road” in navigation, in practice we commonly discuss “social navigation” and how “breaking these rules” is not always a bad thing. One main conflicting issue is that these rules were developed for ship-on-ship encounters (i.e. two vessels). This is seldom the case in complex traffic situations and navigation procedures may become more about inter-party negotiations (both verbal and non-verbal) regarding the “work to be done”. While the wording of the COLREGs is sufficiently precise, actions depend on a navigator’s ability to use common sense (“good seamanship”) to not only determine if a situation currently applies, but also to exploit flexibility in the actions prescribed in a rule [16].

The following scenario highlights the lack of specificity in collison avoidance, although all possible outcomes fall within the regulatory conventions. In good visibility when two power-driven vessels are crossing thus creating a risk of collision, the (“give way”) vessel A which has the other on her own starboard (right) side shall keep out of the way and shall, if the circumstances of the case admit, avoid crossing ahead of the other (“stand on”) vessel B (COLREG rule 15). Every vessel which is directed to keep out of the way of another vessel shall, so far as possible, take early and substantial action to keep well clear (COLREG rule 8 and 16) and where one of two vessels is to keep out of the way the other shall keep her course and speed (COLREG rule 17). In the traffic scenario depicted in Fig. 1, Ship A will, in most situations, change her course to starboard (indicated by the dotted line) and pass astern of Ship B which will keep her course and speed expecting that Ship A will follow the COLREGs and change her course in ample time and pass astern of her.

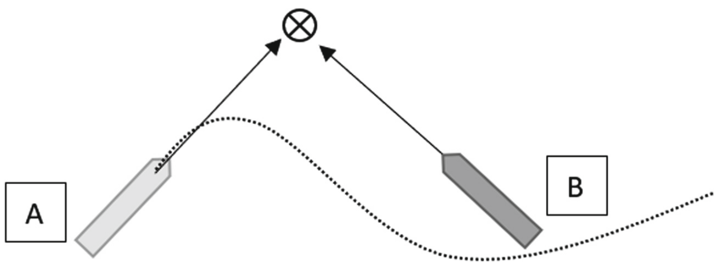


Fig. 1. Crossing situation between two ships.

According to IMO Resolution A.893 [17] all ships are to establish a passage plan before commencing a voyage. A voyage plan (or passage plan) is a comprehensive, berth to berth guide, developed and used by a vessel’s bridge team to determine the most favorable route, to identify potential problems or hazards along the route and to ensure the vessel’s safe passage. After being checked for navigational hazards within operator set parameters, the route is uploaded to an onboard navigation system and monitored on the ship’s Electronic Chart Display and Information System (ECDIS). Any significant deviation from the monitored route may take the ship into waters which have not been checked and deemed safe to navigate. Inserting a possible monitored route (dashed line) for Ship A in a crossing situation is depicted in Fig. 2.

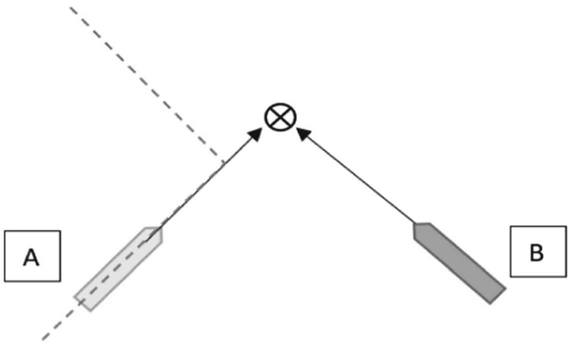


Fig. 2. Crossing situation between two ships with one route displayed.

In the decision process on how to avoid a close quarters situation with Ship B, the officer of the watch (OOW) on Ship A may choose the following actions (see Fig. 3):

a) “disregarding” the monitored route and strictly follow the COLREGs as depicted in Fig. 1 and later turns to join the route.
b) Taking early action and turn to port, i.e. making a “shortcut” in the monitored route.
c) Following the monitored route but calling Ship B on VHF to inform her about the planned action.
d) Slowing down her speed and continue on the monitored route (i.e. letting Ship B pass ahead).

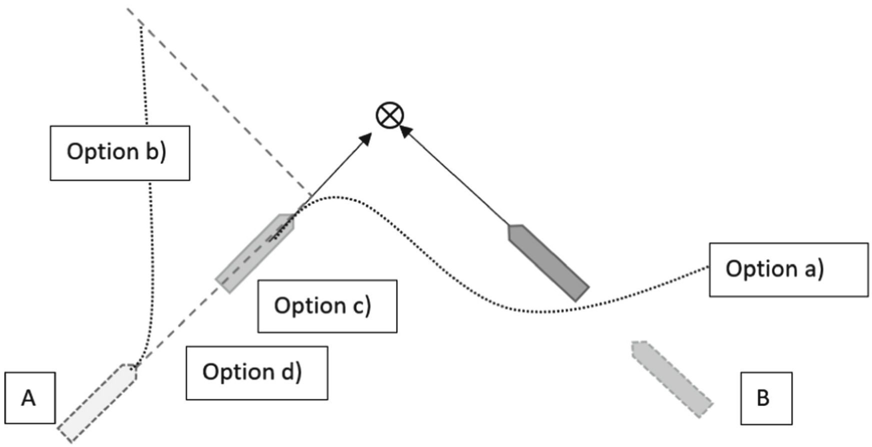


Fig. 3. Possible actions taken by “give way” Ship A

This decision tree is predicated on the assumption that Ship B will, as a “stand on” vessel, keep her course and speed and does not have any information on the monitored route of ship A). None of the possible actions described above explicitly contradict the

COLREGs and deciding on which action to take, the OOW will consider (among other parameters):

I. If a deviation from the route is safe with regard to the proximity of any navigational hazards.
II. The traffic density and the possible consequences of the chosen action in future situations.
III. The maneuverability and speed of the own vessel.
IV. The state of visibility.
V. The state of wind, sea and current and their effects on the ship.
VI. The draught in relation to the available depth of water.
VII. Characteristics of ship encountered (type of ship, speed, size, etc.).

In the situations described in Figs. 1, 2 and 3 several navigational solutions are plausible. A problem with the COLREGs is that these are not value specific so individual interpretations will be also influence by the context of the situation or even local practices. This conundrum will become even more problematic as traffic increases, weather degrades, conflicting competing goals and especially if navigators solve these problems using different heuristic approaches.

4 Can Automation Solve COLREG Ambiguity?

Little research has examined the impact of automation on navigation practices. The Sea Traffic Management (STM) Validation Project (https://www.stmvalidation.eu/documents/) reports how low level automation (LLA) impacted the adherence to COLREGs in taking ship handling decisions. These LLA functions provides the own ship navigator with a route segment consisting of the next 7 waypoints of the monitored route of another vessel. Route segments are broadcasted through Automatic Identification System (AIS) and give additional information to the presently available data acquired by current onboard technologies (e.g. radar). It should be noted that nothing in the ship-to-ship route exchange information exonerates the navigator from applying the COLREGs. Three crossing traffic scenarios involving 3 vessels (72 total trials) were used to compare navigation decisions with and without the route exchange function. In the baseline condition (i.e. no route displayed) there were 2 breaches of COLREGs while in those trials in which participants had access to the function, 8 breaches of COLREGs occurred [18]. Although COLREGs were violated, none of the situations lead to a close encounter between the ships.

While several other proxy measures of relative ship-to-ship safety were collected, perhaps more questions arise from these findings than were answered: Why were more COLREGs breached when more information was available to the navigator? Does automation (i.e. decision-support systems) change navigator (or even system) behaviour or the manner in which navigation conflicts are solved?

5 Conclusions

Actors within a complex navigation environment are required to find common ground [19], share goals and avoid hazardous situations. Automation, implemented as technologies supporting decision-support/making, have the potential for positive changes in situation awareness, workload, risk of collision and overall system safety [20]. Unfortunately, these technologies are deployed into operations often with little consideration for operator training, suitable AI robustness and technology/system readiness.

In order to prepare for increasing levels of automation within the shipping industry certain development matters have to be addressed:

1. Machine learning: Deep machine learning requires significant data. Unlike the automotive industry, in situ data is difficult to obtain. Perhaps simulation technologies can address this gap, however, naturalistic behaviors may be lost in this approach.
2. Technology standardization and regulation: AI and automation will arise from various technology providers. Being based in algorithm development, there is likely going to be a need for oversight in both standardization and regulation of these technologies and their use in navigation practice.

Notwithstanding these technologies questions, large philosophical and ethical issues need to be addressed. Can machines obtain a level of human intelligence that is sufficient for fully unmanned, autonomous systems? And, how do we advance to this reality as we journey through levels of automation that, while linearly defined, are more iterative and concentric in its evolution?

References

1. Vinge, V.: The coming technological singularity: how to survive in the post-human era. In: Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace, NASA Publication CP-10129, pp. 11–22 (1993)
2. IMO: The e-navigation strategy implementation plan. In: MSC 94 (2014)
3. Lloyd's Register: LR Code for Unmanned Marine Systems (2017)
4. WMU: Transport 2040: Automation, Technology, Employment - The Future of Work. World Maritime University, Malmö, Sweden (2019)
5. MacKinnon, S.N., Lundh, M.: Gaps in regulations, pedagogic needs and human/automation interactions in the shipping industry (2019). https://www.lighthouse.nu/sites/www.lighthouse.nu/files/autonomy_webb_slutversion_0.pdf
6. Schwab, K.: The Fourth Industrial Revolution. Crown Publishing Group, New York (2017)
7. Marr, B.: Artificial Intelligence in Practice (2019). ISBN: 978-1-119-54898-0
8. Hutchins, E.: Cognition in the Wild. The MIT Press, Cambridge (1995)
9. Kahneman, D., Klein, G.: Conditions for intuitive expertise a failure to disagree. Am. Psychol. **64**, 515–526 (2009)
10. Salmon, P., Stanton, N., Jenkins, D.: Distributed Situation Awareness: Theory: Measurement and Application to Teamwork. CRC Press, Boca Raton (2009)

11. Endsley, M.: Situation awareness misconceptions and misunderstandings. J. Cogn. Eng. Decis. Making **9**(1), 4–32 (2015)
12. Mueller, J., Massaron, L.: Artificial Intelligence for Dummies. John Wiley & Sons, Hoboken (2018)
13. Jones, A.J.: Game Theory: Mathematical Models of Conflict. Woodhead Publishing, Cambridge (2000)
14. Chauvin, C., Lardjane, S.: Decision making and strategies in an interaction situation: collision avoidance at sea. Transp. Res. Part F Traffic Psychol. Behav. **11**(4), 259–269 (2008)
15. IMO: Convention of the International Regulation for Preventing Collisions at Sea, 1972 (2003). ISBN-10:92-801-4167-8
16. Benjamin, M., Curcio, J.A.: COLREGS-based navigation of autonomous marine vehicles. In: Proceedings of the IEEE/OES Autonomous Underwater Vehicles (2004)
17. IMO: Resolution A.893(21): Guidelines for voyage planning (2000)
18. Aylward, K.: Automated Functions: Their Potential for Impact Upon Maritime Sociotechnical Systems. Licentiate Dissertation, Chalmers University of Technology (2020)
19. Clark, H., Brennan, S.: Perspectives on socially shared cognition. American Psychological Association (1991). ISBN 1-55798-376-3
20. Zhou, X., Huang, J., Wang, F., Wu, Z., Liu, Z.: A study of the application barriers to the use of autonomous ships posed by the good seamanship requirements of COLREGs. J. Navig. **73**(3), 1–16 (2019)

The Anchoring Effect of Technology in Navigation Teams

Vítor Conceição[1,2(✉)], Carlos Teles[1], and Joakim Dahlman[2,3]

[1] CINAV – Portuguese Navy Research Center, 2810-001 Almada, Portugal
placido@chalmers.se
[2] Chalmers University of Technology, SE - 412 96 Göteborg, Sweden
[3] VTI, 402 78 Göteborg, Sweden

Abstract. Modern technology revolutionised marine navigation, reducing errors and increasing navigation safety. However, the same technology has been associated with critical accidents and navigators' errors. On the other hand, expert mariners have proved to manage complex situations, adapting to unforeseen events successfully. To better understand the effects of new technologies and how work is currently done, the Portuguese navy promoted a study about navigation team performance. The results suggest that navigation technology appears to have a strong anchoring effect on team activity. While sensemaking and intuitive judgements complement the shortfalls of the decision support system (DSS), it was found that the combination of high automation influence with lack of coordination leads to a collaborative biased perception of the situation.

Keywords: Team performance · Sensemaking · Navigation · Decision-Making · Intuition

1 Introduction

Digitalization increased the capability to deal with a large amount of information, facilitating system integration and ultimately assisting the navigation tasks. However, it has also brought unforeseen critical consequences, contributing to unsafe practices, or even to accidents or incidents. Those failures have been correlated with over-reliance on automation, over-confidence in the presented data, lack of understanding of inherent weaknesses of automated control systems, ergonomic design considerations, human-computer interface, situation awareness, and information overload. Official reports of three marine accidents occurred in highly technological navies reveal problems regarding the management and use of the available resources, human or technological. On the other hand, operators are also a source of success, due to their unique capability to be adaptive, to learn, to collaborate, to be responsible, to be creative, and to intuitively decide, even under stressful situations. Concerned with navigation teams' safety performance, the Portuguese navy endorsed a study to understand the effects of new technologies and how work is currently done. Data collection was conducted onboard Portuguese naval vessels, sailing in high-density areas of navigation and proximity to hazards. Ethnographic observations of the navigation team activity were complemented

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 436–443, 2020.
https://doi.org/10.1007/978-3-030-50943-9_55

with Eye Tracking data. The results suggest that digital systems have a strong effect on team activity. We also found successful cases of intuitive decision-making and collaborative biased perception associated with misleading pattern recognition.

1.1 The Navigation Plan in Support of Intuition and Teamwork

Experts commonly make good intuitive decisions, grounded in heuristics acquired through experience [1, 2]. They also reveal greater anticipatory thinking skills [3]. However, pondering the strengths and weaknesses of both deliberate and intuitive decision-making, effective use of intuition must be based "*on prelearned insights into specific situations stored in so-called expert schemas. These expert schemas allow pattern matching; that is, they allow the decision-maker to sub-consciously relate perceived cues and decision situations to patterns found in comparable, previously experienced situations and thus produce situation evaluations resulting in high-quality decisions*." [4] (p. 38). Domain expertise is an important factor in the effectiveness of intuitive decision making [5]. Maritime navigation comprises three main functions [6]: forming a goal, defining strategies and moving. In the maritime domain, common ground is established in two dimensions: control systems (enabling actions over the ship) and shared domain (supporting safe interaction with other stakeholders at sea) [7]. Navigation planning helps to create a common cognitive map that supports distributed control. Individual's spatial navigation ability involves sensory cues (environmental and self-motion), computational mechanisms (spatial computations and executive processes) and spatial representations (online and offline) [8]. The navigation plan (NAVPLAN) adds both rational and intuitive process since it is a comprehensive analysis that supports the elaboration of several mechanisms that are used by the navigation team and for DSS parametrization. The design of the NAVPLAN integrates explicit knowledge built on navigator's best practices. Thus, it enables navigators in controlling specific situations, regardless of variability and uncertainty [9]. Anticipatory thinking ability is a kind of sensemaking that helps facing problems that are not clearly presented. All three types of anticipatory thinking [3] are frequently applied: pattern matching to react to individual cues, trajectory tracking to react to trends, and conditional anticipatory thinking to react to the implications of combinations of events. A plan suits several purposes: solving problems, teams coordination, predicting events and assisting improvisation [2, 10]. Team performance and situation awareness are also enhanced by the plan, as it fosters shared cognition and smooth management of priorities and goals. However, that can only be achieved if team members interact dynamically, interdependently and adaptively [11]. Common ground, relevant knowledge, beliefs and assumptions, sustain the interpredictability of team members doing a joint activity facilitating coordination and preventing its breakdown [12]. Team leadership, mutual performance monitoring, backup behaviour, adaptability, and team orientation were seen as key factors in teamwork [13]. The effectiveness of team performance depends on both taskwork and teamwork. Salas *et al.* [14] proposed that teamwork has nine critical considerations, comprised by six processes and emergent states (cooperation, conflict, coordination, communication, coaching, and cognition) and three influencing conditions (composition, culture, and context).

1.2 Methodology

Given the complexity and variability of the situations found in ships bridges while navigating in confined waters, direct observation of the phenomenon was considered the best approach. Nevertheless, also to improve the feasibility of the study, quantitative data complemented the ethnographically informed observations. Data collection were performed by two researchers; one of them also an experienced navy navigator. The concepts used to guide the study were the setting, decision making, participant activities, and teamwork. Observations were made onboard Portuguese navy ships while practising Lisbon port. Ships were carrying out their normal navigation planning; that is, they were not engaged purposely to this study. All participants were in their usual working environment, with minimal influence from all equipment that was assembled and worn. Notwithstanding the study is still undergoing, the results and analysis presented in this paper refer to the first four observation periods, in two vessel class.

2 Preliminary Results and Discussion

In both vessels, each team comprised the roles described in Table 1. Tasks executed in each role (Table 2) are quite similar between vessels since there are under the same organizational culture and norms. Captains leave their team to follow the plan and goals set in the briefing. However, they simultaneously oversee the information flow, look at the surrounding pattern and thinks about situations ahead of those taken by the navigator officer (NAVO). NAVO and lookouts use schemas and visual bearings. The NAVO use it to make sense of the surrounding cues, to help in calculations and coordinate recurring situational reports. Lookouts' schemas are to guide landmarks identification and encourage proactivity and reactivity by understanding the needs of the NAVO and chart. If proficient, they would not wait for request and give the bearings as they feel they are important for the team. The chart operator uses the charted plan with the visual bearing taken by lookouts and additional ship data available in displays. A very proficient operator, in good coordination with lookouts, will perform all the required calculations every three minutes. The radar operator monitors radar's information and plans with the help of schemas. The ECDIS operator monitors the NAVPLAN set in the system, along with additional displayed information. Since all calculations are computed by the ECDIS and relying on the GPS positioning, the operator has the shortest report cycle capability (updates every second) with very high accuracy. Radar operators take advantages of easy access to the external situation, looking through the window. They do this to complement and clarify the radar picture, being actively engaged on validating a radar detection, sometimes speaking the nearest lookout or checking a visual detection not being tracked on the radar (conflicts in shared situational awareness). These proactive and reactive attitudes happen with no direction from the navigator and work as a stand-alone sub-set of the navigation team.

Table 1. Navigation teams organization (some are not indicated like Helman or telegraph).

Role	Rank	Task (see Table 2.)											
		1	2	3	4	5	6	7	8	9	10	11	12
CO	Lieutenant commander	x		x	x		x		x				
Navigator	Lieutenant junior grade	x	x	x	x			x	x	x	x	x	x
Officer of the watch	Lieutenant junior grade		x			x	x						
Radar operator	Lieutenant junior grade, Ensign							x	x	x	x	x	x
ECDIS operator	(A) Lieutenant junior grade/Ensign, (B) Petty Officer/Senior seaman[a)]							x	x	x	x	x	x
Chart operator	Chief petty officer		x					x	x	x	x		x
Lookouts	Seaman										x	x	x
Echosounder	Petty officer 3rd class												x

[a)] One class has a junior officer (ship A) while the other has an enlisted rating (ship B)

Table 2. Tasks, supporting artefacts and systems associated with navigation execution

Task	Artefacts	Systems [a)]
1. Team leader	Schemas	gyro, speed log, rudder angle
2. Teamwork coordination	Schemas	Headset comms, procedures
3. Define, adjust priorities, goals	Schemas	
4. Avoid navigation hazards or vessels		Helm, telegraph
5. Ship's safety		
6. Communication with other vessels and shore stations		VHF comm, visual and sound signals, AIS [c)] text message
7. Determine current, future and desired position	Speed/Time/Dist.Calculator, Hand notes, schemas, nautical chart	Gyro, radar, ECDIS, GNSS, INS[b)], speed log, AtoN[d)]
8. Determine corrections to course, orientation, and speed to attain the desired position		
9. Determine available safe space/time to manoeuvre		
10. Detect, Identify, classify and track navigational dangers, other than vessels	Binocular	Gyro, radar, AIS, AtoN
11. Detect, Identify, classify and track other vessels		
12. Take some measurements	Schemas	Gyro, echo sounder. Speed log, rudder angle, ECDIS

[a)]In different ways, time information is required for all task
[b)]INS – Inertial Navigation System
[c)]AIS – Automatic Identification System
[d)]AtoN – visual and radar Aids to Navigation (external aids)

Notwithstanding the availability of continuous flow of automated information, radar operators still keep looking to the outside. Periodically, they briefly report the surface picture to the navigator. However, the connection that emerges between those two seems also to induce lower engagement of the lookouts. As the radar operator increases his performance level, the NAVO only requests the support of lookouts when what he is looking for is out of his view, or it is prompted by a rule implemented in a procedure. Thus as the navigator and CO become comfortable with the picture provided by those means, lookouts seem to feel unuseful, only acting by request. A similar effect also occurs in chart work, as they lag behind the performance of ECDIS and radar. When they start to feel outperformed, they adopt confirmatory attitudes or even stop trying to report, showing signs of frustration and lack of confidence. In another experimental study in the Portuguese navy, about the effect of GNSS spoofing, a chart operator successively adjusted the visual fix to the erroneous GPS position, and only after a while, he considered the possibility that the GPS was not working properly. Keeping the chart operator with the same role as he used to have before the introduction of ECDIS seems to lead to this type of phenomenon. If it is important to secure resilient processes, we must adjust his role in accordance with the new setting. The changes made were to put him doing the same situational report as the ECDIS and radar in last as he needs more time for the chart work.

The procedure to coordinate situational reports follows a sequence based on the required time to analyse, compute and report in conjunction with the member hierarchy in the team. The situational report cycle starts with the NAVO's report and is followed by the radar, ECDIS, echosounder's operator and paper chart. The ECDIS operator basically only have to read the information computed by ECDIS, which relies on other navigational sensors (speed log, gyroscope), systems (GPS, AIS, radar) and database (navigational chart). The radar operator needs to perform some calculations and measurements, which are based on the radar navigation plan and distances to coastal features. The chart work is performed with the support of position fixes based on visual bearings to landmarks taken by the lookouts, eventually combining distances given by the radar operator and depth readings from the echosounder. Between fixes, he determines estimated positioning taking into account the vessel speed, heading and drift. The NAVO closes the reporting cycle with a concluding report, reviewing his initial perception of the ships' position, and if required, setting adjustments to the navigation plan or goals.

Observations indicate that the ECDIS and radar play a large influence in the navigator role since he frequently asks for information prior to his initial report. Increasing reliability of the teamwork depends on the NAVO's ability to combine his spatial mental model, built on his experience, knowledge and recognition of external cues with the information given by the rest of the team. Critical thinking and integrity validation may be undermined by the large influence of trusted automated sources. On the other hand, from the field of decision-making, we should understand the effect of the influence of high-low competence source, and the adoption of confirmatory of disconfirmation strategies. The effect on lookouts and chart operators suggest that they do not recognize their value in the team when competing with the perceived competence of ECDIS, radar and AIS. Confidence and performance assessment effectiveness is greater in deliberative thinkers, as they are conscious of both deliberative solution and the intuitive alternative.

Once, while approaching a crossing river ferry route, the radar operator reports that a ferry is leaving the north pier (port side). The NAVO and captain look out of the window and confirm the detection, all agreeing that it is heading to the south pier (starboard side). According to the rules, the ferry must give way to the vessels closing on the starboard side, which was the case. The captain orders that if the ferry does show evidence of changing course or speed, they should contact him and request port to port passage. Meanwhile, the radar operator also checks the ferry's AIS information, confirming previous assessment. All theses perceptions were coherent with the contextual pattern activity, as it was early morning with ferries traffic between both sides of the river at the busiest time. At some point, the NAVO orders to communicate with the ferry to clarify intentions and manoeuvres. At the same time, the ferry turn to starboard, the NAVO said that there is no need to communicate as it is manoeuvring according to the rules. As the ferry get close to the ship, the team realize that it was not a ferry, but old ferry used for sightseeing cruises tour. Despite preserving safe navigation, the team collaboratively created a biased understanding and perception of the situation. All checks had the tendency to confirm the first perception: a ferry crossing the river in a usual working day morning. However, the ferry was following the sightseeing cruise route, which was to sail close to the north side, and his behaviour was perceived as a normal anticollision manoeuvre. Team members were not actively engaged in search of relevant information that could conflict the prevailed estimate. A distinctive feature of these ferries is their fluorescent colour that could only be identified visually. Additionally, AIS data should have information about their voyage, revealing a different destination. This case highlights the emergence of confirmation, attentional and anchoring biases. The initial perception was mostly validated from confirmatory automated sources. When looking outside, attention was focused on shape and behaviour, disregarding the unique painting. As the team was confident about the situation, not only the lookout did not take any initiative to check as he was not engaged by the NAVO. The lookout stood still all over the event. The team common perception drifted to the previous knowledge of ferries traffic, subsequently confirmed by identified reliable automated sources. As a result, the team awareness of the implications of combinations of events involving the ferry was biased, impairing the anticipatory thinking ability.

In another case, the inbound ship had a tight window to enter the naval base, while another ship was due to depart a few minutes before their ETA. After updating information between both ships, the navigator proposed a speed adjustment to keep the ETA. The Captain recommended a different approach, ordering to reduce speed and wait. His judgment was based on his estimate about the other vessels condition, just coming from a long period of maintenance and starting a training program. He was not confident that they would comply with their ETD. A few minutes later, they received confirmation that there were dealing with technical problems, requesting 15 min more. The captain experience, organizational knowledge and feeling that something might easily go wrong, was crucial to avoid the risk of close proximity during critical manoeuvres. All DSS and formal approach to the navigation plan could not make sense of the situation, and as suggested by Dreyfus and Dreyfus [15], DSS can not replace the role of an intuitive judgement of decision-makers.

3 Conclusions

How to create conditions of conflict to induce processes of disconfirmation and critical thinking within each individual? At one hand, we have the hypothetical influence of perceived competence of the human source. On the other hand, we have the influence of automated source over the operators. This last might also propagate its influence within the team elements. From another perspective, the type of task performed by the operators also shapes the individuals' forms of judgements. Some tasks are more prone to generate rational thinking while others will call for intuitive processes. Moreover, it might not be the case of enabling one or other decision-making process, but to inhibit one of them. Once we modify the working settings with the introduction of automation, we need to understand the emerging effects on individuals and teams. We need to better understand the perceived influence of power and Artificial Intelligence in decision-making processes [16]. Procedures and strategies that used to work in previous arrangements may not fit the requirements after changing part of the working setting. Any change in the teamwork ought to be preceded by a thorough study on the possible effects and required adaptation in the design of the working setting. The way forward is to test the hypothesis of source influence to identify the best sequence for the situational reports. Another approach is addressing the design of collaborative DSS, looking at new forms to enhance teamwork performance.

Acknowledgements. We appreciate the availability of the Portuguese Navy in facilitating access to the fleet's ships and to the commanders and their crew that participated in this study.

References

1. Klein, G.A., Shneiderman, B., Hoffman, R.R., Ford, K.M.: Why expertise matters: a response to the challenges. IEEE Intell. Syst. **32**, 67–73 (2017). https://doi.org/10.1109/MIS.2017.4531230
2. Klein, G.A.: The Power of Intuition: How to Use Your Gut Feelings to Make Better Decisions at Work. Currency/Doubleday, New York (2003)
3. Klein, G., Snowden, D., Pin, C.L.: Anticipatory thinking. In: Mosier, K.L., Fischer, U.M. (eds.) Informed by Knowledge: Expert Performance in Complex Situations, pp. 235–246. Taylor and Francis, Abingdon (2011). https://doi.org/10.4324/9780203847985
4. Steigenberger, N., Lübcke, T., Fiala, H., Riebschläger, A.: Decision Modes in Complex Task Environments. CRC Press, Boca Raton (2017). https://doi.org/10.1201/9781315116303
5. Dane, E., Rockmann, K.W., Pratt, M.G.: When should I trust my gut? linking domain expertise to intuitive decision-making effectiveness. Organ. Behav. Hum. Decis. Process. **119**, 187–194 (2012). https://doi.org/10.1016/j.obhdp.2012.07.009
6. da Conceição, V.P., Carmo, M.B., Dahlman, J., Ferreira, A.: visualization in maritime navigation a critical review. In: Stanton, N.A. (ed.) Advances in Human Aspects of Transportation, Advances in Intelligent Systems and Computing, pp. 199–212. Springer, Los Angeles (2018). https://doi.org/10.1007/978-3-319-60441-1_20
7. da Conceição, V.P., Dahlman, J., Navarro, A.: What is maritime navigation? unfolding the complexity of a sociotechnical system. In: Human Factors and Ergonomics Society Annual Meeting, Austin, Texas, USA, pp. 267–271 (2017). https://doi.org/10.1177/1541931213601549

8. Wolbers, T., Hegarty, M.: What determines our navigational abilities? Trends Cogn. Sci. **14**, 138–146 (2010). https://doi.org/10.1016/j.tics.2010.01.001
9. Klein, D.E., Woods, D.D., Klein, G.A., Perry, S.J.: Can we trust best practices? six cognitive challenges of evidence-based approaches. J. Cogn. Eng. Decis. Mak. **10**, 244–254 (2016). https://doi.org/10.1177/1555343416637520
10. Weick, K.E., Sutcliffe, K.M.: Managing the Unexpected: Resilient Performance in an Age of Uncertainty. Jossey-Bass, San Francisco (2007)
11. Salas, E., Guthrie Jr., J.W., Ilson-Donnelly, K.A., Priest, H.A., Burke, C.S.: Modeling team performance: the basic ingredients and research needs. In: Rouse, W.B., Boff, K.R. (eds.) Organizational Simulation, pp. 185–228. John Wiley & Sons, Inc., Hoboken (2005)
12. Klein, G.A., Feltovich, P.J., Bradshaw, J.M., Woods, D.D.: Common ground and coordination in joint activity. In: Rouse, W.B., Boff, K.R. (eds.) Organizational Simulation, pp. 139–184. John Wiley & Sons, Inc, Hoboken (2005). https://doi.org/10.1002/0471739448.ch6
13. Salas, E., Sims, D.E., Shawn Burke, C.: Is there a “big five” in teamwork? Small Gr. Res. **36**, 555–599 (2005). https://doi.org/10.1177/1046496405277134
14. Salas, E., Shuffler, M.L., Thayer, A.L., Bedwell, W.I., Lazzara, E.H.: Understanding and improving teamwork in organizations: a scientifically based practical guide. Hum. Res. Manag. **54**, 599–622 (2015). https://doi.org/10.1002/hrm.21628
15. Dreyfus, H.L., Dreyfus, S.E.: Mind Over Machine. The Power of Human Intuition and Expertise in the Era of the Computer. The Free Press, New York (1986)
16. Fast, N.J., Schroeder, J.: Power and decision making: new directions for research in the age of artificial intelligence. Curr. Opin. Psychol. **33**, 172–176 (2020). https://doi.org/10.1016/j.copsyc.2019.07.039

Validation of Virtual Command Bridge Training Environment Comparing the VR-Training with Ship Bridge Simulation

Jenny Lauronen[1,2](✉), Werner Ravyse[1], Mirva Salokorpi[2], and Mika Luimula[1]

[1] Turku University of Applied Sciences, Turku, Finland
jennylauronen@novia.fi, {werner.ravyse, mika.luimula}@turkuamk.fi
[2] Novia University of Applied Sciences, Vaasa, Finland
{jennylauronen, mirva.salokorpi}@novia.fi

Abstract. In this validation exploration, we have studied a virtual reality ship command bridge against the standards and regulations that maritime training simulators must adhere to. We created a virtual reality replica of a command bridge with limited functionality that underwent user testing with 16 experienced ship officers. The results show that our training application did not meet all simulator criteria, but we point out that each of the shortcomings can be overcome with new generation hardware and expanded virtual reality programming. Our conclusion is that VR is a valid, affordable and efficient tool for command bridge simulator training.

Keywords: Virtual reality · Maritime education · Ship bridge simulation

1 Introduction

Decision-making on ship bridges is a challenging professional skill. In many seafaring traffic situations, complex permutations affecting safe ship navigation can develop within minutes. Such challenging traffic situations, combined with fairway limitations, require a clear understanding of decision priorities. Simulated reality offers a realistic hands-on training environment for maritime students and professionals where they can practice situational decision-making in a protected environment [1]. Simulation training is also able to impart or improve application methods and motor skills [2] and although simulation training is a well-established educational method in seafaring, in its traditional form of room-scale simulators, it is expensive and significantly constrained for training officers at sea. The International Chamber of Shipping estimates that the maritime industry currently employs 1, 7 million seafarers, of which 0, 6 million are officers requiring simulator training. Excessive simulator costs are the leading cause that globally, many maritime officers are not able to benefit from traditional simulator training.

Virtual reality (VR) head-mounted display (HMD) training environments, on the other hand, present an affordable, measurable and repeatable training alternative in a variety of disciplines [3]. The latest VR headsets no longer require beacons for player orientation within the virtual space, making them highly portable and suitable for

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 444–451, 2020.
https://doi.org/10.1007/978-3-030-50943-9_56

delivering sea officer instruction anywhere. Other pertinent benefits include a high training immersion [4] and the possibility to include the attributes and behaviors of multiple ships on a single VR training scenario. The question is whether VR HMD training is able to meet the learning outcomes as expected from traditional simulators?

This paper presents a partial validation of a ship bridge VR HMD environment against a ship bridge simulator. The validation objective was to determine whether the further development of command bridge VR headset training is warranted. This paper also defines a set of validation criteria and indirectly establishes suitable verification guidelines for VR headset ship navigation scenario development. Our validation environment was designed to acquaint captains with ship bridge systems and train safety-critical decision-making for collision avoidance. We focused our validation efforts on the basic functional fidelity of steering and navigating a ship by measuring both usability and training relevance. This paper does not include technical verification, but rather emphasizes the system and content suitability for maritime training in comparison to known simulator training.[1]

2 Current Simulator Training Practices

The International Maritime Organization (IMO) drives the obligation for continuous training and upkeep of seafaring skills and since practical skills can be learned and understood best through hands on practice, simulators are a focal training method in seafaring. Some of the core seafaring skill standards of competence to be learned by simulation training are prescribed by the Standards of Training, Certification and Watchkeeping for Seafarers (STCW). These standards form the basis of simulator validation and therefore, also the VR environment we set out to test.

The validation in this study is based on the STCW code for operational level requirements, explained in the DNVG-ST-0033 *Standard maritime simulator systems* by De Norske Veritas and the IMO Model Course 6.10 *Train the simulator trainer and assessor*. These references describe the training requirements for the four different simulator classes (A, B, C and S)[2]. Table 1 shows the competences for each of the simulator classes and indicates the competences used for our VR validation. The development of ship maneuverability (our primary case for validation) for VR caused incidental (but partial) development of: (a) planning and conducting a passage and determining position; and (b) maintaining a safe watch. Hence, limited aspects of these competences also cropped up in our validation testing.

Maritime training simulators consist of computer software systems that mimic the dynamics of a real-time environment and a physical simulator bridge comprising real consoles and instrumentation. Simulators are almost exclusively placed indoors with the simulated environment projected on several large monitors, affording trainees with a 120 to 180° view of the scenery. The software realistically simulates physical ship

[1] For the further duration of this paper, VR will refer to VR HMD and the term simulator will reflect traditional (or current) simulators.

[2] Class A (Full mission); Class B (multi-task); Class C (limited task); Class S (special tasks, performance defined case by case).

behavior in response to navigational input from trainees and environmental factors such as weather conditions and water effects. Simulators additionally allow for the inclusion of system failures and other incidents.

Table 1. Simulator competences used for VR simulator validation (adapted from DNVG-ST-033).

STCW reference	Competence	Class A	Class B	Class C	Class S	VR sim
Table A-II/1.1	Plan and conduct a passage and determine position	Yes	Yes		Yes	Limited
Table A-II/1.2	Maintain a safe navigational watch	Yes	Yes		Yes	Limited
Table A-II/1.3	Use of radar and ARPA to maintain safety of navigation	Yes	Yes	Yes	Yes	
Table A-II/1.4	Respond to emergencies	Yes	Yes	Yes	Yes	
Table A-II/1.5	Respond to distress signal at sea	Yes	Yes	Yes	Yes	
Table A-II/1.8	Maneuver the ship	Yes	Yes	Yes	Yes	Yes

The IMO describes simulator training sessions as a four-step process that begins with an instructor briefing where officers are introduced to the scenario surrounding their navigational challenge. After the brief, officers must plan the ship's course before moving to the command bridge simulator where the scenario unfolds on the screens and instrument panels. The instructor, who is in a different room, now manipulates the various inputs that influence the scenario while maintaining continuous communication with the officer. Upon completion of the session, officers once more meet with the instructor for an in-depth debriefing.

The trainer's role in simulation training is formidable because the laws and regulations of seafaring are largely subjective. This implies that the instructor must be able to interpret all trainee decisions and actions within the context of the scenario. The result is that trainers must be skillful enough to convey complete understanding of the scenario outcome to the officer [5]. A second major challenge trainers face is to scaffold the scenario appropriately. That is, in order to eliminate much of the subjectivity, trainers can simplify the scenario by removing variables or reducing their effect on the navigation task. This would prohibit officers from potentially learning incorrect responses but runs the risk of over-simplification that could result in a false sense of know-how. Trainers must be conscious of this scaffolding balance throughout the training session.

3 Experiment Design

The VR ship bridge for validation is a general model of a 170-m containership. The bridge facilities include steering, speed control and typical ship bridge equipment, such as Electronic Chart Display and Information System (ECDIS), radar and control screens. The VR environment (Fig. 1) was created using the Unity game engine and utilizes HTC Vive as the VR hardware.

The VR task was to safely maneuver the containership through a section of the strait of Denmark and avoid grounding or colliding with another vessel. This other vessel (a fishing boat) in the scenario was not visible from the electronic charts because the automatic identification system (AIS) was not connected. Participants could only see the fishing boat by using radar or looking out the window. However, the radar was set for rainy weather, which hides fairway signs and small objects and participants were expected to adjust the radar settings at the start of their session.

Fig. 1. Command bridge in VR scenario.

To understand the VR bridge maturity, user testing was conducted with 16 sea officers. They were master students, teachers or the trainees of Aboa Mare maritime academy with at least four years of officer experience.

Since 13 participants were completely new to VR, the exercise started with a short familiarization whereby all VR functionalities and the controllers were tested—this lasted about 90 s. Soon the users felt comfortable and the exercise could start. Navigating the vessel to clear waters took between 9–13 min, depending the speed used. The users were not advised about the speed range they should maintain, but they knew the normal speed in such traffic situations would be 12–18 knots. Our VR headset scenario hydro-dynamic model has a limited water-to-vessel interaction speed range of 5–20 knots.

The validation results were obtained through observation, questionnaires and interviews. The observation phase resulted in a set of notes from the research team and the session trainer. The researchers and trainers analyzed and discussed the notes and the key observations served as input for the free-form interviews that concluded the experiment. The participants filled system usability scale (SUS) questionnaires immediately after their VR session. This was done to verify the extent that usability concerns may have influenced the validation results. The researchers summarized the interview notes and wrote the conclusion in collaboration with the trainers.

4 Findings and Results

Observation showed that as long as the participants were uncertain about the VR technology, their gaze locked onto the control panels, but as soon they became familiar with the VR equipment, they started to behave as in real life by looking around and maneuvering as required. The VR controllers appeared to be ungainly and participants struggled to use all the command bridge instruments. Levers and large wheels were easy to handle, but dials with small turning angles and sensitive feedback systems posed problems for the controllers, as did closely spaced buttons. Also, participants were forced to look at the virtual equipment in order to align them with the controllers, hampering habitual hand movements that would normally be able to blindly find ship controls. When turning the ship, five of the participants reached for a support in the virtual world to prepare them for the tilt of the vessel. Since this support was not physically present, these participants slightly lost their balance infringing upon their scenario immersion. No-one claimed cyber sickness during or after the exercise.

The interviews were open discussions with the objective of forming an opinion on the validity of using VR as collision avoidance training to be used in off times during an actual sea voyage. Eleven participants claimed that it was a positively interesting experiment and that they would be willing to test it again after the next development stage. The other five participants found it strange or uncomfortable and did not believe it could be useful for training as it is. There was consensus among the participants that this VR training would be most suited for basic education, rather than advanced scenarios. All participants were delighted with the VR solution's high visual fidelity. The problems they pointed out were centered around the controllers and how this would affect user and training experience.

The SUS is a barometer indicating how much work is left before the system in question could be considered usable—a score of 68/100 is viewed as sufficiently usable with some work to be done [6]. Our VR command bridge scored 66/100. The areas where participants were most satisfied included the system consistency (73/100), the low level of complexity (78/100), onboarding (81/100) and learnability (72/100). On the other hand, the system integration was considered weak (52/100) and many participants felt they would need technical assistance at some point (56/100).

5 Discussion: VR Simulation Validation and Verification Criteria

Simulation pedagogy highlights that physical fidelity, behavioral fidelity and the operating environment are influential in VR solutions [7].

We addressed the physical fidelity by creating a virtual space that accurately follows a real ship bridge and further strengthened the onboard experience by including vessel traffic service radio. Our efforts were well-received and evidenced in the direct participant comments indicating appreciation for the physical realism of our VR setting.

Although our system's behavioral fidelity was not faulted in the direct maneuvering of the ship, it is compromised through a limited hydro-dynamic model. This is highlighted in a general sentiment that VR would only be useful in basic training. Such remarks were not unexpected since the VR scenario for this study was only a partial representation of the current simulator capabilities and the simple scenario was not aligned with the participant level of skill and experience. Nevertheless, to rectify the current misgivings in the target use of VR in ship navigation training: (a) the vessel response to various weather conditions must be refined; (b) the speed range where the ship behavior is natural should be increased; (c) fully functional ECDIS and automatic radar positioning aid systems must be present; and (d) autopilot must be available.

The challenges in the operating environment are somewhat more complex to overcome, as they are related to hardware (rather than programming) limitations. The control system restricts training potential through unwieldy controllers that are unable to adjust finer dials or press closely spaced buttons. Moreover, the controllers obstruct the training by disqualifying habitual movement, which can potentially lead to facilitating bad habits during training that will manifest in unsafe ship navigation. An additional operating environment stumbling block that our VR training scenario revealed, was the HMD's low resolution. Low resolution made it difficult for our participants to observe finer details, such as gauge readings and ECDIS communication. During the course of this validation test, a new generation of high-resolution VR headsets that include finger tracking were released. These headsets could potentially address both crucial operating environment limitations our VR scenario exhibits.

Current practices highlight that a primary strength of simulator training lies in the close instructor-learner collaboration to mimic the teamwork essential in successful ship navigation. The fidelity of an embedded teacher avatar, as a scenario character, is a proven learning facilitator in serious games [8], leading us to conjecture that either asymmetric or conventional multi-player VR scenarios could address teamwork aspects.

Debriefing sessions, known to be a crucial component of the learning process [5], are well-integrated into simulator training and usually happen on the basis of trainee actions recorded during a scenario and discussed post-exercise. VR scenarios should therefore ensure that a there exists a sub-system that collects, stores and reproduces the learning analytics required for meaningful debriefing sessions.

Simple scenarios work well in familiarization and part-task learning. Simulator training sessions are hallmarked by dynamic scaffolding to maintain optimal learning opportunities. By gradually increasing the fidelity and realistic complexity of the tasks, the simulation training goes to higher levels of learning [2]. Research and development of dynamic manipulation of variables in VR scenarios is in its infancy, but given the potential of VR in maritime and other fields utilizing dynamic simulation training, this is certainly worth investigating.

VR's propensity for high immersion also reaches beyond the test scenario, as demonstrated by our participant response of reaching for a stabilizing table in the virtual world. This places trainees in situations where the physical environment could be hazardous and should therefore, be addressed when considering VR simulation training.

6 Conclusion

In this paper we set out to validate whether VR would be a suitable alternative for traditional simulator training. We did this by developing a VR experience that focuses on the ship maneuvering competency from the STCW. User testing with 16 experienced sea captains led us to uncover various shortcomings in our VR training environment. These include unnatural controllers, low HMD resolution, limitations in ship behavior, a lack of interaction with the trainer, missing documentation for meaningful debriefing sessions, omitted possibility to dynamically change environment variables and jeopardized physical safety of the trainees. However, we point out that these drawbacks do not invalidate VR as a simulator training opportunity as each of the challenges can be resolved, be it with next-generation headsets or more elaborate software engineering. The listed shortcomings should be viewed as a verification checklist for VR command bridge simulator development.

We have found the VR simulation environment immersive, interactive and useable, making it a valid tool for command bridge simulation training. VR offers, over traditional simulators, an affordable and efficient training tool that allows officers to practice critical decision-making across a full range of variables and situations. User experience and usability testing among highly experienced seafarers are essential for determining an authentic training interface and in collaboration with maritime education specialists and VR software engineers, VR has the potential to provide much sought-after simulator training to a considerably larger ship officer audience than at current.

References

1. Tichon, J., Burgess-Limerick, R.: A review of virtual reality as a medium for safety related training in mining. J. Health Saf. Res. Pract. **3**, 33–40 (2011). http://burgess-limerick.com/download/a54.pdf
2. Salakari, H.: Käytännön taitoja virtuaalisesti. Licentiate Tampere University (2004). https://trepo.tuni.fi/bitstream/handle/10024/76389/lisuri00022.pdf?sequence=1
3. Saunders, J., Bayerl, P., Davey, S., Lohrmann, P.: Validating virtual reality as an effective training medium in the security domain. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces. Sheffied Hallam University (2019)
4. Feng, Z., Gonzales, V., Amor, R., Lovreglio, R., Cabrera-Guerrero, G.: Immersive virtual reality serious games for evacuation training and research. Comput. Educ. **127**, 252–266 (2018)
5. Rall, M., Manser, T., Howard, S.: Key elements of debriefing for simulator training. Eur. J. Anaesthesiol. **17**(2000), 516–517 (2000)

6. Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean. J. Usability Stud. **4**(2), 114–123 (2009). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.177.1240&rep=rep1&type=pdf
7. Lee, A.: Vehicle Simulation, Perceptual Fidelity in the Design of Virtual Environments. CRC Press, Boca Raton (2018)
8. Ravyse, W.S., Blignaut, A.S., Leendertz, V., Woolner, A.: Success factors for serious games to enhance learning: a systematic review. Virtual Reality **21**(1), 31–58(2017). Sylipsis Corp. Business Opportunities in the global maritime industry. N.D. https://www.sylipsis.com/business-opportunities-in-the-global-maritime-industry/

Addressing Gaps in Offshore Emergency Egress Training Using Virtual Environments

Jennifer Smith(✉), Mashrura Musharraf, and Brian Veitch

Faculty of Engineering and Applied Science,
Memorial University of Newfoundland, St. John's, NL, Canada
{jennifersmith, mashrura.musharraf, bveitch}@mun.ca

Abstract. Operators and regulators of offshore and maritime domains should adopt evidence-based safety training to prepare the workforce for emergency egress. This paper uses pedagogical frameworks and data mining tools to identify training gaps in mandatory offshore safety training, and offers evidence-based virtual environment (VE) training solutions. A VE training setting was used as a human behavior laboratory to provide trainees with artificial experience and record their learning progress in the context of evacuating a virtual offshore petroleum platform during a series of credible emergencies. A longitudinal study was conducted to collect data at three critical learning stages: skill acquisition, retention, and transfer. The empirical evidence identified strengths and deficiencies in the VE training. The modeling provided a more comprehensive assessment of the VE training and demonstrated the utility of data-mining tools for future adaptive training applications.

Keywords: Offshore Emergency Egress · Virtual training · Decision trees

1 Introduction

Offshore petroleum platforms' and merchant vessels' operations are characterized by their isolation and harsh marine weather. As a result, emergency evacuations in these environments are often time-sensitive and complex to manage. To help prepare the workforce for emergency egress, operators and regulators should adopt evidence-based safety training to guarantee workers achieve and maintain demonstrable competence.

Due to regulatory, logistical, and financial constraints, conventional offshore egress training often uses fixed-time instruction through video-tutorials and platform orientation walkthroughs. However, the inherent gap in this training is that it lacks the structure and resources to assure competence is acquired and maintained. Specifically, it assumes all trainees achieve competence even though this form of training does not consider individual learning styles or paces. Consequently, the training produces a workforce with unknown or variable competence, which is a safety concern. The expectation is that existing training gaps will be addressed through on-the-job training, such as participating in mandatory muster and evacuation drills offshore. However, regulations do not provide evidence-based recommendations on how to administer training or how often to schedule training maintenance.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 452–460, 2020.
https://doi.org/10.1007/978-3-030-50943-9_57

This paper demonstrates the use of pedagogical frameworks and data mining tools to address these gaps in offshore safety training, and offers evidence-based virtual environment (VE) training solutions in the context of offshore and maritime safety. A VE training setting was used as a human behavior laboratory to measure crew competence and evaluate new training interventions. A longitudinal study was conducted, combining VE technology with a pedagogical framework called simulation-based mastery learning (SBML) [1], to provide trainees with artificial experience in evacuating a virtual offshore petroleum platform during a series of credible emergencies. To provide evidence on how to deliver and maintain egress training, the impact of SBML-guided VE training on trainee competence was investigated at critical learning stages, specifically during the acquisition, retention, and transfer of egress skills. This paper used empirical and modelling approaches to examine the training gaps and evaluate the efficacy of VE training. The empirical approach analyzed performance metrics to determine if a VE trained group of participants achieved and maintained competence. To build on this analysis, a data-mining method called decision tree (DT) modeling was used. DTs offer a pattern recognition lens to assess the VE training efficacy that goes beyond performance metrics. Human performance data collected during the virtual training can also be modelled using DTs to inform the training efficacy at the three learning stages. In this paper, the DTs were used to diagnose the strengths and weaknesses the VE training at a systemic level.

2 Methods

Figure 1 depicts the longitudinal study used to evaluate VE training at three critical learning stages: 1) a skill acquisition phase; 2) a skill retention assessment and retraining phase; and 3) a transfer of training to novel situations phase.

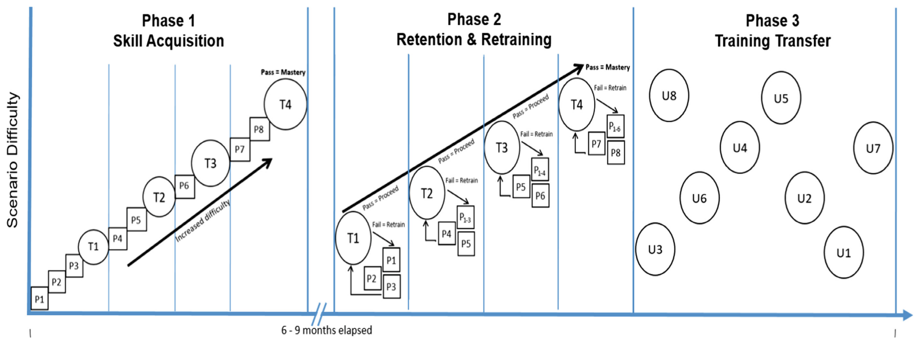


Fig. 1. Timeline for the longitudinal experiment. This schematic shows the learning phases and the corresponding practice and test scenarios used in each phase. The practice, test, and transfer scenarios are denoted as P1–8, T1–4, and U1–8, respectively.

This section describes the participants, the simulator, the approach used to deliver the training and assess the participants' competence at the learning stages, and the data-mining method used to develop the DT models from the participants' performance data.

Participants. Thirty-eight participants completed all three phases of the experiment. Twenty-eight participants were male and ten participants were female. Participants ranged in age from 19 to 54 years (M = 28 years, SD ± 8.7 years). Participants were naïve to the experimental design, had no prior experience working offshore, and had no exposure to the simulator prior to the study.

AVERT Simulator. The emergency preparedness training simulator called AVERT was used for all phases of the experiment. AVERT is a desktop VE designed as a virtual lab to train people in basic offshore emergency duties, such as how to navigate the platform and muster at their designated stations. AVERT allows participants to interact with the virtual offshore platform using a gamepad controller (Xbox). The VE depicts a realistic representation of an offshore Floating Production Storage and Offloading (FPSO) vessel [2]. Participants can move onboard the FPSO by controlling a first-person perspective avatar of an offshore worker.

Learning Phases. The first phase of the experiment (denoted the acquisition phase in Fig. 1) used a SBML pedagogical framework to teach participants the skills to safely evacuate an offshore platform [3]. The SBML training gradually familiarized participants with the platform layout, emergency alarms, egress routes, safety protocols, and mustering procedures. Participants were required to repeat the practice and test scenarios until they demonstrated competence in all performance criteria. The participants' ability to recognize and follow their egress routes was tested in scenarios of escalating difficulty, from benign muster drills (T1–2) to complex emergency evacuations (T3–4).

In the second phase (denoted the retention phase in Fig. 1), participants returned after a period of 6 to 9-months and were tested on their ability to respond to the same egress scenarios (T1–4). Participants who had trouble remembering the egress procedures were provided retraining on deficient skills (e.g. practice scenarios P1–8). Participants were required to complete retraining for any deficiencies before moving on to the third phase.

The third phase (denoted the training transfer phase in Fig. 1) involved testing the participants' ability to apply their skills to new emergencies. The scenarios in the transfer phase were designed to vary the near and far-transfer proximity [4, 5]. Thus, the emergency scenarios in this phase differed from the exercises in the two earlier phases by varying conditions such as the participants' familiarity of the scenario's starting locations on the platform, the placement and severity of hazards on the platform, and the availability of information (or lack thereof) about the situation. The first set of scenarios (U1–4) closely resembled the conditions and knowledge covered in the training scenarios and thus represented near-transfer (e.g. had the same starting location that participants had trained). The second set of scenarios (U5–8) were more different from the conditions of the training scenarios and therefore represented far-transfer (e.g. had a new start location different from where participants had practiced but a location they had seen before). The transfer scenarios gradually increased in difficulty from U1 to U8. Participants performed the transfer scenarios in a random order.

Decision Tree Development. DT modeling is based on supervised learning theory and uses classification to visualize data patterns from data collected from VE training [6]. To develop DTs, a two-dimensional data matrix was created from the programmed attributes of the scenarios and the participants' performance data collected at each phase of the VE training (e.g. their actions in responds to cues from a situation). An individual data matrix was developed for each participant. Table 1 provides an example of a data matrix for a sample participant after finishing the skill acquisition phase. Each row contains the different programmed attribute values for the scenario and the corresponding route choice taken by the participant.

Table 1. Sample participant data matrix developed from data in the skill acquisition phase.

Scenario	Scenario attributes						Participant route choice
	End location	Alarm type	Route by PA	Hazard	Blocked route	Previous route	
P1	Muster	None	1st	No	None	N/A	Primary
P3 (F1)	Lifeboat	None	1st	No	None	1st	Primary
P3 (F2)	Muster	None	2nd	No	None	1st	Secondary
T1	Lifeboat	None	None	No	None	2nd	Secondary
P4	Muster	GPA	1st	No	None	2nd	Primary
P5	Muster	GPA	None	No	None	1st	Primary
T2	Muster	GPA	1st	No	None	1st	Primary
P6	Lifeboat	PAPA	1st	No	2nd	1st	Primary
T3 (F1)	Muster	GPA	None	No	None	1st	Primary
T3 (F2)	Muster	GPA	2nd	No	1st	1st	Secondary
P7	Lifeboat	PAPA	2nd	Yes	1st	2nd	Secondary
P8 (F1)	Muster	GPA	1st	Yes	2nd	2nd	Primary
P8 (F2)	Lifeboat	PAPA	1st	Yes	2nd	2nd	Primary
T4 (F1)	Muster	GPA	None	Yes	None	1st	Primary
T4 (F2)	Lifeboat	GPA	2nd	Yes	1st	1st	Secondary
T4 (F3)	Lifeboat	PAPA	2nd	Yes	1st	1st	Secondary

Musharraf et al.'s [7] decision tree algorithm was applied to the data matrix to identify the participants' egress route selection strategies. As an example, Fig. 2 shows the DT strategy of the sample participant created from the data matrix in Table 1. Full details on the methods used to develop DTs are described in [7] and [8].

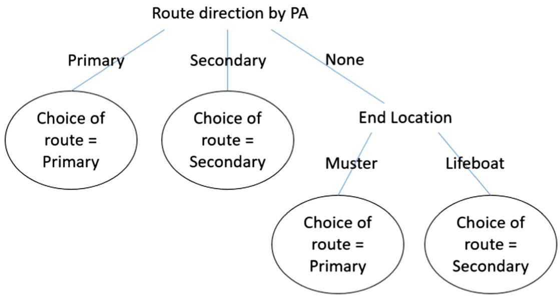


Fig. 2. Sample participant DT developed from the data matrix shown in Table 1.

The collection of DTs provided a depiction of the participants' route selection strategies during an emergency. To assess the efficacy of the VE training, the participants' DTs were compared to the training's intended decision strategies.

3 Results and Discussion

Empirical Evidence. The empirical results offered a conventional means to evaluate the efficacy of VE training. Figure 3 provides a boxplot representation of the performance scores at all three phases.

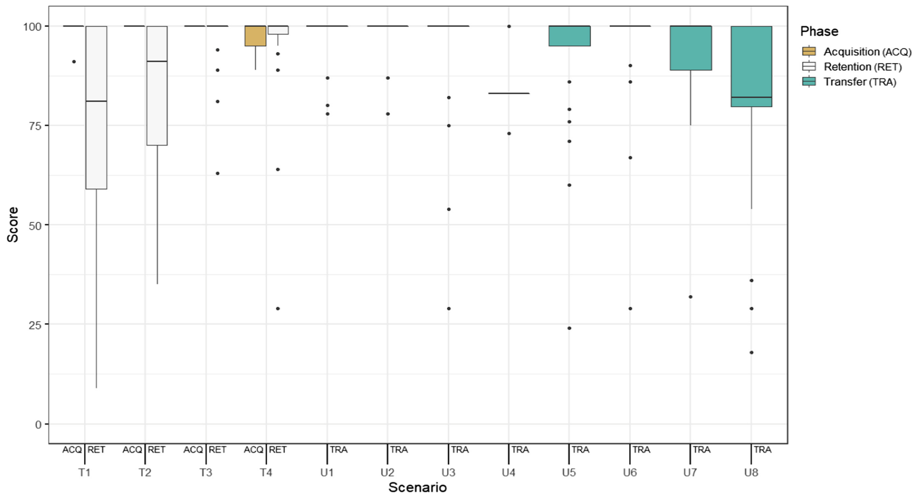


Fig. 3. Boxplot of the performance scores for the skill acquisition (ACQ), Retention (RET), and Transfer (TRA) phases.

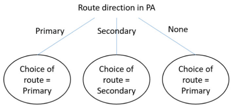
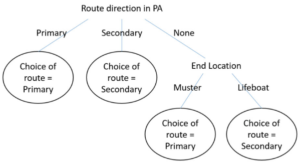
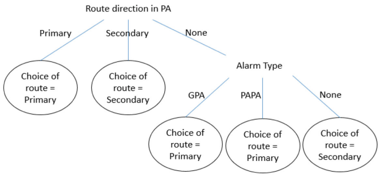
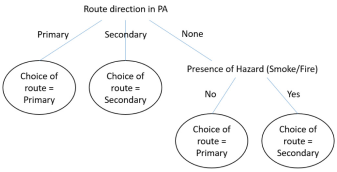
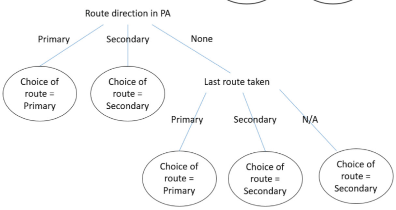
At the skill acquisition phase (ACQ in Fig. 3), the performance results indicated that the SBML-guided VE training addressed performance variability amongst the participants and assured all participants achieved competence. At the retention phase (RET in Fig. 3), the performance results indicated that emergency egress skills taught using the VE training were susceptible to skill decay over a period of 6 to 9-months. The decay in performance was largely due to the degradation of spatial knowledge (e.g. remembering vital egress routes) and reduced compliance with procedures. As such, a shorter retention interval is recommend to help trainees maintain egress skills. To address the skill degradation, an adaptive matrix assigned participants corrective training exercises based on the specific errors they made [9]. This approach accommodated the different learning paths and paces of participants. As a result, the retraining was successful in quickly bringing all participants back to demonstrable competence.

At the transfer phase (TRA in Fig. 3), the VE training provided participants with sufficient artificial experience to apply their knowledge and skills to multiple emergencies. The performance results in the near-transfer scenarios (U1–4) showed that participants were able to safely egress in emergences under similar conditions. However, the performance results in the far-transfer scenarios (U5–8) revealed that there were limitations to the extent participants could transfer their training to novel emergencies.

Modeling Evidence. The DT models provided a more comprehensive assessment of the VE training at the three learning stages. Table 2 provides the percentage of participants that formed different types of decision trees at the end of the VE training and compares the participants' DTs to the intended decision strategies from the training.

From a diagnostic perspective, the DTs provided a visual representation of the participants' strategies that helped detect when the participants received sufficient training to be competent (e.g. when DTs converged on the intended decision strategies and represented safe behaviours). At the end of the VE training, the majority of participants' DTs (91%) converged to the intended egress strategies, which was an indication that the SBML training sufficiently taught participants egress skills. However, for 9% of participants, the DT also showed that some participants developed strategies that did not match the intended training. This result indicated that given the same training, trainees developed their own heuristics for responding to emergencies. This was a useful finding because DTs provided a better understanding of how trainees develop learning strategies or heuristics, which is important for addressing individual variability and assuring demonstrable competence in the workforce.

Table 2. Types of DTs observed after participants completed the VE training (ACQ and RET).

Type	Decision Tree	% Participants	DT Comparison
1		68%	Correct
2		10%	Correct
3	 One participant had similar DT but reverse rules for PAPA and None	13%	Correct
7		3%	Incorrect
8		3%	Incorrect
*No behavioural pattern or strategy identified		3%	Trainees not suitable

4 Conclusion

The results of this research demonstrated that pedagogically designed and data-informed VE training technology can enhance conventional training. This work makes a case for operators and regulators in offshore and maritime domains to support the adoption of VE training to improve workers' overall competence and compliance during emergencies.

The empirical evidence from this experiment demonstrated the utility of VE training in providing structure, standardization, and accountability to offshore egress training. The SBML framework applied to the VE training was effective during the skill acquisition phase as it helped address individual variability in competence. Some limitations were observed at the skill retention and transfer phases of the experimental program. The results indicated that emergency egress skills are susceptible to skill decay and there are limits to how much trainees can apply their egress skills to novel emergencies. These findings challenge the convention of how egress training and training maintenance are administered at fixed-time and fixed-intervals regardless of individual learning and retention tendencies. Employing VE training in industry has the potential to improve the overall safety of offshore operations.

Further, the DT modeling evidence from this experiment revealed that DTs were useful tools for the design and assessment of VE training because they offered a visual representation of an individual's heuristics that was easy to interpret. Comparing the changes in trainees' decision-making patterns, in response to different training interventions, to the intended learning strategies helped identify systemic strengths and weaknesses in the delivery of the VE training.

Incorporating DT modeling into VE technology can improve VE training by highlighting its shortcomings and by challenging instructional designers to develop training that prepares people for emergencies. This analysis identified when the prescribed VE training did not prepare all participants for the difficulty level of the transfer scenarios. This finding is especially important because the goal of VE training is to prepare the workforce for a wide variety of emergencies. Pairing DTs' diagnostic tools with VE training offers the flexibility to provide people with practice, assessment, and corrective feedback on-demand and at a customized schedule that meets the needs of each learner.

Acknowledgments. The authors acknowledge with gratitude the support of the NSERC/Husky Energy Industrial Research Chair in Safety at Sea.

References

1. McGaghie, W., Issenberg, S., Barsuk, J., Wayne, D.: A critical review of simulation-based mastery learning with translational outcomes. Med. Educ. **48**, 375–385 (2014)
2. Smith, J., Doody, K., Veitch, B.: Being prepared for emergencies: a virtual environment experiment on the retention and maintenance of egress skills. WMU J. Maritime Aff. **18**(3), 425–449 (2019)
3. Smith, J., Veitch, B.: A better way to train personnel to be safe in emergencies. ASCE-ASME J. Risk Uncertainty Eng. Syst. Part B **5**(1), 011003 (2018)
4. Barnett, S., Ceci, S.: When and where do we apply what we learn? a taxonomy for far transfer. Psychol. Bull. **128**, 612–637 (2002)
5. Ford, J., Baldwin, T., Prasad, J.: Transfer of training: The known and the unknown. Ann. Rev. Organ. Psychol. Organ. Behav. **5**, 201–225 (2018)
6. Han, J., Kamber, M., Pei, J.: Data Mining: Concepts and Techniques. Elsevier, Amsterdam (2012)
7. Musharraf, M., Smith, J., Khan, F., Veitch, B.: Identifying route selection strategies in offshore emergency situations using decision trees. J. Reliab. Eng. Syst. Saf. **19**, 106179 (2020)

8. Smith, J., Musharraf, M., Veitch, B.: Data informed cognitive modelling of offshore emergency egress behaviour. In: Proceedings of the 2017 International Conference on Cognitive Modeling, Coventry, UK (2017)
9. Doody, K.: Evaluation of performance in emergency response scenarios: a virtual environment skill retention study. Master's Thesis, Memorial University of Newfoundland (2018)

Maritime Resource Management in the Marine Engineering and Nautical Science Education – Attitudes and Implication for Training and Evaluation

Gesa Praetorius[1,2(✉)], Carl Hult[1], and Jan Snöberg[1]

[1] Linnaeus University, Kalmar Maritime Academy, 39182 Kalmar, Sweden
{gesa.praetorius, carl.hult, jan.snoberg}@lnu.se
[2] Department of Maritime Operations, University of South-Eastern Norway, Campus Vestfold, 3184 Borre, Norway

Abstract. This study presents a survey that has been conducted as part of a larger research project focused on crew resource management in the maritime domain. As research focused on this type of training is currently limited, the Ship Management Attitude Questionnaire (SMAQ) developed by a Swedish marine insurance company, has been adopted to explore NTS knowledge of fourth year students in a maritime education program. Thirty-one students within the maritime academy's Nautical Science (n = 21) and Marine Engineering (n = 10) programs participated in the survey. The age of the participants ranged between 22 and 46 years (M = 26.6, SD = 5.79). The results show that the questionnaire as is, is maladapted to explore NTS. It is also indicated that practices trained are not always encountered in the work onboard. The article concludes with a discussion on how to potentially improve the evaluation and assessment of NTS in maritime degree programs.

Keywords: Crew resource management · MRM · Safety training · Maritime safety · Non-technical skills

1 Introduction

Shipping can be regarded as a high-risk domain with a large complexity in operations. Accidents and incidents may involve serious outcomes for seafarers and passengers, as well as for the environment and society at large. Education and training therefore play a crucial role for the safe operation of ships. While technical skills have been at the core of a mariner's skillset, non-technical skills (NTS) have become increasingly important for the safe conduct of merchant vessels. Further, knowledge in NTS has become a mandatory requirement for officers serving on board. Non-technical skills (NTS) can be defined as "*the cognitive, social and personal resources skills that complement technical skills, and contribute to safe and efficient task performance*" [1] p.1. These skills are normally split into seven areas; situation awareness, decision-making, communication, teamwork, leadership, as well as the ability to managing stress and coping with fatigue [1].

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 461–467, 2020.
https://doi.org/10.1007/978-3-030-50943-9_58

As an international industry, shipping is regulated through a legal framework developed through the International Maritime Organization (IMO). Among different codes and regulations is the so-called International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW) [2]. STCW addresses the requirements for training and certification of mariners, with a specific focus on the master and officers onboard. The aim of the convention is to establish the preconditions for comparable training standards world-wide [2]. Since its latest revision in 2010, knowledge in leadership and resource management are identified as mandatory requirements for all officers and other personnel in leadership positions aboard. This has led to an increase in focus on training NTS in maritime operations.

NTS are normally trained within maritime adaptions of Crew Resource Management. These are referred to as Maritime Resource Management (MRM), Bridge Resource Management (BRM), or Engine-room Resource Management (ERM). However, as pointed out by [3] and [4] there are several deficiencies in the way that NTS are trained and assessed. Among others, the authors note that training courses are usually only evaluated on the first or second level of the Kirkpatrick learning taxonomy [5] thus lacking evidence for successful transfer of knowledge into operational setting [3]. Further, maritime adaptions seem to largely consist of translations of aviation courses without thorough underlying training needs analyses [5]. As a result, it remains unclear if these types of courses offer an actual increase in safety, or if they are a mere measure for certification.

This article sets out to explore a questionnaire which has been used to assess MRM skills in mariners. The study is part of the MaRes-project (Maritime Training for Resilience), which is a 3-year project funded by the Swedish Mercantile Marine Foundation. The project focuses on exploring current resource management and safety training in the maritime domain to identify if and how resilience may help to improve current training approaches.

2 Background

2.1 Non-technical Skills Training in the Maritime Domain

Non-technical skills training in the maritime domain received an amplified amount of attention during the 1990s after the aviation domain attributed increased safety to the successful implementation of crew resource management (CRM).

The shipping industry started to adopt the CRM approach and focused initially on bridge operations. The IMO introduced the concept of Bridge Resource Management, the effective use and allocation of all resources available on the bridge, into the STCW Code in 1995 [6]. Since the late 1990s, the concept has been expanded to include other departments onboard, and sometimes even shore-based personnel in shipping companies. Since 1 June 2017 [2], officers in all departments onboard need to provide proof of their knowledge of BRM/ERM principles; including the allocation, assignment and prioritization of resources, effective communication between and within teams, assertiveness and leadership, teamwork and having situational awareness [2]. As in many other settings, the shipping industry is currently undergoing large changes through increased complexity of technology, which pushes seafarers farther away from the original core processes, such as navigating or operating machinery, towards a more supervisory role. Seafarers become part of a socio-technical system incorporating people, technology and organizational settings [7]. Thus, technical and non-technical skills are critical to be able to maintain the safety of people and cargo onboard and guarantee that resources (people and technology) are used and allocated in an effective way.

Although resource management training has become mandatory, research on NTS and training in the maritime domain is considered as sparse [8, 9]. The published work has so far mostly focused NTS training in relation to bridge operations, not engine-room or other departments onboard. The length of training and training modality, such as classroom-based, or simulator-based training and the evaluation and assessment techniques differ in between the reported research [8, 9]. In some instances, NTS are trained in combination with advanced technical training, such as part of an anchor handling course [10]. Three articles address the evaluation of attitudes towards NTS through adapted versions of the Flight Management Attitude Questionnaire [4, 11, 12], but only [12] distributed the questionnaire prior and after the course to determine changes in the participants' attitudes.

In summary, research within how to assess and evaluate the effects of NTS training in the maritime domain is sparse. Most research has focused on bridge operations and the published worked show a large variety in course lengths, content, instructional method, and assessment approach. As several articles [4, 11, 12] have applied maritime adaptions of the FMAQ for the evaluation, this article has tried to explore if this would be a valid method to be applied within the context of maritime educational programs.

2.2 Training NTS at Kalmar Maritime Academy

The BRM and ERM training at Kalmar Maritime Academy consists of classroom-based lectures, seminars with group discussions, and simulator exercises. The main part of the specific BRM/ERM training is performed during the last term in the four-year long education for students in nautical science and marine engineering. Prior to the training, the students have undergone 2½ of undergraduate studies and at least 10 months of seagoing service as cadets.

The classroom-based activities address BRM/ERM tools and underlying theoretical concepts, such as *Communication, Briefing, Challenge and Response, Authority and Assertiveness, Cultural Awareness, Fatigue Management, Short Term Strategy*, and *Situation Awareness*. These topics are discussed in small groups in the seminar to enable students to share their experiences from the time as cadets. In addition, the students have access to 15 computer-based training modules, which can be used to learn at their own pace.

The simulation exercises aim to provide the students with experiences from serving in a crew in different positions, and making use of the different tools from the classroom-based education. During three exercises (15 h in total), the nautical science and marine engineering students are trained jointly in teams of five (three bridge officers and two engine-room officers). Each simulation is followed-up by a debriefing.

3 Methodological Approach

3.1 Participants

Thirty-one students within the maritime academy's Nautical Science (n = 21) and Marine Engineering (n = 10) programs participated in the survey. The age of the participants ranged between 22 and 46 years (M = 26.6, SD = 5.79). All of them had spent at least 7 months at sea working on a range of different vessels, such as bulk carrier, cruise ship, ferry or tanker.

3.2 Questionnaire

The Ship Management Attitude Questionnaire (SMAQ) has been distributed through Swedish Club, a marine insurance company, to evaluate non-technical skills attitudes. The questionnaire is an adaption of the Flight Management Attitude Questionnaire originally developed for the aviation domain. Some of the items not applicable to the students, i.e. items relating to time off work, the relationship to the company and employee appraisal, were removed from the questionnaire before it was handed out. The SMAQ contained a total number of 100 items in the following categories; Satisfaction with own knowledge (n = 11), Teamwork perception (n = 9), Organizational Issues (n = 33), Ship Management (n = 26), Leadership (n = 3), Work Goals (n = 15), Automation (n = 6), and Multicultural aspects of work (n = 9). Each category, but leadership, offer respondents score on a 5-point-Likert scale. The scaling differed in between the different categories. In some questions, participants were asked to express their degree of agreement (e.g. items related to ship manages), in others their level of knowledge (e.g. items related to teamwork perception), or the degree of importance (e.g. work goals).

In addition, the questionnaire also contained a section for demographics and four free-text questions, where the participants were asked to provide their perception of how work environment and safety could potentially be improved, as well as their perception of maritime authorities and other assistance services. To gain insights in which of the tools and practices associated with MRM the students had experienced and used during their practical work experience, a table with the different MRM modules was added at the end of the questionnaire. The participants were asked to indicate how often these tools and techniques have been applied in real work settings on a 5-point-Likert scale ranging from never to always.

Initially, the study was designed to distribute the questionnaire before and after the course to explore potential changes in the attitudes of the students. As the participants included students from the marine engineering and nautical science degree program, the aim was further to explore if there were any difference in the attitudes of the two student groups. However, many of the questionnaires were either incomplete returned empty or incomplete, or respondents had chosen more than one alternative in their answer to an item. Therefore, only descriptive statistics were used to explore the data.

We have therefore chosen to focus more on the indication for the design of assessment and a potential questionnaire for NTS than on the obtained results than the results as such in the following sections.

4 Results

All mean values below relate to the highest possible value of 5. The responses indicate that the students are fairly confident in their own knowledge to with regards to basic seafaring knowledge (M = 3.97, SD = 0.605), their ability to handle normal operations (M = 3.65, SD = 0.608), and their knowledge about safety manuals onboard (M = 3.5, SD = 0.630). The teamwork was perceived to have the best quality among deck officers and engineers. This could be explained by the limited experience that students normally have with the shore-based departments and agents.

With regards to organizational issues and ship management, the responses show that checklists and briefings are not always conducted as planned or prescribed. Some respondents indicated that checklists were omitted to a large extent in the daily work. Communication, and especially the ability to speak up when errors are detected, show to be important for the respondents. Monitoring and supporting each other, as well as the ability to detect signs of stress are considered important. Further, there is an indication for a need to train crewmembers better in how to cope with fatigue (M = 3 SD = 1.155).

The participants responses reveal that most of the techniques and tools related to resource management modules are neither often encountered, nor practiced by the students. Most frequently encountered were the Communication (M = 3.73, SD = 1.337), Situation Awareness (M = 3.3, SD = 1.119), Briefing (M = 3.23, SD = 1.165), and Closed Loop Communication (M = 3.13, SD = 1.033), which indicates that these are seldom to sometimes used. Attitude Management Skills (M = 2.2, SD = 0.887), Automation Awareness (M = 2.37, SD = 1.008) and Short-term Strategy (M = 2.43, SD = 0.935) received the lowest ranking, meaning that these are encountered more seldom.

Only few students chose to respond to the free text questions. With regards to improved safety, responses foremost highlighted more safety drills and education, better briefing and debriefing, increased number of crew and more attention and respect towards the crew's concern as potential measures. To improve the efficiency and job satisfaction responses emphasized the need for more teambuilding, communication across departments, good living and working conditions, as well as the need to have varying work tasks and instill pride in being in a certain company.

5 Discussion

This study intended to measure attitudes regarding NTS in maritime degree programs. However, the number of omitted answers in the questionnaire indicate that the survey as such is not particularly suited to explore students' attitudes. This can partially be explained by the large number of items which require a certain amount of practical work experience. It was especially salient in the responses relating to teamwork quality, organizational issues and ship management. Another aspect, which may have affected the quality of the data, has been the use of different scales for different items. In some categories, values associated to scales changed up to three times throughout the category.

Further, many of the items focused on a general notion of safety and aspects of operations, which do not clearly match non-technical skills such as communication, decision making, or leadership, and which were rather maladapted for the evaluation and assessment of the NTS of undergraduate students in maritime degree programs.

However, based on the results of a descriptive analysis, several areas of concern could be identified. The responses showed that many of the modules, techniques and tools taught in the resource management course were not encountered during the work onboard and thus were applied rather seldom in real-life settings. This gives raise for concern as knowledge in MRM and the practice of effective resource management is required by the STCW. This may mean that knowledge from the courses is not being transferred into work settings, or that such knowledge is not considered as not important for work on board.

With regards to evaluation and assessment techniques, this study shows that an adapted version of the original FMAQ may not be a proper tool to assess NTS attitudes in maritime degree programs. NTS should be clearly defined and operationalized. The number of items should be limited and specific NTS should be targeted. It might also be important to develop a better notion of how to tailor theoretical knowledge and computer-based training modules to student needs, especially in a maritime context so that what is learned in the classroom becomes relevant for practical exercises and work onboard. Further research may also target to successively develop behavior markers, which can be used to explore the relationship between attitudes and behavior to be able to identify training needs in student exercises that would further allow to tailor courses better to the participants instead of the provision of generic knowledge about NTS.

6 Concluding Remarks

This study only marks the beginning of a research project exploring NTS training and assessment. While we did not succeed to analyze and compare attitudes of undergraduate students in maritime education programs, the article still points towards the need for better assessment and evaluation techniques. Non-technical skills in maritime operations, as well as the transfer from training to operational settings are not well-understood. Questionnaires can be a first step to identify training needs and can become a valuable tool to develop training and assessment practices further.

Acknowledgments. The work presented here is part of the MaRes-project financed by the Swedish Mercantile Foundation. Further, we would like to acknowledge the kind support of students and staff at the Kalmar Maritime Academy. Thank you!

References

1. Flin, R.H., O'Connor, P., Crichton, M.: Safety at the Sharp End: A Guide to Non-Technical Skills. Ashgate, Farnham (2008)
2. IMO. STCW International Convention on Standards of Training, Certification and Watchkeeping for Seafarers: Including 2010 Manila Amendments: STCW Convention Ans STCW Code: International Maritime Organization (2017)

3. Havinga, J., de Boer, R.J., Rae, A., Dekker, W.S.: How did crew resource management take-off outside of the cockpit? a systematic review of how crew resource management training is conceptualised and evaluated for non-pilots. Safety **3**(4), 26 (2017). https://doi.org/10.3390/safety3040026
4. O'Connor, P.: Assessing the effectiveness of bridge resource management training. Int. J. Aviat. Psychol. **21**(4), 357–374 (2011). https://doi.org/10.1080/10508414.2011.606755
5. Kirkpatrick, D., Kirkpatrick, J.: Evaluating Training Programs: The Four Levels. Berrett-Koehler Publishers, San Francisco (2006)
6. Chauvin, C., Lardjane, S., Morel, G., Clostermann, J.-P., Langard, B.: Human and organisational factors in maritime accidents: analysis of collisions at sea using the HFACS. Accid. Anal. Prev. **59**, 26–37 (2013). https://doi.org/10.1016/j.aap.2013.05.006
7. Kataria, A., Praetorius, G., Schröder-Hinrichs, J.U., Baldauf, M.: Making the case for Crew-Centered Design (CCD) in merchant shipping. In: Paper presented at the 19th Triennial Congress of the IEA, Melbourne (2015)
8. Wahl, A.M., Kongsvik, T.: Crew resource management training in the maritime industry: a literature review. WMU J. Maritime Aff. **17**(3), 377–396 (2018). https://doi.org/10.1007/s13437-018-0150-7
9. Fjeld, G.P., Tvedt, S.D., Oltedal, H.: Bridge officers' non-technical skills: a literature review. WMU J. Maritime Aff. **17**(4), 475–495 (2018). https://doi.org/10.1007/s13437-018-0158-z
10. Håvold, J.I., Nistad, S., Skiri, A., Ødegård, A.: The human factor and simulator training for offshore anchor handling operators. Saf. Sci. **75**, 136–145 (2015)
11. Röttger, S., Vetter, S., Kowalski, J.T.: Ship management attitudes and their relation to behavior and performance. Hum. Factors **55**(3), 659–671 (2013). https://doi.org/10.1177/0018720812461271
12. Röttger, S., Vetter, S., Kowalski, J.T.: Effects of a classroom-based bridge resource management training on knowledge, attitudes, behaviour and performance of junior naval officers. WMU J. Maritime Aff. **15**(1), 143–162 (2016). https://doi.org/10.1007/s13437-014-0073-x

Development of a SAGAT Query and Simulator Experiment to Measure Situation Awareness in Maritime Navigation

Hui Xue, Bjørn-Morten Batalden(✉), and Johan-Fredrik Røds

Faculty of Science and Technology, Department of Technology and Safety, UiT The Arctic University of Norway, Pb 6050, Langnes, Tromsø 9037, Norway
{hui.xue, bjorn.batalden, johan-fredrik.rods}@uit.no

Abstract. Many ship collisions and groundings occur due to navigators' erroneous situation awareness (SA). The objective of this study is to develop a method to measure SA for maritime navigation and collision avoidance (SA-MA). This study uses the Situation Awareness Global Assessment Technique (SAGAT) as a basis and tool to assess SA. Both interviews with experts and simulator experiments are used. Ten participants, five navigators with extensive experience, and five second-year students at a nautical science program participate in the simulator experiment. Hierarchical Task Analysis (HTA) is used to map the navigation and collision avoidance tasks as input to the SA queries. The objective measurements collected from the simulator and subject matter experts are used for the SAGAT score. A well-developed SAGAT query and simulator experiment results in a difference in the SA-MA between the experienced navigators and the students with less experience. The study found it is difficult to measure SA-MA, especially for level 2 and 3 SA.

Keywords: Situation Awareness (SA) · SAGAT · Maritime · Navigation · Hierarchical Task Analysis (HTA) · Simulator

1 Introduction

Many ship collisions and groundings occur due to navigators' erroneous situation awareness (SA). In particular, unsafe acts and preconditions for unsafe acts are important causes for ship collisions and groundings [1]. For both ship collisions and groundings, decision-errors and resource management are the two most frequent causes [1]. Grech et al. [2] found that 71% of navigators' errors are SA related problems.

The concept of SA is based on the interaction between the operator and the surrounding environment [3]. There is an underlying assumption that the situation in the working environment can be changed in different ways. For example, it can change fast or slow, significantly or not significantly, obviously or concealed, repeatedly or not repeatedly, planned or unplanned [4]. In the maritime domain, during the different tasks related to sailing a ship, the navigator should be able to adjust and adapt the performance based on the current situation or the change of the situation, taking future development into account. SA of navigators can be improved by training and practising

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 468–474, 2020.
https://doi.org/10.1007/978-3-030-50943-9_59

[5, 6]. Therefore, in order to improve the SA of navigators, developing reliable and valid measures of SA has been the focus of this experiment.

The objective of the research is to develop and assess a Situation Awareness Global Assessment Technique (SAGAT) Query to measure SA-MA.

2 Theory of SA and SA in Maritime

Endsley and Jones [7] explain situation awareness as "being aware of what is happening around you and understanding what that information means to you and in the future". It relates to what is important for a task or goal. Several definitions of situation awareness exist [8], but in this paper, we choose to use the definition "Situation Awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future." [9].

Through human sensory systems, being either vision, audition, vestibular system, the somatic sensory system, gustation or olfaction, it is possible to perceive information about the elements (Proctor and Proctor, 2006) [10]. The use of these sensory systems will vary depending on the domain and type of job or task, and the data necessary to achieve or address level 1 SA can be hard to map in many domains [7]. Endsley [9] identify that SA challenges mostly relate to level 1 SA.

Level 2 SA is to convert the sensory information to create an understanding of the current situation. Regarding level 2 SA for a deck officer on a ship, Sharma et al. (2019) [8] present information elements that refer to the parameter Closest Point of Approach (CPA) and Time to CPA (TCPA) as examples. The novice operator may achieve the same level 1 SA as the experienced operator but not being able to convert this information into level 2 SA, achieving a lower level 2 SA [7].

The last level from the definition, level 3 SA, is how an operator manages to translate the information gathered, and understanding of the current state, into a future state. Endsley and Jones [7], state that good level 3 SA can only be achieved from an operator having a sound level 2 SA and an understanding of the "functioning and dynamics of the system they are working in". To achieve a good level 3 SA requires sound domain understanding and time spent on achieving good level 3 SA is often extensive among experts Endsley and Jones [7]. Sharma et al. [8] refer to extending vectors from targets and radar trials as methods to assist in achieving good level 3 SA for deck officers. Insufficient mental capacity and insufficient knowledge of the domain are two possible reasons for not achieving a good level 3 SA [7].

3 Methodology

This study uses both interviews with experts and a simulator experiment to develop and assess the SAGAT for SA-MA. The interviewees have extensive experience as navigators in both the merchant fleet and the navy. The study uses Hierarchical Task Analysis (HTA) to list the navigation and collision avoidance tasks as input to the SA queries. The simulator experiment uses ten participants, five navigators with extensive experience, and five second-year students at a nautical science program with little experience. The 240° view simulator used for the experiment is equipped with the K-sim Navigation software from Kongsberg Digital. The vessel-model used in the experiment is called BULKC11 (length overall of 90 m and a moulded beam of 14 m). The procedure of developing SAGAT Queries is shown in Fig. 1.

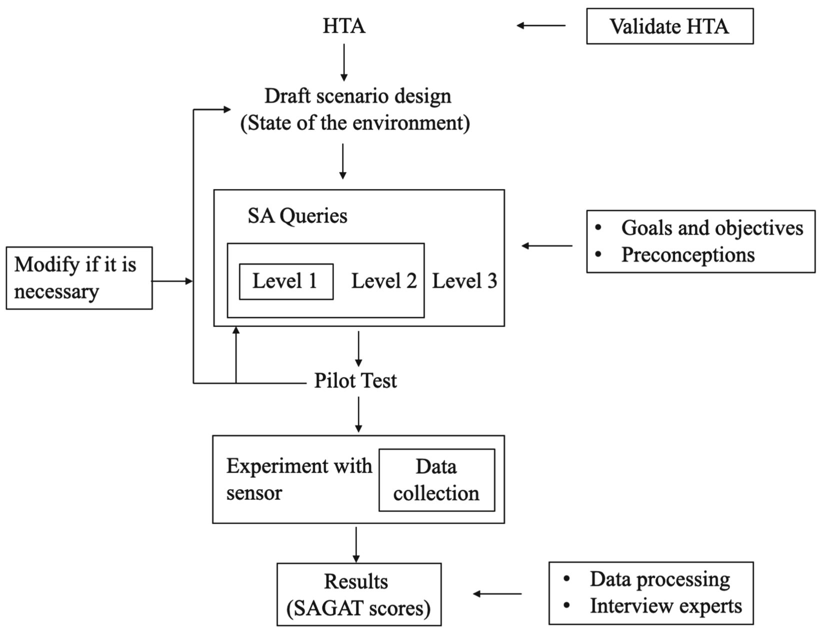


Fig. 1. The procedure of developing the SAGAT Queries.

In this study, the HTA was developed in the navigation tasks. Each task has a goal to achieve. The carry out of an HTA can be adapted to different situation and needs [11]. Before starting the process of developing the HTA, a literature study was conducted to enhance the knowledge regarding the field of research. A draft HTA was then discussed with a subject matter expert for input. The revised HTA was presented to a second subject matter expert and further adjusted and finalised.

SAGAT was initially developed to assess the SA of pilots across the three levels of SA [6, 7]. The procedure of SAGAT comprise of ten steps: (1) Define tasks, (2) Development of SA queries, (3) Selection of participants, (4) Brief participants, (5) Pilot run, (6) Task performance, (7) Freeze the simulation, (8) SA query administration, (9) Query answer evaluation (by a subject matter expert), (10) SAGAT score calculation [5].

The results of the SA requirements analysis are used to develop a set of SA queries for the experiment in the simulator [12]. As a global measure, SAGAT includes queries about all operator SA requirements, including Level 1, Level 2 and Level 3 components [12, 13]. Participants were briefed regarding the purpose of the study and the voyage plan. There were four stops where the simulator was frozen, handing out the queries based on all three levels of SAGAT. For the four stops, each of the stops was conducted within a fixed range on the course line (Fig. 2). In total, it is an approximate forty minutes voyage. An expert completed the same SA Queries with the correct answer on the simulator. The participants' answers were compared to the results of the expert.

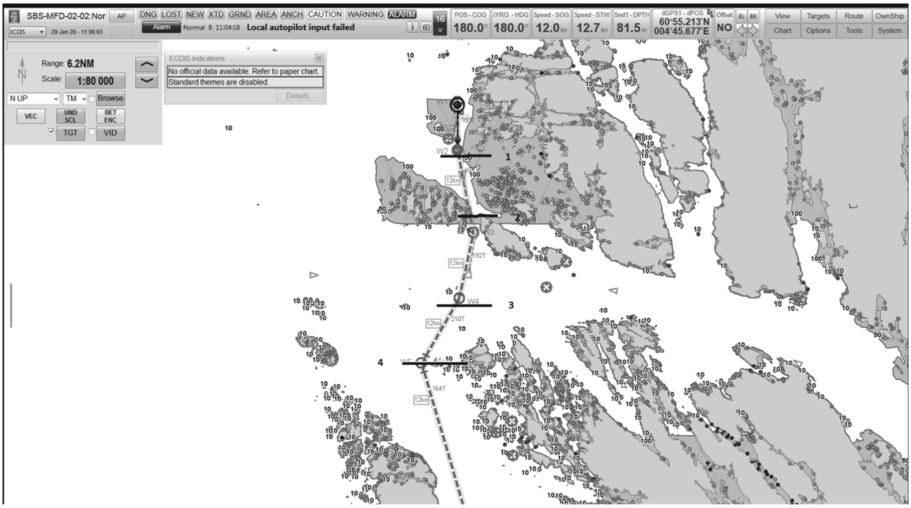


Fig. 2. Chart layout with the route and horizontal lines indicating the four SAGAT stops.

Five expert participants were interviewed face to face (Mean age = 41.8, Stand deviation = 14.0). The interviews were semi-structured with the purpose to elicit expert knowledge and experience from the SAGAT queries and simulator experiment [8, 14]. The average length of the interviews was around 10 min. The participants have an average of 9.7 years' experience as navigators with the longest being 18 years and the shortest being one year.

4 Results

The results were collected from both the students and experts SA scores and interviews with the experts. Each of the results will be presented respectively. The analysis identified that the students and experts have different performance in each of the three SA levels. The results and analysis of the SA scores are presented in Table 1. It was possible to achieve in total a maximum of 45 points on all three levels together. The results of the SA scores show that the students have 42.7% correct in total, while the experts get 46.2% in total. Students score higher than the experts do on level 1 SA while the experts score higher on level 2 SA. For level 3 SA, the students and experts score close to the same, the students scoring 1.8% points above the experts. The scores show that students got better scores than experts on level 1, experts got better scores on level 2, and they have similar scores on level 3.

Table 2 breaks the results further down and show each stop for level 1, 2 and 3 SA. There is no difference between the expert and student on level 3 SA except for the last stop. For level 2 SA, experts scored noticeably higher than the students on all four stops.

Table 1. Results of the SA scores at different levels and different stops.

Role	Level 1	Level 2	Level 3	1st Stop	2nd Stop	3rd Stop	4th Stop	Overall
Student	67.1%	25.9%	30.9%	37.3%	38.5%	48.6%	60.0%	42.4%
Expert	58.8%	44.7%	29.1%	42.7%	46.2%	50.0%	46.7%	46.2%

Table 2. Results of the SA scores at all levels of SA for all four stops.

Role	1st Stop			2nd Stop		
	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
Student	72.0%	16.0%	24.0%	48.0%	36.0%	26.7%
Expert	64.0%	40.0%	24.0%	56.0%	48.0%	26.7%
Role	3rd Stop			4th Stop		
	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
Student	76.7%	26.7%	30.0%	80.0%	20.0%	80.0%
Expert	60.0%	46.7%	30.0%	40.0%	40.0%	60.0%

The interviews with the experts gave valuable insight into possible strengths and weaknesses with the SAGAT queries. The development of SAGAT queries will be scenario- and time-dependent. As an example, two different candidates conducting the same simulator experiment but with slightly different location or time when the query is conducted may have shifted what needs to be focused on by the candidate. Sometimes the time to arrive a certain position is important while for others the position, as a result of speed and time, is important. The experts meant that it is more important to know where two vessels meet than to know the time when they meet. The candidates need to focus on many information sources, and it is less likely that they remember details from all sources when the query is handed out. Most of the experts agreed that for different situations, some information is not important to remember. It is also not always important to remember how many target ships there are when there is not too much traffic.

5 Discussion and Concluding Remarks

From a maritime point of view, it is found to be difficult to measure SA, especially for level 2 and 3. The participants' answers from the queries are analysed by a subject matter expert, which decides if the answers are correct, or within the acceptable range. For level 1 SA, the information necessary to analyse the answers from the queries can be collected directly from the simulator. For level 2 SA and level 3 SA, it is more complex to analyse the answers from the queries. Experts may have to decide if the answers are within an acceptable range based on their experience in the given situations. This may be a challenge since different experts may have different experiences and different opinions. Further, the experts have to combine the information and their experience in order to analyse the situation.

Several explanations may be given to why students score higher on level 1 SA. Firstly, the students may be more focused on the information on the screens but have a lower capacity to understand and utilise the information to comprehend the situation. Secondly, experts have more experience in selecting important information in a situation. Experts may be better at selecting relevant information while the SA query tries to measure too much information. This is in line with Sharma et al. [8]. Thirdly, better memory might also help the student group to get higher scores on level 1 SA since research shows that memory loss is age-related [15, 16]. Fourthly, the students are familiar with the instruments since they practice very often on the simulator while most of the experts are less familiar with the instruments, and rely more on looking out and just collect basic information.

When it comes to level 2 SA, experts got markedly higher scores than students, which may indicate that experts are better in converting the sensory information to the understanding of the current situation. This is in line with the results of Endsley and Jones [7], that novice are not being able to convert this information into level 2 SA. Another explanation could be that some part of the information asked for in the queries did not affect the experts understanding of the current situation. Based on the results of the leve1 3 SA, it indicates that the students and experts have a similar capacity of translating the current state into a future state. This is an unexpected result. It might be that the existing queries are not good at measuring level 3 SA. This is supported by the difference detected from level 1 SA to level 2 SA.

SAGAT queries should perhaps focus more on the required situation rather than specific parameters such as CPA and TCPA. Based on the interviews and the experiments, it is necessary to spend more time on making clearer questions in the queries. The participants need to understand the questions correctly. A well-developed SAGAT query and simulator experiment should result in a difference in the SA-MA between the experienced navigators and the students with less experience. For future studies, it is recommended to create queries that enable utilisation of all tools in the simulator to reduce the subjective assessment. A tested SAGAT for the SA-MA will be a valuable tool for future studies related to maritime navigation and collision avoidance and as a tool to assess the training of the nautical students. With an overall score of between 42.4 and 46.2, there is a need to further develop the SAGAT query for SA-MA.

Acknowledgments. The authors would like to acknowledge the assistance from Ingar Lorentsen and Øyvind Haugseggen at UiT the Arctic University of Norway during data collection on the simulator.

References

1. Yıldırım, U., Başar, E., Uğurlu, Ö.: Assessment of collisions and grounding accidents with human factors analysis and classification system (HFACS) and statistical methods. Saf. Sci. **119**, 412–425 (2019)
2. Grech, M.R., Horberry, T., Smith, A.: Human error in maritime operations: analyses of accident reports using the Leximancer tool. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Sage Publications, Los Angeles (2002)
3. Endsley, M.R., Garland, D.J.: Situation Awareness Analysis and Measurement. CRC Press, Boca Raton (2000)
4. Koester, T.: Situation awareness and situation dependent behaviour adjustment in the maritime work domain. Hum.-Centered Comput. Cogn. Soc. Ergon. Aspects **3**, 255 (2019)
5. Salas, E., et al.: Does crew resource management training work? An update, an extension, and some critical needs. Hum. Factors **48**(2), 392–412 (2006)
6. Stout, R.J., Salas, E., Fowlkes, J.E.: Enhancing teamwork in complex environments through team training. Group Dyn. Theory Res. Pract. **1**(2), 169 (1997)
7. Endsley, M.R., Jones, D.G.: Designing for Situation Awareness: Understanding Situation Awareness in System Design. CRC Press, Boca Raton (2011)
8. Sharma, A., Nazir, S., Ernstsen, J.: Situation awareness information requirements for maritime navigation: a goal directed task analysis. Saf. Sci. **120**, 745–752 (2019)
9. Endsley, M.R.: Situation awareness global assessment technique (SAGAT). In: Proceedings of the IEEE 1988 National Aerospace and Electronics Conference. IEEE (1988)
10. Proctor, R.W., Proctor, J.D.: Sensation and perception. In: Handbook of Human Factors and Ergonomics, pp. 51–88 (2006)
11. Annett, J.: Hierarchical task analysis. Handb. Cogn. Task Des. **2**, 17–35 (2003)
12. Stanton, N.A., et al.: Human Factors Methods: A Practical Guide for Engineering and Design. CRC Press, Boca Raton (2017)
13. Endsley, M.R.: Direct measurement of situation awareness: validity and use of SAGAT. In: Situational Awareness, pp. 129–156. Routledge (2017)
14. Kokar, M.M., Endsley, M.R.: Situation awareness and cognitive modeling. IEEE Intell. Syst. **27**(3), 91–96 (2012)
15. Small, G.W.: What we need to know about age related memory loss. BMJ **324**(7352), 1502–1505 (2002)
16. Luszcz, M.A., Bryan, J.: Toward understanding age-related memory loss in late adulthood. Gerontology **45**(1), 2–9 (1999)

Assessing Situation Awareness Across Different Submarine Control Room Layouts

Kiome A. Pope(✉), Aaron P. J. Roberts, Daniel Fay, and Neville A. Stanton

Human Factors Engineering, Transportation Research Group, Faculty of Engineering and Physical Sciences, University of Southampton, Boldrewood Campus, Burgess Road, Southampton SO16 7QF, UK
{K.A.Pope,aprlc13,D.T.Fay,N.Stanton}@soton.ac.uk

Abstract. Studies on contemporary submarine control room configurations revealed a bottleneck of information transition. The co-location of operators dependent on each other for task-relevant information relieved this, however, the impact of operator co-location on subjective Situation Awareness (SA) was not assessed. In the current work, half of the teams from the baseline study and half of the teams from the co-location configuration were evaluated on two SA questions: (1) "Rate the awareness of the total number of vessels surrounding the submarine" and (2) "How many vessels did the entire command team encounter during the scenario?". Participants completed high and low demand Return to Periscope Depth scenarios and responded to the SA questions immediately after finishing each scenario. Results indicate that operator SA decreased in the high demand scenarios regardless of control room configuration type. Furthermore, operator SA was greater in the co-location configuration than in the baseline study (contemporary configuration).

Keywords: Submarine control room · Situation Awareness · Teamwork

1 Introduction

Submarine control rooms are highly advanced sociotechnical systems, having evolved across many decades of operation, however, this does not mean that they cannot be improved [1, 2]. Requirements for submarine command teams of the future include processing greater volumes of data of increased complexity from sensors with improved capabilities and new sensors (e.g. Unmanned Underwater Vehicles), while maintaining or potentially reducing crew sizes [3–5]. A key challenge will be to ensure Situation Awareness (SA) is maintained at a suitable level for optimal performance, while effectively managing increased volumes of data [6].

A submarine control room is an excellent example of an environment where SA is distributed among multiple agents, both human (operators) and non-human (technological) and must be communicated dependent on operators' respective goals [7, 8]. A previous study by Stanton et al. [9] empirically examined submarine command teams' ways of working while completing a Return to Periscope Depth (RTPD) operation. A bottleneck of information was identified between the Operations Officer

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 475–482, 2020.
https://doi.org/10.1007/978-3-030-50943-9_60

(OPSO) and Sonar Controller (SOC) [9–12]. In a submarine control room, Sonar Operators (SOPs) analyse acoustic and visual cues to detect and identify contacts around the submarine [2, 13]. This information is provided to Target Motion Analysis (TMA) operators via the SOC and OPSO, in order to calculate solutions concerning contact behaviour and position [2, 13, 14]. By integrating this disparate information, the Officer of the Watch (OOW) must make tactical decisions to maintain the three tenets of submarine operation: remain safe, remain undetected, and complete the mission [15, 16]. A recommendation to alleviate the bottleneck between the OPSO and SOC was the co-location of operators dependent on each other for task relevant information [9–12, 14, 17, 18]. Results from Event Analysis of Systemic Teamwork (EAST) analysis indicated that the co-location of operators relieved the bottleneck between the OPSO and SOC, leading to a greater number of communications and increased task completion [14, 17, 18]. However, this work did not examine operators' subjective ratings of SA. As Distributed Situation Awareness (DSA) is "*concerned with how knowledge is used and parsed between agents*" it is important to consider how any changes to social, information, and task networks may affect operator ratings of SA [7].

In the current work, operators' subjective SA is examined during a high and low demand RTPD scenario. This is presented for the contemporary baseline configuration and a novel co-location configuration [9, 18]. Due to the disparate nature of information in a submarine control room, a DSA approach is used in the current work [6, 7].

2 Method

2.1 Participants

Ten teams of eight participants and ten teams of seven participants were recruited opportunistically for the baseline study (contemporary configuration) and the co-location configuration respectively. In the baseline study 71 males and nine females participated with an age range of 18-55 (Mean = 26.83, Standard Deviation = 8.69) and in the co-location study 59 males and 11 females participated with an age range of 20-48 (Mean = 28.94, Standard Deviation = 7.02). The study protocol received ethical approval from the University of Southampton Research Ethics Committee (Protocol No: 10099) and Ministry of Defence Research Ethics Committee (MoDREC) (Protocol No: 551/MODREC/14).

2.2 Equipment and Materials

Subject Matter Experts (SMEs) from the Royal Navy (RN) and relevant industry partners informed the design and construction of a submarine control room simulator to be representational of a currently operational RN submarine [19]. The simulator had nine networked stations, each installed with the simulation engine Dangerous Waters allowing the simultaneous completion of submarine command team tasks. The composition of the command team in the study was informed by SMEs, and was comprised of an OOW, an OPSO, a SOC, two SOPs, two TMAs, a Periscope (PERI) operator, and a Ship Control (SHC) operator. The role of the OOW was assumed by a member of the experimental team in the baseline and co-location studies in order to guide teams tactically with relevant task fidelity. In the co-location configuration, the role of SHC was also assumed by a member of the experimental team.

Two RTPD scenarios (high and low demand) were designed and programmed in Dangerous Waters. Scenario demand was manipulated by the number of contacts presented in the scenario and contact behaviour. Four contacts were presented in the low demand scenario, and 13 contacts were presented in the high demand scenario.

A day long tutorial package was designed in order to train novice participants to be representative of a submarine command team. Video tutorials included essential concepts such as bearing, course, and range, as well as operator specific videos (e.g. sonar tutorial). A communications game was also developed to train participants in the collation of information from disparate sources, as well as correct military verbal protocol.

A range of domain-specific measures of SA were developed with SMEs. This included awareness of the total number of vessels surrounding own ship, awareness of own ship parameters, and awareness of length of the scenario. Participants were also asked to state the number of vessels the command team encountered during the scenario.

2.3 Procedure

The testing procedure remained the same between the baseline and co-location studies. Participants attended for two days: a training day and a testing day. Upon arrival on the training day, informed consent was obtained from participants and operator roles were randomly assigned. In the morning, participants took part in the communications game and watched tutorials on essential concepts. The afternoon consisted of workstation specific tutorials, before both parts of the day were brought together in practice scenarios as part of a command team.

A final practice scenario was completed on the testing day, monitored by the experimental team, to ensure adequate performance of all tasks. Participants then completed all scenarios. After the end of each scenario, participants immediately completed the SA questionnaire digitally via a Portable Document Format (PDF) form. At the end of the final scenario, participants were given a full debrief and thanked for participating.

2.4 Analysis of Data

Upon completion of data processing, a repeated measures Analysis of Variance (ANOVA) will be conducted to examine the interaction between operator role and scenario demand on ratings of SA. To date, SA for five teams each from the baseline ($n = 40$) and co-location configuration ($n = 35$) during low and high demand RTPD have been processed and analysed. As the role of SHC was assumed by an experimenter for the co-location study the results from this operator have been removed from the baseline study for ease of comparison. In the current work, two questions from the SA questionnaire are considered (1) "Rate the awareness of the total number of vessels surrounding the submarine" and (2) "How many vessels did the entire command team encounter during the scenario?”. Responses to question (1) were given on a scale of 1 (having extremely low awareness) to 6 (having extremely high awareness). The results of this are presented as an average. Responses to question (2) were numerical, in the first instance they were scored as correct or incorrect based on the true number of contacts presented in the scenario (low: 4, high: 13). A second stage of analysis calculated the difference between the true number of contacts and the participant reported number of contacts, to provide an indication of participant error.

3 Results

In both studies, subjective ratings of awareness of the number of vessels surrounding own ship were lower in the high demand scenario (Table 1). Furthermore, when comparing the true number of vessels encountered in the scenario to the number reported by participants, a correct responses were higher in the low demand scenario than in the high demand scenario (Table 2). Similarly, the error difference between the true number of vessels and participant reported number of vessels was a larger absolute value in the high demand scenarios (Table 1 and Fig. 1).

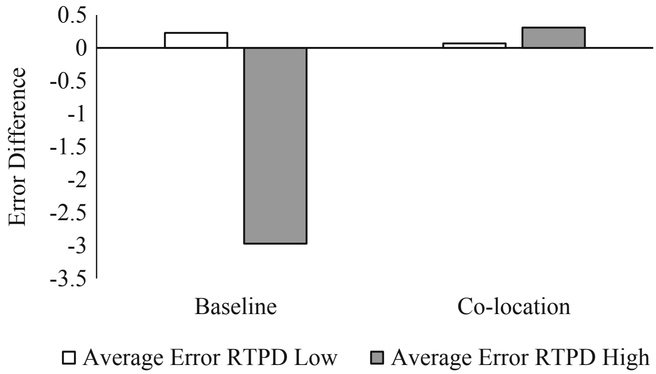


Fig. 1. Error difference between true number of vessels and participant reported number of vessels.

When comparing the two configurations, subjective ratings of awareness of the number of vessels surrounding own ship were greater in the co-location configuration than in the baseline study (Table 1). The correct responses to the number of vessels encountered also increased in the co-location configuration (Table 2). Additionally, the error difference between the true number of vessels and participant reported number of vessels was a smaller absolute value in the co-location configuration than in the baseline configuration. Furthermore, in the high demand scenario, where results indicated an underestimate of contacts in the baseline study, this changed to an overestimate in the co-location configuration (Table 1 and Fig. 1).

Table 1. Awareness ratings and error difference for vessels surrounding own-ship.

	Baseline		Co-location	
	Low	High	Low	High
Awareness of vessels	5.11	3.83	5.34	4.01
Error difference for number of vessels	0.26	−2.86	0.07	0.32

Table 2. Counts of correct and incorrect answers to "How many vessels did the entire command team encounter during the scenario?".

	Baseline				Co-location			
	Low		High		Low		High	
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
OPSO	4	1	0	5	4	1	1	4
SOC	4	1	0	5	5	0	1	4
SOP1	4	1	0	5	5	0	0	5
SOP2	5	0	0	5	5	0	2	3
TMA1	4	1	0	5	4	1	0	5
TMA2	4	1	0	5	4	1	0	5
PERI	5	0	0	5	4	1	0	5
Total	30	5	0	35	31	4	4	31

Breaking the responses to question (2) down by operator role reveals correct and incorrect responses were similar across operator role in the low demand scenarios. Between the two configurations, patterns of correct and incorrect responses were also similar. In the co-location configuration high demand scenario, an increased number of correct responses compared to the baseline study were observed for the OPSO, SOC, and SOP2 (Table 2).

4 Discussion

The current work aimed to evaluate subjective operator SA in a contemporary submarine control room configuration and in a novel co-location configuration. Results indicated that operator SA decreased in the high demand scenarios regardless of control room configuration type. This was to be expected due to the greater number of contacts operators were required to process in the high demand scenario. The difference in scenario demand also had a drastic impact on the number of correct responses to question (2) with zero correct responses in the baseline configuration, and four in the co-location configuration, compared to 30 and 31 correct responses in the low demand respectively. Nevertheless, responses were more accurate in the co-location configuration scenarios than in the baseline scenarios.

Out of all operators who responded correctly in the co-location configuration high demand scenario, two were from the sound room. These operators must directly interact with the sonar system in order to detect, designate, and identify contacts, thus their workstations provide the number of contacts encountered by the submarine in their rawest acoustic/visual format [13]. This is an excellent example of where DSA whereby information is held by a technological agent and must be interpreted and communicated by human operators [7]. It is also encouraging that one OPSO provided one of the correct answers in the co-location configuration high demand scenario. The OPSO facilitates the formation of the tactical picture and supports the OOW in their decision making [14, 16]. As such, it is imperative that they have a high awareness of the number of contacts surrounding own ship in order to support the OOW in maintaining the three tenets of submarine operation [15].

The current work builds upon previous work investigating the co-location of operators dependent on each other for task relevant information in a submarine control room [14, 17, 18]. The observed increase in communications and task completion appears to have facilitated greater operator subjective ratings of SA [14, 17, 18]. This indicates that, despite their highly evolved state, submarine control rooms can be optimized to improve operator SA [1, 2].

4.1 Conclusions and Future Work

Submarine command teams of the future will be required to process greater volumes of data from sensors with increased capabilities, while maintaining, or even reducing, crew sizes [3–5]. It is important that any new technologies, control room layouts, or ways of working are thoroughly evaluated from a sociotechnical systems perspective in order to understand potential impacts on command team performance [3, 16]. Once all data analysis is complete, the current work will provide a comprehensive understanding of operator SA in contemporary submarine control rooms and provide an understanding of the impact of a novel co-location configuration. An observable effect with half the data indicates an optimistic prospect of findings with high statistical power. Once statistical analysis is completed, it is anticipated this work will offer insight into how submarine control rooms may be improved to maintain or increase operator SA.

Acknowledgments. This work was funded by the Underwater Systems Programme of the UK Ministry of Defence Science and Technology Programme. Any views expressed are those of the authors and do not necessarily represent those of the Ministry of Defence or any other UK government department. The authors would like to thank Joy Richardson from the University of Southampton for their help with the collection of data.

References

1. Stanton, N.A.: Representing distributed cognition in complex systems: how a submarine returns to periscope depth. Ergonomics **57**, 403–418 (2014)

2. Fay, D., Stanton, N.A., Roberts, A.P.J.: All at sea with user interfaces: from evolutionary to ecological design for submarine combat systems. Theor. Issues Ergon. Sci. **20**, 632–658 (2019). https://doi.org/10.1080/1463922X.2019.1582115
3. Roberts, A.P.J., Stanton, N.A., Fay, D.: The command team experimental test-bed phase two: assessing cognitive load and situation awareness in a submarine control room. In: Stanton, N.A., Landry, S., di Bucchianico, G., Vallicelli, A. (eds.) Advances in Human Aspects of Transportation. Advances in Intelligent Systems and Computing, pp. 427–437. Springer, Cham (2017)
4. Bateman, S.: Perils of the deep: the dangers of submarine proliferation in the seas of east Asia. Asian Secur. **7**, 61–84 (2011)
5. Mansell, T., Tynan, A., Kershaw, D.: Investigating the network enabled conventional submarine: an Australian perspective. Undersea Def. Technol. Pacific. **436** (2003)
6. Dominguez, C., Long, W.G., Miller, T.E., Wiggins, S.L.: Design directions for support of submarine commanding officer decision making. In: Proceedings of 2006 Undersea HSI Symposium: Research, Acquisition and the Warrior, pp. 6–8 (2006)
7. Stanton, N.A., Stewart, R., Harris, D., Houghton, R.J., Baber, C., McMaster, R., Salmon, P. M., Hoyle, G., Walker, G.H., Young, M.S.: Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. Ergonomics **49**, 1288–1311 (2006)
8. Walker, G., Salmon, P., Green, D., Stanton, N., Baber, C.: Toward a theory of agent-based systemic situational awareness. In: Vincenzi, D.A., Mouloua, M., and Hancock, P.A. (eds.) Proceedings of the 2nd Human Performance, Situation Awareness and Automation Conference (HPSAAII), LEA, Mahwah, NJ, 22–25 March 2004, pp. 83–87. Daytona Beach (2004)
9. Stanton, N.A., Roberts, A.P.J., Fay, D.T.: Up periscope: understanding submarine command and control teamwork during a simulated return to periscope depth. Cogn. Technol. Work **19**, 399–417 (2017). https://doi.org/10.1007/s10111-017-0413-7
10. Roberts, A.P.J., Stanton, N.A., Fay, D.: Land ahoy! understanding submarine command and control during the completion of inshore operations. Hum. Factors **59**, 1263–1288 (2017)
11. Roberts, A.P.J., Stanton, N.A., Fay, D.: Go deeper, go deeper: understanding submarine command and control during the completion of dived tracking operations. Appl. Ergon. **69**, 162–175 (2018)
12. Stanton, N.A., Roberts, A.P.J.: Examining social, information, and task networks in submarine command and control. IEEE Trans. Hum.-Mach. Syst. **48**, 252–265 (2017). https://doi.org/10.1109/THMS.2017.2720659
13. Kirschenbaum, S.S., McInnis, S.L., Correll, K.P.: Contrasting submarine specialty training: sonar and fire control. Dev. Prof. Exp. 271 (2009)
14. Roberts, A.P.J., Stanton, N.A., Fay, D., Pope, K.A.: The effects of team co-location and reduced crewing on team communication characteristics. Appl. Ergon. **81**, 102875 (2019). https://doi.org/10.1016/j.apergo.2019.102875
15. Mack, S.: Perisher: Submarine command training in the Royal Navy. Undersea Warfare (2003)
16. Hamburger, P., Miskimens, D., Truver, S.: It is not just hardware and software, anymore! human systems integration in US submarines. Nav. Eng. J. **123**, 41–50 (2011). https://doi.org/10.1111/j.1559-3584.2009.00198.x
17. Stanton, N.A., Roberts, A.P.J.: Better together? investigating new control room configurations and reduced crew size in submarine command and control. Ergonomics **63**, 1–17 (2019). https://doi.org/10.1080/00140139.2019.1654137

18. Stanton, N.A., Roberts, A.P.J.: Block off: an examination of new control room configurations and reduced crew sizes examining engineered production blocking. Cogn. Technol. Work **22**(1), 29–55 (2019). https://doi.org/10.1007/s10111-019-00542-x
19. Roberts, A.P.J., Stanton, N.A., Fay, D.: The command team experimental test-bed stage 1: design and build of a submarine command room simulator. Procedia Manuf. **3**, 2800–2807 (2015)

Agent-Based Approach to Ship Officer's Navigational Behavior Modeling for Maritime Traffic Analysis

Hongtae Kim[1(✉)], Younghoon Yang[1], and Seung-Kweon Hong[2]

[1] Korea Research Institute of Ships and Ocean Engineering, 171 Jang-Dong, Yuseong, Daejeon, South Korea
hongtae.kim@krsio.re.kr
[2] Department of Industrial and Management Engineering, Korea National University of Transportation, 27469 Chungju-Si, South Korea
skhong@ut.ac.kr

Abstract. Modeling and simulation of a system, such as maritime traffic, with high complexity and a wide range of requirements including ships, navigators, control centers, shipping companies, meteorological systems, and geographic information systems requires human involvement. For maritime traffic simulation in a digital environment, modeling of human factors, a major element, is necessary, and for the reproduction and prediction of realistic maritime traffic situations, it is reasonable to model and reflect the navigator's cognitive processes, behavior patterns, navigational expertise, navigational errors, etc. To develop an intelligent ship agent for maritime traffic analysis, we will analyze the navigator's ship navigation-related cognition and behavior, and introduce the status of the development of an agent-based cognitive and behavioral model as similar to the navigator's behavior.

Keywords: Maritime traffic analysis · Cognitive model · Behavior model · Agent-based simulation · Human element

1 Introduction

The existing maritime traffic simulations, wherein other ships follow the route set at the beginning regardless of the movement of our ship, do not reflect the interaction resulting from the movement of the ships in the real maritime traffic environment.

To solve this problem, we will provide a realistic marine traffic simulation environment based on maritime traffic information, and develop agent-based navigator cognitive and behavioral models for the emergence of new-concept ships such as unmanned and autonomous surface ships, and micro traffic simulation.

In this study, to develop an intelligent ship agent for maritime traffic analysis, we analyze the navigator's cognition and behavior related to ship navigation, and introduce the status of the development of an agent-based cognitive and behavioral model as similar to the navigator's behavior.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 483–489, 2020.
https://doi.org/10.1007/978-3-030-50943-9_61

2 Overview of Related Technologies

Modeling and simulation (M&S) of a system, such as maritime traffic, with high complexity and a wide range of requirements including ships, navigators, control centers, shipping companies, meteorological systems, and geographic information systems requires human involvement. That is, for maritime traffic simulation in a digital environment, modeling of human factors is necessary, and for the reproduction and prediction of realistic maritime traffic situations, it is reasonable to model and reflect the navigator's cognitive processes, behavior patterns, navigational expertise, navigational errors, etc.

A systematic analysis of the cognitive work performed by ship officers in the course of the ship navigation is the basis for various safe navigation methods (e.g., improvement of ship bridge equipment, development of training methods, and identification of potential human errors). In addition, in areas with a multitude of studies on human errors (e.g., aircraft and nuclear power plants), different analysis techniques have been developed to analyze the cognitive work.

To compare and analyze the types of cognitive workload caused by automation of navigational equipment and the use of existing equipment, [1] adopted the operator function model (OFM), and to evaluate the equipment design and training, [2] suggested how to analyze cognitive work concerning offshore work.

Moreover, [3] analyzed the crew's cognitive work on the premise that the "determination of crew size is based on the difficulty of cognitive work performed by the crew," and [4] developed simulations to evaluate risks in the course of ship navigation.

In Korea, [5], in the "Effect of Navigational Expertise on Ship Officer's Situational Awareness," tried to identify the level of situational awareness according to the degree of navigational expertise (type of navigation practice and degree of actual navigational experience) by comparing the level of situational awareness.

Furthermore, [6], and [7] developed a cognitive work analysis framework capable of analyzing the cognitive work occurring on the ship bridge during ship navigation and conducted a navigation task-related case analysis using the foregoing.

On the other hand, [8], using a model capable of assessing a ship's collision risk from the perspective of a ship navigator, developed a risk prediction module that can predict, in advance, collision risk for each traffic situation by adjusting the ship's speed, course, etc.

Studies on agent-based cognitive model have been conducted in earnest in the maritime field, but [9], in the "Development of Aggressive Driving Detection Techniques Reflecting Driver's Driving Patterns by Use of Multi-Agent Driving Simulation," conducted a study on drivers' driving patterns to predict the aggressiveness of car driving in advance.

In the defense field, Computer Generated Force/Semi-Automated Force (CGF/SAF) is attracting attention as the DM&S (Defense Modeling & Simulation) technique proposed to express autonomous forces such as human beings. Currently, CGF/SAF has reached the level of demonstrating simple autonomous behaviors according to established rules, being widely utilized for the military modeling of DM&S [10].

3 Agent-Based Analysis of Navigator's Cognition and Behavior

In this study, we analyze the navigator's cognition and behavior related to ship navigation, and use the information to make an agent-based cognitive and behavioral model similar to the navigator's behavior.

3.1 Navigational Cognition and Behavior Analysis Technology

The navigational cognition analysis technology finds and analyzes information such as the purpose, thinking process, and knowledge carried out in the navigator's brain during navigation. On the other hand, navigational behavior analysis technology finds and analyzes work and environmental factors (workload, stress, alcohol, fatigue, etc.) or individual factors (experience, knowledge, ability, etc.) that affect the behavior via the navigator's cognitive process.

3.2 Agent-Based Navigator Cognition and Behavior Modeling Technology

For the simulation of realistic maritime traffic situations, it is necessary to develop the agent-based navigator cognition and behavior modeling technology, which enables agent-based simulation by modeling the navigator's cognitive processes, behavior patterns, navigational expertise, navigational errors, etc. This technology quantifies the factors affecting the ship officer's cognition and behavior and, thus, transmits the cognitive and behavioral modeling information so that the ship agent can behave similarly to actual ships.

3.3 Analysis of Ship Officer's Behavioral Relationships and Behavioral Errors

Through the analysis of general ship officer tasks, major navigation task data can be examined as follows [11, 12].

- Navigation task: Maintaining course, establishing a navigational plan, and checking ship speed (point of course change, time, speed, RPM)
- Avoidance of navigational hazards: Recognizing the marine environment, identifying navigational hazards, preventing impulse, and avoiding stranding (CPA/TCPA, distance/bearing)
- Maritime communication: GMDSS, MSI, inter-ship communication, VTS communication (reporting location, regulations, etc.)

Moreover, the analysis of the causes of marine accidents over the past five years based on marine accident statistics data of the Korean Maritime Safety Tribunal shows that ship officer's marine accident-related behavioral errors can be categorized into the following four types, and it is necessary to quantify the behavioral errors [13].

– Poor naked-eye monitoring of the surroundings
– Failure to perform proper identification by radar
– Navigation-related violations in passing other ships
– Improper navigation related to ship-handling

By identifying navigation task-related data through the analysis of the ship officer's tasks and analyzing the mechanisms of marine accident occurrence, we identified the work and environmental factors (workload, stress, alcohol, fatigue, etc.) or individual factors (experience, knowledge, ability, etc.) affecting navigational behavior via the navigator's cognitive process.

For the quantification of the foregoing factors affecting ship officer's behavioral errors, it is necessary to set the quantification items and criteria for determining each behavioral error as in Table 1, and this study quantifies the time of behavior and the amount of control through surveys and simulator experiments.

Table 1. Quantification methods, by ship officer's behavioral error

Factors causing navigational errors	Quantification
Look-out	Ship speed
	Distance with target ship
Ship maneuvering	Initial action distance
	Change of course to avoid collision
	Change of speed to avoid collision
	Closest Point of Approach (CPA)

4 Designing and Developing the Ship Officer Agent Model

As in Fig. 1, the ship officer agent interacts with the ship navigation agent and the controller agent to generate and transmit navigational behavior information for the ship navigation agent to behave similarly to actual ships.

The ship officer agent's key function is to receive the input of the ship officer and ship characteristics, decide encounter situations and whether to give way, and transmit the initial action distance (time), change of course, and change of speed to the ship navigation agent according to the existence of navigational errors and control intervention.

That is, the ship officer's final behavior is to control the initial action distance (time), rudder (angle of course change) and speed (engine), and such behavior has a different time of behavior and amount of control depending on the encounter situations, navigational conditions, ship officer and ship characteristics, navigational errors, control intervention, etc. In addition, Fig. 2 is the input/output diagram of the ship officer's agent model.

The ship officer agent receives ship officer and ship characteristic information and navigational error information from the user, and control intervention information from the controller agent. After the reflection of the foregoing to recognize the encounter

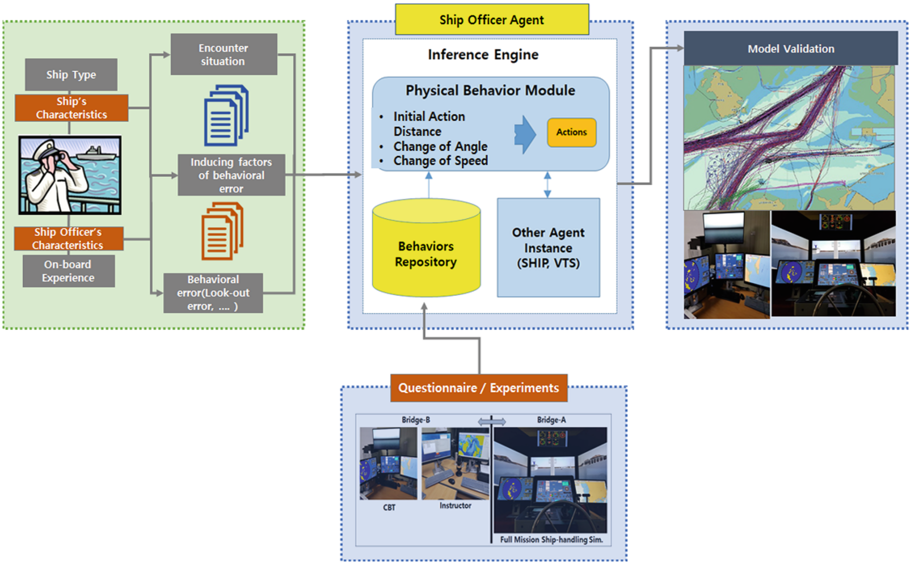


Fig. 1. Diagram of ship officer's agent model

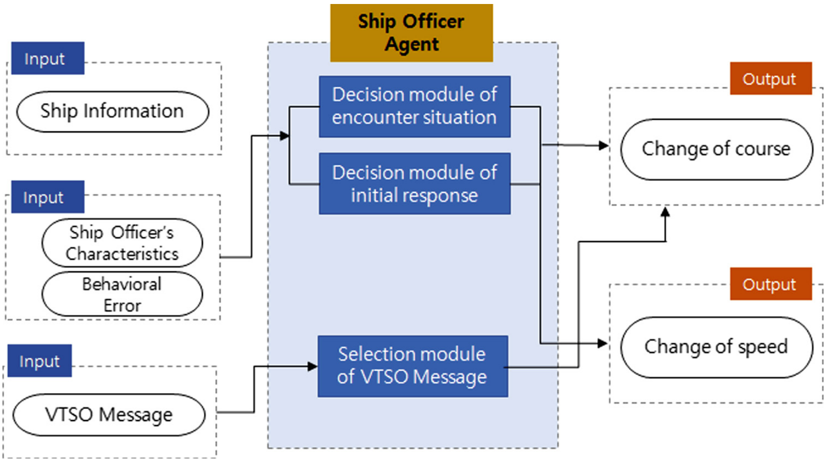


Fig. 2. Input/output diagram of ship officer agent model

situation and decide whether to give way, the point of give-way (distance), change of courses, and change of speed for control of the ship are transmitted to the ship navigation agent via the interlocking agent. As in Fig. 3, the "user" provides the simulation input information to the "interlocking agent" in the use case.

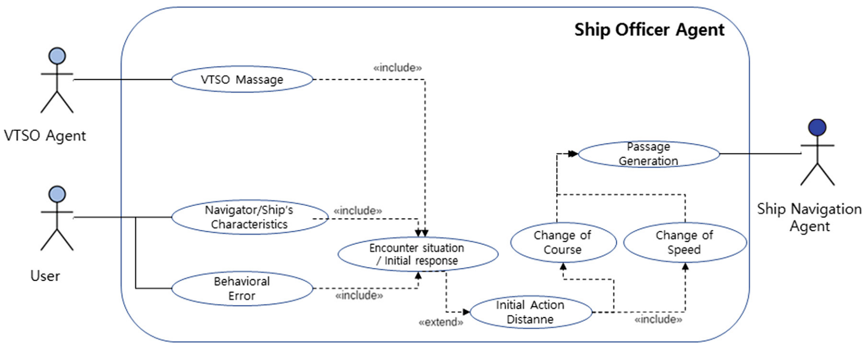


Fig. 3. Use case diagram of ship officer agent

As the ship officer agent, for the verification of ABMS interlock, has to decide the ship encounter situation and whether to give way based on the position information of our ship and other ships, select the change of course/speed, and provide the selected change to the ship navigation agent. It is necessary to verify the adequacy of the ship encounter situation, time of give-way, and change of course/speed as in Table 2.

Table 2. List of verification elements for ship officer agent's interlock

Verification items	Description
Initial action distance (time of give-way)	After deciding circumstances requiring give-way, decide the appropriate time of give-way
Change of course (or command)	Appropriate change of course according to the give-way situation
Change of speed (or command)	Appropriate change of speed according to the give-way situation

5 Conclusions

In this study, to develop an intelligent ship agent for maritime traffic analysis, we analyzed the navigator's cognition and behavior related to ship navigation, and introduced the status of the development of an agent-based cognitive and behavioral model as similar to the navigator's behavior.

In Korea and overseas, 60%–80% of marine accidents are reported to be attributable to human factors, and various approaches for quantitative modeling of the human factors are being studied. An accurate simulation of maritime traffic in a digital environment can contribute to the reduction of the marine accident rate resulting from human factors on the part of the ship and marine-related workers and to the formulation of safety management measures for shipping companies and crew.

In the automotive and defense fields, attempts are being made to build autonomous agent systems out of complex human behaviors such as tactics, strategic knowledge, and driving behaviors, but they have not reached the stage of field application because of the lack of a sophisticated behavior model and linkage among multiple platforms. Although autonomous surface ships and smart marine logistics are currently being discussed in earnest, marine traffic assessment studies reflecting the risk behavior factors of ship navigators are at an early stage locally and abroad.

The results of this study will be used for the development of a maritime traffic reproduction and prediction simulation platform, in conjunction with the ship agent and controller agent.

Acknowledgments. This work was funded by grants from Korea Research Institute of Ships and Ocean Engineering (Grant – PES3600).

References

1. Lee, J.D., Sanquist, T.F.: Augmenting the operator function model with cognitive operations: assessing the cognitive demands of technological innovation in ship navigation. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **30**(3), 273–285 (2000)
2. Sanquist, T.F., Lee, J.D., Rothblum, A.M.: Cognitive analysis of navigation tasks: a tool for training assessment and equipment design. CG-D-19-94, U.S. Coast Guard Research and Development Center (1994)
3. Lee, J.D., McCallum, M.C., Maloney, A.L., Jamieson, G.A.: Validation and sensitivity analysis of a crew size evaluation method. ADA331589, Battle Seattle Research Center (1997)
4. Itoh, K., Yamaguchil, T., Hansen, J.P., Nielsen, F.R.: Risk analysis of ship navigation by use of cognitive simulation cognition. Technol. Work **3**, 4–21 (2001)
5. Kim, B.A., et al.: The effect of navigation officer's expertise on situation awareness. Korea J. Ind. Organ. Psychol. **20**(4), 497–509 (2007)
6. Kim, H.T., Hong, S.K.: Situation awareness and maritime accident. J. Maritime Saf. Aut. (2013)
7. Hong, S.K.: Development of a cognitive work analysis method for safe ship navigation and a case study. J. Navig. Port Res. **37**(3), 231–237 (2013)
8. Park, Y.S., Park, S.W., Cho, I.S.: A basic study on prediction module development of collision risk based on ship's operator's consciousness. J. Navig. Port Res. **39**(3), 199–207 (2015)
9. Kim, Y.J., et al.: Development of a methodology for detecting intentional aggressive driving events using multi-agent driving simulations. J. Korean Soc. Transp. **36**(1), 51–65 (2018)
10. Shin, S.H.: Agent-based SAF modeling tool for DEVS M&S. J. Korean Soc. Simul. **22**(4), 49–55 (2013)
11. ICS. Bridge Procedures Guide, 5th Edition (2016)
12. IMO. Resolutions A.285 (VIII) Recommendation on Basic Principles and Operational Guidance Relating to Navigational Watchkeeping (2016)
13. MOF. Statistics of Maritime Accidents, Korean Maritime Safety Tribunal (2020)

Exploiting Contemporary Technology in Flight Deck Design to Improve Flight Safety

Envisioning Mixed Realities on the Flight Deck

James Blundell(✉), John Huddlestone, Charlotte Collins, Steve Scott, Rodney Sears, and Anastasios Plioutsias

Coventry University, Coventry, UK
{james.blundell, john.huddlestone, charlotte.collins, steve.scott, rodney.sears, ad3903}@coventry.ac.uk

Abstract. Conformal 3D symbology presented on a head tracked head-mounted display (HMD) has the capability to enhanced pilot situation awareness, performance and workload by providing an unlimited field of view of operational hazards. In recent years, a body of research has emerged highlighting the technical advancements (i.e. HMD encumbrance and optical enhancements) that could enable the unique capabilities and benefits of HMDs to be realized on commercial and business flight decks in the near future. The current paper provides a systematic review of the HMD-related pilot benefits (performance, workload, situational awareness, and usability) that have been reported in the literature. We conclude by highlighting the operational contexts where HMDs might enhance pilot performance, flight safety and efficiencies.

Keywords: Head-Mounted displays · Conformal symbology · Pilot performance

1 Introduction

The potential of three-dimensional (3D) displays has attracted considerable interest from the aviation community as a means of further enhancing the intuitive nature in which information is presented to the pilot. Extensive research has documented the situational awareness (SA) benefits afforded by 3D displays. For example, compared to a two-dimensional (2D) orthogonal representations of flight information (e.g. altitude and geographical position on a navigation display (ND)), a 3D representation is able to depict an intuitively understandable spatio-temporal representation of the aircraft's current situation [1]. This is particularly the case for forward perspective 3D displays with the capability of generating virtual "conformal" symbology that can be accurately mapped to geographically locations within the forward perspective scene [2, 3]. Conformal symbology can be "truly conformal", where the imagery directly overlays real objects that exist within the outside scene (e.g. the horizon line or a runway overlay as seen on a head-up display (HUD)). Or the symbology can be "virtually conformal", where the imagery can represent an entity in space that has no real physical properties (e.g. the pilot's perspective flight-path [1]). The combination of the above symbology conformity types presented onto an outside scene, in a sense, generates a mixed-reality flying environment.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 493–499, 2020.
https://doi.org/10.1007/978-3-030-50943-9_62

A large body of evidence exists showing that conformal symbology, presented on a HUD or on a perspective, synthetic vision (SV) head down display, improves flight path tracking and the detection of changes in symbology or traffic (see Fadden, Ververs and Wickens [2] for a meta-analysis). A HUD is a glass-mounted panel fixed within the pilot's near visual field. It allows the pilot to remain "head-up and eyes-out" as near-domain 2D flight information can be overlaid against far-domain 3D information belonging to the external scene [4, 5]. A SV display, on the other hand, presents the pilot with a 3D graphical rendering of a synthetic out-the-window perspective view. In the cockpit, a SV display can be located at either head-up or head-down. Whilst both HUD and SV displays can present information simultaneously in 2D and 3D formats, the coupling of airspace hazard information is more difficult due to the relatively narrow field of view available with the forward perspective view inherent in these displays [6].

Head-mounted displays (HMD), that can be coupled with low latency head-trackers, provide a viable solution to the previously mentioned field of view limitation by granting an unlimited field of regard [7, 8]. Indeed, in the business domain Thales plans to introduce an upgraded variant of their TopMax HMD that has the capability to present both 2D traditional flight references (e.g. airspeed, altitude and power) and 3D conformal imagery. A recent review by Arthur et al. [9] summarized the past 30 years of National Aeronautics and Space Administration (NASA) research of optimizing the collimated optics, head tracking, latency, and weight of HMDs. In terms of operational capabilities, they highlight the potential benefits of a HMD as a key enabler of multiple future air traffic concepts, for example, supporting the safety of simultaneous parallel runways operations where off-boresight traffic monitoring will be important [10]. However, it was concluded that in order for HMDs to see wide adoption in current business and commercial operational contexts HMDs will need to demonstrate equivalence, in terms of both performance and safety, to HUDs.

In this paper, we review literature from the aviation literature from the past 20 years to highlight the performance, workload (WL), SA and usability benefits associated with HMDs. A number of reviews exist describing the perceptual and human factors issues associated with HUDs [2, 4] and 3Ds displays more broadly [11, 12]. In regards to HMDs specifically, in the past 5 years there have been several reviews published by NASA [9, 13] and the German Aerospace Centre (DLR) [8] that detail the findings from their respective research programs on the application of HMDs in commercial and business aircraft and rotorcraft. However, these reviews predominantly adopt a technical focus (i.e. describe HMD encumbrance and optical enhancements) with minimal human factors related discourse. Hence, the purpose of the current paper is to provide a review of the existing empirical HMD evidence with a greater focus on highlighting the associated human factors issues. Furthermore, we try to examine the potential of "HUD equivalence" by only reviewing papers that compare HMD symbology with other symbology types. Be that symbology on the head-down or, preferably, head-up location. The review is structured according to two phases of flight to describe the value of HMDs within different aviation contexts. These include 1) approach/landing and 2) taxiing/ground operations.

2 Methods

Inclusion criteria for experiments required that experiments report at least one measure of performance (e.g. path deviation), WL (e.g. subjective scales, physiology measure), SA, or usability. Experiments were required to contrast HMD symbology to other symbology types (e.g. HUD and/or head down display). Each experiment was classified into one of two different types depending upon the phase of flight the HMD was being implemented in: 1) approaching/landing and 2) taxiing/ground operations.

3 Results

Eight experiments from the aviation human factors literature were identified. Two experiments were described in a single study by Arthur et al. (2014) [14]. Outcomes of the eight experiments were classified as: 1), HMD outcomes were found to be better than alternative symbology formats (e.g. HUD or head down ILS), 2); outcomes between the symbology types was equivalent, or 3); a HMD was found to be inferior to an alternative symbology formats. A summary of the literature review results is presented in Table 1.

Table 1 Results of the literature review for comparing HMD symbology to alternative symbology formats on performance, workload (WL), situational awareness (SA) and usability on approach/landing and taxiing tasks. Asterisks (*) are used to mark findings that were statistically verified.

Author	Date	Craft	Compared	Sample (*N*)	HMD outcomes			
Approach/Landing					**Performance**	**WL**	**SA**	**Usability**
Lorenz, Helmut & Schmerwitz [15]	2005	Fixed Wing	HDD	18	HMD Better*	Equivalent*	HMD Worse*	Equivalent
Arthur et al. [14]	2014	Fixed Wing	HUD	12	Equivalent*	Equivalent*	Equivalent*	Equivalent
Doehler et al. [19]	2015	Rotary	HDD	6	HMD Better	-	-	HMD Better
Schmerwitz et al. [20]	2015	Rotary	HDD	18	HMD Better	Equivalent	-	HMD Better*
Doehler et al. [21]	2012	Rotary	HDD	12	Equivalent	Equivalent	Equivalent	HMD Better*
Taxi/Ground Ops								
Bailey et al. [22]	2007	Fixed Wing	HUD HDD	8	-	Equivalent*	Equivalent*	HMD Better
Arthur et al [23]	2009	Fixed Wing	HUD HDD	8	Equivalent*	Equivalent*	Equivalent*	-
Arthur et al. [14]	2014	Fixed Wing	HUD HDD	12	Equivalent*	Equivalent*	Equivalent*	Equivalent

3.1 Approach and Landing

The literature review identified five human-in-the-loop experiments examining the benefits of a HMD device during approach and landing. A single study, Arthur et al. [14], directly compared a HMD to a HUD, in a scenario that involved a straight approach, landing, and taxi with a fixed-winged commercial aircraft. The two HMD symbology variants included a version where the PFD symbology was fixed to the pilot's gaze direction and a version that was fixed to the HUD combiner glass location (creating an "artificial HUD"). Pilot performance, workload and situational awareness was equivalent across the three symbology types. One possibility for this outcome is that the complexity of a straight approach task did not lend itself to the larger field of regard benefits that are afforded by a HMD. This is supported by an earlier study by Lorenz et al. [15] where tracking performance during a curved approach was enhanced by presenting a perspective flight-path on a HMD, compared to a head-down display. The advantage of a HMD for following curved trajectories would be expected based on the work by Mulder [16]; The reduced field of regard provided by a fixed perspective display (i.e. HUD or head down SV) can make it more difficult to follow curved trajectories. Taking these results together, a more detailed understanding of HMD benefits could be achieved in future studies examining different HMD symbology variants (e.g. the "artificial HUD" from Arthur et al. [14]) in operational contexts that require greater off-boresight monitoring. Interestingly, in the same study by Lorenz, SA with the HMD was worse compared to following the ILS on the head-down display, whereby detection of an unexpected event (detecting a runway incursion on approach) occurred significantly later with the HMD. This is an example of attentional capture, which is a common findings of studies investigating HUD presented perspective flight-path displays [5, 17]. However, it should be noted that the sample used in the study consisted entirely of trainee pilots with no HUD experience, potentially making them more susceptible to attentional capture effects [5].

The remaining three studies all concern a comparison between a HMD and head down display. These studies are based upon DLR research efforts to develop a 3D conformal helicopter landing symbology set that can be presented either on a HMD or head-down display [18]. Two of the studies, Doehler et al. [19] and Schmerwitz et al. [20], demonstrated that pilot landing performance was enhanced by presenting conformal 3D symbology representing the target landing zone on a HMD. Specifically, lateral drift near to touchdown was reduced when using the HMD. This is an expected benefit of a HMD, as the landing pad remains viewable on a HMD display at all times, in contrast to the head-down alternative that requires switching of attention between the head-down ND and being eyes-out. These enhanced performance findings were complimented by pilots higher usability ratings of the HMDs. Unfortunately, neither study reported the statistical results for these performance findings. Whilst this is understandable with the smaller sample size of six pilots included in Doehler's experiment [19], there existed no reported statistical analysis of the performance results in the experiment by Schmerwitz [20]. In contrast, statistical results were reported for the workload and HMD usability findings in the same study. The remaining study by Doehler [21] presented a comparison between a binocular HMD, a monocular HMD and a head down primary flight display (PFD) during a low-visibility landing task.

While pilots reported significantly higher acceptance for the binocular HMD, there were no significant workload or SA differences between symbology types. Possibly due to the employed landing task not being demanding enough. It was noted, however, that the lack of any significant performance findings could have been attributed to the insufficient handling qualities of the study's simulator platform. Unfortunately, similar to the previous two studies, no statistical analysis was presented for the pilots landing performance findings.

3.2 Taxiing/Ground Operations

Three experiments were identified that compared the benefits of a HMD, a HUD and a head-down display during taxiing operations. One of these experiments was conducted alongside one of the experiments described above (Arthur et al. [22]). Similar to the comparison between a HMD, HUD and a head-down display during approach, no significant difference was found in performance (i.e. centerline tracking), WL, SA and usability whilst taxing. The remaining two studies revealed that a SV HMD could enhance safety and improve ground operation efficiencies in a future air traffic environment across a range of visibility conditions. In both studies, 3D conformal symbology was presented on either a SV HMD or HUD to depict taxi routing and traffic information. Pilots were better able to perform the taxi task and reported significantly higher SA with the SV HWD concept compared to a head-down moving map or paper charts. However, whilst there was a preference shown by pilots for the HMD, there were no observed performance benefits of the HMD over the HUD.

4 Discussion

The purpose of the current paper was to review the empirical evidence that HMDs could enhance the safety and efficiency of flight operations in future or current air traffic environments. In terms of performance, firstly, HMD related performance enhancements were consistently observed within flying scenarios where off-boresight monitoring was required. For example, scenarios that required tracking a landing zone or curved trajectory during an approach in a helicopter [19, 20] of fixed-wing aircraft [15], respectively. Secondly, HMD performance was equivalent to alternative symbology types during scenarios where the majority of task-related information was located within the boresight location (e.g. flying a straight approach [14] or following a taxi route [23]). Together, the HMD performance findings across studies are overall positive, particularly in the absence of any reported HMD related performance decrements. This could be attributed to the technical advancements seen in HMDs over the past two decades (e.g. encumbrance and optical enhancements) which could enable HMDs to see wide adoption on the business and commercial flight deck in the near future.

The overall statistical reporting quality of experiments included in the current review was lacking. For example, of the three studies that reported a HMD related performance advantage only one validated the finding with a statistical analysis. Furthermore, where experiments did provide a statistical analysis, not one included the associated effect size. While it might be expected that the effect size of an analysis comparing different display formats (e.g. tracking performance with a HMD versus and head-down display) at this stage of the technology's maturity will yield large effect sizes, the evaluation of more mature HMD concepts will undoubtedly depend upon detecting effect sizes that are far smaller. An example, determining the optimal set size and organization of HMD symbology in a way that does not risk operational safety by inducing attentional tunneling. Technically, this could be difficult to achieve with the mid-level fidelity requirements (e.g. collimated displays) that were used in the majority of experiments reviewed in this paper. A potential solution to this obstacle could be the implementation virtual reality (VR) as a lower fidelity platform to support the evaluation of future HMD 3D conformal symbology concept prototypes.

The studies reviewed in the current paper present a promising foundation of HMD human factors research. Future research will require a greater emphasis on the evaluation of how 3D conformal symbology presented on a HMD can support pilot performance, and enhance safety and efficiency of across a greater range of flight phases. In particular, minimum symbol sets should be derived for each phase of flight and any additional information presented should be justified against well-defined requirements.

Acknowledgements. This article is based on work performed in the programme: Open Flight Deck, which has received funding from Innovate UK programme, the UK's Innovation Agency, with support from the UK Aerospace Technology Institute (grant reference number 113108). The views and opinions expressed in this paper are those of the authors and are not intended to represent the position or opinions of the Open Flight Deck consortium or any of the individual partner organisations.

References

1. Mulder, J.A., Van Paassen, M.M., Mulder, M.: Perspective guidance displays show the way ahead. J. Aerosp. Comput. Inf. Commun. **1**(11), 428–431 (2004). https://doi.org/10.2514/1.14069
2. Fadden, S., Ververs, P.M., Wickens, C.D.: Costs and benefits of head-up display use: a meta-analytic approach. Proc. Hum. Factors Ergon. Soc. **1**, 16–20 (1998). https://doi.org/10.1177/154193129804200105
3. Snow, M.P., French, G.A.: Effects of primary flight symbology on workload and situation awareness in a head-up synthetic vision display. In: AIAA/IEEE Digital Avionics System Conference - Proceedings, vol. 2, pp. 1–10 (2002). https://doi.org/10.1109/dasc.2002.1052981
4. Fadden, S., Ververs, P.M., Wickens, C.D.: Pathway HUDs: are they viable? Hum. Factors **43**(2), 173–193 (2001). https://doi.org/10.1518/001872001775900841
5. Wickens, C.D., Alexander, A.L.: Attentional tunneling and task management in synthetic vision displays. Int. J. Aviat. Psychol. **19**(2), 182–199 (2009). https://doi.org/10.1080/10508410902766549
6. Wickens, C.D., Ward, J.: Cockpit displays of traffic and weather information: effects of 3D perspective versus 2D coplanar rendering and database integration. Int. J. Aerosp. Psychol. **27**(1–2), 44–56 (2017)
7. Yeh, M., Wickens, C.D., Seagull, F.J.: Effects of frame of reference attentional issues with helmet mounted displays (1998)

8. Peinecke, N., Schmerwitz, S., Döhler, H.-U., Lüken, T.: Review of conformal displays: more than a highway in the sky. Opt. Eng. **56**(5), 051406 (2017). https://doi.org/10.1117/1.oe.56.5.051406
9. Arthur, J.T.J., et al.: Review of head-worn displays for the next generation air transportation system. Opt. Eng. **56**(5), 051405 (2017)
10. Kramer, L.J., Bailey, R.E., Ellis, K.K.E., Norman, R.M., Williams, S.P.: Enhanced and synthetic vision for terminal maneuvering area nextgen operations. In: Proceedings of SPIE (2011). https://doi.org/10.1117/12.883036
11. R. J. Stone and F. J. Knight, "3D displays: a human-centred review. Unclassif. Bae Syst. Propr., p. 73 (2012)
12. Mcintire, J., Mcintire, J.P., Havig, P.R., Geiselman, E.E.: Stereoscopic 3D displays and human performance: a comprehensive review Stereoscopic 3D displays and human performance: a comprehensive review. Displays **35**(1), 18–26 (2014). https://doi.org/10.1016/j.displa.2013.10.004
13. Arthur, J.T.J., et al.: A review of head-worn display research at NASA Langley. In: Proceedings of SPIE, vol. 9470, pp. 1–15 (2015). https://doi.org/10.1117/12.2180436
14. Arthur, J.T.J., et al.: Performance comparison between a head-worn display system and a head-up display for low visibility commercial operations. In: Proceedings of SPIE (2014)
15. Lorenz, B., Tobben, H., Schmerwitz, S.: Human performance evaluation of a pathway HMD. In: Proceedings of SPIE, vol. 5802, pp. 166–176 (2005). https://doi.org/10.1117/12.603280
16. Mulder, M.: An information-centered analysis of the tunnel-in-the-sky display part two: curved tunnel trajectories. Int. J. Aviat. Psychol. **13**(2), 131–151 (2003). https://doi.org/10.1207/s15327108ijap1302
17. Prinzel, L., Risser, M.: Head-Up Displays and Attention Capture, vol. 213000 (2004)
18. Schmerwitz, S., Lueken, T., Doehler, H.-U., Peinecke, N., Ernst, J.M., da Silva Rosa, D.L.: Conformal displays: human factor analysis of innovative landing aids. Opt. Eng. **56**(5), 051407 (2017). https://doi.org/10.1117/1.oe.56.5.051407
19. Doehler, H., Schmerwitz, S., Lueken, T.: Visual-conformal display format for helicopter guidance (2015). https://doi.org/10.1117/12.2052677
20. Schmerwitz, S., Knabl, P.M., Lueken, T., Doehler, H.-U.: Drift indication for helicopter approach and landing. In: Proceedings of SPIE, vol. 9471, pp. 1–10 (2015). https://doi.org/10.1117/12.2177816
21. Doehler, H., Knabl, P.M., Schmerwitz, S., Eger, T., Klein, O.: Evaluation of DVE landing display formats. In: Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications IX, vol. 8360, pp. 1–12 (2012). https://doi.org/10.1117/12.922459
22. Bailey, R.E., Arthur III, J.J., Prinzel III, L., Kramer, L.J.: Evaluation of head-worn display concepts for commercial aircraft taxi operations. In: Head- Helmet-Mounted Displays XII Design Applications, vol. 6557, p. 65570Y (2007). https://doi.org/10.1117/12.717221
23. Arthur, J.T.J., Prinzel, L.J., et al.: Synthetic vision enhanced surface operations with head-worn display for commercial aircraft. Int. J. Aviat. **19**(2), 158–181 (2009)

Supporting Astronaut Autonomous Operations in Future Deep Space Missions

M. Natalia Russi-Vigoya[1](✉), Donna Dempsey[2], Brandin Munson[3], Alonso Vera[4], Bernard Adelstein[4], Shu-Chieh Wu[5], and Kritina Holden[6]

[1] KBR, Houston, USA
MariaNatalia.Russi-Vigoya@nasa.gov
[2] NASA- Johnson Space Center, Houston, USA
[3] University of Houston, Houston, USA
[4] NASA- Ames Research Center, Moffett Field, USA
[5] San Jose State University, San Jose, USA
[6] Leidos, Houston, USA

Abstract. Future deep space missions will present new physical and cognitive challenges that could increase risks to astronaut performance. In addition, crews may no longer be able to depend on timely support from NASA Mission Control Center (MCC) due to distance from the Earth. Astronauts will have to work autonomously using onboard resources, while still maintaining high performance. It is critical to understand the type of support MCC currently provides in order to replicate that type of support onboard. A Contextual Inquiry was performed to characterize the support that MCC currently offers to space crews. An expert Focus Group was convened to understand the types of tasks astronauts will do on a future Mars mission, as well as the challenges associated with those tasks. Results from these activities elucidate the types of intelligent tools and capabilities that will be needed for autonomous crews on future deep space missions.

Keywords: Human Factors · Human-Systems Integration · Contextual Inquiry · Space Exploration · Autonomous Missions

1 Introduction

Future long-duration space missions beyond Low Earth Orbit (LEO) will require fundamental changes in crewed spaceflight operations. Today, there is heavy dependence on ground support for completing onboard tasks and for ensuring overall mission success. Time delays and periodic communication blackouts associated with deep space missions will necessitate that certain activities currently performed by Mission Control Center (MCC) personnel will need to be performed partially or entirely onboard. Instead of relying on real-time interaction with a support team on the ground, the crew will be required to operate more autonomously. Developing and planning for these new operations concepts will require the use of intelligent systems. The specific capabilities required of onboard intelligent systems will be driven specifically by the

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 500–506, 2020.
https://doi.org/10.1007/978-3-030-50943-9_63

planned deep space operational concepts. MCC currently supports both the execution of nominal operations on, and problem solving for, the International Space Station (ISS). While MCC will still provide these resources to some extent for deep space missions, it is clear that crew will not continue to have the present level and kind of real-time support (especially for time-critical activities). Moreover, crew will need to operate more independently, especially during periods of communication loss.

In order to begin addressing this challenge, the full scope of capabilities that MCC currently provides to ensure mission success must be characterized and understood. The key tasks, roles, and issues faced by MCC must be identified, as well as something about the individual MCC personnel who solve these problems. With that understanding, one can begin to consider what tasks the crew may have to perform autonomously, what types of intelligent tools will be needed onboard to accomplish those tasks without MCC assistance, and what standards and guidelines need to be developed to support the development of those intelligent tools and capabilities. Problem-solving tasks are of particular concern, since one cannot predict with certainty what problems will arise on these unprecedented missions.

The aim of the present study is to synthesize human factors methods to begin assessing current capabilities and future needs to support the types of tasks astronauts may have to perform autonomously in the future. The output of the study includes findings and lessons learned from methods such as Contextual Inquiry, Focus Groups, Technical Interchange Meetings (TIMs), interviews, and extensive reviews of standards and guidelines. Results will be used to help identify the types of intelligent tools and capabilities that will be needed to aid future autonomous crews in independently solving emergent or unanticipated problems that have potential “Loss of Crew” or “Loss of Mission” outcomes. This study was performed as part of a larger line of research at NASA, called *Human Capabilities Assessment for Autonomous Missions (HCAAM)*.

2 Methodology

A number of activities have been completed as part of the HCAAM effort. Work began with a gap analysis of existing NASA standards for human spaceflight, and review of requirements, standards, and guidelines from NASA and non-NASA sources related to performance support (e.g., intelligent systems, virtual assistants, and augmented reality). An Autonomous Crew Operations Technical Interchange Meeting (TIM) was held with NASA and non-NASA experts to discuss challenges and new technologies that may help support astronauts in operating autonomously from MCC. Contextual Inquiry interviews and site visits at MCC were completed to understand the support MCC provides current crews. Finally, an expert Focus Group was held to discuss crew tasks and challenges on a future deep space mission.

2.1 Review of Standards and Guidelines

To identify new candidate standards and guidelines for future deep space missions, the following activities were performed [1]:

- A gap analysis of "NASA-STD-3001 Volume 2: Human Factors, Habitability, and Environmental Health", which contributed to the development of a gap keyword list;
- Identification of pertinent industry and other U.S. Government (i.e., Department of Defense, Department of Energy, Federal Aviation Administration) standards and guidelines documents;
- Keyword-based search (using the gap keyword list) of the identified government and industry standards documents; and
- Consensus meetings to select final candidate standards and guidelines for inclusion in NASA standards.

Selection criteria were clearly defined at the outset of the keyword-based search. While some initial searching was done manually, automated keyword searching was also accomplished using the *Atlas.ti* (Version 8.3.16) software [2]. Rationale for the adoption or rejection of individual standards and guidelines, and associated notes were recorded. A modified Nominal Group Technique [3] was employed to obtain consensus for each final standard adoption or rejection.

2.2 Technical Interchange Meeting (TIM)

The Autonomous Crew Operations TIM took place at the NASA Ames Research Center [4]. The goal of the meeting was to gather input from five NASA centers, industry, academia, and branches of the Department of Defense (DoD) to address how intelligent technologies can be applied to support crew anomaly response. The TIM hosted 59 attendees from these various institutions, and featured 24 presentations. Attendees shared thoughts about autonomy gaps, and provided potential solution ideas on a collaborative wallboard. The wallboard kept engagement high and promoted significant discussion.

2.3 MCC Contextual Inquiry

The aim of the Contextual Inquiry [5] was to characterize the types of activities currently performed by NASA International Space Station (ISS) Flight Controllers to identify capabilities that would need to be shifted onboard in order to enable crew to resolve emergent, unanticipated problems autonomously [6]. The Contextual Inquiry observations and interviews were performed at the ISS MCC and the Space Station Training Facility (SSTF) at NASA Johnson Space Center (JSC). Researchers worked with Chief Training Officers (CTOs), instructor leads, and instructors from the following seven disciplines: Robotics (ROBO); Environmental and Thermal Operating Systems (ETHOS); Biomedical Engineering (BME); Operations Support Office (OSO); Extravehicular Activity Office (EVA); Spacecraft Communications (CAPCOM); and Station Power, Articulation, Thermal, and Analysis (SPARTAN). Researchers also interviewed flight controller instructors from two other disciplines, Behavioral Health and Performance (BHP) and the Visiting Vehicle Office (VVO). In addition, the researchers received input from a flight director, two Mission Evaluation Room (MER) managers, and representatives from the JSC Flight Safety Office. Through observations and interviews, the research team collected information to characterize the roles, tools, data flows, teaming, and training elements for how tasks are actually performed today. A list of guiding questions was developed to collect information during the observations and interviews.

The Contextual Inquiry in MCC was challenging due to the specialized domain knowledge, many acronyms, and communication system on which many flight controllers talk simultaneously on the same and on different communication loops. To address some of these challenges, the team interviewed flight controllers before the observations (when possible). This provided a foundation of knowledge for the research team, an environment where basic questions could be asked easily, and a preview of the upcoming simulation and faults to be observed. This was invaluable in preparing for an observation [7].

2.4 Expert Focus Group

In addition to the flight control-related activities identified in the Contextual Inquiry, there are many other work and daily living tasks that astronauts will have to perform on a deep space mission (e.g., going to Mars). The goal of the expert Focus Group was to develop a validated high-level list of tasks that crew would need to be able to perform autonomously in the future in order to be successful on future long-duration missions [8]. The results of this effort will be critically important for informing and focusing research and development teams, because the Focus Group discussions were structured to characterize challenges that future autonomous crew will face, and highlight solutions that need to be developed.

Focus Group participants were experts in their respective fields, selected to represent the following domains/perspectives: crew, flight/mission control, mission planning, medicine, behavioral health, and training. Each of the Focus Group participants had more than twenty years of experience at NASA in spaceflight training, operations, or research. The Focus Group was asked to discuss their most serious concerns related to autonomous operations. They were also asked to gain consensus on and validate an autonomous task list, with priority on problem-solving tasks that will be particularly challenging without support from MCC. The experts were asked to brainstorm and discuss their autonomous task concerns before seeing the draft task list. This primed the experts to begin thinking deeply about the topic and provided an opportunity to determine whether their thoughts and concerns about tasks aligned with the draft materials presented. Several scenarios were also drafted to describe autonomous operations.

3 Results

3.1 Review of Standards and Guidelines

The gap analysis identified twenty areas in NASA human-system integration standards and guidelines documents that will need to be addressed in order to support future autonomous crew operations for Gateway missions. The gaps included topic areas such as adaptive systems, automation and robotics, autonomy, cognitive aids, decision support, and augmented reality.

To help fill these gaps, 59 standards and 427 guidelines were identified from the review of 164 documents sourced from existing U.S. Government and industry standards documents. The identified standards and guidelines have been made available to multiple NASA program and document development teams, as well as to HCAAM researchers who are planning to conduct studies in the identified gap areas. This information will inform future need areas and ensure that research aims are targeted to the highest need areas [1].

3.2 Technical Interchange Meeting (TIM)

To address how intelligent technologies can be applied to augment onboard capabilities to support crew anomaly response, the TIM gathered over 110 submissions of ideas on the wallboard. Results showed that there are several areas of need that are insufficiently covered by existing technologies, such as the need for knowledge management throughout the steps of problem solving, particularly resolution documentation, and work force management. The TIM showed that the most pressing need for autonomous crew operations is support for the troubleshooting of rare and unanticipated onboard anomalies that might potentially threaten crew and vehicle safety. For example, in the early days of ISS, high priority anomalies required extended investigation. Rare events can have big consequences. They can also be difficult to train, and the most difficult for intelligent systems to handle [4].

3.3 MCC Contextual Inquiry

The Contextual Inquiry in MCC [6] served to characterize the support that MCC currently offers to space crews. A high-level comparison of what was learned about flight controllers versus astronauts is shown in Table 1 below.

Table 1. A comparison of problem-solving enablers

Flight controllers	Astronauts
Flight controllers are highly trained in their specialty area	Astronauts currently receive limited training in many specialty areas
Flight controllers have many detailed displays, trend graphs, calculation tools, and reference databases	Astronauts have a few high-level displays, and limited tools/reference databases onboard
Flight controllers have a problem-solving culture that has been trained, i.e.: - What is the next worst failure? - Verbal review of the situation - Sanity checking/confirming/vetting with multiple levels of experts	Astronauts have less training in troubleshooting and problem solving

Based on the observations and interviews, researchers identified the following key requirements for future deep space missions:

a. *Advanced training systems* to provide training on core vehicle systems, as well as generic troubleshooting and problem solving. It should provide refresher training, Just-in-Time training, and stress-inducing simulation training.
b. *Intelligent information systems* to provide core system engineering schematics, design documents, job aids (e.g., spreadsheets, equations, calculations, event critical times/durations, tracking grids, "cheat sheets,") and electronic procedures needed to perform, diagnose, maintain, repair, replace, and create new core system components.
c. *Problem solving systems* to provide intelligent assistance for fault management, troubleshooting, problem solving, and decision-making. Systems should include flight rules for communication (including time-delayed communication); mindset (i.e., "what-ifs," & "next worse failure" thinking); consideration of integration impacts; handover of a problem; strategies to "divide and conquer;" strategies to maintain situation awareness; and strategies to document events and outcomes for future use.
d. *Design for autonomy,* that is simplify system designs to reduce the number of components that can fail, and promote ease of training, ease of use, and ease of repair.

3.4 Expert Focus Group

Focus Group discussions about the autonomous task list revolved primarily around abilities to respond to medical or behavioral health events, abilities to respond to unanticipated major vehicle or habitat malfunctions, piloting and navigation tasks, and installation/activation/inspection of vehicle/habitat systems. The participants validated the task list (with minor revisions) and raised the following three significant concerns: (1) lack of preparation and experience for autonomous crew operation; (2) insufficient intelligent system development and testing; and (3) inadequate training approaches for long-duration missions. The autonomous task list and output from the Focus Group will help researchers, engineers, and others understand the tasks most likely to be influenced by missions requiring greater levels of autonomy [8].

4 Discussion and Conclusions

Future deep space missions will present new challenges for astronauts, and increased risks to human performance due to the stress, fatigue, radiation exposure, and isolation that characterize these missions. In addition, crews may no longer be able to depend on timely support from MCC due to distance from the Earth, but will have to work autonomously, while maintaining high performance. This greater crew autonomy will increase dependence on automated systems. Problem-solving tasks are of particular concern, since one cannot predict with certainty, which problems will arise on these unprecedented missions. Consequently, crews will not be able to be trained on all potential anomalies. Communication delays will exacerbate these challenges, especially when unanticipated anomalies needing prompt resolution occur.

The research activities that have been described here represent the first steps in beginning to understand and address these future challenges. Results indicate that the key capabilities needed for future missions include advanced training systems, intelligent information systems, problem-solving aids, and autonomy-focused design. It will be important to focus near-term efforts on these areas in order to be ready to support crew on future deep space missions.

Acknowledgements. This work was funded by the NASA Human Research Program (contract number NNJ15HK11B), and the study was performed at the NASA Johnson Space Center. The authors would like to thank E. Vincent Cross II and Brent Beutter for their technical contributions. The authors would also like to thank the NASA Flight Operations Directorate (FOD) for their assistance in conducting the Contextual Inquiry, and all of the participants of the Contextual Inquiry and Expert Focus Group.

References

1. Holden, K., Russi-Vigoya, M.N., Adelstein, B., Munson, B.: Human Capabilities Assessment for Autonomous Missions (HCAAM). Phase I: Human Performance Standards and Guidelines. In: Internal NASA Report prepared for the Human Factors and Behavioral Performance Element of the Human Research Program (2019)
2. Muhr, T.: ATLAS/ti—a prototype for the support of text interpretation. Qual. Sociol. **14**(4), 349–371 (1991)
3. Fink, A., Kosecoff, J., Chassin, M., Brook, R.H.: Consensus methods: characteristics and guidelines for use. Am. J. Public Health **74**(9), 979–983 (1984). https://ajph.aphapublications.org/doi/pdf/10.2105/AJPH.74.9.979
4. Wu, S.C., Vera, A.H.: Supporting crew autonomy in deep space exploration: preliminary onboard capability requirements and proposed research questions. (Report No.: NASA/TM—2019–22034) (2019). https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20190032086.pdf
5. Whiteside, J., Bennett, J., Holtzblatt, H.: Usability engineering: our experience and evaluation. In: Helander, M. (ed.) Handbook of Human-Computer Interaction, pp. 791–817. Elsevier Science Publishing, New York (1988)
6. Vera, A., Holden, K., Dempsey, D., Russi-Vigoya, M., Wu, S., Beutter, B.: Contextual inquiries and interviews to support crew autonomous operations in future deep space missions: preliminary requirements and proposed future research. In: Internal NASA Report prepared for the Human Factors and Behavioral Performance Element of the Human Research Program (2019)
7. Martin, L., O'Keefe, W., Schmidt, L., Barshi, I., Mauro, R.: Houston, we have a problem-solving model for training. J. Organ. Psychol. **12**(1), 57–68 (2012)
8. Holden, K., Munson, B., Russi-Vigoya, M. N., Dempsey, D., Adelstein, B.: Human Capabilities Assessment for Autonomous Missions (HCAAM). Phase II: Development and Validation of an Autonomous Operations Task List. In: Internal NASA Report prepared for the Human Factors and Behavioral Performance Element of the Human Research Program (2020)

Habitability Study on Space Station Colour Design

Ao Jiang[1(✉)], Xiang Yao[2(✉)], Irene Lia Schlacht[3], Giogio Musso[4], Tang Tang[1], and Stephen Westland[1]

[1] University of Leeds, Leeds, UK
sdaj@leeds.ac.uk
[2] Xiangtan University, Xiangtan, China
31558950@qq.com
[3] HMKW University Berlin, ILEWG at ESA ESTEC, Berlin, Germany
[4] Thales Alenia Space Italy, Turin, Italy

Abstract. Various stressors such as microgravity, vibration, radiation, restriction, and isolation in manned spaceflight environments can cause a variety of negative psycho-physiological effects. At the emotional level, for example, they may provoke anxiety and depression, which affects the astronauts' operational efficiency and overall mission performance. The colour design of a spaceflight environment could positively affect a person's emotional level and thus help to counteract such negative psycho-physiological effects. This paper presents a new model for validating the colour design of spaceflight environments at the psycho-physiological and emotional level in order to increase the quality of emotional habitability and support efficiency and performance. Psycho-physiological experiments were tested on six coloured light in a dedicate physical mockup of a specific spaceflight environment. In particular the sanitary area of the space station was used as a case study. As result the highest quality of emotional habitability was achieved in a yellow coloured light environment, that is very close to the natural solar condition. *Note:* In order to support the confidentiality in this paper is not mentioned the name of the space station.

Keywords: Emotional habitability · Colour design · Human factors · Space station sanitary area

1 Introduction

1.1 Space Colour Design Based on Emotional Habitability

In any space environment, the amount of system design that takes into consideration human factors, i.e., the habitability factor of the environment, is very low. To support the performance of long-term space missions, increasing emotional habitability is a prerequisite [1]. The first step in human factors research based on human-centred design is user analysis, where gathering output from people who will be using the product or service is an important part. In the space environment, the environmental lighting requirements and the colour matching of the visual space in the cabin are important factors affecting emotional habitability. Proper consideration of these factors

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 507–514, 2020.
https://doi.org/10.1007/978-3-030-50943-9_64

in the design can well improve people's psychological identity and stimulate work efficiency; otherwise, they will feel uncomfortable, their work efficiency will at least be reduced, and in severe cases they will even make operational errors and face safety problems [2]. Especially due to the relatively small space of the passenger cabin and the special environment, the design of the cabin colour will also affect the astronauts' space positioning, information acquisition and judgement, and psychological feelings. Therefore, it can be said that the reasonableness of colour matching design in the cabin layout is related to human ergonomics and safety [3].

1.2 Case Study: Sanitary Area of the Space Station

When mankind envisions building a permanent human habitat in space, it is necessary to consider constructing various functional guarantee systems with stable, reliable and safe performance in each functional section of the habitat. The space sanitary area system is the basic guarantee system for fulfilling the survival needs of astronauts in each functional division. It is closely related to the astronauts' life, functional safety, physical and mental health, as well as to efficient work. According to relevant reports published by NASA [4], the basic needs system for astronauts in future space habitats must be the subject of a reliable, stable and long-term target study. Due to the sanitary area's importance in supporting the basic survival guarantee for astronauts, the high degree of matching with humans, the important features of its complex functions, multiple technical interfaces and strong systemicity, the sanitary area system must have spatial emotional habitability. According to related anecdotal reports published by the Russian (former Soviet) space agency and NASA and relevant interviews with astronauts, the design and usability of the space station sanitary area are not good, as shown in Fig. 1. Complaints include: 1. The space is small and closed; 2. colour and light are unfavourable for astronauts to operate in; 3. the use of hardware is complicated and fault tolerance is low; 4. the location and shape of the fixing device and the handrail device mean that they cannot be used well [5]. These related factors lead to abnormal discharge, causing basic physiological disorders such as constipation, and even serious problems such as psychological and mental depression, insomnia, headaches and worsening interpersonal and social relationships [6].

Fig. 1. Existing U.S. and Russian space station and aerospace laboratory sanitary areas.

2 Method

2.1 Research of the Colour Perception Model Based on Emotional Habitability

The influence of colour factors on astronauts in terms of emotional habitability mainly originates from physiology and psychology. The channels for transmitting human colour vision information are light sources, coloured objects, eyes and brain [6]. These four elements not only make people feel colours, but also allow them to accurately analyse colours. If one of these four influential factors is inaccurate or biased, the astronaut cannot accurately analyse the effects of light and colour. The radiation effect of light sources and the reflection effect of objects belong to the discipline of opto-electronic physics. Therefore, colour perception is a concentrated reflection of coloured light, the human visual perception system and a person's mental state. Therefore, making use of the physiological and psychological processing involved in colour perception, a colour perception model based on emotional adaptation was constructed, as shown in Fig. 2.

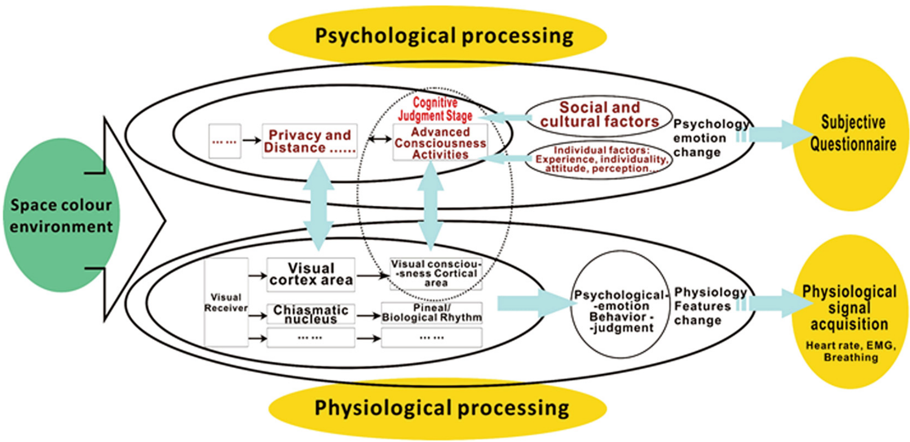


Fig. 2. Colour perception model for spatial emotional habitability

2.2 Development of the Mockup to Test the Sanitary Area

To investigate how to increase the quality of emotional habitability by selecting the best colour design configuration (Fig. 3), a physical mockup of the space station was built (Fig. 4). The mockup included a simulated environmental factor system, a light control system, and a data monitoring and acquisition system. The simulated environmental factor system mainly simulated the temperature and humidity environment of the space station's sanitary area, the noise environment, and the closed environment of the sanitary area to ensure the reliability of the test.

To better understand and simulate the user interaction, interview with specialists, videos and pictures of the ISS's sanitary area operated by the NASA's astronaut Sunita Williams' as well as interaction improvement from NASA and ROSKOSMOS (reported on paragraph 1.2) were analysed and implemented in this study [7].

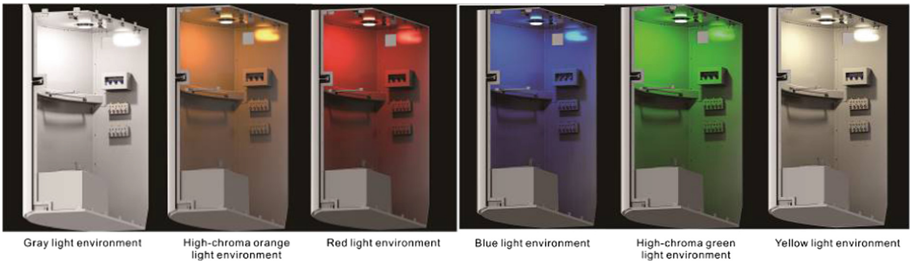


Fig. 3. Colour design study of the space station's sanitary area

2.3 Determination of Colour Specimens Based on Colour Matching Standards

Table 1. CIE LAB colour attributes of 6 colours using the CIE light source D65/1964 colourimetry observer combination [8]

	Colour centre	L^*	a^*	b^*	C_{ad}	H_{ab}
1	Grey	61.1	−3.2	3.2	4.5	135
2	Red	41.0	33.2	25.5	41.9	38
3	High-chroma orange	60.3	33.0	64.3	72.2	63
4	Yellow	84.1	−6.7	50.4	50.9	98
5	High-chroma green	56.0	−45.7	5.7	46.1	173
6	Blue	37.0	−1.3	−27.9	28.0	267

Six CIE LAB colours were investigated in this study according to the International Space Station (ISS) space colour matching standard SSP 41000 designed to support the best psycho-physiological health conditions in space missions. The colours were distributed inside the mockup as coloured lights. The six colours tested were grey, red, high-chroma orange, yellow, high-chroma green and blue. Table 1 lists the CIE LAB values for the selected colours. Six coloured light bulbs (Fig. 4) were used to illuminate the environment of the simulated space station sanitary area.

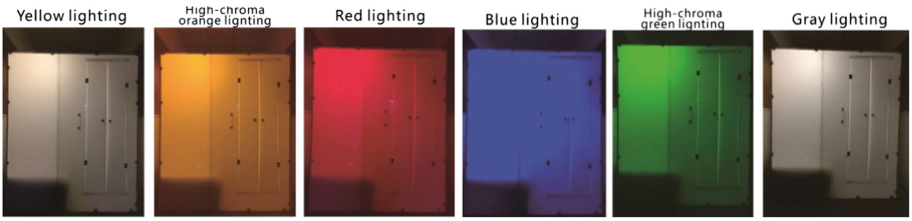


Fig. 4. Colour design test of the entrance of the mockup of the space station's sanitary area (picture of the interior were forbidden for ethical reason)

2.4 Establishment of Test Plan and Countermeasures

During the test the participants were asked to keep their mental state stable, their attention needed to remain focused on the system interaction. The colour changes in the simulated environment were only used for stimulating the physiological state of the participants, which did not affect their test activities. The test set-up was as follows:

A. 40 participants (28 male, 12 female), 30–40 years old, with strong physiques and good physical fitness. They had regular daily routines and did not drink or take drugs.
B. Six 18 W coloured light bulbs (with six different colours of light: grey, red, high-chroma orange, yellow, high-chroma green and blue; the six colours' RGB values were all within the standard value range used by the International Illumination Commission for colour discrimination). At the same time, this study simulating the system environment of the space station sanitary area used the CAPTIV-L7000 human factor data acquisition system to measure the participants' heart rate, breathing rate and myoelectric signals.
C. The test process comprised the following steps:
C1. The participants needed to be fully rested the day before the experiment, had not performed any bowel movements and maintained a calm and good condition. They were made familiar with the test environment and the operation and use procedures of the sanitary area before the test started in order to eliminate any effects of changes in their psychological state in an unfamiliar environment.
C2. First, experiments were performed in a natural indoor bathroom environment, that is, in a spacious and bright bathroom.
C3. After resting for five minutes, the participants entered the simulated test environment to perform the test. Using the sequence of grey, red, high-chroma orange, yellow, high-chroma green and blue, the participants switched the coloured light bulb to a different colour every ten minutes, i.e., the environment colour of the closed simulated sanitary area remained the same for a period of ten minutes each. While switching the colour environment, the participants rested for five minutes. Then they strictly followed the sanitary area operation and use procedures of the "Astronaut Biographies Home Page" published by NASA and the European Aviation Authority. Before entering the environment of the simulated space station, T-Sens breathing frequency sensors and T-Sens heart rate sensors were attached to them. A Sens

surface EMG sensor was also connected to each participant. Once they were ready, this physiological data was collected and stored via the wireless data logger T-log. C4. During the whole experiment, the participants were relaxed, simulated the operation procedures of the space station sanitary area and used the sanitary area normally

3 Result

3.1 Method for Processing Physiological Signal Data

The differences in human physiological states in different environments were compared by analysing the Euclidean distances of the physiological signal data in the different environments. Specifically, the Euclidean distance was calculated for the 40 groups of physiological signals in the simulated experimental environment with grey, red, high-chroma orange, yellow, high-chroma green and blue light, and for the 40 groups of physiological signals in the natural indoor environment. Due to space limitations, only four participants' signal acquisition results are listed in this article, as shown in Figs. 5, 6 and 7.

3.2 Data Processing Results and Analysis

The experimental results show that in the simulation experiment environments where blue, high-chroma orange, red and high-chroma green light was used, the participants' physiological signals were farther away from their physiological signals in the natural bathroom; that is, the difference was greater. In the grey light simulation experiment environment, their physiological signals were close to the Euclidean distance of their physiological signals in the natural bathroom; that is, the difference was small. In the yellow light simulation experiment environment, the Euclidean distance between their physiological signals and those in the natural bathroom was the closest; that is, the difference was the smallest. This shows that in a closed and narrow simulation

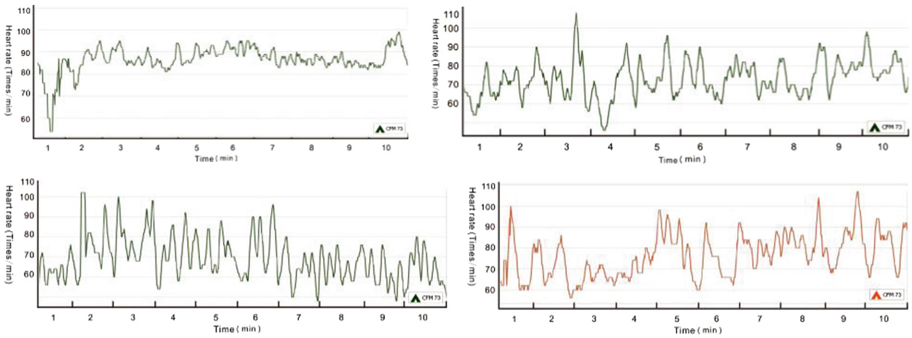


Fig. 5. Heart rate test frequency data

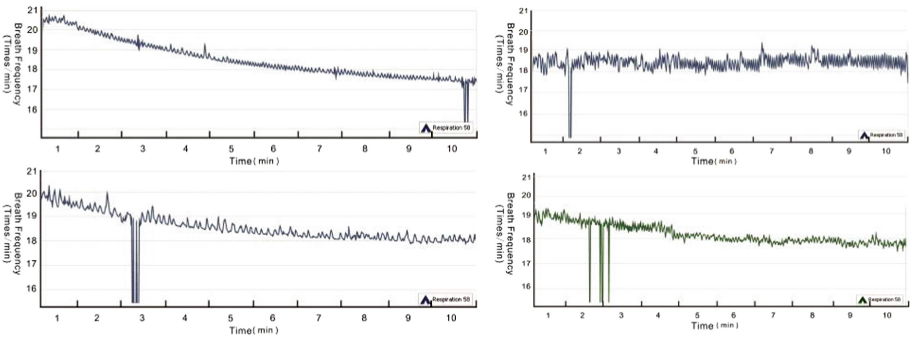


Fig. 6. Breath test frequency data

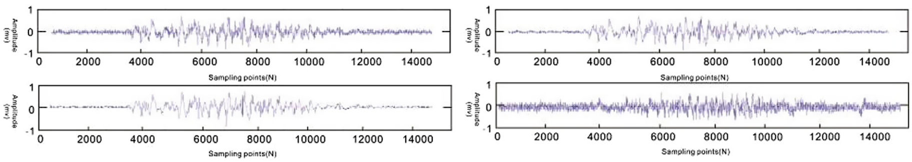


Fig. 7. EMG test frequency data

experiment environment, choosing a yellow environment would be of great help to the mental state of the crew when using the sanitary area. Moreover, yellow is the light colour that resembles the natural environmental solar light the closest.

4 Conclusion

This project involves the study of colour perception models based on emotional habitability applied to space station design. Through the application of a model to the study and evaluation of physiological data and subjective feelings, six different referential light colours were tested in a physically simulated environment based on the space station. In particular, considering the key relevance for supporting the basic survival needs, the environment of the sanitary area was selected as a case study. The results show that the highest quality of emotional habitability was achieved in a yellow coloured light environment, which is the light colour that resembles the natural environmental solar light the closest.

A more profound future study could pay particular attention to three factors: 1. the selection of the participants and the size of the sample to enable a better match of the physical quality and psychological characteristics of the participants with those of astronauts. 2. the influence of microgravity on astronauts' use of sanitary areas and the surrounding environment, could be tested in the future on ISS and parabolic flight as this could impact the physiological signals and subjective feelings. 3. Finally the colour sample variety could be implemented to increase the quality of the results.

Acknowledgements. This work is supported by a scholarship from the China Scholarship Council and the University of Leeds (No. 201908430166), a scientific research project of the Hunan Provincial Department of Education (No. 19B568), as well as a research project of the China Astronaut Research and Training Centre (No. 2018111400419).

References

1. Stuster, J.W.: Space station habitability recommendations based on a systematic comparative analysis of analogous conditions (1986)
2. Schlacht, I.L.: SPACE HABITABILITY: integrating human factors into the design process to enhance habitability in long duration mission. TU-Berlin. Berlin (2012)
3. Durao, M., Favata, P.: Color considerations for the design of space habitats. In: AIAA Space 2003 Conference & Exposition, p. 6350 (2003)
4. Link Jr., D.E., Broyan Jr., J.L., Philistine, C., Balistreri Jr., S.F.: International space station USOS waste and hygiene compartment development. SAE Trans. **116**, 119–124 (2007)
5. Kitmacher, G.H.: Reference Guide to the International Space Station. CreateSpace Independent Publishing Platform, Scotts Valley (2006)
6. Mahnke, F.H.: Color, Environment, and Human Response: An Inter-Disciplinary Understanding of Color and Its Use as a Beneficial Element in the Design of the Architectural Environment. Wiley, New York (1996)
7. NASA. Space Potty (2013). www.youtube.com/watch?v=5WSlGRBTFNI
8. Robertson, A.: CIE guidelines for coordinated research on color-difference evaluation. Color Res. Appl. **3**(3), 149–151 (1978)

Human Factors in Transportation: Aviation and Space

What-if Analysis in Total Airport Management

Reiner Suikat(✉), Sebastian Schier-Morgenthal, Nils Carstengerdes, Yves Günther, Sandro Lorenz, and Florian Piekert

DLR, Institute of Flight Guidance, Lilienthalplatz 7, 38108 Brunswick, Germany
{Reiner.Suikat, Sebastian.Schier, Nils.Carstengerdes, Yves.Guenther, Sandro.Lorenz, Florian.Piekert}@dlr.de

Abstract. This paper addresses design requirements for a what-if tool for airport management. First it provides an overview of the challenges in airport management and the need for a what-if tool. Following this motivation, theoretical and existing solutions of what-if systems are presented to conclude on an approach for a what-if system design specifically for the Total Airport Management (TAM) concept. TAM requires the cooperation of relevant stakeholders at an airport working towards different company specific goals yet trying to optimize overall airport operations. As the interactions between the various processes are highly complex the effects of an individual action on overall operations are not easily understood. Hence there is a need for a process to predict the overall effects of any action and to allow a comparison between several options. This will be provided by a what-if analysis tool, which provides a major component of a cooperative decision support system.

Keywords: Total airport management · What-if analysis · Decision support

1 Motivation

Air transportation is an essential part of the modern world and inefficiencies are generating high costs. This holds true especially for airports. Their limited capacities are an increasing challenge for the air transportation system [1, p. 43]. The often-cited study of Cook and Tanner [2] underpins this view by some remarkable figures: single airports can account for 50% of the total air traffic flow management (ATFM) delay.

Airports are an essential part of the system with high optimization potential due to the various interdependencies and interconnections of the conducted processes. Approach, landing, turnaround and departure processes need to be precisely timed, but are owned by independent stakeholders (airspace users, airport authorities, air navigation service providers, ground handling). Currently, each stakeholder plans his actions according to his individual goals and plans, but mostly without consideration of intentions and actions of other airport stakeholders. A job and task analysis of the German Aerospace Center [3] reveals optimization potential in information sharing and coordination processes between the stakeholders. Due to a lack of coordination, it is a challenge for the human operators to assess the consequences of actions of other stakeholders on own plans and vice versa.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 517–523, 2020.
https://doi.org/10.1007/978-3-030-50943-9_65

2 Problem Statement

The existing procedure of airport collaborative decision making (A-CDM, [4]), being in operation on the major European airports, addresses the alignment of multiple stakeholder actions on a limited scope. Basic flight information is shared and the timing of the departure process is agreed among the stakeholders (pre-departure sequencing). Thereby, other stakeholders gain awareness of delays, but they are not informed about the cause of the delay, the intentions and additional actions to be expected.

Currently TAM [5] is developed within research to provide the stakeholders with a global perspective on the airport. Generally speaking, TAM is an operational concept that brings together the main stakeholders at an airport to collaboratively develop a plan for future airport operations, the so called Airport Operations Plan (AOP). This AOP contains scheduled and target times as well as resource plans for all flight operations related events such as on- or off-block times, runway scheduling or ground handling resource planning. In case of disruptive events at the airport, for example adverse weather affecting operations, industrial actions or security incidents, the stakeholders will collaboratively develop a strategy on how to best react on the event such that the overall impact on operations is minimized.

TAM was initially defined in 2006 by German Aerospace Center and EUROCONTROL [5]. The Single European Sky ATM Research Program SESAR delivered detailed use cases and provided first results of the usefulness of the TAM concept approach [6, 7] and [8]. Within the successor program SESAR 2020, these ideas have been developed further and were validated in different activities [9].

The implementation of TAM demands for multiple elements to support the coordination of the stakeholders. Beside a common airport operation center (APOC, [10]) and the integration of landside stakeholders (e.g. terminal management, security control or baggage system management [5, p. 15]), a so called what-if tool is the major element to create a common awareness about the consequences of disruptions and counteractions [5, p. 40]. The main task of such a what-if tool is to predict the behavior of the airport system for a specified period of time based on one or more defined sets of input parameters, called scenarios. Design and implementation of this what-if tool is a great challenge as it faces two partly opposing objectives: On one hand the what-if tool must be able to predict the execution of the complex airport processes as precisely as possible. On the other hand, the stakeholders using the tool must get a clear and easy to understand picture of the airport situation encountering multiple disruptions and actions, and the prediction must be available within a reasonable time. The question, how to best approach the design of a what-if tool and find a reasonable compromise between the two objectives is addressed in this paper.

3 Possible What-if Solutions in Total Airport Management

Theoretical studies of what-if or decision support systems (DSS) reveal that a common solution for all applications has so far not been found [11, p. 8]. In particular, different objectives and different levels of detail are complex to be integrated into one common concept.

Regarding the objectives, there are two main purposes for what-if analysis - exploration and selection [12, p. 366]. Exploration focusses on the “option finding” phase of the human decision making process (phase two in the GOFER process [13]). Therefore, the what-if system is used to predict possible outcomes of a disruption (e.g. snow front entering the airport). The stakeholders are then able to define possible counteractions (e.g. close airport or apply snow removal). The selection focusses on the “consideration of effects” phase of the human decision making process (phase four in the GOFER process [13]). Within this phase the what-if system supports the stakeholders in the assessment and selection of the counteractions (for instance by the delay caused by an airport closure or snow removal).

Beside the objectives, the level of detail varies from what-if case to case [11, p. 4]. For the theoretical view this is described as the number of attributes each predicted alternative needs to provide [12, p. 366]. With regards to airport management especially the timeframe for the decision making, the participating stakeholders, the input parameters, and the performance indicators determine the number of attributes. The timeframe is an element to be considered as multiple stages of traffic management exist [5, p. 13f]. Depending on that timeframe and the number of scenarios to be considered stakeholders might plan based on aggregated parameters such as capacity and demand values or parameters for individual flights.

Hence, during the exploration phase where a quick evaluation of a potentially large number of scenarios is required, a what-if prediction should also be based on a limited set of aggregated input parameters and rapidly provide an estimation of the development of the main Key Performance Indicators (KPI) over time for each scenario under consideration. This will enable the human operators to make a pre-selection of potentially good solutions for the problem at hand, without spending a lot of time on detailed planning.

During the following selection phase the number of scenarios will be much smaller and at the same time the required level of detail of the prediction will be higher. Here a prediction is needed that takes more detailed planning parameters down to the individual flight level into account and produces detailed data for each flight as well in addition to the aggregated KPI values. This will then support the operators in their final decision for the best solution.

The relevant timeframes and attributes of the what-if scenarios will be defined by the participating stakeholders. All relevant processes of those stakeholders need to be taken into account by the what-if system to predict the attributes of each alternative. Last but not least the KPIs used for the evaluation of the scenarios and computed by the what-if tool have to be selected depending on the airport, the stakeholders’ individual objectives and the given situation.

Finally a what-if process can also be categorized into an individual (single) or multi-stakeholder what-if. In the first case an individual stakeholder is using the what-if tool to explore various options he might have. The stakeholder will only be able to modify data that are under his control and can compute within his own system a prediction of the effects of these actions. A multi-stakeholder what-if is performed within a group of stakeholders involved in a decision process. Each of the participating stakeholders can modify the data under their own control, and the tool will compute a prediction based on their combined inputs, which will be shared by all of them.

These fundamental types may be combined in parallel or in sequence to form complex chains of what-if analyses, which require a highly sophisticated scenario management within the APOC what-if environment. For example, a group of stakeholders can start a multi-stakeholder what-if, and within the resulting scenario individual stakeholders might branch off into single-stakeholder what-ifs to evaluate further actions they might take. These can remain purely within their own systems or shared with the others, thereby creating a new scenario for everyone to evaluate.

A number of issues can arise when combining what-if analyses. When individual what-ifs are performed in parallel, then combining them actually creates a new multi-stakeholder what-if and the prediction needs to be redone using the new combined input parameters. This is necessary as each what-if was based on the original scenario and did not account for the parameter changes in the other what-ifs. This shows that the management of the what-if scenarios can become rather complex. It must be assured that for any what-if scenario analysed by the stakeholders it is absolutely clear which input parameters are actually contained in the scenario.

Concluding these findings, it becomes evident that what-if in airport management needs to consider multiple degrees of freedom. It is expected that a what-if tool for airport management will have to be tailored specifically to the airport, its stakeholders, its specific processes and problem situations. This tailoring process will not be possible without direct involvement and contribution from the stakeholders at the target airport.

4 Existing What-if Solutions

The evaluation of already existing what-if systems is based on a structured literature review. Therefore, four online libraries (Google Scholar, German National Library, European Library and Springer press) have been searched. As a search string "[airport management] and [what-if]"[1] was selected. Analyzing this collection, a set of four relevant systems following the what-if definition given in chapter two were found. The following section describes those systems briefly, characterizing them by objective, input, method and output.

Airport Movement Area Closure Planner. The company ATAC provides the Airport Movement Area Closure Planner (MACP) [15]. MACP is a decision support tool for airport, air traffic control and airline staff to evaluate the effect of taxiway or runway closures. MACP generates a dataset for an airport out of the airport/ airspace structure, flight schedules, NOTAMS[2] and traffic data. The users can manually enter various closure decisions. MACP builds a Bayesian network representing the airport. Out of the traffic data, a routing is derived and the flights are introduced. Finally, this airport system is simulated using ATAC's fast time simulation tool SIMMOD [16]. As a result, the delay for each flight is calculated and presented to the user.

[1] Defined in the contextual query language [14].

[2] NOTAM = Notices to airmen, information about temporary or permanent changes to Aeronautical Information Publications, required for orderly, safe and secure and flowing air traffic.

The tool can be used in multi-stakeholder what-if situations, as several stakeholders may be involved in the decision process. It covers the core part of a what-if DSS tool, the prediction of the operational impact of stakeholder decisions.

TOP. The Total Operations Planner (TOP) is a development by DLR that performs an optimized pre-tactical planning for airport operations given operational conditions at the airport and the demand schedule [17]. It allows for multi-stakeholder interaction (ATC, Airline, Stand/gate planning), where each stakeholder can enter operational constraints into the system. Both, individual and multi-stakeholder what-if analyses are supported by TOP. The tool performs a detailed optimization of the flight plan, taking into account parameters entered by the stakeholders and presents them with a prediction of relevant KPIs such as punctuality or delay. The stakeholders can then compare different approaches to solve a capacity shortfall and select the best suited option.

Currently the KPI prediction is solely based on a prognosis of the flight data from the TOP planning, a coupling with a fast-time simulation to improve the prediction quality for the what-if analysis has not yet been implemented. TOP includes a workflow and scenario management to handle single and multi-stakeholder what-if analyses.

Airport Planning What-if Prototype. Boeing and Architecture Technology Corporation, together with the Ohio State University, implemented a what-if prototype for airport planning [18]. This prototype uses the flight schedule, runway capacity values and traffic management actions (e.g. departure clearance times, ground delays, airspace flows, severe weather avoidance) to evaluate how traffic management initiatives will affect the traffic. To calculate this impact a NASA fast-time queuing simulation is used. A logical network within the simulation represents the infrastructure of the airport and the airspace. As a result, runway throughput, runway queue length, average taxi times and gate delays are provided.

This system is designed to support a single stakeholder what-if performed by the Departure Reservoir Coordinator (DRC) to issue Departure Management Programs (DMPs) to mitigate the negative effects of predicted demand-capacity imbalances.

Spade DSS. The "Supporting Platform for Airport Decision Making and Efficiency Analysis" (SPADE) "Decision Support System" (DSS) was developed within a European research project [19]. SPADE DSS evaluates the impacts of operational interventions on airport performance. Thereby airport decision makers are assisted. SPADE DSS is a framework, which covers multiple simulation modules as well as multiple use cases (change of infrastructure, operational procedure or traffic). As such, the necessary input data depends on the selected use case and modules. To work with SPADE DSS, the user initially selects a scenario, the performance indicators and the use case. Afterwards the necessary simulation modules are triggered. Finally, the results are presented in a graphical interface. SPADE DSS is capable to provide indicators for capacity, delay, level of service, noise, safety and security.

Results are presented for a strategic, i.e. long term airport capacity planning case, supporting a cooperative what-if decision support with a group of multiple stakeholders.

5 Conclusions

In this paper the authors address the question how to best approach the design of a what-if tool. Such a tool is needed to implement the TAM concept. Thereby an agreement among the stakeholders on common actions to exploit operational optimization potential within the airport processes is enabled.

The section on possible solutions for a what-if tool clearly shows that there is a broad bandwidth of options which need to be adapted to the needs of the specific airport. Several prototype tools and ideas have been presented as a result from a literature search, but they each only cover certain parts from the spectrum that is needed to be used in an operational context at an airport. The presented systems either only address a single process at the airport (like taxiing) or a single stakeholder, a complex chain of what-if predictions by single or multiple stakeholders has not been addressed yet. Also it is unclear to what extent such existing tools can address the fact that the scenario is developing over time, so the underlying parameters may even change during the calculation of the prediction.

Nevertheless, all analyzed systems follow the same basic design scheme: at the heart of the DSS a prediction engine computes the relevant KPIs based on the current situation, flight schedule, and current input parameters from the various stakeholders. This prediction can be an optimization algorithm or a simulation tool, depending on the required accuracy and time to produce the result. Also required is a user interface allowing a common view on the KPIs for all stakeholders involved in the decision making, but also to input stakeholder specific data. Finally, support is needed to structure the what-if workflow and to deal with multiple what-if scenarios, so that KPIs for different decision options can be compared and the best suited can be selected.

This basic scheme will be taken up in the SESAR Total Airport Management project (PJ.04) and developed into a system that addresses the requirements laid out in Sect. 3. In this project, a what-if tool for snow events will be designed together with Oslo airport. After multiple design workshops and an implementation phase a validation with the operational stakeholders will be carried out early in 2022 using the airport management simulation platform [20] of the German Aerospace Center.

References

1. EUROCONTROL: Seven Year IFR Flight Movements and Service Units Forecast 2018 – 2024, European Organisation for the Safety of Air Navigation, Brussels, Belgium (2018)
2. Cook, A., Tanner, G.: European airline delay cost reference values - Final Report (Version 3.2), EUROCONTROL (2011)
3. Papenfuss, A., Carstengerdes, N., Günther, Y.: Konzept zur Kooperation in Flughafen-Leitständen. 57. Fachausschusssitzung Anthropotechnik, Rostock, Germany (2015)
4. EUROCONTROL: Airport CDM Implementation – The Manual, version 5.0, European Organisation for the Safety of Air Navigation, Brussels, Belgium (2017)
5. Günther, Y., Inard, A., Werther, B., Bonnier, M., Spieß, G., Marsden, A., Temme, M., Böhme, D., Lane, R., Niederstrasser, H.: Total Airport Management: Operational Concept & Logical Architecture. DLR, Eurocontrol, Braunschweig, Bretigny (2006)

6. Bogers, H., Linde, M., Matas Sebatia, J.R., Álvarez Escotto, I., Martin Espinosa, J.I., Ciprián Tejero, B., Ubeda Torres, J.A.: D16 – OFA 05.01.01 Consolidated OSED Edition 3, Part 1. Brussels, Belgium (2015)
7. Bogers, H., Linde, M., Matas Sebatia, J.R., Álvarez Escotto, I., Martin Espinosa, J.I., Ciprián Tejero, B., Ubeda Torres, J.A.: D16 – OFA 05.01.01 Consolidated OSED Edition 3, Part 2. Brussels, Belgium (2015)
8. Carstengerdes, N., Papenfuss, A., Suikat, R., Schier, S., Günther, Y., Piekert, F., Marsden, A.: EXE-06.03.01-VP-757 Validation Report. D142. Brussels, Belgium (2016)
9. Rousseau, F., Desart, B., Marsden, A., Mikloš, E., López, C., Rahatoka, C., Salomon, R., Suikat, R.: D3.2.150 SESAR Solution PJ04-02: Validation Report (VALR), Brussels, Belgium (2020)
10. Spies, G., Piekert, F., Marsden, A., Suikat, R., Meier, C., Eriksen, P.: Operational concept for an airport operations center to enable total airport management. In: 26th International Congress of the Aeronautical Sciences, Anchorage, Alaska (2008)
11. Golfarelli, M., Rizzi, S., Prolli, A.: Designing what-if analysis: towards a methodology (2006)
12. Philippakis, A: Structured what if analysis in DSS models. In: Proceedings of HICSS, pp. 366–370 (1988)
13. Mann, L., Harmoni, R, Power, C.: The GOFER course in decision making. In: Baron, J., Brown, R.V. (eds.) Teaching Decision Making to Adolescents, Lawrence Erlbaum Associates, Hillsdale, NJ. pp. 61–78 (1991). ISBN 978-0805804973. OCLC 22507012
14. OASIS Standard. searchRetrieve: Part 5. CQL: The Contextual Query Language Version 1.0 (2013)
15. Keller, B., Saraf, A., Popish, M., Luch, N., Tan, P.: Airport Movement Area Closure Planner (MACP), Airside Simulation and Performance Assessment Group Meeting, May 2019
16. Bertino, J., Boyajian, E.: 21st century, fast-time airport and airspace modeling analysis with Simmod. Managing Skies **9**(3), 21–23 (2011)
17. Günther, Y., Pick, A., Kern, S., Lorenz, S., Gerz, T., Keis, F., Köhler, M.: Improved airport operations planning by using tailored forecasts of severe weather. In: Conferences in Air Transport & Operations, TU Delft, Delft (2015)
18. Timar, S., Peters, M., Davis, P., Smith, P., Lapis, M., Wilson, I., van Tulder, P.: A what-if analysis tool for planning airport traffic. In: 2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC), pp. 1–9 (2017). https://doi.org/10.1109/DASC.2017.8101982
19. Zografos, K.G., Madas, M.A., Salouras, Y.: A decision support system for total airport operations management and planning. J. Adv. Transp. **47**, 179–189 (2013)
20. Schier, S., Timmermann, F., Pett, T.: Airport management in the box – a human-in-the-loop simulation for ACDM and airport management. Deutscher Luft-und Raumfahrt Kongress 2016, Braunschweig (2016)

Pilot Tailored Helicopter Systems

Christian A. Niermann(✉)

German Aerospace Center (DLR), Institute of Flight Guidance,
Lilienthalplatz 7, 38108 Brunswick, Germany
christian.niermann@dlr.de

Abstract. An online questionnaire about assistance systems in helicopters was answered by 153 professional pilots from all areas of civil and military aviation. The results indicate that pilots generally appreciate the systems provided in their helicopters but see potential for improvement in the usability. One notable fact is the high proportion of visual display concepts. The human auditory channel is supported by the systems to a comparatively small extent. Pilots would like more possibilities to directly influence the systems. The multimodal approach is considered promising for the future. The results suggest that assistance systems and their forms of display to the pilots must be designed differently depending on the current flight phase.

Keywords: Human factors · Helicopter · Survey · Assistance systems · Professional pilots · Visual · Auditory

1 Introduction

In modern helicopter cockpits, pilots receive most of the information provided from the aircraft visually. At the same time, flying a helicopter is a combination of visual monitoring of reference points to control position and movement of the aircraft and noting the information presented visually and acoustically by the cockpit's systems. For a successful mission, flight crews rely on training, modern cockpit instrumentation, and the support of other crewmembers to retain situational awareness and perform even under a high workload and poor weather conditions. Based on current publications and accident reports, loss of situational awareness with a subsequent loss of control has caused most helicopter accidents in recent years [1–3]. To avoid future accidents due to a loss of situational awareness, innumerable assistant systems have been developed and integrated into modern helicopter cockpits. However, one must be careful when introducing additional assistance systems into the cockpit since each new system not only contributes more important information but also attracts the crew's attention, potentially negatively affecting safety-critical decisions. Instead of reducing the workload in the cockpit, the new systems increase pilots' workloads and foster attention tunneling to a specific system. This leads to a disruption of the actual distribution of the fundamental flying prioritization: Aviate-Navigate-Communicate-Manage Systems. Analyses of accidents in the past have demonstrated that pilots have been so focused on a system or issue that they have not sufficiently followed the basic aviate tasks [4, 5]. To address the problem of overloaded cockpits, this work assesses

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 524–530, 2020.
https://doi.org/10.1007/978-3-030-50943-9_66

which assistance systems helicopter pilots currently use and how pilots are operating and judging them in their daily flying.

For a more pilot-tailored helicopter system, it is vital to understand current systems and how they are used and rated by pilots. To address this, an online questionnaire was designed. The questionnaire consisted of 54 questions. Besides demographic and personal flying related questions, personal sensory abilities and current systems used in the helicopter combined with ratings of their usability and benefits were explored. Specific questions about warning systems and a dedicated part on audio systems were included. Additionally, the auditory and visual capacities of the pilots were examined based on pilots' self-assessment according to the different flight phases.

2 Questionnaire and Participants

Related work as known to the author is based only on small number of pilots, limited flying experience or only certain helicopter manufacturers or flying missions. This work provides a broader view to exclude possible effects caused by the missions or helicopter types to be flown. The results in this work are based on the answers of 153 professional helicopter pilots. Of the respondents, 97% fly professionally several times a week with typical flight time between two and four hours a day. The average work experience was 21 years and total flight time more than 2,500 h (78% of the pilots). Most participants fly predominantly under visual flight rules (VFR) during daylight (86%); only 7% fly VFR at night, and 7% fly under instrument flight rules. All answers are based on flying experience in the following countries in descending order of the number of pilots participating: Germany, Switzerland, Austria, the USA, Italy, the United Kingdom, Australia, and Iceland (Fig. 1).

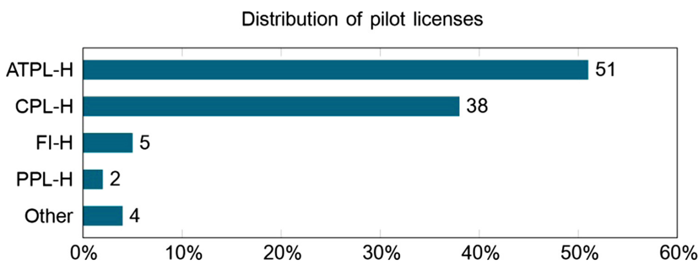


Fig. 1. Distribution of pilot licenses as answered in the questionnaire. Four participants stated under "other" their military licenses.

The high number of participants with significant experience leads to robust results and gives a detailed view about actual helicopter systems used. In addition, the system-related answers were based on a wide variation of 27 different helicopter models. Pilots flew light-intermediate single-engine helicopter like the Robinson R22/R44, Airbus H120 or MD Helicopters MD 500; medium twin-engine helicopters like the AgustaWestland AW109, Airbus H135/H145 or Bell Helicopter Bell 430; and heavy twin-engine helicopters like the Boeing AH-64 Apache, Airbus AS332 Super Puma or Sikorsky UH-60 Black Hawk. The distribution of the helicopter classes is presented in Fig. 2. The light-medium twin-engine class, commonly used for Emergency Medical Service (EMS) and Search and Rescue (SAR) work is, as expected, disproportionately represented.

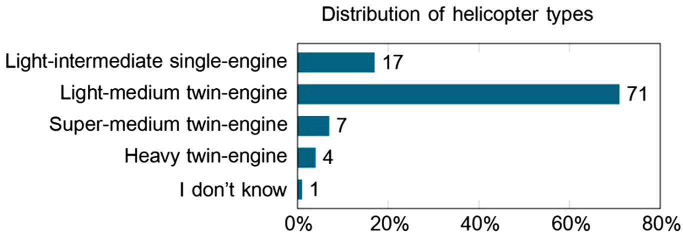


Fig. 2. Distribution of helicopter types as answered in the questionnaire. The Airbus H135 and H145 represent 54% of all named helicopters over all helicopter types.

As illustrated in Fig. 3, the majority of the pilots fly EMS/SAR missions (in total 69%). All other working fields are equally distributed. The location of use is evenly distributed over service in cities, the countryside and mountains. This distribution in the survey reflects the helicopter missions flown and forms a balanced basis for the following assessments of daily work in the helicopter cockpit.

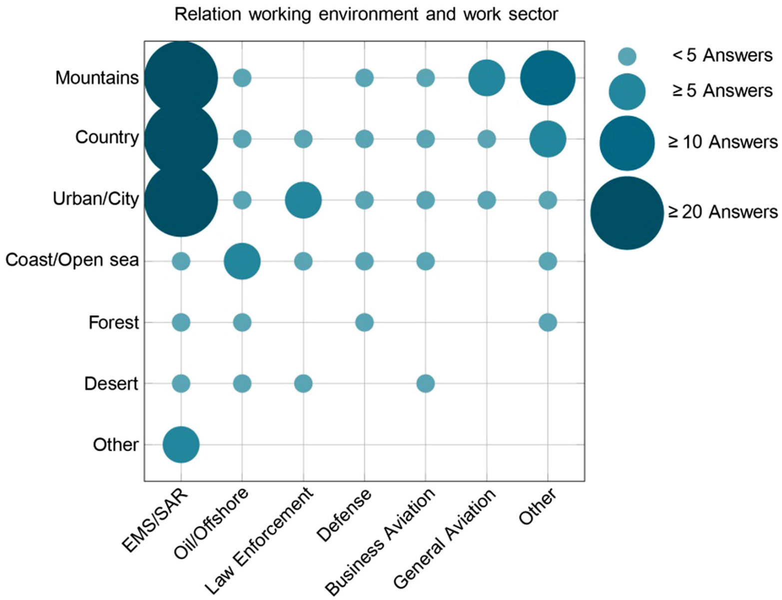


Fig. 3. Relationship between the working environment and the work sector of all participants. Participants that selected "Other" have commented that they do not fly predominantly in a specific environment.

In order to get an impression of how the pilots assess their visual and auditory exhaustion in the cockpit, they were asked to rate their visual and auditory stress independently. Both types of stress are assessed comparably on a scale from 0 (very underutilized) to 100 (very exhausted). Auditory exhaustion was specified as m = 53.21 (SD = 21.44) and visual exhaustion as m = 55.24 (SD = 18.38). Although the questions were asked at different times in the questionnaire, the majority of the respondents rated their personal visual and auditory stress equally.

It was asked in detail how exhausting the different flight phases of take-off, en route, maneuvering, hovering, approach, and landing are with respect to visual and auditory sensory perception. Pilots report that during en route and maneuvering, the visual and auditory load is equal. However, during take-off and approach, auditory exhaustion is higher because at airports pilots need to talk to air traffic control, and during off-airport operations pilots must talk with other crewmembers. Hovers and landings are flight phases with high demand on fine motor skills and simultaneous visual recognition of the movement of the helicopter and require full concentration. In this context, the participants stated that their visual channel is disproportionately highly loaded, while the auditory channel was rated as relatively free.

3 Helicopter Systems

One purpose of the questionnaire was to determine with which assistance systems current helicopters are equipped. The pilots were asked what systems are available in their regularly flown helicopters and how they rate these individual systems. Figure 4

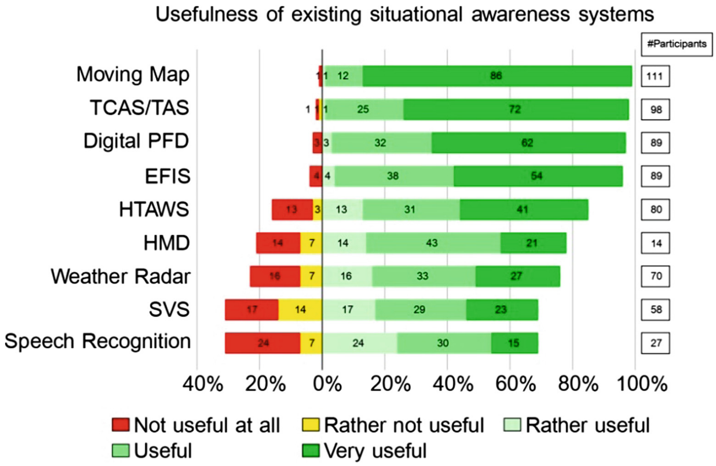


Fig. 4. Assistance systems in regularly flown helicopters and pilots' rating of usefulness. Multiple answers were possible. The percentages refer only to the answers of the respondents in whose helicopters the respective system is equipped. Traffic Alert and Collision Avoidance System (TCAS), Traffic Advisory System (TAS), Primary Flight Display (PFD), Electronic Flight Instrument System (EFIS), Helicopter Terrain Awareness and Warning System (HTAWS), Helmet-Mounted Display (HMD), Synthetic Vision System (SVS).

demonstrates that pilots mostly rate available systems from useful to very useful. Except for the speech recognition system, the most frequently mentioned systems all use the human visual channel. Only a few systems are combined with an acoustic signal in addition to the visual display.

All the systems that have been evaluated as being predominantly useful provide an auditory signal in addition to the visual display except for the moving map. Of the pilots polled, 99% stated that they always want an auditory indication to supplement the visual indication. Seventy-three percent of the respondents found a *voice-warning* that indicates the actual fault (e.g., "RPM low," "Engine out") helpful. A simple indication sound like the well-known Master Caution or Master Warning *gong/beep* were only favored by 32% as high cognitive work needs to put into the solution. Further, 30% of the pilots would like to receive *voice-advice* as direct order of what to do (e.g., "Collective down!" "Don't sink!") instead of naming the actual problem. If the audio signal were a combination, pilots would like to have first a tone as attention trigger, followed by a voice-warning with the actual problem and finally voice-advice on what to do.

One problem with advice or warnings that do not remain on a display as a message but are implemented as a short tone or a warning light is that they can be overlooked or forgotten. To prevent this, a repetition of the alert is usually used. For auditory warnings, 44% of the pilots were in favor of a quick repetition within one second and 50% for repetition after about 5 s. Only 6% favored a one-time alarm without repetition. The responses for visual warnings were comparable. Here 54% of the pilots favored a repeated flashing alarm compared to 39% for permanent light. Only 7% voted for an alarm that only flashes once.

Due to the larger number of assistance systems in the cockpit and their sometimes very specific areas of use, 98% of pilots agree that it is necessary for pilots to be able to activate/deactivate systems individually. Of these pilots, 43% want a system to be switchable at any time, and 49% think that a system must only be switchable in safe conditions. If pilots can turn systems off, consequently there must be a strategy to turn them on again. Of the pilots surveyed, 52% agree that a warning system must turn itself on again automatically if a new warning occurs. Re-activating after a certain amount of time was favored by 23%, and 18% stated that the system should switch back automatically to its standard mode only if the helicopter is back in normal and safe condition.

As pilots in previous work indicated that not every important system in current helicopters meets pilots' expectations, the questionnaire asked how they would like to redistribute the current representations. As presented in Fig. 5, the majority would like to change general cautions from visual to a combination of visual and auditory notification. This combination is also requested for warnings in general. Additional visual information instead of only auditory information is requested for terrain warnings and autopilot messages.

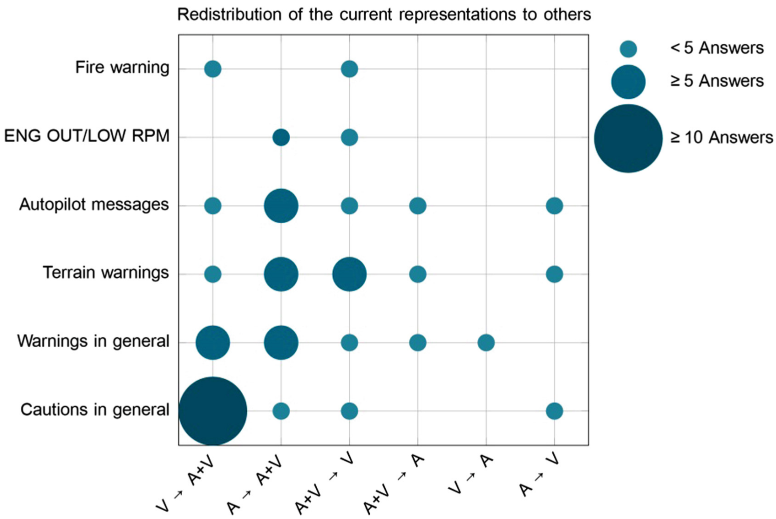


Fig. 5. Redistributing from the current display of warnings to other display types. The redistribution from "not presented" to another warning display was not considered. V ≙ Visual, A ≙ Auditory, A + V ≙ Auditory and Visual.

The majority of pilots would like to see a transition from single warnings to multimodal warnings. This wish is understandable. Nevertheless, it also entails the risk of over-stimulation instead of reducing the pilot's workload. Since this survey only addressed individual systems, the wish for multimodality is understandable. In the future, however, this will have to be investigated for each system individually and then combined with the other assistance systems in the helicopter.

4 Conclusion

The results of the survey are based on the answers of 153 professional and experienced helicopter pilots from the typical working environments of commercial helicopter aviation. Although the pilots fly different models/types from small light single-engine helicopters, medium to heavy twin-engine helicopters to military helicopters, the answers were clear and comparable.

Visual and acoustic stress is comparably equal during the duration of a flight. As expected, there are peaks during certain flight maneuvers. All pilots were satisfied with the basic support by assistance systems in modern helicopter cockpits. At the same time, they saw potential for improvement in the use of multimodal elements. The results also make it clear that a singular view of a system is not target-oriented and that the effects on the overall helicopter system must always be considered.

In addition to the already intensively researched and used visual interface between human and machine, the auditory component offers significant potential. At the same time, auditory elements must also be investigated with the same care as visuals, otherwise the visual overload from the past will be followed by auditory overload in the future. The author's ongoing work will continue in this direction.

References

1. U.S. Joint Helicopter Safety Analysis Team: Calendar Year 2006 Report: To The International Helicopter Safety Team (2010)
2. European Helicopter Safety Team: Analysis of 2006-2010 European Helicopter Accidents (2015)
3. United States Helicopter Safety Team: Comparative Report Volume 1: U.S. JHIMDAT Data to U.S. JHSAT Data (2014)
4. Miller, J.D., Godfroy-Cooper, M., Wenzel, E.M.: ARSAD: an augmented-reality spatial auditory display for obstacle avoidance during all phases of flight. In: Proceedings of the 74th Annual Forum, AHS International (2018)
5. Godfroy-Cooper, M., Szoboszlay, Z.P., Kahana, A., et al.: Terrain and obstacle avoidance displays for low-level helicopter operations in degraded visual environments. In: Annual Forum Proceedings - AHS International, vol. 2, pp. 1118–1134 (2016)

Operational Complexity in Performance-Based Navigation Arrival and Approach Flight Operations

Divya Chandra(✉), Andrea Sparko, Andrew Kendra, and Janeen Kochan

United States Department of Transportation Volpe Center, 55 Broadway, Cambridge, Massachuttes 02142, USA
{Divya.Chandra, Andrea.Sparko, Andrew.Kendra, Janeen.Kochan}@dot.gov

Abstract. We studied how pilots handle operational variations during arrival and approach instrument flight procedures (IFPs), focusing on factors that may be related to performance-based navigation (PBN). PBN is a key enabler of the Next Generation Air Transportation System (NextGen). We developed a factor rubric based on an iterative review of events in the Aviation Safety Reporting System (ASRS) public database and prior research. We coded 164 ASRS reports selected for relevance to PBN. We identified where each event occurred relative to the route of flight, tallied the coded factors and event outcomes, and gathered data on crew actions that indicated resilience to operational variations such as unexpected behavior of aircraft automated systems. We conclude that PBN appears to magnify the effects of operational complexity for these events. Pilots would benefit from training that provides opportunities to experiment with new situations they could encounter in PBN scenarios.

Keywords: PBN · RNAV · RNP · STAR · OPD · ASRS · NextGen · Resilience

1 Introduction and Background

Performance-Based Navigation (PBN) is a key component of the Next Generation Air Transportation System (NextGen). It enables pilots to fly terminal-area instrument flight procedures (IFPs) whose lateral and vertical routes are more precise. It is based on area navigation (RNAV) and required navigation performance (RNP). PBN has been implemented gradually over the past thirty years, beginning with Instrument Approach Procedures (IAPs) using the Global Positioning System (GPS), and more recently, IAPs that use RNP. Optimized Profile Descents (OPDs) are a new type of PBN arrival route. They have many altitude constraints and waypoints to allow for a relatively continuous descent that is fuel efficient and low noise. PBN IFPs are in daily use across the United States and in other countries.

The Volpe Center has led a multi-year research effort to provide the Federal Aviation Administration (FAA) with data and analyses to understand and address the complexities of flight deck operations associated with NextGen airport terminal IFPs

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 531–538, 2020.
https://doi.org/10.1007/978-3-030-50943-9_67

such as RNAV Standard Terminal Arrival Routes (STARs) and IAPs. This study is a follow-on effort to our prior research on PBN [1–3].

PBN IFPs have the potential to increase the safety, predictability, and efficiency of airport terminal operations when flown as designers intended. However, many factors can increase the complexity of these IFPs for flightcrews, resulting in an increased number of tasks that pilots have to manage and prioritize. Therefore, factors that induce complexity can affect the crew's ability to fly the route as intended.

Operational complexity is a term we use to describe the many day-to-day variations that pilots handle during normal flight operations. Operational complexity is not new, but it has not been studied formally in this context. We define operational complexity in terms of extra *tasks* for the pilot, either cognitive or physical. Extra cognitive tasks could require additional memory or attention resources. Extra physical tasks could be, for example, more button pushes. In this concept, extra tasks require pilots to update and manage their task agenda [4] effectively. Operational complexity has been studied in manufacturing, where it is associated with uncertainty, variety, and unpredictability in the degree of connectivity and interaction among system elements [5]. The same characteristics apply to operational complexity in flight operations.

We are interested in understanding the relationship between PBN and operational complexity. Past studies have hinted at possible connections. We know from [1] that the complexity of an IFP is not easily measured objectively. We also know that the complexity of IFP designs is reflected in the visual complexity of their depictions (i.e., aeronautical charts) [1–3]. Chandra and Markunas interviewed line pilots to learn more about their perspectives on the complexity of charts and IFP designs [3]. Their study identified many factors that affect the complexity of IFP design and charting (e.g., the energy profile of the route, constraints, and the number of notes). They identified five categories of operational complexity factors for pilots: Air Traffic Control (ATC) interventions, aircraft equipment or performance, crew factors, environment factors, and operator factors. IFP designers and chart producers cannot control or eliminate operational complexity. Pilots have to manage operational complexity in real-time.

One goal for this study was to refine the list of operational complexity factors from [3]. A second goal was to develop and test a method to differentiate between operational complexity factors related to PBN from unrelated operational complexity. We also wanted to explore how operational complexity factors interact and play out in real events. Real events often have an element of time pressure and acute stress that can make cognitive processes more error prone, as discussed in [6] and [7]. Finally, we were interested to know whether we could find examples of crew behaviors that improved safety as well. *Resilient* crew behaviors are identified and discussed in [8].

2 Method

We analyzed 164 reports from the National Aeronautics and Space Administration (NASA) Aviation Safety Reporting System (ASRS) submitted in 2016 and 2017. The events were selected for relevance to PBN arrivals and approaches using keywords such as PBN, OPD, RNAV, and variants. Most of the reports (120) were submitted for Title 14 Code of Federal Regulations (CFR) Part 121 (scheduled air carrier) operations;

29 were submitted for Title 14 CFR Part 91 (general aviation) operations. Sometimes more than one person reported the same event (e.g., both crewmembers or a pilot and a controller). NASA combines related submissions into a single report.

ASRS report limitations are well understood. For example, the events are self-reported, subjective, and written from memory. The narratives can be incomplete and difficult to interpret. They can also be biased because of difficulty in observing one's own behavior. The frequency of events in the database cannot be assumed to represent the frequency of occurrence in actual operations. Finally, determining whether factors were related to PBN or not was subtle; this judgment could require making inferences.

We developed two tools and a process for reviewing each report. One tool was a review template and the other was a rubric of factors. The review template had sections for basic information about the event (e.g., who reported it and where it occurred), a custom synopsis of the event describing the outcome, contextual information, and a record of issues and coded factors. The template also had sections for the reviewer to justify the factor coding and explain any inferences. The reviewer referred to the rubric of factors and their definitions to code the factors.

Table 1 lists the seven operational complexity factors in the rubric and their sub-factors. We also coded three other factors (IFP design, IFP-induced chart issues, and chart-specific issues) because they were often present and interacted with the operational complexity factors. We developed this rubric initially based on the results of [3].

Table 1. Operational complexity factors from the rubric.

Main factor	Sub-factors	
ATC & PBN issues that affect pilots	• Interventions related to PBN • Phraseology related to PBN	• Controller knowledge or training related to PBN
ATC issues only	• Aircraft sequencing • Internal ATC coordination	Generic ATC error
Flightcrew factors related to PBN	• Crew resource management (CRM) related to PBN • Lack of familiarity (with terrain, local area, or local IFPs) • Lack of knowledge or training related to PBN	• Confusion related to PBN • Lack of flight path awareness on PBN IFP • Time pressure related to PBN IFP
Flightcrew factors not related to PBN	• Distraction unrelated to PBN • Time pressure unrelated to PBN • Crew physical condition • Non-normal situation unrelated to PBN • Communication with ATC unrelated to PBN	• CRM unrelated to PBN • Decision-making unrelated to PBN • Confusion unrelated to PBN • Generic crew error

(*continued*)

Table 1. (*continued*)

Main factor	Sub-factors	
Aircraft/equipment factors related to PBN	• Unexpected behavior of automated system related to PBN • Aircraft flight performance	• Flight Management System or Mode Control Panel programming or setup, or autoflight configuration/operation
Environment	• Terrain related to PBN IFP • Terrain unrelated to PBN IFP • Man-made structures unrelated to PBN IFP • Man-made structures related to PBN IFP	• Airspace • Airport • Traffic • Weather (related to PBN or not) • Nighttime
Operator	• Dispatch • Clarity of pilot roles	• Clarity of standard operating procedures

We then refined the rubric iteratively based on events in the ASRS dataset. Examples and definitions for the factors are in the full report [9].

Each report was reviewed by two reviewers, one primary and one secondary. All the reviewers had expertise in aviation human factors and/or flight experience. The primary reviewer filled out the template first and the secondary reviewer reviewed the draft. If there were discrepancies or questions, the two reviewers discussed and reconciled the factor coding. They could also consult with the full team of four reviewers if needed.

3 Findings

The search criteria effectively identified events that had PBN-related factors. Most of the reports (148 out of 164, or 90%) had a coded PBN-related factor, or occurred along a PBN IFP. There were three reports for which we inferred a PBN-related factor.

3.1 Results

Locations of Events. The events in the dataset occurred at several different locations or regions. The most frequently occurring locations were Atlanta, Georgia (22 reports); Northern California, near San Francisco (21 reports); Southern California, near Los Angeles (18 reports); and Denver, Colorado (13 reports).

Seventy-two events occurred on STARs, most of which (at least 63) were OPDs. Fifty-four events occurred on approaches, including 27 on RNAV (GPS) approaches, 14 on RNAV (RNP) approaches, and 13 others (e.g., six on visual approaches with RNAV backup). Twenty-four events occurred on transitions, of which 23 were transitions from the arrival onto the approach.

Factor Tallies. Figure 1 illustrates how often the 10 *main* factors (seven operational and three IFP/chart factors) appeared, as a percent of the 148 reports that had a coded PBN-related factor, or occurred along a PBN IFP. Flightcrew factors related to PBN were the most common (occurring in 47% of the events), but five other factors occurred in at least 30% of events. IFP-induced chart issues, chart-specific issues, and operator issues were not commonly identified. Sometimes reporters confused IFP design issues with chart issues, but the reviewers coded these separately. All of the most common sub-factors tallied were related to PBN. These include ATC interventions related to PBN; flightcrew lack of knowledge/training related to PBN; CRM related to PBN; unexpected behavior of an aircraft automated system related to PBN; lack of flight path awareness on a PBN IFP; IFP design with constraints related to PBN; and finally, confusion related to the PBN IFP. Some factors were clustered, meaning that they tended to co-occur. In particular, crew lack of flight path awareness on a PBN IFP, crew lack of knowledge/training related to PBN, and ATC interventions tended to co-occur.

Event Outcomes. Table 2 lists the frequencies of different types of outcomes. (Note that events could have more than one outcome, such as deviations in both altitude and speed.) In agreement with [1], altitude deviations were the most common outcome for the arrivals and approaches. Overall, 60 of the 148 events (41%) had altitude deviations. In [1], lateral deviations were more common on departure procedures. Similarly, we found just nine lateral deviations for the arrivals and approaches in this study. Of the 38 events that had *No Deviation*, at least 21 mentioned pilot actions that prevented the deviation. *Miscellaneous (Other)* deviations included, for example, encountering wake turbulence, hard landings, and unexpected terrain alerts.

Resilient Crew Behaviors. We identified resilient crew behaviors in at least 50 events. Events where the outcome was *No Deviation* typically mentioned a resilient behavior. We suspect these behaviors happened in other events too, but were not always reported. Some examples of resilient *responses* include the crew disengaging an aircraft automated system, effective CRM (e.g., splitting or reallocating tasks), notifying ATC and requesting assistance, or saying *"unable"* to an ATC request. One example of good crew *monitoring* was when a crew determined they were landing at an unintended airport by noticing that the runway lighting configuration was not what they had briefed. Crews also sometimes *anticipated* outcomes such as unstable approaches, and took an action to go-around instead of land. In some events, crews practiced the PBN IFP under visual conditions, helping them learn what to expect.

3.2 Recommendations

Based on this analysis, we recommend that pilots should have opportunities to develop adaptive expertise [10] in order to be able to apply knowledge to the novel or atypical situations they might encounter with PBN. Pilots also need opportunities to experiment with new situations in safe training environments to promote deep learning. Training should reinforce basic skills and knowledge related to flying with automated systems under PBN. Pilots also should practice CRM in PBN scenarios, including task agenda management and decision-making skills. In particular, pilots should learn, through practice, when and how to intervene or say "*unable*" to an ATC request.

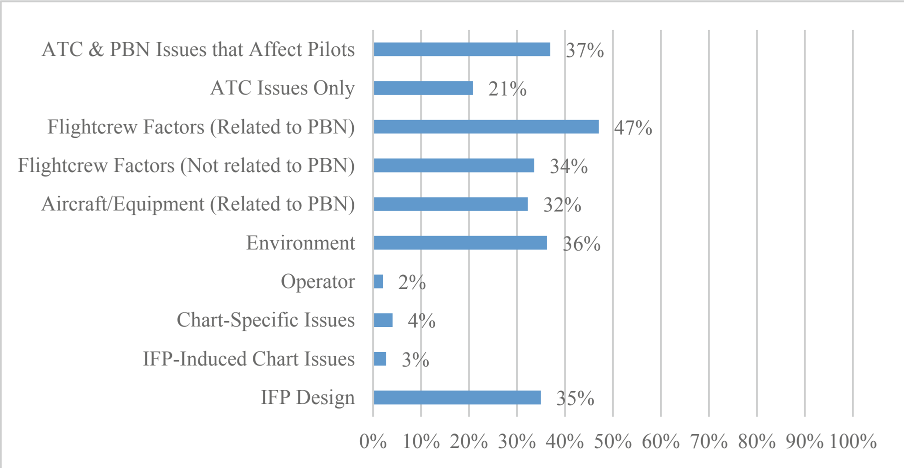


Fig. 1. Prevalence of the coded main factors in the event dataset.

Table 2. Event outcomes.

Outcome	Count
Altitude deviation	60
Speed deviation	5
Loss of separation from other aircraft	12
No deviation	38
Miscellaneous (other)	36
Insufficient information	12

4 Summary

This study was conducted to understand the relationship between PBN and flight deck operational complexity in real events. We developed a comprehensive rubric to code ASRS events. The rubric was based on a long history of research on PBN complexities from a flightcrew perspective. Because of this past research, we were able to separate (at least to some extent) the impact of PBN from operational complexity in general. We were not able to determine patterns of factor interactions from this analysis. A larger dataset might, or might not, uncover patterns. Such patterns may not exist, given the uncertainty and variability associated with operational complexity.

We noticed in the analysis that there are degrees of relatedness to PBN; it is not a simple yes or no answer. Sometimes factors were clearly either related or unrelated to PBN, but sometimes we had to make informed, consistent judgments or inferences about the relationship of an event or factor to PBN. For example, we sometimes considered whether an event would have unfolded the same way had the flight been on a conventional (not a PBN) IFP. If yes, then the factor was generic, not specific to PBN.

We also found several events with *No Deviation* where crew actions prevented a possible deviation, demonstrating resilience. We found examples of these behaviors even though the narratives have many limitations as noted earlier. This finding gives us hope of discovering more about how pilots increase the safety of flight operations.

Our analysis identified many operational complexity factors related to PBN. Many of these factors were related to flightcrew and ATC behaviors. We conclude that PBN appears to magnify the effects of operational complexity in the reports we analyzed. Whenever pilots have incomplete understanding, PBN operations can reveal those weaknesses. Because pilots are more reliant on automated systems to fly PBN IFPs, for example, it is important for them to develop a detailed mental model of how those systems work under many different PBN scenarios. Pilots should learn, for example, how to deal with anomalous system behavior.

Our results verify what is already well known among experienced pilots. Most of the time, pilots gain this expertise by simply doing their jobs, flying to different airports and using PBN every day. Well-informed pilots know, for example, that the transition from an arrival to the approach is an area of risk. STARs connect to different types of approaches in different ways and there is little room for error because the aircraft's energy profile must be managed carefully during this transition. Pilots know this because they have developed a detailed understanding of their automated systems and the route of flight, and they understand how PBN is designed to work. These pilots are more likely to understand the factors they need to consider to determine whether to accept an ATC instruction or say *"unable"*. All pilots would benefit from training that helps them achieve this level of understanding, especially for situations that are new, because NextGen PBN IFPs will continue to evolve over the years.

Acknowledgements. This document was prepared for the FAA under the FA6YCE Interagency Agreement (Human Factors of Flight Operations and Technologies) with funding from the NextGen Human Factors Division (ANG-C1) via the Fiscal Year 2017 RE&D System Development – NextGen Air/Ground Integration Project Level Agreement, Task 12. The Volpe Center thanks the FAA technical sponsors, Dr. Kathy Abbott and Jeff Kerr, and the program manager, Sabreena Azam. We thank NASA for selecting the reports in the study. We also thank Dr. Mark Shelhamer of the Volpe Center for his insights and feedback on this effort. The views expressed herein are those of the authors and do not necessarily reflect the views of the Volpe National Transportation Systems Center or the United States Department of Transportation.

References

1. Butchibabu, A., Midkiff, A., Kendra, A., Hansman, R.J., Chandra, D.C.: Analysis of safety reports involving area navigation and required navigation performance procedures. In: Proceedings of the International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero 2010), Cape Canaveral, FL (2010)
2. Chandra, D.C., Grayhem, R.J., Butchibabu, A.: Area navigation and required navigation performance procedures and depictions, DOT-VNTSC-FAA-12-10; DOT/FAA/TC-12/8. U.S. DOT Volpe National Transportation Systems Center, Cambridge, MA (2012)
3. Chandra, D.C., Markunas, R.: Line pilot perspectives on complexity of terminal instrument flight procedures. In: 35th Digital Avionics Systems Conference, Sacramento, CA (2016)

4. Funk, K.: Cockpit task management: preliminary definitions, normative theory, error taxonomy, and design recommendations. Int. J. Aviat. Psychol. **1**(4), 271–285 (1991)
5. Wu, Y., Frizelle, G., Efstathiou, J.: A study on the cost of operational complexity in customer-supplier systems. Int. J. Prod. Econ. **106**, 217–229 (2007)
6. Lazarus, R.S., Folkman, S.: Stress, Appraisal, and Coping. Springer, New York (1984)
7. Dismukes, K.R., Kochan, J.A., Goldsmith, T.: Flight crew errors in challenging and stressful situations. Aviat. Psychol. Appl. Hum. Factors **8**(1), 35–46 (2018)
8. Hollnagel, E.: The four cornerstones of resilience engineering. In: Hollnagel, E., Dekker, S. (eds.) Resilience Engineering Perspectives: Preparation and Restoration, vol. 2, pp. 117–133. Ashgate, Farnham (2009)
9. Chandra, D.C., Sparko, A., Kendra, A., Kochan, J.: Operational complexity in performance-based navigation (PBN) arrival and approach instrument flight procedures (IFPs), DOT-VNTSC-FAA-20-02. U.S. DOT Volpe National Transportation Systems Center, Cambridge, MA (2020)
10. Hatano, G., Inagaki, K.: Two courses of expertise (1984). http://hdl.handle.net/2115/25206

Towards a Glossary of Aviation Communication Factors

Simon Cookson(✉)

Aviation Management Department, J. F. Oberlin University,
3-23-1 Hyakunin-cho, Shinjuku-ku, Tokyo 169-0073, Japan
scookson@obirin.ac.jp

Abstract. Air transport is now so common that airlines carried 4.3 billion passengers in 2018 [1]. Most passengers travel safely. However, the air transport system is complex and as a result accidents continue to happen. Some accidents involve communication breakdowns between pilots and air traffic control (ATC). The aim of this project is to construct a glossary of factors that may lead to pilot-ATC communication breakdowns. The glossary is based on information collected from: (1) a review of literature in various research fields, and (2) interviews with pilots at two major airlines in Europe and the Middle East. The project is ongoing. This paper describes the methodology and preliminary results from the pilot interviews. It also includes a list of 68 communication factors as well as one sample glossary entry. The glossary will contribute to aviation safety by raising awareness of communication factors that cause accidents.

Keywords: Airline accidents & incidents · Aviation safety · Code switching · Communication breakdown · Pilot-ATC communication · Pilot training

1 Introduction

"Communication breakdown" is a phrase used in aviation to cover all manner of communication problems. There are, for example, more than 17,800 incidents involving communication breakdowns in NASA's ASRS incident database [2]. However, communication breakdown is a general construct that does not explain *specific* communication processes [3].

The aim of this project is to compile a glossary that breaks the concept of communication breakdown into discrete communication factors. Some factors are small-scale features of speech such as hesitation or repetition. Others are larger-scale phenomena that impact communication such as startle effect. This glossary is an alphabetic list of core terms in the domain of aviation. The terms are defined and explained, with examples from actual accidents and incidents.

The glossary is intended to raise awareness of communication factors that contribute to accidents. It provides a vocabulary for discussing communication breakdowns, and bridges a terminology gap between linguists on the one hand, and flight crew and air traffic controllers on the other. It is written in a non-technical style to make

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 539–546, 2020.
https://doi.org/10.1007/978-3-030-50943-9_68

the contents easily accessible. The target audience includes frontline operators – pilots and controllers – and also researchers interested in aviation communication.

The project is ongoing. This paper outlines the methodology used to compile the glossary, and describes the current state of the project. It includes: a list of 68 communication factors that have been identified so far; one sample entry from the glossary (code switching); and preliminary results from interviews with pilots.

2 Method

Figure 1 shows the 3-stage cyclical process used to compile the glossary. In stage one, an initial list of communication factors is compiled based on an extensive literature review. In stage two, a glossary entry template is designed, and entries are created for each communication factor. In stage three, pilot interviews are conducted to collect feedback about the list of factors, the template and the glossary entries. The cycle is then repeated, and the list of factors and the entries are revised based on the feedback from pilots.

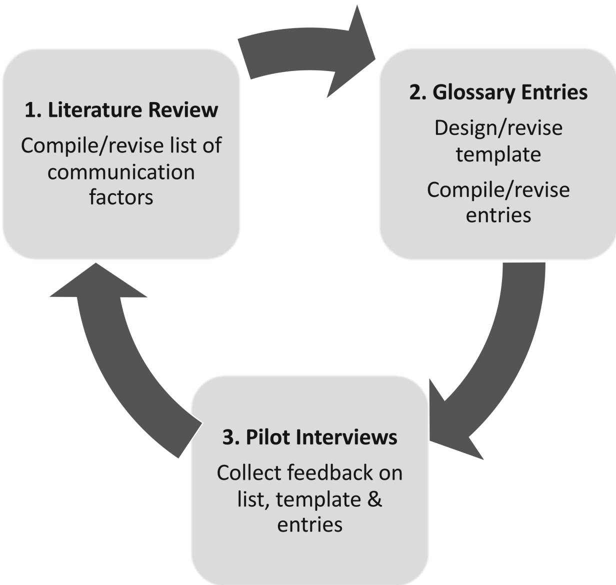


Fig. 1. Process for compiling glossary.

The literature review in the first stage brings together information from fields including sociolinguistics, cognitive psychology, social psychology, aviation human factors and accident investigation. Based on this information, an initial list of factors was compiled which has subsequently been modified with feedback from pilots. The list includes 68 factors, as shown in Appendix 1. The factors are divided into six categories: the body (e.g., startle effect), interacting with people (e.g., authority gradient), the radio (e.g., message length), ways of speaking (e.g., code switching), words and grammar (e.g., plain language) and the workplace (e.g., noise).

In the second stage glossary entries are compiled. Each entry describes one factor associated with communication problems that pilots and air traffic controllers may experience. The entry includes: a definition; a description of "language indicators" associated with the factor; related accidents or incidents; and references for further reading. In addition, many entries have a brief dialogue illustrating the factor in an aviation context. The dialogues are taken from accident or incident reports, and show the actual language use of pilots and controllers.

The third stage involves semi-structured interviews with pilots at two major airlines based in Europe and the Middle East. The purpose of the interviews is to collect feedback on: (1) the template for glossary entries; (2) the list of factors; and (3) individual glossary entries. They are one-to-one interviews, recorded with the consent of the interviewees. The interviews include the following questions:

Q1. What do you think about the template for glossary entries?
(Comment on its usefulness, usability, format and content.)
Q2. What do you think about the list of communication factors?
(Suggest any changes you think should be made to the list, including additions, deletions and revisions. Also, indicate the factors most relevant to your work.)
Q3. What do you think about the individual glossary entries?
(Suggest any changes you think should be made, for example, to improve the readability of an entry or to rectify omissions. Also, describe any experiences you have had related to the entry.)

3 Sample Glossary Entry

At the time of writing, glossary entries for 27 communication factors have been completed (indicated by an asterisk in Appendix 1). The entry for code switching is reproduced in Fig. 2. First there is a brief definition of the factor, followed by a description of the communication problem it may lead to. The problem is expressed as a factual conditional sentence based on a format used for "findings as to risk" in Transportation Safety Board of Canada accident reports.

The next section, "language indicator", shows linguistic features associated with the factor. For code switching, these features include a change of language (or dialect) and hesitation. A dialogue box shows a short excerpt of authentic pilot communication taken from a transcript in an accident or incident report. Colour-coding links the language indicator to instances of language in the dialogue box.

As shown in the code switching example, there is a list of accidents or incidents involving the factor [4–9], and a "further reading" section describing publications that provide more detailed information [10, 11]. A reference list at the end of the glossary includes the accident reports and sources named in these sections.

Each glossary entry presents a single communication factor. In actual flight operations, two or more factors may occur simultaneously. Indeed, it is widely accepted that accidents invariably involve the interaction of multiple factors [12]. For instance, unintentional *code switching* can occur when a second language (L2) speaker is placed under *time pressure* or *stress*, and it may be accompanied by *hesitation*. To highlight these connections, each entry ends with a list of related factors.

CODE SWITCHING

Definition A speaker alternates between different languages or dialects during a conversation. Code switching may be *intentional* or *unintentional*. Intentional code switching includes a pilot speaking one language (eg: English) and adding a greeting in another language (eg: "buenos dias").

Communication risk If a pilot or controller switches between different languages, listeners may not understand important information.

Language indicator In the dialogue, the controller **changes language** from English to Serbo-Croatian. This unintentional code switching is preceded by **hesitation**.

Dialogue Upper sector controller at Zagreb speaking with Inex Adria Airways flight JP550, on September 10, 1976. L1 of controller and pilots was Serbo-Croatian.[17]

TIME (GMT)	SPEAKER	SPEECH	TRANSLATION
1014:14	Controller	What is your present level?	
1014:17	JP550	327	
1014:22	Controller	**[Stuttering] ... e ... zadržite se za sada na toj visini i javite prolazak Zagreba**	…e… maintain now that level and report passing Zagreb
1014:27	JP550	Kojoj visini?	What level?

Accidents & incidents This factor contributed to:

- 1976 mid-air collision of Inex Adria Airways 550 and British Airways 476 above Zagreb, former Yugoslavia (AAIB, 1977, 1982; see **Dialogue**)
- 2000 runway collision of Streamline Aviation 200 and Air Liberté 8807 in Paris, France (BEA, 2001)
- 2001 runway collision of SAS 686 and Cessna 525A Citation in Milan, Italy (ANSV, 2004)
- 2012 loss of separation between Bombardier BD-700 and NetJets Europe 599U near Ibiza, Spain (CIAIAC, 2014)
- 2018 take-off without clearance of Cessna 525A Citation at Reykjavik, Iceland (RNSA, 2019)

Further reading Unintentional code switching sometimes happens when a second language (L2) speaker is under time pressure or stress. Due to the high cognitive load required to speak the L2, the speaker may revert to the first language (L1). Ganushchak and Schiller (2009) describe the effect of time pressure on code switching. Weston and Hurst (1982) discuss code switching in the 1976 mid-air collision over Zagreb.

Related factors ACCOMMODATION, BLOCKED TRANSMISSION, HESITATION, STRESS, STYLE SHIFTING, TIME PRESSURE

Fig. 2. Glossary entry for code switching.

4 Pilot Interviews

Interviews have started with pilots from a major European airline. Some preliminary results from the interviews are as follows:

- Communication factors – A number of revisions have been made to the initial list of factors. For example, at the suggestion of pilots, hearing loss (especially age-related) and headphone quality were added to the body and workplace categories respectively. The interviews have also highlighted factors that the pilots consider especially salient (e.g., distraction) and terms they are unfamiliar with (e.g., community of practice, diffusion of responsibility, in-group bias, avoidance and interlanguage).
- Template for entries – The template design was modified several times during the initial literature review, incorporating input from a sociolinguist and aviation researchers. Since the start of the pilot interviews, only minor changes have been made, such as simplifying section headings and standardizing the dialogue box format. (The dialogues are taken from transcripts in accident and incident reports. Depending on the investigating agency, various methods are used to express translated text, cockpit sounds, etc., in the original transcripts.)
- Glossary entries – Revisions have been made to several entries. For instance, the entry for call sign confusion initially focused on productive errors made by air traffic controllers. Based on interview feedback, it has now been modified to also include receptive errors made by pilots.
- Variation by region & flight phase – An interesting theme to emerge is the suggestion that some factors are likely to occur in certain regions of the world, or in particular flight phases. A British captain who regularly flies long-haul between Europe, Asia and the Americas, commented that code switching by air traffic controllers was, in his experience, most prevalent in France: "even in TMAs [terminal manoeuvring areas] and airways traffic they [controllers] will still talk to some of the French traffic in French … it's quite frustrating."
- Pilots' experiences – During the interviews pilots have related a number of anecdotes related to specific communication factors. Examples include: use of *plain language* by a tug driver at Heathrow Airport in London; use of *non-standard phraseology* by controllers during the approach to Tokyo Haneda Airport; and *high rate of speech* by controllers in India.

5 Conclusion

This project is compiling a glossary that provides a vocabulary for discussing airline accidents or incidents involving communication breakdowns. The intention is to publish the glossary, which could be used in recurrent pilot training to facilitate discussions of communication problems.

An extensive literature review has been completed and a list of 68 communication factors compiled. A template for glossary entries has been designed and 27 entries have been completed so far. Pilot interviews are being conducted to collect feedback on the template, the list of factors and the glossary entries. This project is ongoing: the remaining entries will be created; interviews will continue with pilots from two major airlines in Europe and the Middle East; and the glossary entries will be revised based on interview feedback.

The primary value of the project is that it assembles information from a range of sources (including the literature review and pilot interviews) into a single easy-to-use volume. At present, a lot of valuable information is difficult to access because it is spread over numerous publications. Accident reports are valuable sources of information about communication problems, but the information can be difficult to find because the reports are often lengthy and each report typically only addresses one factor. (See, for example, the appendix about fatigue in [13]).

During the course of the project, two additional areas have been identified in which this research may be extended in future: (1) mapping the occurrence of factors by region and flight phase; and (2) creating scenario-based training workshops based on anecdotes collected during the pilot interviews.

Acknowledgments. I would like to express my gratitude to the pilots who have participated in the interviews, and also to Professor John Maher at International Christian University, Tokyo, for his considerable input into the design of the glossary.

Appendix 1: Communication Factors

The table below shows the factors being examined in this project. Glossary entries have been created for the factors marked with an asterisk (*) (Table 1).

Table 1. List of communication factors.

THE BODY
1. Fatigue *
2. Hearing loss
3. Illness
4. Medication
5. Startle effect *
6. Stress *

INTERACTING WITH PEOPLE
7. Authority gradient *
8. Call sign confusion *
9. Community of practice
10. Comprehension
11. Courtesy bias
12. Diffusion of responsibility
13. Expectation bias
14. Face
15. Groupthink
16. Information bias
17. In-group bias
18. Intelligibility
19. Overhear
20. Speech community
21. Sterile cockpit *
22. Stigmatization

THE RADIO
23. Appended message *
24. Blocked transmission *
25. Frequency change
26. Message complexity
27. Message length *
28. Microphone clipping *
29. Stuck microphone

WAYS OF SPEAKING
30. Accent
31. Accommodation *
32. Ambiguity
33. Avoidance *

WAYS OF SPEAKING (continued)
34. Code switching *
35. First language interference *
36. Fluency
37. Hesitation
38. Misfire
39. Multiple negation
40. Non-speaking
41. Number transposition *
42. Omission *
43. Politeness
44. Pronunciation *
45. Rate of speech *
46. Repair
47. Repetition *
48. Sarcasm
49. Style shifting *

WORDS & GRAMMAR
50. Abbreviation
51. Argot
52. Colloquialism
53. Expletive *
54. Grammar
55. Homophone
56. Idiom *
57. Interlanguage
58. Loanword
59. Phrasal verb *
60. Plain language *
61. Standard phraseology *
62. Vocabulary

THE WORKPLACE
63. Distraction *
64. Headphone quality
65. Noise *
66. Temperature
67. Time pressure
68. Workload

References

1. International Air Transport Association: IATA Annual Review 2019. IATA, Montreal (2019)
2. Aviation Safety Reporting System (ASRS). https://asrs.arc.nasa.gov/index.html
3. Cookson, S.: The concept of communication breakdown: useful fallacy or time for overhaul? Presentation at the 19th International Symposium on Aviation Psychology, Dayton, OH, 8–11 May 2017
4. Air Accidents Investigation Branch: British Airways Trident G-AWZT, Inex-Adria DC-9 YU-AJR: Report on the Collision in the Zagreb Area, Yugoslavia, on 10 September 1976. Aircraft Accident Report 5/77. AAIB, London (1977)
5. Air Accidents Investigation Branch: British Airways Trident G-AWZT, Inex-Adria DC-9 YU-AJR: Report on the Collision in the Zagreb Area, Yugoslavia, on 10 September 1976. Aircraft Accident Report 9/82. AAIB, London (1982)
6. Bureau Enquêtes-Accidents: Accident on 25 May 2000 at Paris Charles de Gaulle (95) to Aircraft F-GHED Operated by Air Liberté and G-SSWN Operated by Streamline Aviation. Aircraft Accident Report F-ED000525. BEA, Paris (2001)
7. Agenzia Nazionale per la Sicurezza del Volo: Final Report on Accident involved aircraft Boeing MD-87, registration SE-DMA, and CESSNA 525-A, registration D-IEVX, Milano Linate airport, 8 October 2001. ANSV, Rome (2004)
8. Civil Aviation Accidents and Incidents Investigation Commission: REPORT IN-037/2012. CIAIAC, Madrid (2014)
9. Rannsóknarnefnd samgönguslysa: Final Report on Aircraft Serious Incident, 18-007F002, 11. January 2018, Reykjavik Airport (BIRK), Airplane Took Off Without a Takeoff Clearance. RNSA, Reykjavik (2019)
10. Weston, R., Hurst, R.: Zagreb One Four: Cleared to Collide?. Jason Aronson Inc., New York (1982)
11. Ganushchak, L.Y., Schiller, N.O.: Speaking one's second language under time pressure: an ERP study on verbal self-monitoring in German-Dutch bilinguals. Psychophysiology **46**(2), 410–419 (2009)
12. Dismukes, R.K., Berman, B.A., Loukopoulos, L.D.: The Limits of Expertise: Rethinking Pilot Error and the Causes of Airline Accidents. Ashgate Publishing Ltd., Aldershot (2007)
13. National Transportation Safety Board: Aircraft Accident Report: Uncontrolled Collision with Terrain, American International Airways flight 808, Douglas DC8-61, N814CK, U.S. Naval Air Station Guantanamo Bay, Cuba, 18 August 1993. NTSB/AAR-94/04. NTSB, Washington, D.C. (1994)

An Approach to Aerospace Design Integrating Crew Resource Management in Operational Environments

Tiziano Bernard[1(✉)], William A. Tuccio[1], Sebastien Boulnois[2], Aleksandar Tasic[1], and Lucas Stephane[3]

[1] Garmin International, 1200 E. 151st Street, Olathe, KS 66062, USA
{tiziano.bernard,bill.tuccio,aleks.tasic}@garmin.com
[2] Round Feather, 1551 Shelter Island Drive, San Diego, CA 92106, USA
sebastienboulnois@roundfeather.com
[3] Institutt for Energiteknikk, Os alle 5, 1777 Halden, Norway
lucas.stephane@ife.no

Abstract. The integration of crew resource management into the product design process for aerospace systems is discussed. Human-centered design methods - which can include human in the loop simulations, rapid prototyping, problem definition, evaluations, and verifications - often rely on operational knowledge of expert users to address human-systems integration needs. Modern integration efforts attempt to reduce the divide between user intentions and system state allowing for a safer system operation. However, in a multi-crew environment, crew interactions ought to be heavily taken into consideration. For this reason, crew resource management is recommended as a training supplement to both aerospace designers/researchers and expert users.

Keywords: Human factors · Human-systems integration · Crew resource management · Line operational simulations · Aerospace design process

1 Introduction

Product development and design cycles, which are present across disciplines including human-centered design (HCD) and human factors (HF) engineering, rely on solid fundamentals that require the formulation of a problem statement, state of the art research, rapid prototyping, and evaluations. Since rapid prototyping is an iterative process, human in the loop (HITL) simulations are often used. HITL simulations often exist in the form of simulator-based scenarios involving operational experts using and evaluating designs.

These topics have recently been under great analysis as they relate directly to two Boeing 737 MAX accidents. In relation to the 737 MAX accidents, the NTSB recommends, under A-19-11 "design enhancements […], pilot procedures, and/or training requirements, where needed, to minimize the potential for and safety impact of pilot actions that are inconsistent with manufacturer assumptions." Furthermore, the Board recommends the development of robust tools and methods for system design that uses

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 547–553, 2020.
https://doi.org/10.1007/978-3-030-50943-9_69

industry and human factors experts (A-19-13). Under A-19-15, the NTSB also suggests the development of aircraft system diagnostics tools to increase mediation between system state and pilot knowledge [1, 2]. The findings and recommendations in in essence request an enhancement of human-systems integration through the identification of realistic operational scenarios.

Aerospace HCD should not only be based on operational scenarios, but also on appropriate flight deck workflows. As with most life-critical operations, aeronautics and astronautics use checklists extensively to ensure a safe and expedited system utilization. Checklist operations are easily reproduced in a simulator. However, operations involving multi-crew environments almost universally involve a high level of crew-resource management (CRM) that is not easily reproducible without extensive training. This paper describes how CRM is incorporated in the training of usability testing participants and facilitators to enhance designs processes. The use of CRM - not only for evaluating, but also for designing aerospace systems (especially user interfaces (UIs) in avionics) - is pivotal in anticipating actual flight deck operations and outputting a higher fidelity design. Usability testing which appears complex and nonlinear due to unexpected CRM based interactions can be addressed in the design process. This enhancement of HSI necessitates designers and evaluators receive appropriate CRM training and use prototyping tools allowing HSI to be addressed.

2 The Role of Experts in the Design Cycle

HCD of aerospace-related life-critical systems involves the collaboration of research, design, organizational, and operational experts. These collaborations happen at different stages of design. Leifer, Meinel and Plattner proposed the human-centered design thinking process that includes the five stages in their ideal theoretical and linear sequence shown in Fig. 1 (left visual) [3].

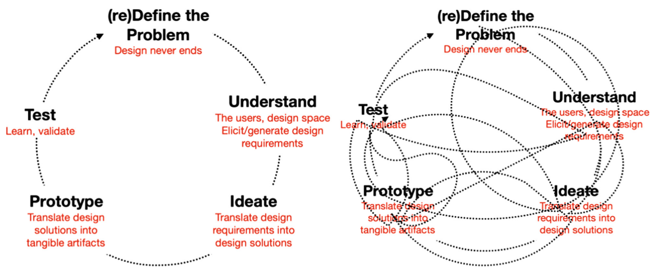


Fig. 1. Theoretical design stages (left) vs. nonlinear design stages (right) (adapted from [1])

In practice though, project-specific elements and the level of expertise of the research team (e.g. availability/accessibility of operational experts, budget/time constraints, extra research needed) lead to performing these steps in a nonlinear agile way (Fig. 1, right visual). An example of sequence is provided for describing the role of each expert in the design steps.

The researchers' role is to unpack operational experts' context, activities and interactions with organizational experts or systems (i.e. "understand" phase), so they can define potential underlying issues (i.e. "define the problem(s)" phase). This can be done by interviewing and/or observing them in their work environment. The researchers present their findings and issues to the designers so that they can together translate results into design requirements and design solutions (i.e. "ideate" phase). The designers collaborate with the researchers to translate design solutions into tangible prototypes (i.e. "prototype" phase). The researchers evaluate the usability of the prototype with organizational and operational experts in context and collect feedback determining what stage has to be performed next. Ultimately, the goal is to improve the maturity of the system so that it can be "engineered".

For optimal design, researchers should elicit the knowledge of organizational and operational experts and test any prototype in their real work environment. In the context of CRM, observing pilots' work, or testing a prototype during a real flight may threaten its safety. Thus, one solution is to use HITL simulations in order to mimic this work environment as best as possible.

3 Simulation Tools for Rapid Prototyping

As usability experts and designers begin to formulate ideas and concepts, drawing tools such as vignettes or digital imaging allow designers and researchers to better understand designs and formulate opinions. Although this is indeed a good start, aviation life-critical systems need to follow pre-determined standards, design philosophies, and certification requirements that limit design possibilities. This limitation may benefit the design process, as it allows designers and researchers to begin "ahead" already having in mind company design philosophies, UI guidelines, etc. What the designer is therefore allowed to do is focus on HSI.

HSI, a discipline that focuses around the concepts of people, organizations, and technology, is highly visible in UI interactions where users constantly manipulate the system to achieve a certain system state [4]. The mediation from the UI to the user is also part of HSI since the system is responsible for advising the user of the system's state [5, 6].

To be able to simulate system response to user interaction, static images or drawings are not sufficient. These do not allow users to realistically experience interactions. Instead, prototyping tools that allow realistic system animations and workflows fill the void and allow not only realistic operational scenarios to be enacted, but to account for CRM.

4 Concerning Crew Resource Management

The most recent FAA Advisory Circular (AC) on Crew Resource Management (CRM) Training defines CRM as the "*effective use of all available resources: human resources, hardware, and information. Other groups routinely working with the cockpit crew, who are involved in decisions required to operate a flight safely, are also essential participants in an effective CRM process*" [7]. In defining CRM training the

FAA says it is "one way of addressing the challenge of optimizing the human/machine interface and accompanying interpersonal activities."

Significant to this paper focusing on design challenges of usable avionics systems is the human/machine interface, situation awareness, and dealing with automated systems. One way of thinking about these interactions for a crew is the idea of shared/distributed cognition. In their ethnographic study of flight deck interactions, Hutchins and Klausen observe information processing is a combination of information stored in the mind of each pilot, communications systems (external or intracockpit), and machine feedback [8]. Each information element is not only its state, but its dynamics throughout the system.

Information elements in the SHELL model consider the interfaces between software, hardware, environment, and the system's livewires (people) [9], which we abbreviate herein to the machine-pilot-copilot, where the crew act as separate, independent minds. How these independents mind achieve distributed cognition (or fail to do so) can be found through experimentation, simulator studies, or accident investigation.

Real world examples of distributed cognitions triumphs and failures at each nexus point of the machine-pilot-copilot system can be found in accident reports. *Machine-pilot-copilot in sync* was shown when a crew dealt with the complex CRM-centric procedures of a failed gear system by following checklists, coordinating resources, and having adequate systems information to accurately gauge the state of the machine [10]. In another gear case, the *pilot-copilot were in sync with each other but out of sync with the machine* as systems state information and checklists were inadequate [11]; as one submission surmised, "The distinctive interaction between humans, humans, aircraft, checklists and the operational environment, makes the checklist problem a true human factors issue" [12]. *Pilot-copilot out of sync with each other, each with their own mental model of the machine*, occurred during a hydraulic failure on final approach leading to the pilot trying to reject the landing and another trying to abort the landing [13]; as one submission surmised, "The decision to land without determining beforehand the availability of braking capability defeated the CRM principal of maintaining situation awareness of the true nature of the problem, and the likely issues that may manifest themselves upon landing. The crew did not discuss alternatives and inoperative systems, including emergency braking procedures, and did not coordinate who was to activate the emergency brakes" [13]. And finally, *all elements can be out of sync with each other*, as occurred when pilot/copilot communications broke down during a complex circling approach; whereby, the pilot violated standard operating procedures and lost situational awareness of the aircraft automation state, while the copilot's mental model was indeterminate, and both were out of sync with what the machine was actually doing [19]. Investigators surmised, "FO [copilot] simply remained a passive bystander in the cockpit and did not participate as an effective team member failing to supplement/compliment or to correct the errors of his captain [pilot] assertively in line with the teachings of CRM due to Captain's behavior in the flight" [14].

5 Approach to Operational CRM Training

As these examples show, crew-machine interactions crucially depend upon professional competence fusing together technical and CRM skills. One important training method contributing to professional competence are gate-to-gate flight simulator scenarios, known generally as Line Operational Simulations (LOS), which have matured since the mid-1970s [15].

The complete evolution of Line Operational Simulations (LOS) has culminated with implementation of Title 14, CFR Part 121, Subpart Y, Advanced Qualification Training Programs (AQP). AQP programs encompass proficiency based formalized string of LOS sessions to form an "all encompassing" training program that evaluates not only technical crew skills but grade and evaluate crews' ability to effectively manage line operations using CRM and Threat and Error Management (TEM) skills [16].

LOS scenarios are developed to meet specific training or evaluation objectives often based upon operational trends collected from such things as Aviation Safety Action Programs (ASAPs) and more broadly, Flight Operational Quality Assurance (FOQA). In the case of avionics manufacturers training their engineers, scenario objectives may develop from trends observed in customer feedback, software issue tracking tools, and plans for future product maturation.

With scenario objectives in hand, scenario details are constructed for a line-representative flight, which will ideally span the lifecycle from preflight thought shutdown. Incorporated in the scenario are natural cues and distractions, rather than facilitator-initiated cues to keep the environment as realistic as possible [15]. During the scenario, all phases of flight should occur in real-time and mistakes should be allowed to play out just as they would in normal flight operations; by so doing, crews can learn and practice error management strategies [15].

Post-LOS debriefs are critical to LOS success. For human factors engineers (HFE) engaging in LOS, the post-LOS session extracts issues where the systems could be improved in the crew-machine interactions.

6 Conclusions

As discussed in this paper, the design cycles (or processes) are dynamic, often non-linear, and with extensive ramifications within an organization's structure. For this reason, HFEs are involved in various stages of the design and implementation process. The application of HCD principles, enhanced through CRM training, have the potential to "linearize" the organizational aspects of design. In fact, while current organizational information systems - and in particular knowledge management systems - enable data storage and retrieval and analysis of HFE interventions, such HFE contributions are in general linked to the specific stage of the design and development process within the given project or program, intra or inter organizational [17]. By consequence, it is sometimes difficult to relate HFE deliverables across existing or future programs and projects. Furthermore, the evolution of human factors tools from document-based evaluation forms (e.g. surveys, rating scales) toward live data capturing (e.g. eye tracking) and finally cutting edge modeling and simulation including digital twins of both operators and their working environments, add to the complexity of HFE missions and tasks [18–20].

Past research revealed through hands-on practice that the design thinking approach enables organizations to span the complexity of HFE work and to clearly address it in well-defined steps for design and evaluation stages [21, 22]. In doing so, the CRM topics mentioned in this paper may be addressed and dealt with more efficiently using an agile design thinking process [3]. In conclusion, the integration of CRM training within the design process for aerospace systems can provide benefits to the design itself and, indirectly, to one's organizational framework.

Acknowledgements. Disclaimer: Any views, opinions, findings, conclusions, or recommendations expressed in this document are the authors' and do not necessarily reflect the views of Garmin International, Round Feather, or the Institutt for Energiteknikk.

References

1. Sumwalt, R.L., Landsberg, B., Homendy, J.: Assumptions used in the safety assessment process and the effects of multiple alerts and indications on pilot performance. Safety recommendation report, National Transportation Safety Board. Washington, D.C. (2019)
2. Moak, L., McDew, D.W., Haskins, G., Hylander, K., Grizzle, D.: Official report of the special committee to review the federal aviation administration's aircraft certification process. Official report, Department of Transportation. Washington, D.C. (2020)
3. Leifer, L., Meinel, C., Plattner, H.: Design Thinking: Understand - Improve - Apply. Springer, Heidelberg (2011)
4. Bernard, T., Stephane, L.: Human-systems integration methods in flight test: a consolidated framework for sociotechnical modeling during critical phases of flight in general aviation. In: Society of Flight Test Engineers 48th Annual Symposium (2017)
5. Bernard, T., Stephane, L., Boy, G.A.: Autonomous stall recovery dynamics as a prevention tool for general aviation loss of control. In: Stanton, N.A. (ed.) International Conference on Applied Human Factors and Ergonomics 2017, vol. 597, pp. 38–47. Springer, Heidelberg (2017)
6. Andre, A.D., Wickens, C.D.: Information processing and perceptual characteristics of display design: the role of emergent features and objects. Technical memorandum, United States Army Human Engineering Laboratory (1990)
7. Federal Aviation Administration: Crew Resource Management Training. Advisory Circular 120-51E. Washington, D.C. (2004)
8. Hutchins, E., Klausen, T.: Distributed cognition in an airline cockpit. In: Engestrom, Y., Middleton, D. (eds.) Cognition and Communication at Work, pp. 15–34. Cambridge University Press, New York (1996)
9. Stolzer, A.J., Halford, C.D., Goglia, J.J.: Safety management systems in aviation. Ashgate, Burlington (2008)
10. National Transportation Safety Board: Aviation Incident Final Report CEN11IA615. Washington, D.C. (2014)
11. National Transportation Safety Board: Aviation Incident Final Report ERA11IA316. Washington, D.C. (2012)

12. NetJets Association of Shared Aircraft Pilots [NJASAP]: Party submission of NJASAP to the NTSB 749QS Stewart-Newburgh, New York, ERA11IA316, National Transportation Safety Board Public Docket (2012)
13. Gulfstream Aerospace: G550 outfitting check flight accident: Gulfstream party submission, CEN11FA193, National Transportation Safety Board Public Docket (2013)
14. Pakistan Civil Aviation Authority [PCAA-SIB]: Investigation Report: Air Blue flight AQB-2020 A-321 Reg AP-BJB Pakistan crashed on 28 July 2010 at Margalla Hills, Islamabad. Pakistan (2011)
15. Federal Aviation Administration.: Flightcrew Member Line Operational Simulations: Line-Oriented Flight Training, Special Purpose Operational Training, Line Operational Evaluation. Advisory Circular 120-35D. Washington, D.C. (2015)
16. Merritt, A., Klinect, J.: Defensive Flying for Pilots: An Introduction to Threat and Error Management. The University of Texas Human Factors Research Project (2006)
17. Stephane, L.: User behavior patterns: gathering, analysis, simulation and prediction. In: Kurosu, M. (ed.) Human Centered Design. LNCS, vol. 5619, pp. 322–331. Springer, Heidelberg (2009)
18. Stephane, L.: Visual patterns in civil aircraft cockpits. In: Proceedings of Human Computer Interaction Aero, Seattle, pp. 20–22 (2006)
19. Stephane, L.: Eye tracking from a human factors perspective. In: Boy, G.A. (ed.) The Handbook of Human-Machine Interaction: A Human-Centered Design Approach, pp. 339–364. Ashgate, Burlington (2012)
20. Hildebrandt, M., Langstrand, J.P., Nguyen, H.T.: Synopticon: a real-time data fusion platform for behavioral research. In: AutomotiveUI 2019: 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Utrecht, Netherlands, pp. 494–497 (2019)
21. Plattner, H., Leifer, L., Meinel, C.: Design Thinking Research: Building Innovators. Springer, Cham (2015)
22. Boulnois, S., Stephane, L.: Human-centered design of a 3D-augmented strategic weather management system: first design loops. In: Congress of the International Ergonomics Association, pp. 555–575. Springer, Cham (2018)

Helicopter Noise Footprint Depiction During Simulated Flight for Training

Anna C. Trujillo[1(✉)], Eric Greenwood[2], and Daniel R. Hill[1]

[1] MS 152, NASA Langley Research Center, Hampton, VA 23681, USA
{anna.c.trujillo,daniel.r.hill}@nasa.gov
[2] Department of Aerospace Engineering, 229 Hammond Building, Penn State, University Park, PA 16802, USA
eric.greenwood@psu.edu

Abstract. To support the Revolutionary Vertical Lift Technology project's goal of addressing source noise and response for vertical takeoff and landing aircraft, the authors have developed a noise training aid. This training aid depicts the predicted ground noise footprint using a noise algorithm and a flight simulator. This paper describes the implementation of the noise training aid, feedback from potential users on improvements, and future plans that will eventually provide an inflight noise display.

Keywords: Human factors · Helicopter noise · Training aid

1 Introduction

The Revolutionary Vertical Lift Technology project is addressing source noise and response to reduce negative acoustic impacts of vertical takeoff and landing (VTOL) aircraft operations. While urban air mobility (UAM) promises to integrate VTOL aircraft into the transportation infrastructure, community acceptance of these operations is a growing concern. To address community noise, the second author has developed real-time helicopter noise modeling (for example see [1, 2]) that predicts the ground noise footprint for a particular helicopter model. At present, models have been constructed for the Bell 407B and Airbus AS350 with additional model development underway for several civil helicopters flown in the 2017 and 2019 NASA/FAA noise abatement flight test campaigns [3].

The next logical step is to make this information useful to the vehicle operator. Thus, this paper describes a depiction of the predicted ground noise footprint and related information on how to reduce the ground noise footprint. The initial methodology was to develop a noise training aid that pilots could use during simulated flight that indicated the predicted real-time noise footprint on the ground dependent on vehicle flight state, attitude, and position. A training aid was initially chosen because it is independent of vehicle cockpit layout and it is a natural precursor to an in-flight display, which will be highly dependent on the vehicle cockpit equipage.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 554–561, 2020.
https://doi.org/10.1007/978-3-030-50943-9_70

2 Architecture

2.1 Flight Simulator

X-Plane© 11[1] (X-Plane) is used for the flight simulator with the Bell 407 model available from xplane.org [4]. X-Plane provides the helicopter's attitude and orientation, velocities, engine rotation speed (Ω), and engine thrust (T) to the Noise Algorithm described below. A Puma Pro Flight Trainer [5] is used for the control inceptor.

Currently, X-Plane is run on a desktop using a single 2.5 GHz Intel® Core™ i7 CPU core. Initially the flight simulator computer was separate from the ground noise footprint display computer; therefore, data from the flight simulator was passed to the noise algorithm and noise footprint display using the NASA developed X-Plane Connect toolbox [6] via a crossover cable.

2.2 Ground Noise Footprint Display

The ground noise footprint display incorporates the second author's FRAME (Fundamental Rotorcraft Acoustic Modeling from Experiments) noise algorithm [2] in order to provide data for the predicted noise footprint, which is shown on a display programmed using Qt© XML version 5.11 [7, 8]. Qt XML was chosen because it is an easy to use cross-platform application framework that has a map plug-in, Qt Location [9].

Initially the noise algorithm and display were run on a laptop with an Intel® Core™ i5 CPU at 2.30 GHz with 16 GB of RAM. The computer is running a Windows™ 64-bit operating system.

Noise Algorithm. FRAME is first used offline to generalize a limited set of measured steady flight acoustic data for a single helicopter to the wide range of rotor operating conditions achievable during normal maneuvering flight. The acoustic state of the rotors are represented as a database of acoustic spheres describing the magnitude and direction of noise radiation as a function of operating condition. Then, in real time, the FRAME-QS (FRAME Quasi-Static) method determines the appropriate acoustic sphere to use based on the helicopter's physical characteristics and the data from the simulator. This data includes measured orientation of the helicopter [1]. FRAME-QS also uses effective flight path angle (γ_e), which is calculated by $\sin\gamma_e = \frac{\vec{v}\cdot\hat{n}_a}{V}$ [10] where V is longitudinal velocity (available from the simulator), $\vec{v}$ is the inertial frame velocity (derived from the simulator velocity vector (V and V_z) and heading, pitch, and roll), and $\hat{n}_a$ is the inertial tip path plane normal vector, which is related to the inertial acceleration. Therefore, $\hat{n}_a(t) = \frac{\vec{a}(t)}{|\vec{a}(t)|}$ where acceleration (a) is derived from changes in velocity (V and V_z). Also needed are advance ratio (μ) and thrust coefficient (C_T). Advance ratio is a function of the velocity vector and engine rotation speed (Ω), both provided from the simulator, in addition to the rotor radius ($\mu = f([V, V_z], \Omega, r)$). The thrust coefficient uses engine thrust (T) and engine rotation speed, all provided from the

1 The use of trademarks or names of manufacturers in this report is for accurate reporting and does not constitute an official endorsement, either expressed or implied, of such products or manufacturers by the National Aeronautics and Space Administration.

simulator, in addition to the rotor radius, area of the rotor, and air density ($C_T = f(T, \rho, \Omega, r)$).

Noise Display. The final noise training aid display is shown in Fig. 1. The display is meant for training; therefore, the information provided consists of a background map, a pitch and bank ladder, and four tapes—longitudinal acceleration (acceleration), speed, altitude, and vertical velocity. Pitch, acceleration, and altitude are displayed because these parameters directly affect the noise generated. Speed and vertical velocity were also included because these are common pilot control parameters and they are directly related to acceleration and altitude respectively.

Any values outside the maximum and minimum maneuvering limits of acceleration, speed, altitude, vertical velocity, and pitch for a particular helicopter model are represented by red on the tapes and pitch ladder. As an example for the Bell 407, the maneuvering limits for pitch are −30°–50° with a ceiling of approximately 18,700 ft [11–13].

When the noise level is greater than 110 dBA, it is depicted as orange. That noise level is described as almost painful [14]. Blue is used to depict noise level equal to 65 dBA, which is described as a moderate noise level [14]. The transition colors from orange to blue are yellow and green. The color palette was chosen so that there is minimal conflict with the alerting colors [12]. The 65 dBA level is chosen as the annoyance threshold because it is a common threshold for transportation noise as it is comparable to daytime background noise levels of suburban neighborhoods.

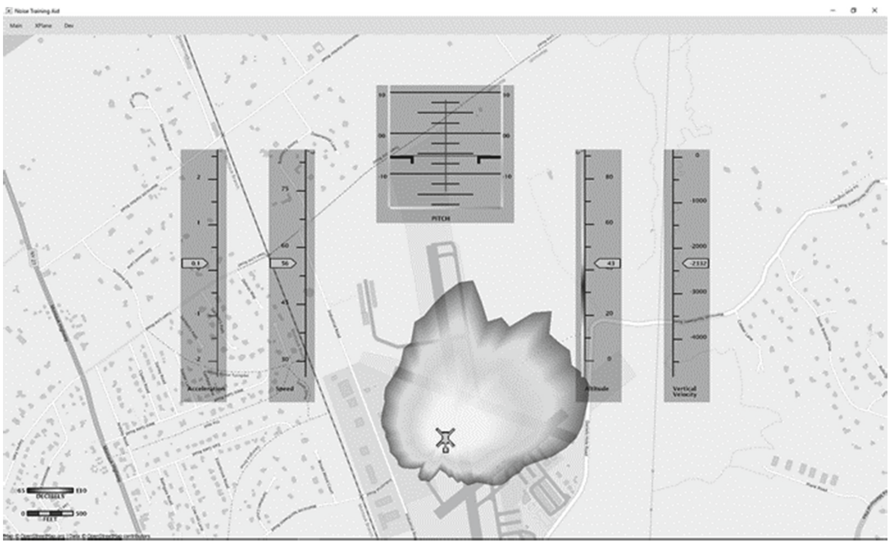


Fig. 1. Noise training aid display. Noise and map scales in the lower right corner were added for VFS 75th Annual Forum and the tape ribbons were added for EAA AirVenture Oshkosh 2019.

The noise footprint range is calculated along every 5° radial from the vehicle center to the distance where the annoyance threshold is reached. The noise level directly below the helicopter is computed and the color contours are linearly interpolated between this level and the noise threshold at the edge of the footprint. The helicopter icon's color matches the loudest noise level color.

The map is supplied by OpenStreetMap® [15]. The map is synchronized to the flight via latitude and longitude data from the simulator. The noise footprint is also scaled to the map based on the map scale. The map is oriented track up.

3 Exhibit Feedback

Feedback on the noise training aid was obtained from exhibiting it at three different forums.

3.1 HAI Heli-Expo 2019

The noise training aid as described above was exhibited at HAI Heli-Expo 2019. At this exhibit, the authors were able to obtain feedback from a population of potential users. This included private pilots and commercial operators. Most of the visitors had positive comments on the display but did offer suggestions for improvement.

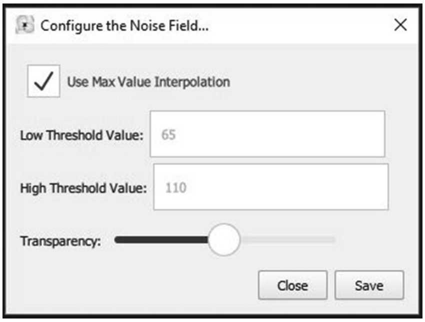


Fig. 2. Noise training aid dialog box to change noise threshold values.

Tour businesses that operate in quiet areas, such as national parks, suggested that the annoyance level should be lower in these regions because the overall noise level is typically very low. Therefore, the authors implemented an option to vary both noise footprint endpoints via a dialog box (Fig. 2).

Another improvement made based off user comments was to add a legend indicating map scale. The map scale can be changed via mouse scroll wheel or touchpad. The noise display now has a map scale legend in the lower left corner (Fig. 3). Additionally, a scale indicating the current endpoints of the noise threshold values was also added.

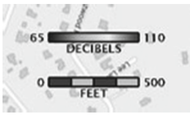


Fig. 3. Map scaling and noise threshold scaling values.

An improvement suggested by several pilots, both private and commercial operators, was to provide a playback capability. This function would allow for ingesting of previous flight data to review that flight's community noise exposure. This has not yet been implemented but work is being done to accomplish this. Tied to this capability is the ability to input a flight plan with waypoints in order to determine the optimal vehicle attitude and velocity to minimize an upcoming flight's ground noise footprint.

3.2 VFS 75th Annual Forum

The noise training aid with the map and noise threshold scaling values (Fig. 3) was exhibited at VFS 75th Annual Forum and Technology Display. At this exhibit, the noise training aid was demonstrated primarily to academicians and students, and to both commercial and governmental engineers and scientists. Once again, most visitors had positive feedback on the system. The primary suggestions for improvements centered around providing indications on the tapes and ladder of the attitude needed to minimize the predicted ground noise footprint and flight practice capability.

The authors were able to implement indications of vehicle attitude needed to minimize the predicted ground noise footprint. The values to minimize noise were calculated by using the FRAME-QS noise model to predict how the ground noise level would change for an incremental change in the corresponding quantity associated with the ribbon. The ribbon uses the same color coding as the noise footprint (Fig. 4). For the bank and pitch ladder, two possible implementations were developed: (i) using a U-shape where the pitch ribbon is shown on both sides of the ladder and (ii) using an L-shape where the pitch ribbon is shown only on the left side of the ladder.

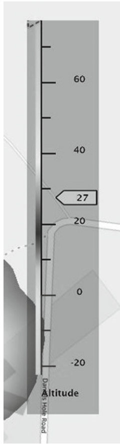


Fig. 4. Tape ribbon indicating the attitude needed to minimize predicted ground noise footprint. This example is for altitude.

3.3 EAA AirVenture Oshkosh 2019

At EAA AirVenture Oshkosh 2019, the noise training aid was exhibited with the ribbons indicating the attitude needed to minimize ground noise (see Fig. 4 for example). Furthermore, the noise training aid and the helicopter simulation were run on a single desktop computer. The desktop computer's operating system is 64-bit Windows™ 10 running two monitors—one for the X-Plane simulation and the second for the noise footprint display.

A variety of participants indicated the utility of the noise ribbons to minimize noise. Various pilots, both fixed-wing and rotor, thought the noise ribbons would be helpful during practice flights to find ways to minimize their noise footprint for flight paths they generally flew. Other participants indicated a need for an overall noise level for the duration of a flight or over sensitive communities.

4 Further Improvements

The next improvements made to the noise training aid are to finalize playback capabilities and the ability to add waypoints with straight leg connections. This will add to the utility of using the noise training aid for flight paths typically flown by being able to compare how different attitudes and velocities over the same flight path affects the ground noise footprint. Additionally, a time-integrated sound exposure level (SEL) for the whole flight based on the maximum noise generated every unit of time is being implemented. Current implementation is to use a summation of SELs [16]

$$L_{AE} = 10\log_{10}\int_{t_1}^{t_2} 10^{\frac{L_a(t)}{10}} \approx 10\log_{10}\sum_{n=0}^{n} 10^{\frac{L_a}{10}} \quad (1)$$

where L_{AE} is the SEL, t_n is time (in sec) and n is the time slice, and $L_a(t)$ and L_n are the noise levels in dBA at time t or time slice n. Lastly, noise exposure levels over defined sensitive areas is being calculated by projecting a noise level for an observer at a bearing and distance from the noise origin. This noise level is then used in Eq. (1). These improvements were showcased at HAI Heli-Expo 2020.

The authors also plan on using this training aid to depict trajectories optimized for noise minimization developed by the second author [17]. This would include bugs on the tapes and ladder indicating the required attitude needed to follow this trajectory. This capability could be used to practice upcoming noise minimized flights. Functionality of the recording capability is also being expanded along with improvements to its associated user interface so that it is more user friendly. This will allow users to playback practice runs on the noise training aid.

5 Noise Training Aid Usefulness and Follow-on Research

To determine the usefulness of the noise training aid, the authors took survey data at another exhibition. This survey also incorporated research that has started on developing a real-time in-flight display. The initial in-flight noise display is focusing on non-glass cockpit rotorcraft. Key requirements for this display are that its dimensions are within the typical mechanical dial footprint [18, 19] for use in non-glass cockpits and for that information to be garnered with just a glance. This requirement is necessary because the goal is to maintain pilot situation awareness outside the cockpit when at low altitudes rather than inside the cockpit.

Further out plans include developing a real-time in-flight display for a glass cockpit rotorcraft. Additionally, the authors would like to expand the work to include UAM operations.

6 Conclusion

To support Revolutionary Vertical Lift Technology project's goal of addressing source noise and response for VTOL aircraft, the authors have developed a noise training aid. This training aid depicts the predicted noise ground footprint using a noise algorithm and flight simulator. The training aid can easily be run on a laptop computer and can communicate with an off-the-shelf aircraft simulator; in this case, X-Plane© 11.

Feedback was obtained on the noise training aid at three exhibits: (i) HAI Heli-Expo 2019, (ii) VFS 75th Annual Forum, and (iii) EAA AirVenture Oshkosh 2019. In all cases, comments were positive with improvement suggestions made by potential end users. The final noise display after these comments were considered displays the predicted ground noise footprint with user modifiable threshold values. A pitch and bank ladder is available in addition to four tapes indicating longitudinal acceleration, speed, altitude, and vertical velocity. The ladder and tapes also include ribbons depicting the attitude needed to decrease noise below the low threshold.

Other improvements being worked on are to provide a playback capability of simulated and actual flights and depicted a noise minimized trajectory. Long-term work includes a survey to determine the usefulness of the noise training aid, adding additional noise models for other rotorcraft—including UAM concepts, and developing an in-flight predicted ground noise footprint display.

Acknowledgements. The authors would like to thank HAI Fly Neighborly and DOT Volpe for allowing us to share their exhibit booth at the HAI Heli-Expo 2019. The authors would also like to thank Gregory Harbert from Wyle Labs at NASA Ames Research Center for his help at the exhibit at VFS's 75th Annual Forum and Technology Display. This work is funded by NASA's Revolutionary Vertical Lift Technology Project under the Advanced Air Vehicles Program.

References

1. Greenwood, E.: Real time helicopter noise modeling for pilot community noise awareness. In: NOISE-CON 2017, p. 9 (2017)
2. Greenwood, E., Rau, R.: A maneuvering flight model for helicopter mission planning. J. Am. Helicopter Soc. **65**, 2 (2020)
3. Watts, M.E., Greenwood, E., Smith, C.D., Stephenson, J.H.: Noise Abatement Flight Test Data Report. NASA, NASA Langley Research Center (2019)
4. Dreamfoil. https://store.x-plane.org/Bell-407-XP11_p_620.html
5. Puma. https://www.pro-flight-trainer.com/pages/our-story
6. Teubert, C., Watkins, J., Bole, B.: X-Plane Communication Toolbox (XPC). Tech Briefs: Engineering Solutions for Desing & Manufacturing 40, 1 (2016)
7. The Qt Company. https://www.qt.io/
8. The Qt Company. https://doc.qt.io/qt-5/qtxml-index.html
9. The Qt Company, Qt Location Open Street Map Plugin (2019). https://doc.qt.io/qt-5/location-plugin-osm.html
10. Gopalan, G.: Quasi-static acoustic mapping of helicopter blade-vortex interaction noise. Ph. D. thesis, University of Maryland, University of Maryland (2004)
11. Bell Helicopter: Bell 407 Model: Rotorcraft Flight Manual (BHT-407-FM-1). Bell Helicopter (2019)
12. U.S. Department of Transportation: Title 14: Aeronautics and Space. In: FAA (ed.) Electronic Code of Federal Regulations. U.S. Government Printing Office (2019)
13. Wikipedia.org. https://en.wikipedia.org/wiki/Bell_407
14. boomspeaker.com. https://boomspeaker.com/noise-level-chart-db-level-chart/
15. Brentner, K.S.: Rotary Wing Technology Acoustics. A Comprehensive Short Course in Rotary Wing Technology, p. 261. Pennsylvania State University, College Park, PA (2019)
16. OpenStreetMap Foundation. https://www.openstreetmap.org

17. Greenwood, E.: Dynamic replanning of low noise rotorcraft operations. In: Vertical Flight Society 75th Annual Forum and Technology Display, VFS, Philadelphia, PA, p. 9 (2019)
18. Avionics Mounts Inc. https://avionicsmounts.com/pdf/ARINC%20408A%20Adapter%20Plates.pdf
19. MSP Aviation Inc. https://msp-aviation.com/product/408a-arinc-housing-4-ati/

Predictive Safety Through Survey Interviewing - Developing a Task-Based Hazard Identification Survey Process in Offshore Helicopter Operations

Felipe A. C. Nascimento(✉), Arnab Majumdar, and Washington Y. Ochieng

Imperial College London, The Lloyd's Register Foundation Transport Risk Management Centre, South Kensington Campus, London SW7 2AZ, UK
f.nascimento09@alumni.imperial.ac.uk, {a.majumdar,w.ochieng}@imperial.ac.uk

Abstract. Offshore helicopters play a vital role in energy production worldwide and must be operated safely. Safety is underpinned by hazard identification, which aspires to be predictive and remain operationally relevant. A process to elicit pilots' operational hazard knowledge in a predictive manner is currently absent. This paper redresses this by developing a Task-Based Hazard Identification Survey Process which, through Talk-Through interviewing, collects data from a statistically representative sample of pilots based in specified regions. A factual and exhaustive hazards' template is formed, to which various statistical methods are applied. Subjected to multiple validation and reliability checks, the process delivers on the aspiration to be predictive on safety.

Keywords: Predictive safety · Survey methods · Hazard identification · Task analysis · Content analysis · Offshore helicopters

1 Introduction

Helicopters play a vital role in the movement of people and cargo to and from the installations of the oil and gas industry at sea, often in challenging environments where accidents tend to have serious impacts. Effective safety management is essential [1].

A key element of safety management is the identification of hazards. As in the wider aviation industry, hazard identification in offshore helicopter operations relies on the analysis of a multiple datasets categorised as reactive, proactive or predictive [2].

Predictive is defined as data collected from the analysis of system components and the environment, with the aim to identify potential future negative outcomes [2]. This open definition alludes to system analysis and qualitative outcome prediction, methods of which are typically hardware-centred and cannot fully identify context-based hazards prevalent in human-centred systems. Moreover, statistical modelling, out of which robust numerical predictions can be truly achieved, is disregarded.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 562–569, 2020.
https://doi.org/10.1007/978-3-030-50943-9_71

The context-based hazard variables necessary for statistical modelling can be elicited from pilots during surveys. However, aviation survey methods remain underdeveloped.

Most aviation surveys lack a factual task-based focus and, therefore, gather opinions and perceptions. Moreover, survey sampling strategies are typically qualitative and opportunistic, and outcome variable information is collected reactively only. Finally, the statistical relationships between hazard variables are not explored, leaving predictive safety as an unfulfilled aspiration.

Aimed at delivering predictive safety and exploiting the full potential of surveys, this paper builds on key past papers of the authors to develop the Task-Based Hazard Identification Survey Process explained below.

2 Process Explained

The process (Fig. 1) combines the human-centred focus of surveys with the factual character of task analysis. It enables a statistically representative sample of pilots to be selected in key offshore areas and, through Talk-Through interviewing, pilots' factual hazard knowledge to be elicited. Hazard data is then categorised by the recursive and exhaustive application of Grounded Theory and Template Analysis, followed by extensive statistical analysis through Content Analysis. This culminates with the fitting of predictive models based on the cases of key, meaningful and context-based outcome and explanatory hazard variables. Multiple rounds of reliability and validity checks ensure the quality of the hazard data and results.

2.1 Select Task, Geographical Scope and Sample Sizes

Well-discriminated phases of flight are amenable to task-based hazard identification [3]. The geographical scope reflects the intention of the survey, i.e., the level of generalisation pursued. The methods for estimation of respondents' sample sizes outlined in [3] should be used. Where statistical sampling methods prove difficult to implement, e.g., due to respondents' availability, quota sampling may be applied. This supports generalisation (and thus statistical inference) when the samples are deemed representative in light of other considerations (e.g., percentage representativeness) [4] and the interviewer controls for selection bias [3].

2.2 Preliminary Hazard Identification

This stage comprises of the review of the literature, applicable taxonomies and available safety data (i.e., accident and incident reports). The outcome is a seminal hazard template, if the knowledge available supports it.

The seminal template formed, if any, is then subjected to an initial qualitative assessment, which is a value judgement that reflects the level of trust in the template's believability, completeness and timeliness [4, 5]. The outcome is to enter either the template development cycle or Content Analysis procedure.

2.3 Template Development Cycle

Hazard data collection begins by using the Talk-Through task analysis technique, when practical issues (e.g., potential biases) should be considered and mitigated.

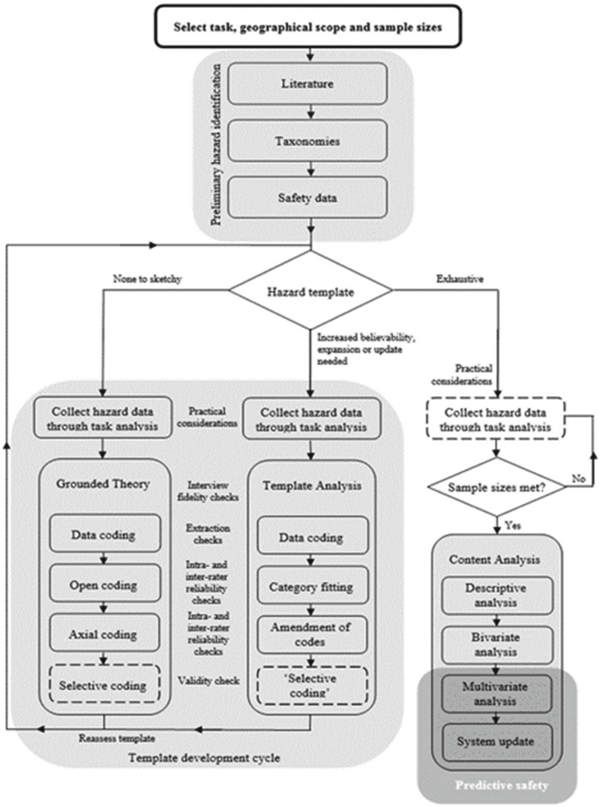


Fig. 1. Task-based hazard identification survey process

Subsequently, Grounded Theory is applied to categorise the data when no template can be drawn in the previous step [6]. With a template, Template Analysis is used [7].

Grounded Theory enables the generation of categories and theory from raw data and includes four steps: data coding, open coding, axial coding and selective coding. Selective coding does not aid the development of the template directly and, thus, was marked by a dashed outline in Fig. 1. However, it can be used for interim template validation. Validity and reliability are also checked through interview fidelity, statement extraction, intra- and inter-reliability checks conducted at various stages of the process.

Template Analysis improves a template's believability, completeness and timeliness. The steps are like those of Grounded Theory, only with terminological differences.

2.4 Template Reassessment

The template produced by Grounded Theory and/or improved by Template Analysis is subject to a new assessment of believability, scope and timeliness. For as long as the template needs improving, Template Analysis is conducted recursively. When the template is deemed exhaustive, more data are collected only where necessary to complete the sample sizes calculated. An exhaustive template is one in which the creation of new categories achieves clear diminishing returns in new rounds of data analysis. Sample sizes secured, quantitative hazard analysis is conducted.

2.5 Content Analysis

Content Analysis produces quantitative summaries of responses and allows for the analysis of the characteristics of those who commented versus those who failed to comment on particular topics (e.g., hazards).

Since there is no established statistical analysis procedure in Content Analysis [8], the steps described in Fig. 1 were established. Descriptive analysis offers an overview of the characteristics of the pilots interviewed, template's categories and codes, and frequencies of pilots who commented on each code and category; raw hazard statements are presented at the code level for clarity.

Bivariate analysis can be based on codes and/or selected main categories, the former producing a high-level overview of the hazards. The relationships between hazards and respondents' demographic variables are explored and codes may be cross-tabulated.

Multivariate analysis is where predictive safety occurs. Regression models are fitted to predict the odds of outcome variables occurring, based on the occurrence of explanatory and the events of demographic variables. Finally, the explanatory relationships found in the predictive models are used to update known system structures (e.g., [1]); however, this is not covered in this paper.

3 Process Implementation

3.1 Selection of Task, Geographical Scope and Sample Sizes

The nighttime is the worst condition for the occurrence of pilot-related accidents in worldwide offshore helicopter operations [4, 5]. The visual segment of instrument approaches is particularly hazardous [3] and was, therefore, the task selected.

Offshore helicopter accidents in the nighttime are endemic worldwide (ideal geographical scope), being more prevalent in the North Sea, Africa and Middle East [4, 5] (reduced scope). Due to practical reasons (e.g., access to respondents), regions sharing similar operational characteristics to those of Africa and the Middle East (as per the regional sampling strategy laid out in [9]) were selected: Brazil and Spain. In the North Sea, UK and Norway were selected because these are the countries with the largest volumes of offshore helicopter operations.

Quota sampling was applied because an element of convenience (related to pilots' limited time availability between successive flights) could not be avoided. Because no pilot demographics were associated with higher levels of risk experienced during nighttime approaches [3], all offshore helicopter pilots in the areas targeted formed the population of interest to which sampling was applied.

Sample size estimation was done on the basis of the number of nighttime deck landings [3]. Setting the ideal and tolerable Z for 95% and 90% confidence intervals, the sample sizes required were, ideally, 50 pilots and, tolerably, 34 pilots per geographical area (i.e., North Sea and non-North Sea, represented by Brazil and Spain).

3.2 Preliminary Hazard Identification and Template Development Cycle

Generic aviation taxonomies cannot accurately describe the hazards of offshore helicopter operations in the nighttime [3–5]. Their usefulness is limited by multiple factors such as poor definitions, non-mutually exclusive categories, incompatibility with helicopter system models, incompleteness with regards to human factors and a lack of predictive ability. Safety data (i.e., from accident and incident reports) are scarce, transparency lacks in many countries and data quality is generally low [4, 5].

The template created in [6] and developed in [7, 9] went through all the stages of the template development cycle. It is formed by exhaustive 103 hazard categories, grouped under 13 codes, segregated into two sections, which capture meaningful explanatory and outcome variables: 'Contextual Factors' and 'Impacts on Crews' [9]. The 'Contextual Factors' section is formed by 78 hazard categories grouped into eight codes, which reflect recognised system structures [1]. The 'Impacts on Crews' section is formed by 25 hazard categories grouped into five codes. Finally, with 39 respondents in the North Sea and 34 away from the North Sea, the sample sizes required for progression into Content Analysis (and predictive safety) were met.

3.3 Content Analysis

Descriptive and Bivariate Analyses. Descriptive analysis is described in [9]. The results of the bivariate demographic analysis are presented in Table 1. It shows that the hazards related to 'training', 'offshore installations', 'environmental conditions' and the 'regulator' were not specific to any pilot group neither experience level. Therefore, addressing these hazards should have an impact in many pilots, with the potential to improve safety considerably.

Regarding the effects of area of operation on the hazard codes, all the significant associations stemmed from a concentration of pilots who identified hazards in the North Sea. On the effects of aircraft category, the significant associations stemmed from a preponderance of comments by heavy twin-turbine (HT) helicopter pilots. All other significant results referred to some measure of pilot experience (e.g., licence, rank, various types of flying hours, number of night deck landings) with a greater number of experienced pilots commenting on the hazard codes.

Multivariate Analysis and Predictive Safety. Binomial logistic regression was used to predict membership to each code of 'impacts on crews' (outcome variables) based on membership of the eight 'contextual factors' codes and the events of the following demographic variables: area of operation, aircraft category, cockpit generation (i.e., dial versus glass) and number of night deck landings (explanatory/predictor variables). This explained the impact of each contextual factor/demographic variable on the odds of identifying each impact on crew when other contextual factors/demographic variables are kept constant. Interactions between predictor variables were not included in the model. The forward stepwise model fitting strategy based on the log-likelihood statistic was used. Neither multicollinearity nor overdispersion were present. The omnibus test of model coefficients revealed that the predictors as a set reliably distinguished between members and non-members of the impacts on crew codes when compared to a constant only model. The contribution of each variable to model prediction was assessed by the Wald criterion.

Table 2 summarises the results for the code 'perception' which, like the results for other codes not presented in this paper, agreed with the results of the bivariate and descriptive analyses. The odds ratio indicated that, when operating in the North Sea, the likelihood of commenting on issues related to 'perception' was 398.5% of the likelihood in other areas. On the other hand, comments on perceptual issues were only 22.2% as likely to happen when pilots did not identify hazards associated with the offshore installation. The regional effect observed might be caused by the notoriously poor weather conditions in the North Sea, which require more concentration from the pilots, consume their attentional resources and are prone to perceptual issues. The change in odds associated with the identification of hazards related to the offshore installation highlight that improving the installation environment has the potential to improve perceptual problems.

Table 1. Bivariate demographic analysis of hazard codes

		Contextual factors								Impacts on crews				
		Training	Procedures	Offshore installations	Aircraft / OEM	Internal factors	Environmental conditions	Organisational issues	Regulator	Handling	Attentional resources	Perception	Decision-making	Crew cooperation
	X^2 or FET													
Area		X	✓	X	X	X	X	X	X	✓	✓	✓	✓	X
Licence		X	X	X	X	X	X	X	X	✓	X	✓	X	X
Rank		X	X	X	✓	✓	X	X	X	✓	X	X	X	X
Aircraft category		X	X	X	X	X	X	X	X	✓	✓	✓	✓	✓
Cockpit generation		X	X	X	X	X	X	X	X	X	X	X	X	X
	MW													
Age		X	X	X	X	X	X*	X	X	X	X	X	X	X
Flying hours (FH), total		X	X	X	X	X	X	X	X	✓	X	✓	X	✓
FH IFR		X	✓	X	✓	✓	X	X	X	✓	✓	✓	✓	✓
FH night		X	X	X	X	✓	X	X	X	X	X	✓	X	✓
Night deck landing		X	X	X	✓	✓	X	X	X	✓	X	X	X	X
Time as an offshore helicopter pilot in the area		X	X	X	X	✓	X	✓	X	X	X	X	X	X

X^2 or FET Chi-square or Fisher's Exact test, as determined by the expected frequencies at the contingency tables.
MW Independent samples Mann-Whitney U test.
* Parametric data. Independent Samples t-test applied instead of MW. X Non-significant result: retain null hypothesis. ✓ Significant result: reject null hypothesis.

Null hypotheses X^2 and FET: identifying the hazard in the column is independent of membership to the categories of the row variable.
MW and Independent Samples t-test: the distribution of the quantities of the row variable is the same across those who identified and who did not identify the hazard in the column.

Table 2. Prediction of impacts on 'perception'

Omnibus test of model coefficients: $\chi^2(2)=15.029$, $p=0.001$			Prediction success: 68.5% (members: 86.0%; non-members: 43.3%)		
			95% CI for odds ratio		
		b (SE)	Lower	Odds ratio	Upper
Area	North Sea	1.382** (0.533)	1.403	3.985	11.316
Offshore installation	No comment	-1.506* (0.606)	0.068	0.222	0.728
Constant		0.068 (0.392)		1.070	
	$R^2_{Hosmer\ \&\ Lemeshow}= 0.991$; $R^2_{Cox\ \&\ Snell}=0.186$; $R^2_{Nagelkerke}= 0.251$			* $p<0.05$	** $p<0.01$

4 Conclusions

The Task-Based Hazard Identification Survey Process represents a transition from the bottom-up purely exploratory and inductive research of Grounded Theory towards an explanatory and deductive research with Content Analysis. The significant statistical relationships and differences across the pilots' hazard accounts identified are normally not evident from simple descriptive analyses and, therefore, the process provides an invaluable enhancement to present qualitative and quantitative safety survey methodologies, and a tool for predictive safety.

References

1. Nascimento, F.A.C., et al.: Fundamentals of safety management: the offshore helicopter transportation system model. Saf. Sci. **85**, 194–204 (2016)
2. ICAO, Safety Management Manual (SMM). ICAO, Montréal (2013)
3. Nascimento, F.A.C., et al.: Nighttime offshore helicopter operations: a survey of risk levels per phase of flight, recency requirements and visual approach technique. Aeronaut. J. **119** (1222), 1475–1498 (2015)
4. Nascimento, F.A.C., Majumdar, A., Ochieng, W.Y.: A 15-year multivariate analysis of worldwide offshore helicopter accidents. In: AHS 69th Annual Forum and Technology Display. AHS: Phoenix (2013)
5. Nascimento, F.A.C., Majumdar, A., Ochieng, W.Y.: Helicopter accident analysis. J. Navig. **67**(1), 145–161 (2014)
6. Nascimento, F.A.C., Jarvis, S., Majumdar, A.: Factors affecting safety during night visual approach segments for offshore helicopters. Aeronaut. J. **116**(1175), 303–322 (2012)
7. Nascimento, F.A.C., Majumdar, A., Jarvis, S.: Nighttime approaches to offshore installations in Brazil: Safety shortcomings experienced by helicopter pilots. Accid. Anal. Prev. **47**, 64–74 (2012)
8. Mostyn, B.: The content analysis of qualitative research data: a dynamic approach. In: Brenner, M., Brown, J., Canter, D. (eds.) The Research Interview: Uses and Approaches, pp. 115–145. Academic Press, London (1985)
9. Nascimento, F.A.C., et al.: A multistage multinational triangulation approach to hazard identification in night-time offshore helicopter operations. Reliab. Eng. Syst. Saf. **108**, 142–153 (2012)

Aircraft Accommodation for People Living with Obesity: A Call for a Review of Existing Seating, Safety and Emergency Regulations

Kayla Daigle[1], Dawson Clark[1], Chantal Trudel[1], and Shelley Kelsey[2](✉)

[1] School of Industrial Design, Carleton University, 1125 Colonel by Drive, Ottawa, ON, Canada
{kayla.daigle,dawson.clark, chantal.trudel}@carleton.ca

[2] National Research Council of Canada, 1200 Montreal Road, Bldg. U-61, Ottawa, ON K1A 0R6, Canada
Shelley.Kelsey@nrc-cnrc.gc.ca

Abstract. The commercial air travel industry has incentivized strategies that maximize the number of bodies that can fit into the cabin. This demand for increased capacity in a limited space has led to economy class cabins becoming increasingly cramped. This lack of space particularly impacts travelers who are living with obesity. To develop an understanding of this issue, we evaluated the current space within the cabin and compared it to anthropometric data of both the general population and a cohort of people living with obesity. The lack of accommodation for those living with obesity may have severe impacts on the health and safety of all passengers. Upon review of the data, it is clear that there remains more work to be done to address the accommodation of passengers that better reflect the anthropometry of the flying public. Identifying a solution that balances economics with passenger health, safety, and comfort will require a collaborative effort between carriers, manufacturers, regulators, and passengers.

Keywords: People living with obesity · Passenger experience · Commercial cabin interiors · Safety · Ergonomics · Anthropometry

1 Introduction

It is no secret that commercial aircraft economy class cabins feel increasingly cramped. Higher load rates (a greater proportion of seats filled) and capacity maximization strategies (an emphasis on fitting as many seats in the cabin as possible) have both contributed to the current state of commercial air travel [1]. The impact of these trends is a cabin struggling to accommodate passengers. This lack of accommodation is particularly egregious in the case of passengers who live with obesity. Industry demands dictate a cabin design that relies on outdated and insufficient ergonomic standards [2], leading to uncomfortable and potentially unsafe conditions onboard commercial flights. The purpose of this literature review was to examine the current ergonomic conditions facing travelers who live with obesity. The literature revealed a

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 570–575, 2020.
https://doi.org/10.1007/978-3-030-50943-9_72

lack of accommodation throughout the aircraft, specifically within seat and aisle dimensions, contributing to unsafe and unhealthy cabins. This paper presents a review of the current state of standards related to aircraft seat width and pitch, as well as the impact a more representative anthropometry of current population body sizes may have on safety issues such as ingress and egress.

2 Seat Dimensions

As a part of this literature review, we sought to clarify and define useful definitions for the two most used measurements to evaluate passenger space aboard aircraft, namely, seat pitch and seat width. For the purposes of this discussion and subsequent analyses, when referring to seat pitch we will rely on the Federal Aviation Administration (FAA) definition as stated by Porta et al. [3] where seat pitch is the distance between a point on one row of seats to the same point on the next row of seats. Seat width, as defined by the FAA, is the distance between two adjacent arm rests within a row. We sought out data regarding current seat widths and pitches across the airline industry via literature [1] and the publicly available SeatGuru database [3]. This review focused on seat width, however pitch is also important and will be discussed in Sect. 3 'Ingress and Egress'.

This analysis demonstrated that the average seat width in economy class across both short and long-haul flights is approximately 444 mm (17.5″). Comparing this information with anthropometric data from a cohort of people living with obesity [4], we found that current aircraft dimensions may not accommodate passengers with higher BMI's (Class I Obesity and above).

For the purposes of this paper, accommodation is defined as the passenger's seated hip breadth + 25 mm being less than the distance between the armrests (seat width). In this case, the critical dimension that allows for the accommodation of passengers is the seated hip breadth of a 95th percentile Dutch female, obtained from the CAESAR database, at 486 mm (19.13″). With a current seat width of 444 mm, only the 50th percentile Dutch female is properly accommodated in a seat width of 445 mm (17.53″). To accommodate the 95th percentile female, aircraft seats would need to increase to at least 511 mm (20.12″). A lack of accommodation for modern adults has also been noted by Porta, Saco-Ledo, and Cabañas who studied Spanish adults and recommended seats be increased to between 50.2 cm (19.75″) and 52.3 cm (20.60″) [5].

A more pressing issue arises when evaluating accommodation with a cohort of people living with obesity. Wiggerman et al. [4] developed a methodology for collecting anthropometric dimensions from individuals living with obesity. Their data includes measurements for the seated waist/hip breadth for 208 females living with obesity [4]. This work demonstrates a mean seated hip/waist breadth of 545 m (21.46″) with a SD of 58.5 mm (2.30″). The current average seat width of 444 mm would only accommodate the 1st percentile of the cohort measured by Wiggerman and colleagues [4]. To accommodate the 95th percentile of Wiggerman's cohort identified in this study, representing a seated hip/waist breadth of 646 mm (25.4″), the seat width would have to expand to 671 mm (26.4″). Table 1 summarizes the accommodation disparity between these two sources of data.

Table 1. Accommodation disparity across data sources

Data source	95th percentile seated hip breadth	Required seat width for accommodation (seated hip breadth + 25 mm)	Current percentile of cohort accommodated by 17.5″ seat width
Wiggerman et al. Cohort [4]	646 mm (25.4″)	671 mm (26.4″)	1st percentile
CAESAR – [5]	486 mm (19.13″)	511 mm (20.12″)	50th percentile

Aisle widths pose a challenge and safety hazard for passengers, regardless of their size and mobility. The FAA mandates that, aircrafts with 20 or more passengers, have a minimum 15″ wide aisle anywhere below 25″ above the floor, and a minimum width of 20″ at any height above 25″ from floor level. The minimum widths of the aisle may restrict the standing position of passengers while also hindering access and direction of travel within the aircraft [6]. Passengers living with obesity have increased difficulties when navigating through these narrow aisles and are at risk of facing the social and physical prejudice of stigma from other passengers and crew.

The disparity in accommodation between the CAESAR anthropometric database [5] and the measurements provided by Wiggerman et al. [4] indicate the need for further research to develop an accurate anthropometric picture of airline travelers across the globe. While the data indicates that aircraft cabins do not accommodate an appropriate range of passengers, the question as to how much more space is needed still remains to be answered.

3 Ingress and Egress

Passenger ingress and non-emergent egress play a major role in the flight experience and interaction with the cabin's physical environment. The ease of movement among aisles and seats greatly affects the safety considerations and standards within the air cabin environment. As such, it is important to consider both the turnaround process for airlines and its effects on the ingress and egress of passengers, along with the physical structures impeding or influencing this process, such as the seating arrangement and aisle widths.

Turnaround times are highly dependent on the aircraft type, the number of passengers, amount of loaded and unloaded cargo, and the business model of the airline operator [7]. The increasing number of flying travelers and the changes to baggage and service strategies have greatly affected the amount of time needed to turnaround a plane. Furthermore, changes in demographics and the rising population of travelers living with obesity have influenced the time and processes associated with boarding and deplaning. Schmidt estimates that current ingress and egress times could be retained despite an estimated increase of 20% in the number of passengers boarding, with incremental changes and adaptations to this process [7]. Accommodating for passengers with reduced mobility and higher BMI would facilitate more efficient turnaround times and would reduce risks associated with mobility for passengers of all sizes and abilities.

Due to the increasing number of larger bodied air travelers, it is essential to consider how this affects ingress and egress and how passengers interact with the physical structures within the cabin. Weight and waistline dimensions have been proven to affect the flow and speed of passenger boarding and deplaning processes [8]. Further research is required to establish evidence-based safety standards that accommodate people living with obesity as they enter and exit the aircraft.

Based on a survey conducted in 2001, the design of arm rests, seat base, and seat back were all rated as important features pertaining to ease of access to and from seats, in addition to the seat pitch, width, and foot clearance [9]. The form and strength of arm rests has been noted to impede and/or assist in exiting seats for a range of passengers. Further research is needed to address their significance in ingress and egress during normal and emergency situations for those living with obesity. Schmidt predicted that there will be an annual 4.6–4.9% growth in passenger traffic in the next 20 years, meaning that airports and aircraft need to be designed in a way that accommodates our growing passenger population and associated requirements [7].

4 Aircraft Safety

The essential function of an airline is to safely transport its passengers [10]. But a review of the literature revealed that most safety standards are designed and implemented without considering passengers of all anthropometrics. This includes aircraft evacuation, structural considerations, and passenger safety procedures.

The current FAA specification for aircraft evacuation at full seating capacity, including crewmembers, is 90 s with half of the emergency exits being unavailable [11]. However, there is no requirement for the distribution of body types or range of (dis)abilities in evaluating the efficacy of these evacuation times. Efforts are being made to create evacuation models that include waist size, mobility issues, gender and age among other demographics in simulated trials [12]. Through a comparative study, Liu et al., found that the variation in physical characteristics of passengers does cause a considerable impact on evacuation times [12]. What these models show is that stakeholders must consider a more varied and realistic population set to ensure the 90 s standard is met.

Aircraft seating is designed and manufactured to meet safety standards as set out by regulatory bodies such as the FAA for crash resistance. However, the tests that determine if structural seating system standards are met or failed are carried out using a 170-lb 50th percentile (stature) anthropometric human analog, which can account for approximately 50% of the human population [2]. The use of such body size to represent "average" is not an inclusive method for safety testing and excludes not only passengers living with obesity but those who fall outside the 50th percentile for stature. For example, researchers found that the head path for a 95th percentile dummy directly intersects with the rear of the seat in front of the passenger, increasing the risk for neck injuries in the case of an emergency landing. Similar standards are also applied to lap belt performance. An increase in passenger weight increases restraint force loads by 30–35% on a 95th percentile analog [2, 9, 13].

Persons with disabilities (including disabilities that may be related to obesity) are less likely to take action to protect themselves in emergencies. For example, high levels of stress have been related to low levels of action in dangerous situations [14]. Furthermore, passengers with reduced mobility present high risks to cabin safety, posing a threat to their wellbeing and the wellbeing of those around them in the case of an emergency evacuation [15]. Decisions regarding the location and dimensions of emergency exits have also been made without considering passengers with larger body measurements and/or proportions and anthropometric characteristics. This includes strength, posture, and other factors that may influence body measurements and ease of evacuation [16].

Travelers living with obesity must not only contend with difficulties in the normal flight experience, but must also fly knowing that they are at a higher safety risk since they have been excluded from the safety design of aircrafts.

5 Discussion

There is significant work to be done to increase the level of accommodation for people living with obesity in the aircraft cabin. The consequences of not accommodating a larger percentile of passengers is an increased health and safety risk for all travelers. The path towards accommodation will require creative solutions on the part of designers, regulators, manufacturers, carriers, and the other travelers. As an industry that prides itself on a safety culture, those in the world of aviation can no longer afford to ignore this issue as one that is too complex or expensive. More work must be done to understand the anthropometric profile of air travel passengers. This review was the first phase of our study. In the second phase we will obtain data on the air travel experience of people living with obesity via an online survey. The survey data collected will help us to identify pain points which we will then use for the third phase, contextual inquiry and participatory design sessions with end users at the National Research Council. It is our hope that through our efforts and others, that a more inclusive change in seating standards, evacuation tests and accessibility endeavors will more accurately reflect the conditions found on board commercial aircraft.

6 Conclusion

This paper calls for a review and update of existing seating, safety and emergency regulations to accommodate passengers living with obesity to protect the interests of all passengers and crew on board. Current anthropometric data demonstrates that there is a disconnect between the space available on board an aircraft and the actual size of the travelling population. This disconnect is increasingly pronounced in the consideration of travelers living with obesity and has the potential to compromise the health and safety of air travelers, specifically in the case of emergency evacuations. A review and subsequent update of the standards for accommodation within the aircraft that reflects the physical reality of the flying public will serve to protect all passengers.

References

1. Miller, E.L., Lapp, S.M., Parkinson, M.B.: The effects of seat width, load factor, and passenger demographics on airline passenger accommodation. Ergonomics **62**, 330–341 (2019)
2. Bhonge, P.S., Moradi, R., Lankarani, H.M.: Evaluation of dynamic performance of aircraft seats for larger passenger population using finite element analysis. In: Advances in Aerospace Technology, vol. 1, p. 203. ASME (2012)
3. Porta, J., Saco-Ledo, G., Cabañas, M.D.: The ergonomics of airplane seats: the problem with economy class. Int. J. Ind. Ergon. **69**, 90–95 (2019). Compare Airline Seats – SeatGuru. https://www.seatguru.com/charts/generalcharts.php
4. Wiggermann, N., Bradtmiller, B., Bunnell, S., Hildebrand, C., Archibeque, J., Ebert, S., Reed, M.P., Jones, M.L.H.: Anthropometric dimensions of individuals with high body mass index. Hum. Factors **61**, 1277–1296 (2019). 001872081983980
5. Robinette, K.M., Daanen, H., Paquet, E.: The CAESAR project: a 3-D surface anthropometry survey. In: Second International Conference on 3-D Digital Imaging and Modeling (Cat. No. PR00062), Ottawa, Ontario, Canada, pp. 380–386 (1999)
6. FAA: Airworthiness Standards: Transport Category Airplanes. Subpart D: 25.815 (1976). https://www.ecfr.gov/cgi-bin/text-idx?SID=708ac4515cba2c8717f6e7f95968f67e&mc=true&node=se14.1.25_1815&rgn=div8
7. Schmidt, M.: A review of aircraft turnaround operations and simulations (2017)
8. Melis, D.J., Silva, J.M., Yeun, R.C.K.: Impact of biometric and anthropometric characteristics of passengers on aircraft safety and performance. Transp. Rev. **38**, 602–624 (2018). https://doi.org/10.1080/01441647.2017.1396264
9. Quigley, M.C.; Southall, D.; Freer, M; Moody, A.; Porter, M.: Anthropometric Study to Update Minimum Aircraft Seating Standards. Rep. Prep. Jt. Aviat. Authorities. 1–142 (2001)
10. Lynch, D.M.: Heavy issue: weight-based discrimination in the airline industry. J. Air L. Com. **62**, 203 (1996)
11. FAA: ARAC Emergency Evacuation Activities (1993)
12. Liu, Y., Wang, W., Huang, H.Z., Li, Y., Yang, Y.: A new simulation model for assessing aircraft emergency evacuation considering passenger physical characteristics. Reliab. Eng. Syst. Saf. **121**, 187–197 (2014)
13. Reed, M.P., Ebert-Hamilton, S.M., Rupp, J.D.: Effects of obesity on seat belt fit. Traffic Inj. Prev. **13**, 364–372 (2012)
14. Boyd, D.D.: General aviation accidents related to exceedance of airplane weight/center of gravity limits. Accid. Anal. Prev. **91**, 19–23 (2016)
15. Chang, Y.C.: Cabin safety behavioral intentions of passengers with reduced mobility. J. Air Transp. Manag. **25**, 64–66 (2012)
16. Melis, D.J. Silva, J.M., Clothier, R.: The changing size of the commercial aviation passenger and its potential impact on the aviation industry (2015)

Urban Air Mobility Fleet Manager Gap Analysis and System Design

Richard Mogford[1(✉)], Dan Peknik[2], Jake Zelman[3], and Cody Evans[4]

[1] NASA Ames Research Center, Mountain View, CA, USA
richard.mogford@nasa.gov
[2] Plural Designs, Cave Creek, AZ, USA
dan@plural-designs.com
[3] flyUAM, Dallas, TX, USA
jake@flyuam.com
[4] San Jose State University Foundation, San Jose, CA, USA
cody.a.evans@nasa.gov

Abstract. NASA's Urban Air Mobility (UAM) Sub-Project is engaged in research to facilitate the introduction of air taxis into the US National Airspace System. Given the history of conventional aircraft operations, it is clear that dispatcher support will be required for UAM. This paper presents a gap analysis, system requirements, and a workstation design concept for the UAM dispatcher or Fleet Manager (FM) position. The gap analysis focuses on the differences between the tasks of the airline dispatcher and those projected for the FM. FM system capabilities and data requirements are then presented as foundations for software development. An initial user interface concept is provided. The FM software uses a single, large display. The system supports prediction, monitoring, and task execution. This paper is intended to support FM software design for future air taxi systems.

Keywords: Air taxi · Urban air mobility · Dispatcher · User interface design · Human factors

1 Introduction

NASA's Urban Air Mobility (UAM) Sub-Project is engaged in research to support the introduction of air taxis into the US National Airspace System. Such operations will require a range of communication, navigation, and surveillance capabilities. Air vehicles for UAM are under development and will initially have human pilots. Separation from other aircraft, obstacles, and weather may be a pilot responsibility or provided by an operator's ground-based systems. Eventually, air taxis may be flown from the ground or fly autonomously [1].

"An air taxi could be defined as a flying vehicle with a range of 50–120 miles, carrying two to four passengers and cruising at an altitude of 3,000 to 5,000 ft. In the near term, based on the current battery technology, the most common commute might be a 50-mile round trip with two short vertical takeoffs and a 30-min energy reserve on a single battery charge" [2].

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 576–582, 2020.
https://doi.org/10.1007/978-3-030-50943-9_73

2 Background

Given the history of conventional (Part 121 and Part 135) operations, it is clear that dispatch functions will be required for UAM. The FAA may mandate these services, as with passenger and cargo aircraft today.

Airline dispatcher (AD) tasks include:

- Maintain operation control
- Evaluate weather and hazards
- Create and deliver flight plans
- Monitor flights and intervene as needed for:
  - Weather changes
  - Delays
  - Aircraft system problems
  - Fuel reserves
  - Air traffic control actions
- Monitor fleet schedule
- Manage aircraft maintenance

At this time, the role of the UAM ground operator is not defined. For air taxis, there may be advantages in combining remote piloting functions with dispatcher tasks. For now, we will treat the UAM dispatcher as a dedicated position with no other duties. To distinguish this from AD operations, we will call the UAM dispatcher the "Fleet Manager" or FM.
This report presents a gap analysis, capability and data requirements, and workstation design concepts for the UAM FM position. The gap analysis focuses on the differences between the tasks of the AD and FM.

3 Gap Analysis

In the original internal report (available from the first author), AD tasks are listed with the comparable FM task. FM tasks that have significant differences compared to the AD are highlighted. These findings are summarized below.

- FM needs a software tool to support shift change since FMs may be located in different geographical locations
- Flight planning has different parameters (latitudes/longitudes for waypoints, referencing surface features such as buildings or roads for navigation, pre-defined routes, reserved airspace corridors, etc.)
- For flight planning, altitude options are very limited since maximum altitude will be constrained (e.g., 5,000 ft.)
- FM must monitor low-altitude flights using detailed geographical and weather information
- FM must make quick decisions to manage low altitude traffic over and within urban areas
- FM handles a mix of scheduled and on-demand flights

- Payload and weight balance are critical for small, air taxi aircraft
- Standard (recurring) routes must be regularly checked due to the constantly changing low altitude environment
- FM submits flight plan to a private service supplier rather than the FAA
- FM coordinates with a corporate service provider for traffic and separation management (rather than the FAA)
- FM must notify the FAA if vehicles deviate into controlled airspace

4 Functional and Information Requirements

The following lists show the software functions that are required for the FM to perform their tasks.

Operations

- Flight planning
- Flight following
- Hazard alerting
- Flight path manipulator (in the event of a contingency)
- Airspace corridor saturation monitor
- In flight route optimizations
- Weather integration into decision-making
- Aircraft energy/consumption monitor

Communications

- Voice and text links with pilot
- Messaging with passengers
- Voice and data communications with air traffic management and local authorities
- Data communications with service suppliers
- Two-way data contact with aircraft (equipment status, fuel/battery level, etc.)
- Voice and text with vertiports
- Voice and text with aircraft maintenance

System

- Vehicle scheduling
- Vehicle maintenance control
- Personnel role management and delegation
- System health monitor
- Pilot scheduling

The following lists show the information requirements for the FM software.

Shift Briefing

- Current and projected impacts of crew roster
- Aircraft complement and availability

- Gate assignments
- Weather
- Vertiport status
- Company policy changes
- Emergency procedures in progress
- FAA/local authority publication changes
- Current customer status (complaints)
- Public events (e.g., airshow)
- Ground traffic
- Information technology or computer-related issues or changes
- Outgoing FM notes

Map

- All UAM aircraft identifier, type, position, speed, heading, altitude, and flight plan
- Other aircraft displayed as needed
- Flight corridors
- Speed and descent profiles
- Trip legs and waypoints
- Human-made and natural obstructions (e.g., buildings, trees, hills, etc.)
- Temporary flight restrictions
- Special use airspace
- Noise regulations that affect vehicle choices and routes
- Vertiport obstruction areas
- Vertiport detailed map view with traffic
- Vertiport status
- Emergency and safety related notices affecting flights

Weather

- Very detailed, real time surface to 5000 feet weather data
  - Barometric pressure
  - Temperature
  - Relative humidity
  - Dew point
  - Wind speed and direction
  - Precipitation
  - Icing aloft
- Detailed wind flows for urban canyons, structures, topographical areas, and vegetation (e.g., trees) with a few meters resolution
- Localized ceiling and fog

Aircraft

- Fuel/battery capacity, range, and remaining
- Time to charge battery
- Payload:
  - Aircraft capacity
  - Weight

  - Manifest
  - Loading notes
- Human pilot, ground pilot, or autonomous
- Safety procedures
- Repair and maintenance data
- Deferred maintenance items
- Onboard system status

Communications

- Downloads of updated regulations, management advisories, etc.
- Information about crew schedule changes
- Security or emergency items (terrorism threats, reroutes, medical emergencies)
- Company system status (communications, facilities, personnel, etc.)
- Information technology or computer-related issues or changes
- Gate assignments
- Curfew issues
- Noise footprint
- Noise abatement
- Community feedback
- Backup pilot for emergency events

5 User Interface Design

FMs will be challenged by the number of flights they must control and the sheer complexity of low-altitude vehicle management. Therefore, information displays are forward-looking and assist the FM with suggested courses of actions rather than just displaying data.

System capabilities include:

- Predictive interactions as a core design principle
- Optimization of FM to vehicle management ratio
- Decrease training needs of FMs relative to ADs
- Enhanced safety considerations for a low altitude, automated flight environment

The system architecture provides two modes. One mode reveals recommended interface actions based on user tasks, with the computer looking ahead and presenting calculated options. If the user wishes to drill down to understand the process by which the system is making recommendations, or perhaps create other courses of action, a "behind-the-curtain" or detail mode of operation can be accessed. The behind-the-curtain views are designed with the goal of being secondary to the assisted mode.
The FM interface uses a single, large display. Information in the interface is integrated. An example might be that the user chooses a route, the impact is noted as an issue during or after creation of the route, a map graphically shows the issue, and the user can take several actions including a computer-generated option. The user communicates the chosen action to others, logs the issue, and continues to monitor the flight path. Three actions are at the center of the architecture: Prediction, Monitoring, and Execution.

Prediction: This is an assisted automation mode. The system is aware of the user's intended actions and unobtrusively makes recommendations based on a larger set of data than is available to the user, and with greater depth by use of algorithmic processes. For example, a change in a vertiport landing area causes extensive rolling delays. The computer generates efficient responses, weighs the choices for the user, and presents the results.

Monitoring: Situational awareness is maintained pre-flight, in-situ, and post-flight. This allows users to understand how the computer's choices are being carried out and, if needed, change the plan based on factors perhaps outside of the system's scope.

Execution: The user puts a plan in place, such as authorizing a flight the system has recommended or that the user has constructed. The interface makes user decisions and their effects obvious, showing timing and pertinent state data feeding back to Monitoring.

Figure 1 depicts a high-level user interface architecture. The maximum number of windows is four. The user can drag the controller at the center to resize all windows simultaneously. The information flow starts in top left and then moves to top right, bottom left, and finally to the bottom right, with each user selection defining what information populates the next window, thus maintaining data continuity.

Fig. 1. UAM dispatcher software screen layout.

6 Summary

This gap analysis was completed as an initial step toward designing software to support the UAM FM. AD and FM tasks are compared and those FM tasks that differ from conventional FD duties are highlighted. Functions and information requirements are listed that could guide software and user interface design. For additional information on the work on the FM workstation, please contact the first author.

References

1. UBER Elevate: Fast-Forwarding to a Future of On-Demand Urban Air Transportation, 27 October 27. https://www.uber.com/elevate.pdf. Accessed 20 Nov 2019
2. Lavars, N.: 8 flying taxis that are so crazy, they just might work, newatlas.com, para. 15, 12 February 2018. https://newatlas.com/eight-flying-taxis/51726/. Accessed 12 Oct 2019

Investigation of Commercial Aircraft's Cargo Luggage Dash Impact on Passenger During Emergency Evacuating

Li Wen Wu(✉)

Shanghai Aircraft Design and Research Institute,
No. 5188 Jinke Road, Pudong New District, Shanghai, China
usweapon@21cn.com

Abstract. According to the statistics, emergency landing and evacuating case of commercial aircraft usually causes serious accidents, this paper investigates the definition, the airworthiness regulations of commercial aircraft emergency landing, analyzes the impact on passenger evacuating, when the structure and system around cargo are damaged, due to luggage dash during the emergency landing. This paper also introduces methods of avoiding the impact of luggage's dash on passenger evacuating during emergency landing.

Keywords: Emergency landing · Luggage dash · Damage impact · Passenger evacuating

1 Background

Air transport industry is a important portion of transport industry, based on the latest statistics from International Air Transport Association (IATA), global air transport volume of passenger is up to 4.59 billion in 2019, at the same time, the security requirement of air transport is more stricter than any other period.

Air transport accidents concentrate to happen in the phase of taking off, climbing, descending or landing, including the accidents in emergency landing and passenger evacuating.

The emergency landing case means that the commercial aircraft is forced to landing without routine landing procedure, when the aircraft suffers these incidents such as in-flight shutdown of engine, cabin pressure loss, insufficient fuel, adverse weather, etc. Emergency landing case is shown as Fig. 1. These photos show the test of Boeing 727 aircraft's emergency landing taken by Discovery channel.

In the case of emergency landing, first important problem is below:

For the security of passenger sitting on the seat, the passenger seat, belts and the structure connected to them which are in the passenger cabin must successfully endure the inertial load of forward 9 g (also other directions such as downward, upward, sideward, etc.) during emergency landing.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 583–589, 2020.
https://doi.org/10.1007/978-3-030-50943-9_74

Fig. 1. Test of Boeing 727 aircraft's emergency landing

Many investigators have paid much attention to this problem for long years, and the corresponding calculation and test methods have been very mature, for example, Federal Aviation Administration (FAA) has built full scale structure model of ATR42-300 business aircraft by finite element method, calculated the shock response of fuselage floor, passenger seat track and dummy [1]. The investigation of Lillehei and Robinson indicates that the survival probability of air transport accidents will increase 20% by using more appropriate seats and belts [2].

However, there are less investigation related to the impact of cargo luggage dash during the emergency landing and passenger evacuating (or we call this situation as emergency evacuating) relatively, for the background, this paper investigate the structure and system damage impact of cargo luggage dash in the passenger emergency evacuating situation.

2 Airworthiness Regulations for Emergency Landing

There are the key regulations for emergency landing from CCAR25 [3] and FAR25 [4] below:

1. Regulation 25.561(b).
2. Regulation 25.561(c).
3. Regulation 25.787(a).
4. Regulation 25.787(b).
5. Regulation 25.1362(a).

Based on the Regulations above, though regulation 25.787(a) indicates that the inertial loads of regulation 25.561(b) (such as load of forward 9 g) specified in the emergency landing conditions need not be applied to compartments located below, or forward, of all occupants in the airplane, for example, the inertial load need not be applied to the cargo (cargo compartment) which is below the cabin (passenger compartment), shows as Fig. 2.

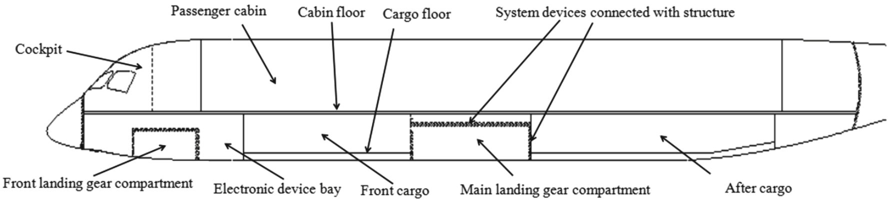


Fig. 2. Cargo's position of commercial aircraft fuselage

However, the cargo of commercial region aircraft usually includes no container, all the luggage are divided by some cargo barrier net, shows as Fig. 3. when the region aircraft and cargo endure inertial load, the cargo barrier net will arrest the luggage, and the net's inertia load coefficient of forward direction is 3 g by design, on the basis of cargo luggage density between the neighboring net. This load is higher than any other load which the aircraft endures in the normally situation of ground or air, but in the emergency landing case, for the security of passenger, the inertia load coefficient of forward direction is 9 g by design (according to regulation 25.561(b)), it is higher than the design value of cargo luggage.

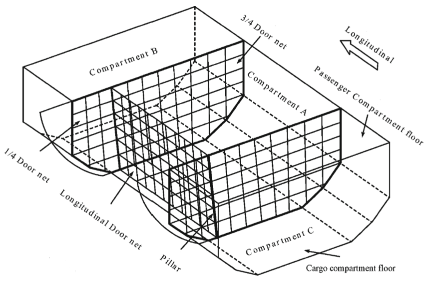


Fig. 3. Cargo barrier net layout

All the parts of aircraft endure the 9 g inertia load of forward direction in the same time, including the cargo luggage, so the luggage will break through the cargo barrier net and dash the structure or system around the cargo. In this situation, partial damage of structure or system is accepted, but according to the regulation 25.561(c), 25.787(b) and 25.1362(a), all the damage should not affect the passenger evacuating environment or the work of evacuating support system, so the airworthiness regulations means that the impact of cargo luggage dash during passenger emergency evacuating must be considered in aircraft design.

3 Structure and System Damage Impact on Evacuating

For most of system devices around cargo are usually installed in the front or after area of cargo, few of them are installed close to the other area of cargo due to lack of space, in addition, the luggage's inertial load of forward 9 g is higher than the load of the other directions, so the inertial load's impact of forward 9 g is more serious than the load of other directions, this paper concentrates to investigate this load's impact on emergency evacuating.

3.1 Thin-Walled Structure Damage Impact

The thin-walled structure in front of cargo usually consists of vertical and horizontal beam, metal panel, or complex material panel, and other parts, shows as Fig. 4. If the strength of the structure is not enough to endure the luggage's inertial load of 9 g forward, then the structure will damage, part of the dash energy will be absorbed by the damage structure, but the luggage will still shock the system devices which are closed to the structure, in this situation, there are risks of fire, explosion, or evacuation supporting devices failure, it must be considered that whether these risks will be occur under the shock load or not.

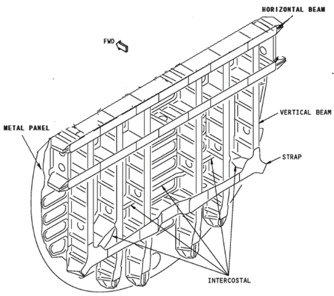


Fig. 4. Thin-walled structure in the front of cargo

3.2 Air Control System Damage Impact

The air control system is usually in the electronic device bay of cargo's forward area. For all the devices of electronic device bay have meet the temperature environment requirement of Federal Aviation Administration (FAA)'s Technical Standard Orders (TSO), such as environmental conditions and test procedures for airborne equipment (DO-160G),the requirement is that all these device must work normally for at least 30 min without being cooled by air control device, when the thin-walled structure has been damaged by cargo luggage and the air control device has been failure to work due to the luggage dash, there is no impact on passenger evacuating if the evacuating time is less than 30 min, but when the time is more than 30 min, the support devices for evacuating in the electronic device bay are failure to work, and passenger will lose some support to evacuate.

3.3 Fuel/Fuel Inertness System Damage Impact

The devices of fuel/fuel inertness system in front area of cargo which may be damaged by cargo luggage dash are fuel feed pipes, if fuel feed shut-off valve is failure to shut off the fuel feed through pipes after emergency landing, then some fuel exist in the pipes, and the fuel will leak from pipes when the luggage dash the pipes and pipes crack. The leaking fuel may be ignited by electric current of damaged electric cable or other high-temperature environment, sometimes may cause explosion. In this situation, passenger evacuating will be very dangerous and difficult.

3.4 Hydraulic System Damage Impact

The devices of hydraulic system in front area of cargo which may be damaged by cargo luggage dash are usually the hydraulic oil feed pipes which connected with landing gears and actuator of flight control system, if hydraulic oil feed shut-off valve is failure to shut off the hydraulic oil feed through pipes after emergency landing, then some hydraulic oil exist in the pipes, and the oil will leak from pipes when the luggage dash the pipes and pipes crack. The leaking oil may be ignited by high-temperature environment. In this situation, passenger evacuating will also be very dangerous and difficult.

3.5 Fire Protection System Damage Impact

The devices of fire protection system in front area of cargo which may be damaged by cargo luggage dash usually are smoke detector device, fire extinguisher, or fire protection electronic controller. If any one of these devices is failure to work due to luggage dash, the aircraft can't detect fire disaster, or can't extinguish the fire, or can't control the fire protection system, the potential fire disaster cannot be controlled, and the passenger in the cabin will be lack of oxygen or be poisoned or be burned, else, evacuating passage way may be cut off.

3.6 Door Signal System Damage Impact

There are many doors in commercial aircraft fuselage, normally, shows in Fig. 5. The devices of door signal system in front area of cargo which may be damaged by cargo luggage dash are usually the electronic controller of proximity sensor, if the device is failure to work due to luggage dash of emergency landing case, then the aircraft crew can't distinguish whether the doors using for evacuating are open or not, in the same time, the structure of doors and fuselage will deform, sometimes very seriously, some doors of fuselage can't be opened successfully, before passenger evacuating, the crew can't quickly determine which doors can be used for passenger evacuate, this problem will reduce the speed of evacuating.

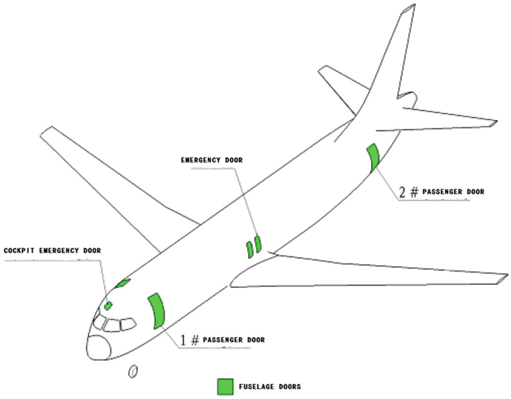


Fig. 5. Doors of commercial aircraft which can be used for evacuating

3.7 Electronic Power Supply System Damage Impact

The devices of electronic power supply system in front area of cargo which may be damaged by cargo luggage dash are usually the main battery, emergency battery (or independent battery for special device), primary power distribution devices, secondary power distribution devices, static inverter, transformer, electric cable, etc.

Normally, the support devices of passenger evacuating which must use electronic power and which are connected with the devices of electronic power supply system in front area of cargo include:

1. Electronic controller of smoke detector.
2. Electronic controller of fire extinguisher.
3. Electronic controller of fuel feed shut-off valve.
4. Electronic controller of hydraulic system overheat shut-off valve.
5. Passenger address and cabin inter-phone.
6. Audio control panel.
7. Recorder independent power supply device.
8. Electronic controller of door's proximity sensor.

Due to corresponding devices of electronic power supply system in front area of cargo are damaged by luggage dash, the support devices of passenger evacuating can't work during passenger evacuating even though they are not in the front of cargo at all and can avoid from luggage dash, the support functions such as detecting smoke, extinguishing fire, shutting off fuel or oil feed, broadcasting evacuating orders, will lose partly or totally, there are multiple obstacles during passenger evacuating, this situation must be avoided by aircraft designer.

4 Methods of Avoiding the Luggage's Dash Impact

To avoid the structure or the system devices failure in the inertial load of 9 g forward, and to protect or support passenger evacuating during emergency landing, these methods could be used for aircraft main manufacturer below:

1. Design aircraft structure in the front of cargo which can directly endures luggage inertial load of forward 9 g.
2. Install the system devices which are related with emergency evacuating out of luggage dash impact area.
3. Analyze or test the system device's function related to emergency evacuating under inertial load forward 9 g of luggage.

5 Summarize

In emergency landing case, the impact on passenger evacuating due to cargo luggage dash is a complex problem, which involves explanation of airworthiness, stress analysis of structure or system devices, function hazard analysis of many system and tests for simulation of the dash impact. This paper provides the analytical methods for the problem, these methods can conduct the integrated design work of commercial aircraft.

References

1. Jackson, K.E., Fasanella, E.L.: Test-analysis correlation of a crash simulation of a vertical drop test of a commuter-class aircraft. In: The Fourth Triennial International Aircraft Fire and Cabin Safety Research Conference, Lisbon, Portugal (2004)
2. Lillehei, K.O., Robinson, M.N.: A critical analysis of the fatal injuries resulting from the continental flight 1713 airline disaster: evidence in favor of improved passenger restraint systems. J. Trauma-Injury Infect. Critical Care **37**(5), 826–830 (1994)
3. CCAR, C.: China Civil Aviation Regulations Part 25, Airworthiness Standards Transport Category Airplanes (2011)
4. FAA, C.: Federal Aviation Regulations Part 25, Airworthiness Standards Transport Category Airplanes (2014)

Research on the Comprehensive Evaluation System of Cabin Comfort of Civil Aircraft

Jian-Ping Chen[1], Jin Wang[2(✉)], Jin-Yi Zhi[2,3], and Li-Li Zhang[1]

[1] The Second Research Institute of Civil Aviation Administration of China, Chengdu 610041, China
[2] School of Architecture and Design, Southwest Jiaotong University, Chengdu 610031, China
28135671O@qq.com
[3] Institute of Design and Research for Man-Machine-Environment Engineering System, Southwest Jiaotong University, Chengdu 610031, China

Abstract. In order to clear the factors and degree that affect the comfort experience of the aircraft cabin, and to make a reasonable strategy for improving the comfort of the aircraft cabin, the comprehensive evaluation system of the comfort in aircraft cabin is studied. An indicators system, weight system, and evaluation method for cabin comfort evaluation are proposed, and the comfort of aircraft cabins is evaluated from four aspects: the indicators of criterion layer, the indicators of attribute layer, the factors of factor layer, and the overall comfort of cabin. Based on the evaluation system, the cabin comfort of a flight was investigated, and 206 valid questionnaires were obtained. Statistical analysis of the data found that the factors and indictors influencing comfort. The comprehensive evaluation system of comfort experience in civil aircraft cabin provides quantitative reference for the selection and optimization of cabin scheme, and provides direction for the improvement of aircraft cabin comfort in future design.

Keywords: Aircraft cabin · Comfort · Comprehensive evaluation

1 Introduction

Cabin comfort is not only one of the important factors for passengers to choose air passenger transport, but also a powerful means for the market competitiveness of airlines [1]. The research on cabin comfort includes three aspects: the influencing factors and degree of cabin comfort; the influencing mechanism of comfort experience and the methods of improving comfort; and the comprehensive evaluation of cabin comfort [2]. At present, studies on cabin comfort in China mainly focus on comfort evaluation, among which the construction method of comprehensive evaluation index system remains to be optimized, the subjective evaluation sample of weight assignment is small and the evaluation result is single.

Based on the questionnaire survey and statistical analysis, a comprehensive evaluation system is constructed in this paper for the cabin comfort of civil aircraft, the research result provides quantitative basis for the improvement of the cabin comfort of civil aircraft.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 590–596, 2020.
https://doi.org/10.1007/978-3-030-50943-9_75

2 Current Status of Comfort Research

Zhang et al. [3] find that discomfort is related to physical factors in the environment, while comfort is related to experience. In the discomfort pyramid proposed by Bubb [4], the unpleasant smell seriously affects the comfort. However, the physical environment and other factors are relatively high in the aircraft, so anthropometry becomes the focus. In the business competition, the service factor shall be located at the top of the pyramid.

In the study of the influencing factors and degree of cabin comfort, Richards and Jacobson [5] show that the factors affecting comfort include leg space, seat characteristics and aircraft motion. Chen [6] believes that the most important factor affecting the service quality of airlines is the staff/facilities, followed by the helpful flight attendants and products, among which the service, communication (e-mail, Internet, etc.) and on-board entertainment equipment are also considered to be very important. Blok [7] discovers that the lowest comfort scores are successively knee space, personal space and seat width. For long-distance passengers (more than 5 h), entertainment facilities and the attention of flight attendants are quite important for comfort evaluation.

3 Composition of Comprehensive Evaluation System

The comprehensive evaluation system of comfort degree mainly includes multi-level index system, weight system and evaluation method.

3.1 Index System

In the research on the selection of evaluation indexes and the construction of index system, there are mainly four selection basis of comfort evaluation indicators, as shown in Table 1.

Table 1. Selection and classification basis of comfort evaluation indicators

Literature	Classification basis	Indicator type
[8]	Based on the human aircraft environment	People
		Cabin
		Cabin environment
[9]	Based on the development of comfort definition	Vibration
		Physical comfort
		Physiological comfort
		Psychological comfort
[10]	Based on literatures constructing the index system, adjusting indicators and structure through cluster analysis, exploratory factor analysis, reliability and validity analysis	

3.2 Weight

Subjective weighting method is the most important method of weight calculation. In the comprehensive evaluation including subjective evaluation and objective evaluation, the weighting of subjective evaluation indicators and objective evaluation indicators is relatively complex. The problem of comprehensive weighting is solved by Chu et al. [11] with data envelopment analysis. Liu et al. [12] propose that the subjective and objective comprehensive weighting method based on the DEMATEL-entropy-weight method may be used to weight the cabin comfort evaluation index of civil aircraft, in which DEMATEL is used for subjective weighting, entropy-weight method is used to calculate the objective weight, and the subjective weight and objective weight are both averaged to get the weight of each index.

3.3 Evaluation Method

The evaluation method of subjective scale is to use the form of questionnaire to evaluate the indexes reflecting the psychological comfort degree in the simulated or real cabin scenes, and then carry out quantitative analysis on the psychological comfort degree through statistical analysis, which is the most commonly-used comfort evaluation method. In order to ensure the consistency of each index dimension in comprehensive evaluation, the evaluation results shall be expressed in a unified scoring system.

4 Construction of Comprehensive Comfort Evaluation System for Civil Aircraft Cabin

4.1 Index System

The factors influencing the comfort of airplane cabin mainly include physical environment, seat, interface, service, catering and entertainment facilities. The indicators can be improved from the physical, physiological and psychological aspects, and the comfort questionnaire can be constructed based on the index system. The questionnaire is in the form of Likert scale, in which the indicator comfort is scaled with the level scale method. The comfort of aircraft cabin indicator is quantified with five satisfaction levels, and the importance of aircraft cabin indicator is quantified with five importance levels, 1–5 score was set accordingly.

A trial investigation is conducted on the completed questionnaire, and the reliability and validity were analyzed. Finally, 96 indicators of cabin comfort are obtained. After deleting through statistical analysis, there are still different indicators reflecting the same comfort feeling repeatedly. The dimension of index layer indicators is reduced through factor analysis, and 80% is taken as the cumulative contribution rate of initial eigenvalue to determine the number of factors after dimension reduction. Finally, 28 criterion level indicators are obtained.

4.2 Weighting System

The questionnaire survey on the importance of cabin comfort index obtains the importance evaluation of each index. The evaluation results reflect the importance degree and importance order of the indicators, but fail to reflect the relative importance degree between the indicators. Therefore, the dimensionless expression is necessary for the importance evaluation according to the hierarchical subordinate relationship of the indicators.

Factors at factor level = (M_1, M_2, M_3, M_4, M_5, M_6); indicators at index level = (A_1, A_2, A_3, …., A_{96}); indicators at criterion level = (N_1, N_2, N_3, ..., N_{28}); and indicators at attribute level = (S_1, S_2, S_3);

$$S_r = (N_1, N_2, N_3, \ldots, N_t);\ (r \leq 3);\ (t \leq 28)$$

$$M_x = \left(N_1, N_2, N_3, \ldots, N_{\mathrm{y}}\right);\ (X \leq 6);\ (y \leq 28)$$

$$N_y = (A_1, A_2, A_3,\ \ldots., A_b,);\ (\mathrm{b} \leq 96)$$

Assuming P is the mean of comfort score, and Q is the weight.

4.3 Evaluation Method and Process

The index system of the comprehensive evaluation system of cabin comfort for civil aircraft is composed of four levels, namely factor level, criterion level, attribute level and index level. Based on the analytic hierarchy process (AHP), the cabin comfort is calculated for four levels. The indicators at criterion level and factors at factor level provide specific factors and indicators influencing the comfort experience, indicating the direction for comfort improvement; the indicators at attribute level provide the path for comfort optimization; and the comprehensive comfort provides quantitative basis for the comprehensive comparison and selection of aircraft cabin comfort.

$$\text{Comfort score at criterion level} : P_{N_y} = \sum_b (P_{A_b} \cdot Q_{A_b}) \tag{1}$$

$$\text{Comfort score at factor level} : P_{M_x} = \sum_y (P_{N_y} \cdot Q_{N_y}) \tag{2}$$

$$\text{Comfort score at attribute level} : P_{S_r} = \sum_t (P_{N_t} \cdot Q_{N_t}) \tag{3}$$

$$\text{Comprehensive comfort score} = \sum_x (P_{M_x} \cdot Q_{M_x}) \tag{4}$$

5 Case Study

5.1 Questionnaire

The research data of medium-range large airliner (2 + 4 + 2 layout) of an airline are studied: 103 satisfaction survey results and 103 importance survey results on cabin comfort indicators. The survey results are entered into an EXCEL form for analysis and processing with SPSS software.

5.2 Evaluation Results

(See Table 2)

Table 2. Evaluation results of aircraft cabin comfort

Factors	Criterion level indicators	Comfort evaluation	Weight of criterion level indicators	Factor weight	Factor comfort	Overall evaluation
Physical environment	Vibration	4.02	0.173	0.169	4.095	3.916
	Noise	3.60	0.178			
	Air pressure	4.18	0.174			
	Temperature	4.30	0.165			
	Humidity	4.19	0.161			
	Illumination	4.31	0.150			
Seat	Upper body comfort	3.853	0.148	0.161	3.825	
	Seat support	3.879	0.133			
	Apparent comfort	3.959	0.140			
	Aesthetics	3.737	0.135			
	Safety and stability	4.099	0.151			
	Lower extremity comfort	3.77	0.150			
	Supporting mobile device	3.492	0.142			
Interface	Toilet	4.013	0.197	0.166	4.000	
	Visual comfort	4.066	0.203			
	Auxiliary facilities	4.062	0.196			
	Guidance information	4.095	0.205			
	Space layout	3.760	0.199			
Service	Steward	4.283	0.369	0.165	4.054	
	Broadcast quality	4.024	0.340			
	Journal	3.80	0.291			

(*continued*)

Table 2. (*continued*)

Factors	Criterion level indicators	Comfort evaluation	Weight of criterion level indicators	Factor weight	Factor comfort	Overall evaluation
Catering	Catering quality	3.651	0.340	0.172	3.968	
	Process	4.21	0.311			
	Convenient sanitation	4.061	0.349			
Entertainment system	Operation performance	3.57	0.247	0.166	3.549	
	Resource application	3.649	0.248			
	Intelligent interaction	3.819	0.242			
	WiFi	3.185	0.263			

5.3 Analysis of Evaluation Results

(1) Evaluation of overall comfort

According to the relevant calculation method for the comprehensive evaluation system of the aircraft cabin comfort, the overall comfort of the investigated flight is 3.916, lower than the "satisfactory" comfort level.

(2) Comfort evaluation at factor level

In the factor evaluation of aircraft cabin, the comfort evaluation indexes of seat, catering and entertainment system are lower than the "satisfactory" comfort level, among which the comfort evaluation value of entertainment system (3.549) is the lowest.

(3) Comfort evaluation of indicators at criterion level

In the comfort evaluation of indicators at criterion level, noise in physical environment factors, upper body comfort, seat auxiliary facilities, apparent comfort, aesthetics, lower extremity comfort, supporting mobile device in seat factors, space layout in interface factors, journal in service factors, catering quality in catering factors, operation performance, resource application, intelligent interaction and WIFI in entertainment system factors all have a comfort lower than the "satisfactory" comfort level and remain to be improved. Among them, noise, supporting mobile device, operation performance and WIFI all have low comfort evaluation and take a priority to be optimized and improved.

(4) Comfort evaluation results of indicators with different attributes

The comfort evaluation values of the physical indicators, physiological indicators and psychological indicators are 3.976, 3.891 and 3.869, respectively, all lower than the "satisfactory" comfort level, among which the comfort evaluation value of psychological indicators is the lowest.

6 Conclusion

In order to improve the cabin comfort of civil aircraft, a comprehensive evaluation system is proposed for the cabin comfort of civil aircraft, which is used to analyze the influencing factors and degree of the cabin comfort of civil aircraft. Such evaluation system can quantitatively evaluate the overall comfort, factor comfort, index comfort and attribute comfort of different indicators for the aircraft cabin, and comprehensively reflect the comfort of all aspects of the aircraft cabin. The research results can not only be used for the comparative study of the overall comfort of the aircraft cabin, but also provide index and priority reference for the comfort optimization of the aircraft cabin.

References

1. Ahmadpour, N., Robert, J.M., Lindgaard, G.: Aircraft passenger comfort experience: underlying factors and differentiation from discomfort. Appl. Ergon. **52**, 301–308 (2016)
2. Wang, J., Zhi, J.Y., Chen, J.P., et al.: Review and outlook of cabin comfort of civil aircraft. J. Mach. Des. **36**(09), 121–126 (2019)
3. Zhang, L., Helander, M.G., Drury, C.G.: Identifying factors of comfort and discomfort in sitting. Hum. Factors **38**(3), 377–389 (1996)
4. Bubb, R.: Sitting comfort. Hamburg: IQPC aircraft interior innovation (2008)
5. Richards, L., Jacobson, I.: Ride quality assessment III: questionnaire results of a second flight programme. Ergonomics **20**(5), 499–519 (1977)
6. Chen, C.F.: Investigating structural relationships between service quality, perceived value, satisfaction, and behavioral intentions for air passengers: evidence from Taiwan. Transp. Res. Part A **42**(4), 709–717 (2008)
7. Blok, M., Vink, P., Kamp, I.: Comfortabel vliegen: comfort van het vliegtuiginterieur door de ogen van de gebruiker. Tijdschrift voor Ergonomie **4**(32), 4–11 (2007)
8. Liu, J., Yu, S.H., Chu, J.J.: Passenger cabin comfort evaluation of civil aircraft based on multi-level fuzzy comprehensive evaluation. Chin. J. Ergon. **23**(02), 60–64 (2017)
9. Wang, H.Y.: Study on the comfort evaluation mechanism and application of high-speed trains. Lanzhou Jiaotong University (2014)
10. Chen, X.: Research of comprehensive evaluation model on riding comfort of high-speed train. Southwest Jiaotong University (2010)
11. Chu, J.J., Pei, H.N., Yu, S.H., et al.: Study on evaluation index system and method of passenger cabin comfort of civil aircraft. Chin. J. Ergon. **23**(3), 50–55 (2017)
12. Liu, J., Yu, S.H., Chu, J.J.: The comfort evaluation in civil aircraft cabin based on objective and subjective weighting method. J. Graph. **02**, 192–197 (2017)

The Importance of Human Factors When Designing Airport Terminals Integrating Automated Modes of Transit

Seth Gatien[1](✉), John Gales[2], Ata Khan[1], and Ariel Yerushalami[2]

[1] Civil and Environmental Engineering, Carleton University, 1125 Colonel by Drive, Ottawa, ON K1S5B6, Canada
{Seth.Gatien,Ata.Khan}@carleton.ca

[2] Civil Engineering, York University, Toronto, Canada
{Jgales,ariely}@Yorku.Ca

Abstract. The fast-paced advancement of ground transportation methods such as connected and autonomous vehicles makes it difficult for professionals to design infrastructure, which includes airports. Due to the amplified variability in both transportation modes and demographics, planners must consider the best alternative, which will service all of their users while integrating an autonomous system. This paper describes mobility activities at the interface between ground transportation and airport terminal, characteristics of automation in airport access, and planning and design requirements with a focus on human factors. Knowledge gaps are identified. A human factors study is introduced for understanding characteristics and behaviour of users of landside facility. This paper aims to assist planners and designers in considering human factors when planning for autonomous systems in airport terminals; specifically focusing on pick up and drop off zones.

Keywords: Access to airport terminal · Automation in driving · Human factors · Pick-up and Drop-off areas

1 Introduction

Connected and autonomous vehicle technology has progressed to the extent that cities are initiating programs to adopt their transportation infrastructure for their safe and efficient use [1, 2]. It is widely known that future proofing ground transportation for use by connected and autonomous vehicles is a challenge faced by professionals [3–5]. This also applies to the landside infrastructure and operations at airports. This paper describes the activities at the interface of ground transportation and an airport terminal in Canada, discusses characteristics and the role of automation in airport access, and defines how planning and design requirements can be met while considering human factors.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 597–602, 2020.
https://doi.org/10.1007/978-3-030-50943-9_76

2 Landside Facilities

The study herein considers a case-study commuter airport in Ontario Canada. Below the authors describe the recent changes, which this (and other airports) are considering in the day-to-day operations with respect to their landside operations. A landside road transportation system includes road access to terminal buildings, rail transit facilities (if applicable), parking, shared mobility facility, holding areas, and car rental facilities. The capacity and level of service offered to users depends upon the type, size, configuration, technology, and operational practices of the airport. Figure 1 provides a functional viewpoint demonstrating the complex and multimodal nature of the transportation system.

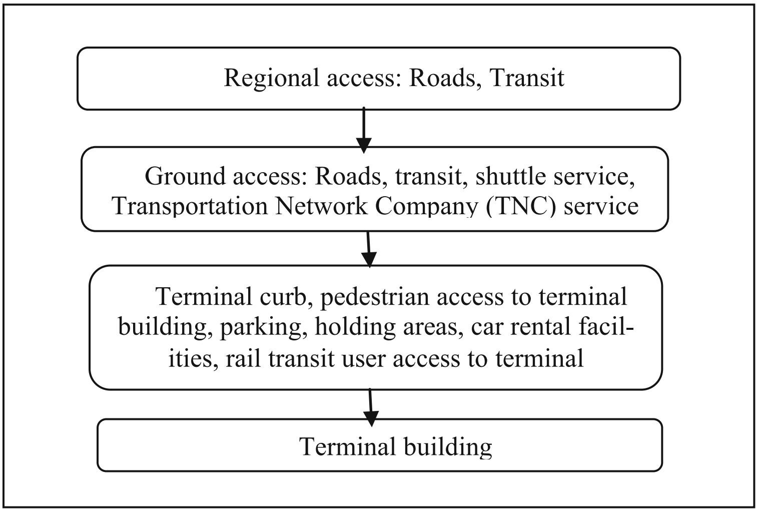


Fig. 1. Functional view of airport landside sub-systems

3 Changes to Landside Facilities and Operations

The recent additions to airport landside services and facilities include: Cell phone parking lot; Cell phone waiting area and pickup; and Transportation network company (TNC) service for pick-up and delivery, which requires a waiting area for TNC drivers. The cell phone waiting area is intended for private motorists who are picking up friends or family. When available, electric vehicles can be charged at this location using high-speed charging stations at no cost to the user in an attempt to reduce congestion in the waiting area [6]. The TNC drivers are provided a ride share area at the airport. According to regulations, this lot is for TNC driver and partners. These drivers will only receive requests at the airport when they are located in the ride share waiting area, at which point they can accept a trip request.

Owing to demand changes, outdated design, and general lack of understanding user characteristics, these components may become crowded, causing delays and inconvenience. Demand characteristics refers to the number of air passengers and their associated human factors, which must be considered in design, and operation of facilities.

4 Human Factors

For modelling purposes, demand characteristics include distribution of passenger arrival over time, modes of travel to and from airport, number of bags carried and checked (if applicable), passengers' age, trip purpose, number of people accompanying the passenger, and whether the passenger has a boarding pass. Airlines conduct research to better understand the most optimal way to fulfill passenger demands through their services. Airport planners, designers, operators and managers are aware that quality of service offered to customers is highly affected by the functioning of the landside ground access system. Generally, airports experience congestion at the terminal curb area due to increasing traffic volume in association with its peaking characteristics and complicated interaction between pedestrians and vehicles.

The airport authorities frequently note that the curbside is a location where congestion occurs. The area immediately outside an airport terminal building beyond the sidewalk is referred to as the processing curbside area. Major activities at the enplaning or deplaning curbside area are the processes of loading/unloading of passengers and their baggage [7].

Airport users are directly impacted by the facility's level of service (LOS). The LOS is of equal importance to airport authorities, in addition to the capacity level. Historically, techniques to addressing LOS and capacity problems have dealt with vehicles rather than persons. Similarly, curb problems are frequently described in terms of vehicle characteristics. However, current management is keen on understanding human factors (i.e. passenger behaviour and characteristics) to support decision-making during design.

Airport management authorities tend to focus on curb operations, hold lot capacity and location, and the impact of renovations and major capital improvements on curb management and customer service. Attention is devoted to balancing changes in mode share, re-assigning staging area locations, and capacity [8].

At a higher level, there is a need for research on level of service criteria and standards for the design and evaluation of landside facilities in both the existing and the future transportation environment. This need arises from the requirement that they reflect technical developments, enhanced knowledge of user preference and bahaviour, and the need to increase the efficiency and cost-effectiveness of the airport system.

The vehicle and pedestrian movement into and out of an airport are assessed in an attempt to analyze the impacts the introducing autonomous systems to ground transportation. Airports serve a wide range of modes and demographics, and are therefore considering the integration of autonomous systems to create more seamless pickups and drop offs thanks to the flexible design of airports, the space is simultaneously used to observe current traffic while also testing the impact of future changes.

5 Connected and Autonomous Vehicles at Airports

Automated driving is quickly advancing both in the technological progression and by attracting the attention of public sector decision-makers. The Society of Automotive Engineers (SAE) [9] has defined levels of automation that range from 1 to 5. Many factors can contribute to complexity in the driving environment. A notable future factor will be the presence of vehicles with automation levels ranging from 1 (no automation) to 5 (fully automated). According to the SAE [9] and the U.S. Department of Transportation [10], the level 5 automation has "The full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver"*. Here, it is understood that the human driver is not distracted and also drives in a non-aggressive manner in terms of maintaining safe distances in the longitudinal and lateral direction as well as accepting a safe gap in traffic for merging or lane change manoeuver. On the other hand, connected autonomous vehicles (CAVs) enable an autonomous vehicle to communicate wirelessly with surrounding vehicles, infrastructure, the cloud/network, and pedestrians. This comprehensive capability is referred to as vehicle to everything. The applications of CAVs at the airport are of interest to landside managers [8].

Airport authorities and other stakeholders are aware that CAVs will become an important part of airport systems [11], and are thus heightening the efforts to incorporate them into modern transportation systems. In fact, the technologies are already undergoing tests at airports around the world. For example, self-driving vehicle trials at Las Vegas airport are already underway [12]. Additionally, autonomous shuttles in traffic-dense environments at and around Charles de Gaulle airport in Paris are also being researched [13]. There are numerous examples of autonomous technologies operating in the terminal and in the apron part of the airport [11, 14], which further adds to the legitimacy of fabricating them in Canadian regions.

Given that airports have to accommodate autonomous technologies, researchers and planners are keen on defining how these advanced technologies can play a role in enhancing safety, security, efficiency, environmental quality, and other objectives of the airport authorities.

At present, airport authorities are interested in knowing how best to meet the objectives of Mobility as a Service (MaaS) for the benefit of airport users, as well as adapting the facilities to accommodate for CAVs. The MaaS, based on digital platforms, enables users to determine their most attractive mobility options in terms of level of service and cost. Both public and private modes are considered.

6 Human Factors in Future Proofing Landside

6.1 Facilities and Services

In addition to serving other functions, autonomous technologies at the airport will frequently interact with people. Thus, it is essential to consider human factors as cost-effective as possible. After developing a foundational knowledge of autonomous vehicles and airport facilities, authorities will be at a more credible position to future proofing landside facilities and services [8, 11].

Examples of autonomous technology applications that are likely to materialize and will require knowledge of human factors in order to succeed are self-driving vehicles (e.g. CAVs) including robo taxis, airport-owned robo taxis with in-vehicle biometric check-in and passport control, self-driving wheelchairs, on-demand autonomous transport pods to access car parks, autonomous shuttles to serve car parks and other locations, and robots in car parks.

Landside facility changes will be required to accommodate autonomous vehicles and systems. Examples include pick-up and drop-off locations designed for privately owned CAVs, robo-taxis, shared CAVs, CAV shuttles, and on-demand autonomous transport pods to access car parks. Also, the following adaptation will be needed: holding and charging areas for TNC-owned CAVs, park and charge facility for privately-owned CAVs, monitoring and control centre for CAV traffic, and re-purposing some car parks near the terminal building.

Further, changes in processors and logistics will be necessary to serve autonomous vehicles and systems. These include CAV request for time slot for pick-up or drop-off., fee payment if applicable, request for battery charging time slot availability at a garage or holding area, request for availability of self-driving wheel chair, information system for advising customers about location for accessing their mode of travel, and availability of MaaS service information.

7 Conclusions

The study of human factors in future proofing airport landside facilities and processors, with a focus on interface between ground transportation and airport terminals leads to the following conclusions:

1. Autonomous vehicles and future airport systems require changes in landside facilities and operations. Changes in physical infrastructure will be essential for CAVs to function properly.
2. The infrastructure and operational changes will enable the airport authorities to achieve efficiency, safety, security and other objectives, including enhancing the user experience.
3. Enhanced knowledge of human factors will assist the planners and designers in making informed decisions regarding the future upkeep of the airport. In support of this objective, landside level of service estimation criteria will be required based on human factors-inspired utility theoretic approach.

It is expected that the findings of research reported in this paper will assist planners and designers to consider human factors in accommodating connected and automated vehicles in the groundside part of the airport terminal.

Acknowledgements. The Natural Sciences and Engineering Research Council of Canada (NSERC) sponsored this research. The views ex-pressed are those of the authors.

References

1. City of Toronto: Automated Vehicles Tactical Plan and Readiness 2022. IE8.7 – Attachment 1, Automated Vehicles Tactical Plan (2019)
2. City of Boston: Boston Transportation Department. Go Boston 2030. Vision and Action Plan, Boston MA (2017)
3. Catapult Transport Systems: Future-Proofing Infrastructure for Connected and Automated Vehicles, Technical report. U.K, February 2017
4. Hallmark, S.: Minnesota Department of Transportation: Preparing Local Agencies for the Future of Connected and Autonomous Vehicles, Research Report 2019–18. Iowa State University, May 2019
5. Hourdos, J.: Minnesota Department of Transportation: How locals need to prepare for the future of V2V/V2I Connected vehicles. research report 2019–35, University of Minnesota, August 2019
6. Pearson International Airport Toronto (2019). https://www.torontopearson.com/en/transportation-and-parking/
7. National Academies of Sciences, Engineering, and Medicine 2010. Airport Passenger Terminal Planning and Design, vol. 1: Guidebook. Washington, DC: The National Academies Press. https://doi.org/10.17226/22964
8. National Academy of Sciences, Engineering, and Medicine: Transportation Network Companies (TNCs): Impacts to the Airport Revenues and Operations. Washington, D.C.: The National Academies Press (2019). https://doi.org/10.17226/25671
9. Society of Automotive Engineers (SAE) International: Summary of SAE International's Levels of Driving Automation for On Road Vehicles. J3016 (2014). Also see SAE: International, J3016_201806, Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Warrendale: SAE International, 15 June 2018). https://www.sae.org/standards/content/j3016_201806/
10. U.S. Department of Transportation (USDOT): Preparing for the Future of Transportation, Automated Vehicles 3.0, Washington, D.C, October (2018)
11. Airports Council International (ACI): Autonomous Vehicles and Systems at Airports Report. Montreal (2019)
12. Autonomous Vehicle International: Craig, T.: Self-driving trials at Las Vegas airport, 9 January 2020
13. International Airport Review: Group ADP to trial autonomous shuttle in traffic-dense environment, 5 April 2018
14. International Airport Review: How airports can prepare for the rise of autonomous vehicles. Author: Gittens, Angela. 27 November 2019

The Flight Scenarios Development Method for Cockpit Design and Evaluation of Civil Aircraft

Hongyu Zhu(✉), Hua Meng, Shasha Lu, and Guangyu Bao

COMAC ShangHai Aircraft Design and Research Institute, No. 5188 Jinke Road, Pudong New Area, Shanghai, China
{zhuhongyu,menghua,lushasha,baoguangyu}@comac.cc

Abstract. The development of cockpit flight scenarios is an important precondition for cockpit design and crew workload analysis. Aiming at the complex dynamic relationship between crew, aircraft, and environment, this paper proposes three dimensions of cockpit flight scenarios. Based on the engineering practice, a cockpit flight scenario matching method is proposed to simplify the scenario development work while ensuring the integrity of the scenario. This paper provides recommendations for flight scenario development in the cockpit development process.

Keywords: Scenario development · Civil aircraft · Cockpit

1 Introduction

The cockpit of civil aircraft is the command center of the aircraft. The cockpit human-machine interface (HMI) is the link between the aircraft and the pilots and a very complex human-machine system.

During the human-centered cockpit HMI design, it is important to perform the human factors requirements capture, human error analysis, and crew workload measurement. To achieve this object, the specific flight scenarios should be selected at first. The actual operation environment and status of the aircraft are complex, changeable and uncertain. Considering the complex dynamic relationships of flight crew-aircraft-environment, it is hard to ensure that the selected flight scenarios are integrated. The cockpit flight scenarios should be developed during the cockpit design, and we must establish a scientific cockpit flight scenarios development methods to ensure the integrity.

Based on the requirements for flight scenarios development in cockpit design and evaluation, this paper proposes the principles and methods for cockpit flight scenarios development.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 603–609, 2020.
https://doi.org/10.1007/978-3-030-50943-9_77

2 The Definition of Scenario

Scenarios are originally applied in the field of software engineering to describe the interaction between human and software [1]. At present, it has also been widely used in the field of aircraft development. The “scenario” in the field of civil aircraft can be defined as: “the expected behavior (or function) of the aircraft in a combination of crew (pilot and flight attendant), external environment (atmospheric, radio, terrain, electromagnetic, etc.) and internal conditions (normal, malfunction)” [2].

The aircraft’s operating environment, crew operation, and aircraft systems are extremely complex. In system engineering, requirements can be divided into aircraft-level, system-level, and component-level. Refer to system engineering, Ling Xie [2] divide flight scenarios into 3 levels: aircraft-level scenarios, system-level scenarios, and component-level scenarios. Because the component-level scenarios are too specific to be involved in the development of the aircraft cockpit, so it will not be explained in this paper. The focuses of aircraft-level and system-level scenarios development are as follows:

a) Aircraft-level scenarios: Taking the aircraft as the research object, consider the operation and maintenance of the aircraft, including the use of aircraft by all stakeholders such as pilots, flight attendants and maintenance personnel, as well as air traffic controllers and passengers, and the operating environment of the aircraft;
b) System-level scenarios: Taking the system as the research object, consider the operation and maintenance of the system. Different from the aircraft-level scenario, the system-level scenario also needs to consider the impact of the status of the interconnected systems, including systems with electrical, mechanical, and energy interfaces.

The cockpit, as the integration of the human-machine interface of the entire aircraft system, involves almost all stakeholders of the aircraft, including: regulators (airworthiness, air traffic controllers), operation support (airports), partners (suppliers), and customers (airlines, rental companies, pilots, flight attendants, maintenance staff), etc. In addition, the cockpit design focuses on the specific human-aircraft-environment interaction behaviors. In many cases, it will involve the system working logic and detailed schemes. The cockpit flight scenario is required to have a finer granularity. Therefore, the flight scenarios related to the cockpit are generally between aircraft-level and system-level, and some are even equivalent to system-level scenario.

3 Cockpit Flight Scenario Development

3.1 Multi-dimensional Matrix of Scenario

In the published research results, the commonly used flight scenario cap-ture methods can be summarized as follows: the elements that affect the scenario are divided into several dimensions, and then, these dimensions are superimposed on each other to establish a multi-dimensional matrix of flight scenario. This paper calls it “multi-dimensional matrix method”.

Minmin Xu [3] regards the flight scenario as a multi-dimensional space, which includes three dimensions: "weather condition", "aircraft factor" and "flight environment", as shown in Fig. 1. By defining the type of each dimension of these subspaces and limiting their value, the flight scenario can be clearly defined.

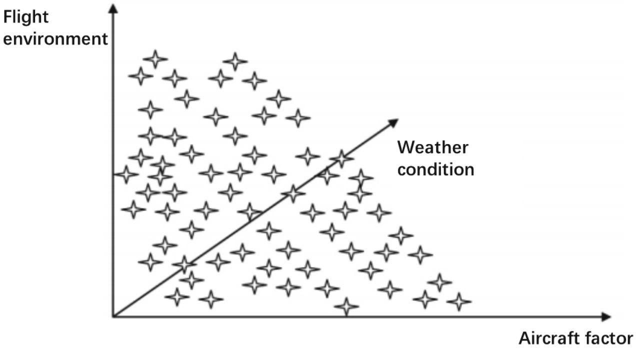


Fig. 1. Flight scenario as a multi-dimensional space (from Minmin Xu [3])

Ling Xie [2] divided the aircraft scenario into 5 dimensions: time dimension, environment dimension, status dimension, task dimension, and stakeholder dimension. Any combination of the first three dimensions, plus the last two dimensions, constitutes a flight scenario. For example, "in route operation, the crosswind is greater than 20kt during the normal landing phase, the pilot operates".

This paper will also use the "multidimensional matrix method" to capture cockpit flight scenarios. Combining the cockpit human-machine interaction characteristics, similar to Ling Xie, this paper divides the cockpit flight scenario into the following three dimensions:

(a) Time dimension: The flight phase of the aircraft, which include all the phases in a complete flight mission profile. Considering the number of scenarios and the granularity of subsequent crew mission analysis, workload analysis, and design requirements capture, we recommend referring to the sub-chapter of the flight crew Standard Operating Procedure (SOP) as the dimension elements, where the aircraft status or spatial position changes significantly, such as cockpit preparation, pushback, climb, parking.
(b) Environmental dimension: Focus on the internal and external operating environment of the aircraft, and describe the operation of the aircraft in the specific environment. It should include at least:
(1) Natural environment: temperature, altitude, visibility, light, etc.;
(2) Human environment: hijackings, aircraft conflicts, etc.;
(3) External facilities: runways (short runways, contaminated runways, etc.), navigation facilities (ground-based, star-based, etc.);
(4) Special operations: polar, RVSM, PBN, ETOPS, etc.

(c) Status dimension: the normal status of the aircraft and the system/crew, and the failures that expected to happen due to internal and external factors during operation. It should include at least single or multiple system failure conditions that in the daily training courses for airline pilots, and may affect the flight crew task. For example: cabin depressurization, one engine fail, hydraulic system fails.

The dimensions and main dimension elements of the cockpit flight scenarios are shown in Fig. 2.

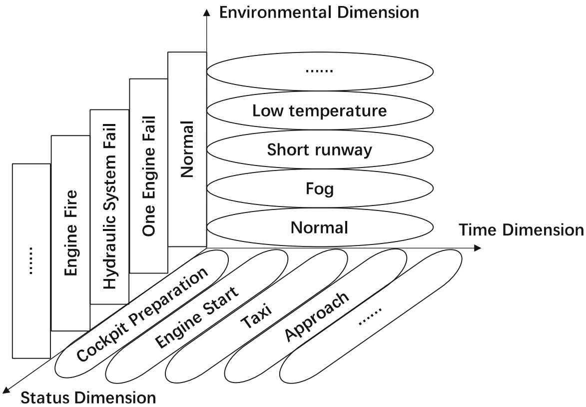


Fig. 2. Three dimensions of the cockpit flight scenario

When enumerating the elements of the scenario dimension, the data sources should include: flight crew operating manual (FCOM), the minimum equipment list (MEL) for aircraft operations, and airworthiness operations regulations. Dangerous scenarios from the aviation accident investigation report should also be acquired as a complement and consummation.

It should be ensured that the dimensional elements can include all situations that expected to happen during the actual operation of the aircraft, so as to ensure the integrity of subsequent flight scenario captures.

3.2 Scenario Dimension Matching Method

By superposing the three scenario dimensions proposed above, the complete flight scenario of the aircraft cockpit can be obtained. However, assuming that there are 20, 40, and 50 elements in the three dimensions respectively, the number of scenarios is 20 * 40 * 50 = 40000, which is huge in magnitude and difficult to be practically applied in engineering practice.

By analyzing the superposition of the three dimensions in combination with engineering practice, it is observed that the aircraft operation has its own characteristics and rules, and the superposition of some dimensional elements is meaningless. For example, during cockpit preparation flight phase, there will be no cabin depressurization; during the cruise flight phase, there will be no high temperature. Therefore, it is necessary to filter out meaningless scenarios and capture only meaningful scenarios when superposing dimensions to capture cockpit flight scenarios.

A meaningful cockpit flight scenario should satisfy the following conditions:

a. It could actually happen, and
b. The flight crew tasks in this scenario are different from all other scenarios'.

Further analysis of the three dimensions shows that the time dimension is continuous, unified, and unidirectional, while the other two dimensions are more isolated and distributed. Therefore, the time dimension can be the main thread of the scenario dimensions superposing, and the environmental dimension and state dimension surround the main thread to capture the scenario. As shown in Fig. 3, the cockpit flight scenario development can be divided in 4 parts or 4 steps.

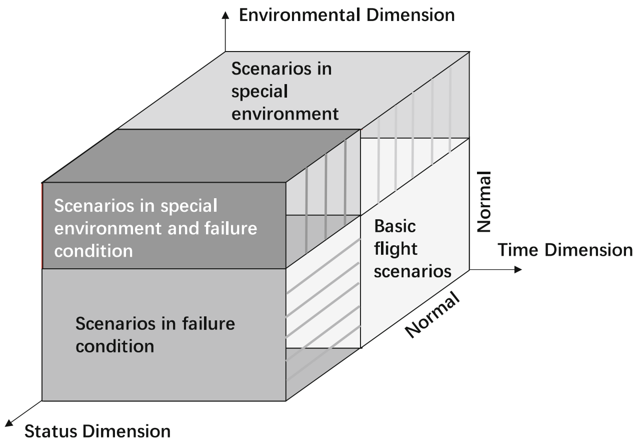


Fig. 3. Flight scenario capture in 4 parts/steps

(a) Basic flight scenarios: Set the environmental dimension to the normal operation environment and the status dimension to normal status. Take each element in time dimension as a scenario, and get the basic flight scenarios.

The basic flight scenarios include most scenarios under normal conditions, such as engine start, normal takeoff, approach and landing. Confirmation and analysis of these scenarios can fully consider the pilot's flight task, human-machine interaction and human error in most cases.

(b) Scenarios in special environment: Based on the basic flight scenarios, and assuming that the aircraft is in a normal status (i.e., no failure). Analyze when the flight operating environment changes in each flight phase (time dimension), if the original flight crew tasks have significant change. If so, they should be captured as flight scenarios, otherwise they are filtered out.

The capture of special flight scenarios can mainly be based on relevant technical data such as civil aviation operation regulations, aircraft flight manuals (AFM), flight crew operation manuals (FCOM), flight crew training manuals (FCTM), and crew resource management. For example, landing (time) in foggy weather (environment), approach (time) in mountain terrain (environment), etc.

(c) Scenarios in failure condition: Still based on the basic flight scenarios, and the environmental dimension is assumed to be normal. Analyze when the aircraft status changes (main function of the system fails) in each flight phase (time dimension), if the original flight crew tasks have significant change. If so, they should be captured as flight scenarios, otherwise they are filtered out.

The capture of flight scenarios in failure condition can refer to relevant technical materials such as abnormal procedures and emergency procedures in FCOM and QRH, MEL manuals, and configuration deviation list (CDL) manuals.

For example, one engine fails (status) after takeoff (time), singe hydraulic system fails (status) during cruise (time), etc., mainly consider the relatively severe failure of the entire system.

(d) Scenarios in special environment and failure condition: Based on the basic flight scenarios, analyze each flight phase (time dimension) when the aircraft is in special operating environment, and there is failure condition (system main function fails), if the original flight crew tasks have significant change. If so, they should be captured as flight scenarios, otherwise they are filtered out.

The capture of this part of flight scenario can refer to abnormal procedures and emergency procedures in FCOM and QRH. In addition, aviation accident cases should be collected to analyze various special scenarios under extremely small probability situations. For example, short runway (environment) landing (time) with automatic brake failure (status).

3.3 Applications

By establishing a multi-dimensional matrix of cockpit flight scenarios and using the scenario dimension matching method that take the time dimension as main thread, the cockpit flight scenario library can be finally obtained. An example of cockpit flight scenario library is shown in Table 1.

Table 1. Example of cockpit flight scenario library

Scenario description	Time dimension	Environmental dimension				Status dimension
		Natural environment	Human environment	External facilities	Special operations	
Normal takeoff	Takeoff	Normal	Normal	Normal	Normal	Normal
Landing in foggy weather	Landing	Foggy	Normal	Normal	Normal	Normal
Single hydraulic system fails during cruising	Cruising	Normal	Normal	Normal	Normal	Single hydraulic system fails
Short runway landing with automatic brake failure	Landing	Normal	Normal	Short runway	Normal	Automatic brake failure
……	……	……	……	……	……	……

4 Conclusion

Cockpit flight scenario development is a necessary precondition for human-machine interface design, flight crew task analysis, flight crew workload measurement, and human error analysis. This paper proposes three dimensions of the cockpit flight scenario: time dimension, environmental dimension and status dimension. In additional, this paper proposes a scenario dimension element superposition method based on the time dimension, and divides the dimension superposition into four steps to easily capture meaningful scenarios and filter out meaningless scenarios. Using this method, the dimensional superposition work can be greatly simplified, meanwhile, ensuring the integrity of the scenarios.

References

1. Min, Y., Xin, Q., Yongyuan, W.: Operational requirement capture Method of civil aircraft based on DoDAF modeling. J. Ordnance Equipment Eng. **6**, 81–85 (2018). (in Chinese)
2. Ling, X., Fang Junwei, X., Zhou, T.C., Lei, Z.: Research on capture and confirmation method for aircraft requirements based on functional scenario analysis. Sci. Technol. Inf. **18**, 89–90 (2015). (in Chinese)
3. Minmin, X., Yuwen, J.: Research of airworthiness certification oriented scenario. Civil Aircr. Des. Res. **2**, 66–69 (2014). (in Chinese)

The Effect of Anticipatory Conditions on Pilot Performance in Encountering Stall: A Flight Simulator Study

Meilisa Hajriani[1(✉)] and Hardianto Iridiastadi[2]

[1] Directorate of General Civil Aviation, Ministry of Transportation Republic of Indonesia, Air Transport Analyst, Medan Merdeka Barat. 8, 10110 DKI Jakarta, Indonesia
meilisa_hajriani@dephub.go.id
[2] Faculty of Industrial Technology, Bandung Institute of Technology, Tamansari 64, 40116 Bandung, Indonesia
hiridias@vt.edu

Abstract. Stall is one of abnormal events in flight which can cause fatal accident if the pilot cannot overcome the critical situation. This research examines the effect of pilot anticipatory condition towards pilot performance when encountering a stall and several other human factor conditions such as stress and sleep duration of previous night. A total of 18 respondents were divided into two groups namely surprised and anticipated against stall scenario. The respondents have passed the stress level test before it was continued with a flight simulation which included a stall scenario in it. Based on the results of the research, surprised and anticipated conditions as well as sleep duration of the pilot affect pilot performance in encountering stall. Pilots who have anticipated stall have a greater chance to have better performance. In addition, pilots with more than six h sleep duration also tend to have better performance in encountering stall.

Keywords: Abnormal events · Stall · Pilot training · Stress · Sleep duration · Anticipatory condition · Logistic regression

1 Introduction

Based on the factors causing the accident investigated by National Transportation Safety Committee of Indonesia (KNKT) [1], there are more than 60% of airplane accidents and incidents caused by humans. According to Belcastro and Foster [2], loss of control in-flight can be caused by improper responses from pilots when facing certain events, especially abnormal events. One of the abnormal events often occur in flights is stall, which can result in fatal accidents of almost 50% compared to non-fatal accidents [3]. According to Kenny, around 20% of fixed-wing aircraft fatal accidents were caused by mistakes when pilots conducted recovery from a stall. In Indonesia, a major aircraft accident involving stall problems occurred in October 2018 in which Lion Air's Boeing 737 Max with flight number JT610 flying from Jakarta to Pangkal Pinang crashed into Karawang waters, West Java after flying for 13 min and killing 189 people [4].

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 610–616, 2020.
https://doi.org/10.1007/978-3-030-50943-9_78

Although pilots have been trained many times specifically to deal with stall, many of them still found difficulty in facing it. In 2013, Casner et al. [5] conducted a research to find out the effectiveness of training for pilots to encounter abnormal events. Therefore, Casner et al. designed an experiment of abnormal events with two conditions for pilots, namely predictable and unpredictable conditions. A similar research was also conducted by Landman et al. in 2017 [6]. Landman et al. suspected that there was an influence of anticipatory condition of a pilot during abnormal events towards the ability to perform recovery. Moreover, flight hours or flight experience also supports the readiness of pilots to carry out their duties, especially in the critical flight phase prone to accidents [7]. In addition to training effectiveness, quoted from the statistical data of Flight Crew negligence factors presented by the Aviation Safety Network [8] from 1950 to 2016, there are 12 main causes which trigger pilot negligence. The highest percentage of 16.5% is due to the consumption of alcohol and drugs to reduce stress. Stress experienced by pilots is caused by high workloads related to passenger safety. Pilot is even categorized as one of Safety Sensitive Jobs.

Based on the background description above, the aim of this research is to find out the relationship between the pilot performance in encountering stall and the anticipatory condition of the pilot (whether the pilot know that the scenario of stall will be given), flight hours, sleep duration and stress level. The assumptions contained in this research is the similarity of the flight functions of Mechtronix Ascent FNPT II aircraft simulator which is similar to Pipper Warrior aircraft. In addition, it is assumed that surprised condition when a stall occurs in one of the groups of respondents can represent stall event on an actual flight.

2 Previous Research

In 2013, Casner et al. [5] designed an experiment to find out whether pilot performance in dealing with previously known abnormal events was as good as pilot performance in encountering abnormal events which had not been known before. Casner et al. assumed that if pilot performance in both conditions was equally good, then the training in dealing with abnormal events carried out by pilots has been effective. Conversely, abnormal events training for pilots is not effective if the performance in both conditions is different. A total of 18 pilots working for airlines were involved in the research conducted by Casner et al. The simulator used was Boeing 747–400 (Level D). The abnormal events included in the scenario were aerodynamic stall, low-level wind shear and engine failure during take off. The participants were then divided into two groups, namely group that aware of upcoming stall and group that did not aware. The parameters used in the research conducted by Casner et al. were reaction time, flight hours, and total altitude lost. The research used ANOVA to analyze the effect of flight hours as well as to compare the average and standard deviation to see the differences between groups which have known and have not known the scenario of abnormal events before the simulation. The results of the research showed there was a significant difference in altitude lost between pilots who experienced stall at a sudden and who expected the stall. Regarding to flight hours, this variable did not affect pilot performance when encountering a stall.

Through this research, Casner et al. suggested to update training methods and include elements of surprise of abnormal events in training.

In 2017, Landman et al. [6] also conducted a similar research to find out whether there was a difference in performance between pilots who anticipated and encountered a stall and who did not. In that research, pilots were divided into two groups where the first group was asked to rate the fidelity level of the aircraft in a stall, while the other group was not asked to do the same thing. Furthermore, the first group was called anticipated group while the second group was called surprised group. Flight parameters used to assess pilot performance when encountering a stall included reaction time, recovery time, air speed, maximal rate of descent, maximal vertical g-load, stick shaker events and total altitude lost. Moreover, there were measurements of physiological conditions, namely the average heart rate and Galvanic Skin Response (GSR) used to find out the level of stress and mental workload. The statistical analysis used was independent t-test without considering outliers. Based on the results of analysis, it was found that there were significant differences in reaction time, air speed, descent rate and altitude lost between the anticipated and surprised groups. In addition, there were differences in stress level between anticipated and surprised groups. Landman et al. suggested to use a more varied scenario to include the surprised condition in abnormal event training.

Based on these two research, this research adopted an abnormal stall event with losing altitude and recovery time performance parameters. In addition, sleep duration was also included in the model referring to Tran et al. [9] who found that sleep deprivation can reduce pilot performance. Furthermore, stress and flight hours were also discussed specifically to find out their relationship with performance before stall occurred between pilots who have and who do not known the stall scenario in the simulation. The analysis in this research was conducted through logistic regression so that the level of pilot performance can be obtained based on the related variables.

3 Research Methods

This research involved 18 final year pilot cadets from Indonesian Aviation High School STPI Curug, who had been licensed as Private Pilot License (PPL). All respondents were male with an average age of 22 years. Based on stall training scenario, this research divided respondents into two groups. The number of respondents who were asked to do flight simulation as co-pilots before the experiment with stall scenario and those who were not asked to do so were 9 respondents each. Those two groups then conducted a simulation with stall scenario as pilot during cruise flight phase with a stage area at an average altitude of 3000 ft.

Stall simulation was carried out in the Simulator Room of Aviation Department of STPI Curug using Mechtronix Ascent FNPT II for Pipper Warrior aircraft. The flight route started from Budiarto Airport, Curug to Serpong area and back to the Budiarto Airport in Curug with a total distance of 7 nautical miles and a total of flight duration of 45 min. The collected data in this research are the profile of respondents, stress level and pilot performance. The profile of respondents data consisting of flight hours and sleep duration of previous night before the experiment are collected through the questionnaires of the respondents before the experiment began. In this research, the stress level is measured before the flight simulation began using saliva collected by using cocoro meter. In 2004, Yamaguchi et al. [10] found that salivary amylase can measure stress more sensitive and faster than cortisol. Moreover, pilot performance is measured when the respondents undergo flight simulation with a predetermined stall scenario indicated by losing altitude and recovery time. When it reaches a height of around 3000 ft. in the cruise phase, stall scenario is performed.

4 The Results of Analysis

The result of experiment are explained in the following table (Table 1).

Table 1. Contingency of respondents' groups with treatment of losing altitude and recovery time

Group	Surprised		Anticipated		Total
LA/RT	<150 ft	>150 ft	<150 ft	>150 ft	
<85 s	2	4	2	1	9
>85 s	1	2	4	2	9
Total	3	6	6	3	18

Based on the Table of Contingency, there are 9 respondents who experiences losing altitude of less than 150 ft., where 67% of them come from the anticipated group. Furthermore, losing altitude of more than 150 ft. is experienced by 6 people from the surprised group and the remaining 3 are from anticipated group. Meanwhile, 9 respondents sucessfully manage to return to the initial altitude in less than 85 s where 67% of them come from the surprised group, and 9 others need more than 85 s.

In this research, losing altitude (LA) was placed as a response variable to see the ability of respondents during recovery after being given stall scenario in flight simulation experiment. The respondents with lower losing altitude were considered to have better abilities. The significance of the parameters will be tested for three explanatory variables, namely pre-experimental treatment, surprised or anticipated condition (SA), flight hours (FH), and sleep duration (SD) which are expected to affect respondents' LA during recovery. The significance of explanatory variables is partially tested through Wald Test with a significance of 10% for each variable after iteration on the logistics model is conducted (Table 2).

Furthermore, from the table above, a mathematical model for binary logistic regression is generated as follows:

$$\hat{\pi}(LA) = 35809.387 - 0.13\ \mathrm{SA} - 0.21\ \mathrm{SD} \tag{1}$$

As a note, the dependent variable is also categorical with reference losing altitude of more than 150 ft. The two variables which are included into the model after two

Table 2. The results of partial wald test

Step	Explanatory variables	B	S.E.	Wald	df	Sig.	Exp(B)
1	SA	−2.03	1.24	2.67	1	0.10	0.13
	FH	0.01	0.02	0.35	1	0.55	1.01
	SD	−1.42	0.92	2.37	1	0.12	0.24
	Constant	8.23	6.63	1.54	1	0.21	3781.60
2	SA	−2.08	1.25	2.78	1	0.09	0.13
	SD	−1.57	0.89	3.06	1	0.08	0.21
	Constant	10.49	5.70	3.38	1	0.06	35809.39

iterations are SA and SD. While, FH variable is not fit to be included into the model after the second iteration. The conclusion of the Wald test is that at the 10% level, whether the respondents experience the stall in the simulation before the experimental simulation (SA) and sleep duration (SD) partially affects the losing altitude in the experimental simulation.

Based on the logistic regression equation, SA variable has a negative slope parameter ($B = -2.08$, $p < 0.1$) which means that respondents who have previously conducted flight simulation as co-pilot and experienced stall have a greater chance to experience losing altitude of less than 150 ft. during recovery in experimental flight simulations with a tendency of 8 times greater (Ex (B) = 0.125) than pilots who have not conducted flight simulation before. The SD variable has negative slope parameter also ($B = -1.571$, $p < 0.1$) which means that respondents with longer sleep duration have a greater chance of experiencing losing altitude of less than 150 ft.uring recovery in experimental flight simulation with a tendency of 5 times greater (Ex (B) = 0.208) than respondents with shorter sleep duration.

By the recovery time, respondents from anticipated group have an average of lower recovery variability (M = 85.4; SD = 8.43) than respondents from surprised group (M = 87.11; SD = 15.68) which indicates respondents from the anticipated group have better ability to manage time in recovery from stall. Furthermore, based on the results regarding the stress level of pilots, respondents with losing altitude of less than 150 ft. and recovery time of more than 85 s tended to experience higher stress level.

5 Discussion

Based on its tendency, pilots who have known that stall will occur during training have a greater chance of performing a better recovery stall or losing altitude of less than 150 ft. In contrast, pilots who encounter sudden abnormal events, according to the nature of the abnormal events themselves, tend to have lower performance with a losing altitude level of more than 150 ft. This shows that pilots are more aware of abnormal events if they have obtained previous information. However, it has already been mentioned, abnormal events contain element of surprise that must be noticed out for because they require greater cognitive readiness.

According to that, Casner et al. [5] and Landman et al. [6] found that pilots experienced more difficulties to conduct recovery under surprised than anticipated condition. Even in anticipated condition, some pilots still found it difficult to provide recovery performance according to criteria. It indicates that the instructions given by flight instructor and a short training before the simulation began was not enough to help the pilots perform recovery stall well.

Ideally, abnormal condition occurred is out of pilot's readiness to be dealt with. In other words, without anticipation in the form of knowing the upcoming abnormality in the flight, the pilot is expected to have good performance when conducting recovery. According to this research, on average, the respondents from anticipated group needed shorter time than the surprised group to recover after experiencing stall. It shows that pilots from anticipated group were able to handle stall situation more efficiently than pilots from surprised group. In addition to saving fuel, time efficiency when encountering abnormal events could give pilots the opportunity to coordinate with ATC to overcome problems after dealing with abnormal events.

According to Hancock and Szalma [11], cognitive abilities are required to encounter them. Kochan et al. [12] also stated that cognitive performance is in the best condition when one's arousal is at the right level. He also reveals that when something unexpected happens and triggers a reaction of surprise, it usually causes a disruption in performance. Therefore, it is recommended to arrange training for new abnormal events that can improve the ability to recognize abnormal events. In 2012, Healy et al. [13] revealed that the ability taught and trained to control abnormal situations on a flight by memorizing during practice did not help the pilots to deal with other abnormal situations compared to how to implant concepts before practice.

This research found that respondents with a longer sleep duration (more than 6 h) have a greater chance of experiencing losing altitude of less than 150 ft. during recovery in an experimental flight simulation with a propensity of 5 times greater than respondents with shorter sleep duration. This findings are similar to the results of the study of Tran et al. [9]), who observed a male pilot with more than 700 h of flight. According to Choo et al. [14], sleep deprivation affects several cognitive functions such as attention, memory and reaction time.

Based on the results regarding the stress level of pilots, respondents with losing altitude of less than 150 ft. and recovery time of more than 85 s tended to experience higher stress level. The higher stress level was allegedly also raised the level of awareness of the surrounding situation so that it resulted to an increase in respondents' performance. This assumption was strengthened by the findings of Tiwari [15] which states that stress can be useful to improve performance. Other allegations are that stress also arises since pilots from anticipated group already knew about stall scenario so that preparation to deal with the event was made earlier before the experiment was carried out where the stress level was also measured at the same time.

This research indicates that the evaluation of pilot training in encountering abnormal situations is needed. It is important to realize that the concept of abnormal events is important to be taken seriously before entering practice in a flight simulator to improve pilot understanding. Trainers must note that memorization practices result in poor learning and give pilots wrong understanding (Shepard [16]; Smith and Fey [17]).

Furthermore, rearrangement including the preparation of the syllabus and the selection of learning methods should also be carried out.

Based on the research conducted, further research are suggested to accommodate more measures of pilot performance such as reaction time and the changes of aircraft angle when pilots perform recovery and pilot's cognitive abilities such as situation awareness and stress level when encountering abnormall situation.

References

1. National Transportation Safety Committee of Indonesia (KNKT). http://knkt.dephub.go.id/knkt/ntsc_aviation
2. Belcastro, C.M., Foster, J.V.: Aircraft loss-of-control accident analysis. In: AIAA Guidance, Navigation, and Control Conference (2013)
3. Kenny, D.J.: Stall and spin accidents: keep the wings flying. Technical report, AOPA Air Safety Institute (2017)
4. BBC News. https://www.bbc.com/news/world-asia-46373125
5. Casner, S.M., et al.: The effectiveness of airline pilot training for abnormal events. Hum. Factors **55**, 477–485 (2013)
6. Landman, A., et al.: The influence of surprise on upset recovery performance in airline pilots. Int. J. Aerosp. Psychol. **27**, 2–14 (2017)
7. Lee, K.: Effects of flight factors on pilot performance, workload, and stress at final approach to landing phase of flight. Electronic Theses and Dissertations University of Central Florida (2010)
8. Aviation Safety Network. https://aviation-safety.net/statistics/
9. Tran, T.Q., Raddatz, K.R., Cady, E.T., Vowels, C.L., Amstutz, B.J., dan Elgin, P.D.: Effect of extreme sleep deprivation on human performance. In: Idaho Academy of Science Symposium and Meeting (2007)
10. Yamaguchi, M., et al.: Performance evaluation of salivary amylase activity monitor. Biosens. Bioelectron. **20**, 491–497 (2004)
11. Hancock, P.A., Szalma, J.L.: Performance Under Stress. CRC Press, Florida (2008)
12. Kochan, J.A., Breiter, E.G., Hilscher, M.B., Priest, J.E.: Pilots' perceptions and retention of in-flight upset recovery training: evidence for review and practice. In: Proceedings of the Human Factors and Ergonomics Society (2005)
13. Healy, A.F., Bourne, L.E.: Empirically valid principles for training in the real world. Am. J. Psychol. **126**, 389–399 (2013)
14. Choo, W.C., Lee, W.W., Venkatraman, V., Sheu, F.S., Chee, M.W.L.: Dissociation of cortical regions modulated by both working memory load and sleep deprivation and by sleep deprivation alone. Neuroimage **25**, 579–587 (2005)
15. Tiwari, G.K.: Stress and human performance. Indo-Indian J. Soc. Sci. Res. **7**, 40–49 (2011)
16. Shepard, L.A.: The role of assessment in a learning culture. Educ. Res. **29**, 4–14 (2000)
17. Smith, M.L., Fey, P.: Validity and accountability in high-stakes testing. J. Teach. Educ. **51**, 334–344 (2000)

Criteria Indicators of the Consistency of Air Traffic Controllers' Preferences on a Set of Characteristic Errors

Oleksii Reva[1(✉)], Volodimir Kamyshyn[1], Andrii Nevynitsyn[2], Valerii Shulgin[2], and Serhiy Nedbay[3]

[1] Ukrainian Institute of Scientific and Technical Expertise and Information, Kyiv, Ukraine
ran54@meta.ua, kvv@ukrintei.ua

[2] Flight Academy of the National Aviation University, Kropyvnytskyi, Ukraine
{nevatse, vashulgin}@ukr.net

[3] Flight School «Condor», Kyiv, Ukraine
s.nedbay@outlook.com

Abstract. Based on ICAO recommendations and statistics on aviation events and incidents during air traffic control, a spectrum of characteristic errors of n = 21 air traffic controllers was generated, which most fully and comprehensively illustrates their erroneous decisions. To establish the coefficients, a mathematical method of ranking priorities was applied. To summarize the views of m = 37 professional air traffic controllers involved in the experiment, regarding the frequency and hazards of errors identified in the ICAO scale, a multiplicative approach was implemented. It has been founded that the error' rate indistinguishability indicator in a group system of preferences is equal $R^*_{LF} = 1,23 \cdot 10^{-2}$, and the error' danger indistinguishability indicator is three times worse and equal $R^*_{LD} = 3,77 \cdot 10^{-2}$. Integrative (holistic) indicator of each error' significance summarizes its significance coefficients in terms of frequency and hazards and is multiplicative. That allowed building the final group system of preferences. The degree of indistinguishability of the errors significance is equal $R^*_g = 0,91 \cdot 10^{-2}$, which testifies to the effectiveness of the proposed approach to the implementation of ICAO recommendations regarding the frequency and the dangers of unwanted events in flight safety management.

Keywords: Flight safety · Human factor · Air traffic controllers · Decision making · System of preferences · Characteristic errors · Frequency and hazard of consequences · Coefficients of significance · Multiplication · Integrative assessment

1 Introduction

Aviation events are usually the result of errors of the "front line" aviation operators (AO), including air traffic controllers (ATC). Therefore, it is relevant to study the ability of the ATC to mentally predict the false consequences of erroneous actions, the formation of skills to distinguish them, recognize, remember, and therefore, prevent.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 617–623, 2020.
https://doi.org/10.1007/978-3-030-50943-9_79

ATC professional activity is the essence of a continuous chain of decisions that are made and implemented in explicit and implicit forms under the influence of many factors of various nature. And of course, the wrong actions are usually the result of the wrong decisions. Therefore, it is equally important to identify the impact of the human factor on the decision-making processes of ATC. Indicators of this impact (major decision-makers, levels of harassment, and fuzzy risk assessments) relate specifically to the human factor and reveal the concept of "attitude to dangerous actions or conditions," which explains the interplay of other components of the ICAO Safety paradigm [1, 2].

However, since preference systems (PS) are an integral part of decision-making, they must be included in the list of components of this "attitude" and therefore investigated.

2 Development of the Algorithm of Personality-Oriented Simulator Training for Air Traffic Control Students

Based on world accidents and incidents statistics at ATC, ICAO recommendations, as well as personal experience of practical ATC, teaching and instructor work of the authors of this publication, a list of n = 21 characteristic errors was generated, which is currently the most complete and comprehensive coverage of the inappropriate actions of air traffic controllers [3–7]:

$Er._1$ – Violation of radiotelephony phraseology
$Er._2$ – Inconsistent entry of the aircraft into the zone of the adjacent ATC
$Er._3$ – Violation of longitudinal course time separation
$Er._4$ – Violation of counter course time separation
$Er._5$ – Violation of separation between aircrafts at crossed courses
$Er._6$ – Address less ATC messaging
$Er._7$ – Error in determining of aircraft call sign
$Er._8$ – Error in aircraft identification
$Er._9$ – Misuse of ATC schedule
$Er._{10}$ – Absence of mark of the control transfer to the adjacent Air traffic control center in the strip
$Er._{11}$ – Absence of mark of the coordination of the entrance of the aircraft to the adjacent ATC area in the strip.
$Er._{12}$ – Violation of coordinated geographic boundary by ATC
$Er._{13}$ – Violation of coordinated time of control transfer at FIR - boundary by ATC
$Er._{14}$ – Negligence in applying to the strip of the letter-digital information (the possibility of double interpretation)
$Er._{15}$ – Non-economical ATC
$Er._{16}$ – Violation of shift handover procedures
$Er._{17}$ – Issued commands to change the altitude or direction of flight are not reflected on the strip
$Er._{18}$ – Attempt to control the aircraft under condition of TCAS system operation in the resolution advisory mode

$Er._{19}$ – Errors in entering information about aircraft into an automated system
$Er._{20}$ – Violation of emergency procedures
$Er._{21}$ – Violations of airspace use

The following methods and strategies of decision-making were applied in establishing the system of advantages of ATC's on the spectrum of characteristic errors: pairwise comparison and part of total intensity; the median of Kemeni; classic decision-making criteria (Wald, Savage, Bayes-Laplace); optimal prediction.

The results obtained are clearly proactive and in line with ICAO's Flight Safety Management Policy.

It was discovered that professional air traffic controllers who were involved in our tests and carried out $n(n-1)/2 = 210$ pairwise comparisons of errors from the Table 1, defining individual LC before simulator training, it was assumed during training for one-third less mistakes than those who were not covered by our research [8–11]. That is, the formation of the same mental skills prediction of the erroneous consequences of erroneous actions through their differentiation, recognition, memorization, and, consequently, the prevention that was covered in the introduction.

However, all the studies concerned building a system of preferences for ATCs' considering only the dangers of the consequences of errors, although it would be advisable to consider their frequency as well. It should be noted that the linguistic indicators of hazards and incidence of adverse events recommended by ICAO have a pronounced comparative significance, which opened the prospects for applying a mathematical method of prioritization to establish adequate coefficients of significance for these indicators [3].

Based on the results of the V-i iteration of the application of the distribution of priorities method, we obtain a nomogram that allows correlation of qualitative and quantitative factors of frequencies and the dangers of errors (Fig. 1).

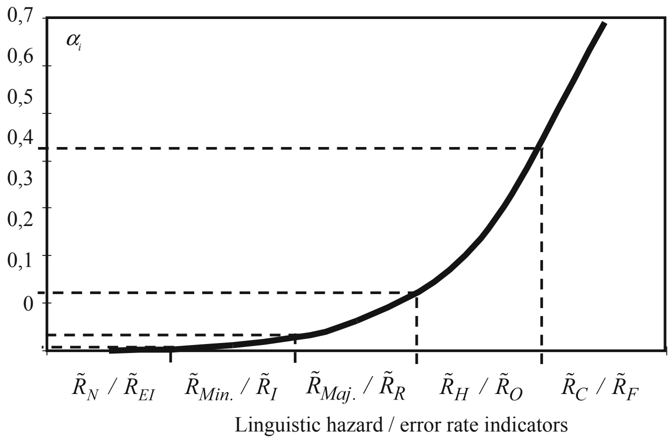


Fig. 1. Nomogram of conformity of linguistic indicators dangers/frequency of errors and their corresponding coefficients of significance.

3 The Use of Indicators of Hazards and Frequencies to Build Systems Preferences of ATCs' on the Spectrum of Characteristic Errors

The research involved m = 37 professional air traffic controllers, who were offered to use their linguistic scales ICAO to express their views on the frequency of manifestations of the typical errors listed in Table 1, and the dangers of their consequences.

A sample of the results is illustrated by Fig. 2, which contains histograms of responses to the error in Error 1 ("Violation of radiotelephony phraseology"), which, according to respondents, air traffic controllers are most often assumed. And errors in Error 18 (Attempt to control the aircraft under condition of TCAS system operation in the resolution advisory mode), which is the most dangerous of all their spectrum.

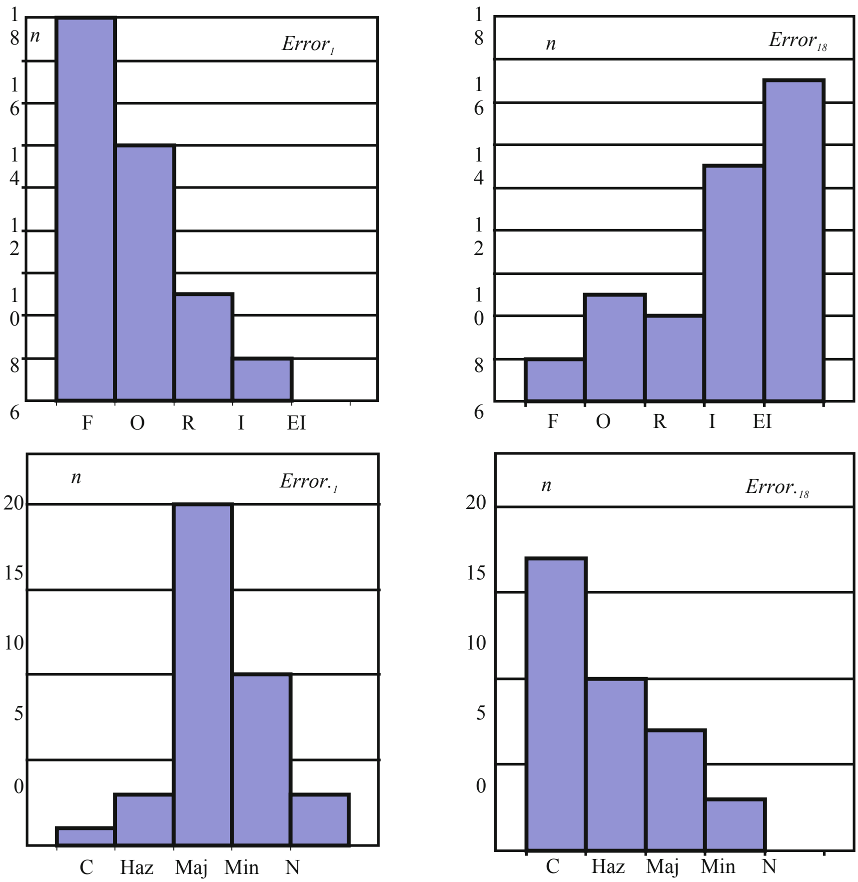


Fig. 2. Histogram of statistics of opinions of experienced air traffic controllers regarding the frequency and danger of manifestation of errors Error 1 and Error 18

A multiplicative approach was applied for obtaining integral (holistic) estimates of each and every error in the LD $\left(\alpha^{LD}_{Er.i}\right)$ and LF $\left(\alpha^{LF}_{Er.i}\right)$ parameters:

$$\begin{aligned}
&Er._{1} \underset{LF}{\succ} Er._{21} \underset{LF}{\succ} Er._{15} \underset{LF}{\succ} Er._{7} \underset{LF}{\succ} Er._{9} \underset{LF}{\succ} Er._{12} \underset{LF}{\succ} Er._{4} \underset{LF}{\succ} \\
&\underset{LF}{\succ} Er._{13} \underset{LF}{\approx} Er._{10} \underset{LF}{\approx} Er._{2} \underset{LF}{\approx} Er._{11} \underset{LF}{\succ} Er._{18} \underset{LF}{\approx} Er._{3} \underset{LF}{\approx} Er._{14} \underset{LF}{\succ}, \\
&\underset{LF}{\succ} Er._{16} \underset{LF}{\approx} Er._{20} \underset{LF}{\succ} Er._{6} \underset{LF}{\approx} Er._{8} \underset{LF}{\approx} Er._{17} \underset{LF}{\succ} Er._{5} \underset{LF}{\succ} Er._{19}
\end{aligned} \tag{1}$$

$$\begin{aligned}
&Er._{18} \underset{LD}{\succ} Er._{5} \underset{LD}{\succ} Er._{20} \underset{LD}{\succ} Er._{17} \underset{LD}{\succ} Er._{15} \underset{LD}{\succ} Er._{14} \underset{LD}{\approx} Er._{10} \underset{LD}{\succ} \\
&\underset{LD}{\succ} Er._{12} \underset{LD}{\succ} Er._{13} \underset{LD}{\approx} Er._{9} \underset{LD}{\succ} Er._{21} \underset{LD}{\succ} Er._{16} \underset{LD}{\succ} Er._{19} \underset{LD}{\succ} Er._{4} \underset{LD}{\approx}, \quad \text{where} \\
&\underset{LD}{\approx} Er._{7} \underset{LD}{\approx} Er._{1} \underset{LD}{\approx} Er._{8} \underset{LD}{\approx} Er._{6} \underset{LD}{\approx} Er._{11} \underset{LD}{\approx} Er._{2} \underset{LD}{\succ} Er._{3}
\end{aligned} \tag{2}$$

$\underset{LF}{\succ}, \underset{LF}{\approx}$ – a mark corresponding to the preference of one error before the other in terms of the frequency of manifestation, as well as the adequacy of errors in the frequency of manifestation.

$\underset{LD}{\succ}, \underset{LD}{\approx}$ – a mark corresponding to the preference of one error before the other in terms of danger, as well as the adequacy of mistake' danger

An expression is proposed to determine the ATC coefficient of indistinguishability R^* that varies within $R^* = \overline{0,\ 1}$ It is established that $R^*_{LF} = 1,23 \cdot 10^{-2}$, and $R^*_{LD} = 3,77 \cdot 10^{-2}$. That is, ATCs recognize the risk of errors three times worse than the frequency of their manifestation.

Using the same multiplicative approach, we obtained integrative indices of each error, which took into account comprehensively both the frequency and the risk of errors. The following final group system of preferences was obtained:

$$\begin{aligned}
&Er._{18} \underset{g}{\succ} Er._{5} \underset{g}{\succ} Er._{1} \underset{g}{\succ} Er._{3} \underset{g}{\succ} Er._{20} \underset{g}{\succ} Er._{15} \underset{g}{\succ} Er._{21} \underset{g}{\succ} \\
&\underset{g}{\succ} Er._{9} \underset{g}{\approx} Er._{10} \underset{g}{\approx} Er._{12} \underset{g}{\succ} Er._{13} \underset{g}{\approx} Er._{14} \underset{g}{\approx} Er._{17} \underset{g}{\succ} Er._{7} \underset{g}{\approx}, \quad \text{where} \\
&\underset{g}{\approx} Er._{16} \underset{g}{\succ} Er._{2} \underset{g}{\approx} Er._{11} \underset{g}{\succ} Er._{4} \underset{g}{\approx} Er._{6} \underset{g}{\approx} Er._{8} \underset{g}{\succ} Er._{19}
\end{aligned} \tag{3}$$

$\underset{g}{\succ}, \underset{g}{\approx}$ – marks respectively the advantages and adequacy of errors in significance in the group system of preferences.

The degree of indistinguishability of the errors significance sets the value $R^*_g = 0,91 \cdot 10^{-2}$

4 Conclusions

1. It was proposed to add to the components that create the attitude of aviation personnel to dangerous actions or conditions and are determined by the influence of the Human Factor on the decision-making processes, system of preferences of "front line" aviation operators on the indicators and characteristics of professional activity.

2. List of n = 21 most typical mistakes, which air traffic controllers make during their professional duties, was formed.
3. Based on the recommendations of ICAO on the use of qualitative indicators of frequency and the dangers of unwanted events, the term-sets of the corresponding linguistic variables “LD” and “LF” are formed.
4. Based on the obvious distribution of LD and LF indicators, a mathematical method of ranking priorities was used to determine the coefficients of the significance of these indicators. That made it possible to get group systems of advantages of m = 37 professional air traffic controllers on frequency and sky-high error rates.
5. The normalized index of error indistinguishability is introduced. It is determined that the indistinguishability index of the hazards of errors is equal to $R^*_{LD} = 3,77 \cdot 10^{-2}$ and three times worse than the index of the indistinguishability of their frequency $\left(R^*_{LF} = 1,23 \cdot 10^{-2}\right)$.
6. Multiplicatively obtained integral indicators of the significance of each error and consequently, aggregated in frequency and risk of errors in the group systems of preferences. The degree of indistinguishability of the errors significance sets the value $R^*_g = 0,91 \cdot 10^{-2}$ and is 1,35 times better than in the group system of preferences on the error rate indicators, and is 4,14 times better than in the group system of preferences in the danger of errors. This testifies to the effectiveness of the proposed approach to the implementation of ICAO recommendations regarding the frequency and risks of undesirable events in flight safety management processes.
7. The scientific results obtained in this way proposed for using in the simulator training of air traffic controllers, as well as the implementation of measures of flight safety management in aeronautical navigation systems.

References

1. Reva, A.N., Tumyishev, K.M., Bekmuhambetov, A.A.: Chelovecheskiy faktor i bezopasnost poletov: (Proaktivnoe issledovanie vliyaniya): monografiya [Human factor and safety: Proactive impact research: monograph]. Almaty, p. 242 (2006)
2. Global Aviation Safety Plan (GASP) 2017–2019: Doc. ICAO 10104. - Montreal, Canada (2016)
3. Nasirov, Sh.Sh.: Pilotne vyznachennia system perevah aviadyspetcheriv Azerbaidzhanu na kharakternykh pomylkakh v protsesi upravlinnia povitrianym rukhom Aviacijno-kosmicna tehnika i tehnologia - Aerospacetechnic and technology, **7**, pp. 124–134 (2010). (In Ukraine)
4. Reva, A.N., Nasirov, Sh.Sh., Mirzoyev, B.M., Nedbay, S.V.: Empiricheskiye modeli otsenki riska-neopredelennosti gruppovykh sistem predpochteniy aviadispetcherov. Elmi məcmuələr: Jurnal Milli Aviasiya Akademi-yasinin, - Baki, iyul – sentyabr 2012. – T.14. - № 3. – C.46 60
5. Reva, O.M., Kamyshyn, V.V., Nasyrov, Sh.Sh.: Mediana Kemeni yak hrupova systema perevah aviadyspetcheriv na mnozhyni kharakternykh pomylok. Aviatsiyno-kosmichna tekhnika i tekhnolohiya: naukovo-tekhnichnyy zhurnal. – KH.: Natsionalʹnyy aerokosmichnyy universytet im. M. YE. Zhukovsʹkoho « KHAI » , - № 4. – S. 106–115 (2012)

6. Reva, O.M., Nasirov, Sh.Sh., Lypchans′kyy, V.O.: Stratehiya optymal′noho peredbachennya u vyznachenni stavlennya aviadyspetcheriv do nebezpek kharakternykh pomylok. Aviatsiyno-kosmichna tekhnika i tekhnolohiya: nauk.-tekhn. zh. – KH.: Natsional′nyy aerokosmichnyy universytet im. M. YE. Zhukovs′koho « KHAI » , - № 8. – S. 126–138 (2017)
7. Reva, A.N., Nasirov, Sh.Sh., Myrzoev, B.M.: Éffektyvnost′ metodov opredelenyya hruppovykh system predpochtenyy dyspetcherov na opasnosty kharakternykh oshybok, sovershaemykh v protsesse upravlenyya vozdushnym dvyzhenyemy. Aviatsiyno-kosmichna tekhnika i tekhnolohiya: nauk.-tekhn. zh. – KH.: Natsional′nyy aerokosmichnyy universytet im. M. YE. Zhukovs′koho « KHAI » , - № 6. – S. 93–103 (2018)
8. Reva, O.M., Borsuk, S.P., Shylgin, V.A., Mirzoev, B.M., Mukhtarov, P.Sh., Nasirov, Sh. Sh.: Stavlennia aviatsiinykh operatoriv « perednoho kraiu » do nebezpechnykh dii abo umov profesiinoi diialnosti – holovnyi chynnyk zabezpechennia bezpeky polotiv [The attitude of the "front edge" aviation operators to dangerous actions or conditions of professional activity – the main factor of flights safety]. Materiali 8 Mizhnarodnoi naukovo-praktychnoi konferentsii « Suchasni informatsiini ta innovatsiini tekhnolohii na transporti (MINTT-2016) » [Proc. 8th International scientificpractical conference « Modern information and innovative technologies in transport MINTT-2016) »]. Kherson, 24–26 May, Kherson State Marine Academy Publ, pp. 90–97 (2016). (In Ukraine)
9. Reva, O., Borsuk, S., Mirzayev B., Mukhtarov P.: New approach to determination of main solution taking dominant of air traffic controller during flight level norms violation. Advances in Human Aspects of Transportation: Proceedings of the AHFE 2016 International Conference on Human Factors in Transportation, 27–31 July 2016, Walt Disney World, Florida, USA. – pp. 137–147 (2016)
10. Borsuk, S.P.: New methods for air traffic controller main solution taking dominant determination concerning their attitude to risk/S.P. Borsuk // Logistics and Transport, - no 1. – pp. 25–29 (2017)
11. Reva, O.M., Borsuk, S.P., Nasirov, Sh.Sh., Seleznov, G.M. Teoretychne obgruntuvannia systemno-informatsiinoi kvalimetrii liudskoho chynnyka v aeronavihatsiinykh systemakh. Materiali 8 Mizhnarodnoi naukovo-praktychnoi konferentsii «Suchasni enerhetychni ustanovky na transporti, tekhnolohii ta obladnannia dlia yikh obsluhovuvannia SEUTTOO-2017» [Proc. 8th International scientific-practical conference «Modern energy installations on transport, technologies and equipment for their maintenance MEITTEM-2017»]. Kherson, Kherson State Marine Academy Publ., pp. 127–132 (2017). (In Ukraine)

High-Level Review Principles for Human-Machine Interface Design of Civil Aircraft Flight Deck

Fei Li(✉) and Kaiwen Chen

Shanghai Aircraft Design and Research Institute, Shanghai 201210, China
{lifei, chenkaiwen}@comac.cc

Abstract. High-level human factor engineering principles, which could be used to develop many detailed review guidelines in the design and evaluation of Civil Aircraft Flight Deck Human-Machine Interface, have been introduced in this paper. These principles only address the human factor engineering aspects of the Flight Deck design, but no other related considerations, such as instrumentation, control or structural design. The human factor engineering review principles are organized into four categories: general principles, awareness principles, workload principles and error management principles.

Keywords: Human factors · Human-Machine Interface · Civil Aircraft Flight Deck

1 Introduction

The process by which civil flight decks are designed is complex, largely unwritten, variable and non-standard [1]. The process also overly relies on the knowledge and experiences of individuals involved in each program [2]. The design of human-machine interfaces (HMI) should support the flight crew's primary tasks of aviation, navigation, communication and system management, without imposing an excessive workload associated with using the HMI. The HMI should also support the tolerance, recognition and recovery from flight crew's errors. Human Factor Engineering design reviews help to ensure that these goals are achieved. The 16 Human Factor Engineering design review principles, which could be divided into four categories: general principles, awareness principles, workload principles and error management principles (summarized in Table 1), are introduced in this paper.

2 General Principles

These principles ensure that the flight deck HMI design supports flight crew safety and is compatible with flight crew's general cognitive and physiological capabilities.

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 624–631, 2020.
https://doi.org/10.1007/978-3-030-50943-9_80

Table 1. Human Factor Engineering Design review principles

No.	Category	Principle
1	General	Flight crew safety and efficiency Cognitive compatibility Physiological compatibility Simplicity of design Consistency
2	Awareness	Situation awareness Task compatibility Human model compatibility Organization of HMI Elements Timeliness Feedback
3	Workload	Mental workload Physical workload
4	Error management	Error prevention Error detection Error recovery

2.1 Flight Crew Safety and Efficiency

The flight crew is directly responsible for the safety of the aircraft and given today's technology, safety is enhanced in a complex and dynamic environment if flight crew make final command decisions. Each decision of the design of Flight Deck HMI should consider overall flight safety and efficiency. Combining flight crew performance and flight deck system performance is more important than partial optimization of the performance of any human or automated component in that system.

2.2 Cognitive Compatibility

The flight crew's role should consist of purposeful and meaningful tasks that enable flight crew to keep familiar with the flight deck and maintain a level of workload that is not so high as to negatively affect performance, but enough to maintain vigilance.

2.3 Physiological Compatibility

The design of the Flight Deck HMI should reflect consideration of human physiological characteristics including visual/auditory perception, biomechanics (reach and motion), characteristics of motivation control and anthropometry.

2.4 Simplicity of Design

The HMI should represent the simplest design consistent with functional and task requirements. The goal is to ensure that the HMI solution is easy to operate in normal, non-normal and emergency conditions to the objective levels of performance by more than 95% of the pilots.

2.5 Consistency

There should be a high degree of consistency between HMI and procedures. Talking about HMI, the way the system functions should always be consistent with how it appears to the flight crew, reflecting a high degree of standardization and consistency with procedures.

Flight deck HMI features mainly include displays, controls, automation, and alerting. In Table 2, these flight deck HMI features are combined to produce high-level review guidelines for consistency.

Table 2. High-level review guidelines for Flight Deck HMI consistency.

Flight deck HMI features	Displays	Controls	Automation	Alerting
Displays	Symbol Format Layout Coding			
Controls	Direction Proportion Movement	Symbol Layout Motion		
Automation	Commanded/actual Display management	Feedback Override Pilot review/ confirmation	Interactions	
Alerting	Content Salience Highlighting Consistency Display access	Acknowledgment Review	Authority limits Faults	Consistency Conflicts Contradictions Combinations Distinction

3 Awareness

These principles support the flight crew's awareness for their tasks of aviation, navigation, communication and system management.

3.1 Situation Awareness

The information presented to the flight crew by the HMI should be correct, rapidly recognized, easily understood and support the higher-level goal of user awareness of the status of the system.

Because situation awareness relies to some degree upon the flight crew being actively involved, the design of the flight deck must make sure that the level of engagement of the flight crew is appropriate for all critical flight functions and tasks for which he or she is currently responsible. If automation performs critical flight functions for which the pilot serves as a backup, a level of involvement must be maintained under normal conditions such that the pilot is prepared to take over the function under non-normal conditions.

3.2 Task Compatibility

The HMI should meet the requirements of flight crew to perform their tasks. Data should be presented in forms and formats appropriate to the task. Control options should encompass the range of potential actions. There should be no unnecessary information or control options.

3.3 Human Model Compatibility

All aspects of the system should be consistent with the flight crew's mental models such as:

- The organization of display functions should support simple and accurate mental models to aid the pilot in figuring out how to access required information.
- The operation of control devices should be consistent with the simple and accurate mental models the flight crew develop in training of the process being controlled.
- Automation should support simple and accurate mental models of its operation so the flight crew can easily determine what automation is doing and anticipate what it is going to do.

3.4 Organization of HMI Elements

The organization of all aspects of the HMI should be based on pilot requirements and should reflect the general principles of organization by importance, frequency and order of use. Critical safety-function information should be available to the entire operating crew in dedicated locations to ensure its recognition and to minimize response.

3.5 Timeliness

The HMI design should consider flight crew's cognitive processing capabilities as well as process-related time constraints to ensure that tasks can be performed within the time required. Information flow rates and control performance requirements that are too fast or too slow could degrade performance.

- The response time of feedback to control inputs should be rapid enough to prevent multiple input attempts.
- Displayed responses to flight control inputs should be rapid enough to prevent pilot induced oscillations.
- Displayed response times should provide positive feedback of all flight crew inputs quickly enough that the pilot does not attempt to repeat the input, thinking that the first attempt was unsuccessful, and quickly enough to prevent the pilot from having to perform tasks more slowly than the natural pace or the pace required by other factors.

- If the system cannot provide the final result of a commanded process within the required time, it should still indicate to the pilot that it is in processing.

3.6 Feedback

To ensure appropriate system feedback, the flight crew should always have access to information about what the various on-board systems are doing and that access should be appropriate to the current pilot responsibilities. The HMI should provide useful information on system status, permissible operations, errors and error recovery, dangerous operations and validity of data.

- Feedback about the current states of the automated systems and the process they are controlling should always be provided. This allows the crew to determine process state from observing or monitoring the behaviour of the control device, and it facilitates graceful transfer of control because the crew can assume control with the device in the proper position.
- Positive feedback should be given for the operation of every control device. Tactile feedback should be provided to indicate that the control has been actuated and positive feedback of system acknowledgment should be provided to prevent the pilot from attempting multiple inputs due to lack of response.
- Control devices should provide feedback as to the effects of automation on the states of the processes they are controlling.

4 Workload

Workload reflects the relationship between the physical and/or mental demand imposed by a task and one's capacity to perform that task. Workload describes how busy the flight crew is, how complex the tasks are that are being performed and whether the flight crew can manage or perform additional tasks [3]. The effect of workload on performance can be described by a U-shaped curve; workload that is too high or too low results in poor performance. When workload is too low, the pilot may become inattentive or bored and devote less concentration to the task at hand. On the other hand, workload that is too high may cause the pilot to miss information, fail to perform tasks in a timely manner or make errors.

4.1 Mental Workload

The information presented by the HMI should be rapidly recognized and understood; therefore, the system should minimize requirements for making mental calculations or transformations and use of recall memory (recalling lengthy lists of codes, complex command strings and information from one display to another, or lengthy action sequences). Automation should not increase task difficulty or mental demand, result in extreme levels of workload, or distract the pilot from other tasks [4].

Equipment operating procedures should be designed to maximize operational suitability, minimize pilot workload and minimize reliance on pilot memory [5].

4.2 Physical Workload

The HMI must allow the minimum flight crew to perform their duties without unreasonable concentration or fatigue and should require a minimum number of actions to accomplish an action; e.g., single versus command keying, menu selection versus multiple command entry, single input mode versus mixed mode. In addition, the system should not require the entry of redundant data, nor the re-entry of information already in the system, or information can generate from already resident data.

There are several ways to measure workload:

- Collect physiological measures, such as heart rate and evoked brain potential.
- Collect subjective measures (e.g. NASA-TLX and the SWAT).

5 Error Management

These principles address the characteristics of the HMI that provide the error management for flight crew, such as providing:

- Error Prevention.
- Error Detection.
- Error Recovery.

5.1 Error Prevention

- The HMI should provide controls and indications that can be used either to reverse an erroneous action directly so that the airplane or system is returned to the original state, or to mitigate the error's effect so that the airplane or system is returned to a safe state.
- The HMI should provide available information so the flight crew can become aware of an error or a system/airplane state resulting from a system action, and make sure a clear relationship between flight crew action and the error so recovery can be made in a timely manner [6].

5.2 Error Detection

- Indication of an error should be provided within 0.2 s of the data entry [7].
- The system should be designed to detect and trap errors (e.g., out-of-range values, invalid alphanumeric characters) as the pilot inputs are entered
- The system should check for input errors after data is keyed in but before they are processed.
- A positive indication of failures of the system, the system hardware components, the loss of a data link connection and the failure to successfully send or receive a message shall be provided.

5.3 Error Recovery

- The system shall permit correction of individual errors without requiring re-entry of correctly entered commands or data elements.
- Error messages should describe the problem and the recommended solution in very specific terms, that is, they should tie the problem to a process, object, action, data entry field or other data element.
- The user should be able to correct errors immediately before the effects of the error propagate through the system.
- The HMI design should not obscure, compound, or exacerbate the presence or effect of an error.
- The HMI should be designed to prevent flight crew's error by providing unambiguous indications of system state, and to allow easy recovery from flight crew's error by ensuring that inputs can be cancelled or reversed.
- For error recovery, the systems should be designed to allow the pilot to easily correct any errors, assuming that the guideline above for rapid detection of errors is followed and errors are detected quickly.

6 Summary and Conclusions

Human-centred design processes will ensure efficient, effective, user acceptable Flight Deck HMI that will be easy to train, operate and maintain. Human Factor Engineering evaluation play a very important role in the design of the Flight Deck HMI. The "high-level" design-review principles here represents the generic HMI characteristics necessary to support flight crew performance. While these principles are not detailed review guidelines, they may serve several purposes. First, they might be used to develop many of the detailed review guidelines for HMI design. Second, as general principles, they could be used to support the preliminary evaluation of aspects of the novel HMI design, such as new display formats or new interactive method etc.

References

1. A crew-centered flight deck design philosophy for high-speed civil transport aircraft. NASA TM-109171, Jan 1995
2. The interfaces between flight crews and modern flight deck systems. FAA, Jun 1996
3. Wickens, C.D., Hollands, J.G., Banbury, S., Parasuraman, R.: Engineering Psychology and Human Performance. Prentice Hall, Upper Saddle River (2016)
4. Ahlstrom, V., Longo, K.: Human factors design standard (HFDS) for acquisition of commercial off-the-shelf subsystems. Non-Developmental Items, and Developmental Systems (DOT/FAA/CT-03/05 HF-STD-001), Atlantic City, NJ: US DOT Federal Aviation Administration William J. Hughes Technical Center (2003)
5. RTCA DO-229D. Minimum operational performance standards for global positioning system/wide area augmentation system airborne equipment

6. Advisory Circular (AC) 25.1302-1. Installed systems and equipment for use by the flightcrew. 3 May 2013
7. Military Standard (MIL-STD) 1472G. Department of Defense Design Criteria Sdtandard: Human Engineering. Washington, DC: US Department of Defense (2012)

Author Index

N. Stanton (Ed.): AHFE 2020, AISC 1212, pp. 633–636, 2020.
https://doi.org/10.1007/978-3-030-50943-9