JYU DISSERTATIONS 156

Hannu Karvonen

User Experience Goals in Human-Centred Design of Safety-Critical Systems

UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF INFORMATION TECHNOLOGY

JYU DISSERTATIONS 156

Hannu Karvonen

User Experience Goals in Human-Centred Design of Safety-Critical Systems

Esitetään Jyväskylän yliopiston informaatioteknologian tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Agora-rakennuksen auditoriossa Aud 3
joulukuun 14. päivänä 2019 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Information Technology of the University of Jyväskylä,
in building Agora, auditorium Aud 3, on December 14, 2019 at 12 o'clock noon.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2019

Editors
Marja-Leena Rantalainen
Faculty of Information Technology, University of Jyväskylä
Timo Hautala
Open Science Centre, University of Jyväskylä

Cover image by intographics & xresch / Pixabay, image editing by Timo Hautala and Hannu Karvonen

This is a printout of the original online publication.
Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-7984-3

ISBN 978-951-39-7984-3 (PDF)
URN:ISBN:978-951-39-7984-3
ISSN 2489-9003

Jyväskylä University Printing House, Jyväskylä 2019

ABSTRACT

Karvonen, Hannu
User experience goals in human-centred design of safety-critical systems
Jyväskylä: University of Jyväskylä, 2019, 141 p. (+appendix and included articles)
(JYU Dissertations
ISSN 2489-9003; 156)
ISBN 978-951-39-7984-3

This dissertation studies the usage of user experience (UX) goals in early-stage human-centred design activities regarding safety-critical systems. The purpose of UX goals is to describe the desired user experiences to be focused on in the design work. The safety-critical technology environments of the empirical research of this thesis include rapid transit systems, container cranes in ports, command bridges of ships, and cars with driver-assistance systems. The empirical cases of the dissertation focus specifically on these four environments and their associated safety-critical systems' human activity analysis, concept/prototype design, and early-phase evaluation from the UX goals perspective.

The studies in the cases have been conducted with different human factors and UX approaches, methods, and techniques. Core-Task Analysis is used as a key method in the cases with the analysis phase study reported in this dissertation. As analysis methods for the user evaluation studies' results included in the thesis, the Usability Case approach, qualitative analysis of UX survey results, and quantitative statistical analysis of user acceptance survey results are used.

The specific research questions and results of this dissertation are related, firstly, to considering the value of experience design as part of safety-critical systems development in general. Secondly, the contribution of user activity analysis methods for the identification of UX goals is studied. Thirdly, the systematic specification, utilisation, and evaluation of UX goals in human-centred design of safety-critical systems is in the core focus of this thesis.

The results of this thesis indicate that UX goals bring additional value to the human-centred design work of safety-critical systems development. UX goals help in taking into account the experiences of users analytically in the concept or prototype design and evaluation stages. Moreover, the systematic use of UX goals makes the human-centred design and early-stage evaluations of safety-critical systems more target-driven, accurate, traceable, and measurable. Finally, UX goals' appropriate usage in design can be seen to contribute to the meaningfulness, safety, effectiveness, and efficiency of human activity in safety-critical technology environments in general.

Keywords: user experience goal, human-centred design, safety-critical systems, experience design, human factors

TIIVISTELMÄ (ABSTRACT IN FINNISH)

Karvonen, Hannu
Käyttäjäkokemustavoitteet turvallisuuskriittisten järjestelmien
käyttäjäkeskeisessä suunnittelussa
Jyväskylä: University of Jyväskylä, 2019, 141 s. (+liite ja artikkelit)
(JYU Dissertations
ISSN 2489-9003; 156)
ISBN 978-951-39-7984-3

Tämä väitöstyö tutkii käyttäjäkokemustavoitteiden käyttöä turvallisuuskriittisten järjestelmien varhaisen vaiheen käyttäjäkeskeisessä suunnittelutoiminnassa. Käyttäjäkokemustavoitteiden tarkoitus on kuvata käyttäjäkokemukset, joihin suunnittelutyössä tähdätään. Väitöskirjan empiiristen tutkimusten turvallisuuskriittisiä kohdeympäristöjä ovat metrot, satamakonttinosturit, laivojen komentosillat ja kuljettajan apujärjestelmällä varustetut henkilöautot. Väitöstyön empiiriset tutkimukset keskittyvät näihin ympäristöihin liittyvien teknologiajärjestelmien käyttäjien toiminnan analyysiin, konsepti-/prototyyppisuunnitteluun ja varhaisen vaiheen arviointiin käyttäjäkokemustavoitteiden näkökulmasta.

Tutkimukset on toteutettu erilaisilla ihmistoiminnan analyysin ja käyttäjäkokemustutkimuksen lähestymistavoilla, menetelmillä ja tekniikoilla. Päämenetelmänä analyysivaiheen tutkimuksissa on käytetty perustehtäväanalyysiä. Käyttäjäarviointien tulosten osalta analysointimenetelminä on käytetty Usability Case -menetelmää, käyttäjäkokemuskyselyiden tulosten kvalitatiivista analyysiä sekä käyttäjähyväksyntää koskevan kyselyn tulosten tilastollista analyysiä.

Tutkimuksen tarkemmat tutkimuskysymykset ja -tulokset liittyvät kokemussuunnittelun arvoon osana turvallisuuskriittisten järjestelmien kehitystä, ihmistoiminnan analyysimenetelmien osuuteen käyttäjäkokemustavoitteiden tunnistamisessa sekä käyttäjäkokemustavoitteiden systemaattiseen määrittelyyn, hyödyntämiseen ja arviointiin.

Väitöstyön tulokset osoittavat, että käyttäjäkokemustavoitteet tuovat lisäarvoa turvallisuuskriittisten järjestelmien käyttäjäkeskeiseen suunnittelutyöhön. Käyttäjäkokemustavoitteet auttavat ottamaan huomioon analyyttisesti käyttäjien kokemukset järjestelmän konseptin tai prototyypin suunnittelun ja arvioinnin eri vaiheissa.

Lisäksi käyttäjäkokemustavoitteiden systemaattinen käyttö tekee turvallisuuskriittisten järjestelmien käyttäjäkeskeisestä suunnittelusta ja varhaisista arvioinneista tavoitekeskeisempää, tarkempaa, jäljitettävämpää ja mitattavampaa. Käyttäjäkokemustavoitteiden tarkoituksenmukainen käyttö turvallisuuskriittisten teknologiaympäristöjen suunnittelussa edistää lopulta myös järjestelmiä hyödyntävien käyttäjien toiminnan mielekkyyttä, turvallisuutta, vaikuttavuutta ja tehokkuutta.

Avainsanat: käyttäjäkokemustavoite, käyttäjäkeskeinen suunnittelu, turvallisuuskriittiset järjestelmät, kokemussuunnittelu, inhimilliset tekijät

Author Hannu Karvonen
Faculty of Information Technology
University of Jyväskylä
Finland

Supervisors Senior Researcher Dr. Tuomo Kujala
Faculty of Information Technology
University of Jyväskylä
Finland

Professor Dr. Pertti Saariluoma
Faculty of Information Technology
University of Jyväskylä
Finland

Thesis advisor Principal Scientist Dr. Jari Laarni
VTT Technical Research Centre of Finland Ltd
Finland

Reviewers Director of User Experience Dr. Elizabeth F. Churchill
Google Inc.
United States of America

Professor Dr. Chris W. Johnson
School of Computing Science
University of Glasgow
United Kingdom

Opponent Professor Dr. John D. Lee
Department of Industrial and Systems Engineering
University of Wisconsin-Madison
United States of America

ACKNOWLEDGEMENTS

It has been a long journey to complete this thesis and I am very happy now that it is finished. This dissertation would have not been possible without the kind support of many extraordinary people. Firstly, I want to thank my thesis supervisors Dr. Tuomo Kujala and Prof. Pertti Saariluoma from the University of Jyväskylä for patiently supporting me throughout the years of writing this thesis. Secondly, I want to express gratitude to my thesis advisor and VTT teammate Dr. Jari Laarni for providing insightful comments and recommendations on the different versions of the thesis. Thirdly, I am really honoured to have Dr. Elizabeth F. Churchill and Prof. Chris W. Johnson as the preliminary examiners of this dissertation and truly appreciate the comments they provided to improve this work. I also eagerly look forward to having Prof. John D. Lee as my opponent in the upcoming defence. Fourthly, I want to thank warmly the former and current members of the human factors research team (you know who you are) in my workplace at VTT for their peer support and for some of you being part of conducting the empirical research, which is included in this thesis. I am also sincerely grateful to my earlier team leaders Dr. Leena Norros and Dr. Paula Savioja who originally introduced me to the scientific theories, approaches, and research methodology in the human factors group at VTT.

Naturally, a big acknowledgement goes to the co-authors in the articles of this dissertation: thank you for each of you. All the anonymous reviewers of the included articles also deserve recognition for their constructive feedback that helped to advance the articles into their final versions. Many professionals have also commented the thesis itself during its writing. I want to thank especially Dr. Eija Kaasinen in this respect. I have also had good discussions, for example, with Hanna Koskinen, Dr. Yichen Lu, Dr. Virpi Roto, and Dr. Jari Varsaluoma about the topic: it has been an honour to share the same research interest in UX goals with you and I want to express my gratitude for providing inspiring viewpoints to the topic all from your own perspectives!

The empirical research of this dissertation was conducted as part of the following projects: AMOVEO – Work Practices and the Transition to Ubicomp (2008–2011) that was partly funded by the Academy of Finland, FIMECC UXUS – User Experience and Usability in Complex Systems (2010–2015) that was partly funded by FIMECC (Finnish Metals and Engineering Competence Cluster Ltd) and Tekes (Finnish Funding Agency for Technology and Innovation), and DisGuard (2013–2014) that was also partly funded by Tekes. I want to thank the funding organisations and all my colleagues in these projects for making it possible to conduct this fascinating research.

I also want to thank some other VTT'ers with whom I have had the pleasure to work with, who have influenced my possibilities to conduct this research, and who have supported me in other ways along the way in finishing this thesis. These people include (in alphabetical order by surname): Johannes Hyrynen, Dr. Jari Hämäläinen, Dr. Kalle Kantola, Dr. Jari Kiviaho, Raimo Launonen, Dr. Simo-Pekka Leino, Timo Lind, Dr. Maaria Nuutinen, Pekka Pohjanne, Ismo Vessonen, and Dr. Riikka Virkkunen. Additionally, the language optimisation of this thesis done by Jason DeBose and Deron Fuller deserves to be mentioned here as well as the support from Sami K., Timo H., Päivi L., Tuisku S., and Sini T.

A big shout-out also goes to my dear friends (in alphabetical order by surname): Ilmari H., Mari H., Maija I., Mervi I., Tommi I., Salla J., Tommi K., Joonas K., Elina K., Anna K., Hannu L., Mikko M., Aino M., Juho N., Neeme O., Pasi R., Santtu R., and others: massive respect! It is always fun with you all when we meet. To my closest relatives like Pia, Mika, Sirkka, Antti, Marika, Sarianne, Lea, Leyla, Ayla, and Samad: I cannot thank you enough for being there. A huge thank you also goes to Laila and Veikko: without your childcare and all other practical help, this thesis would not have been possible – I am truly grateful to you. Additionally, Harri's and Herkko's families must be thanked for their support and good spirit.

I am deeply grateful to my mother and father for their immense support along the way in everything my family and I have been through. I am particularly thankful to my wife Reetta for persistently making it possible for me to do this thesis on free time by taking care of our kids and the household in the meantime: I respect and love you. I also want to thank our dear children, Elsi and Aapo, for bringing balance to my life for the last few years by nudging me to think also about other things in life than work and this thesis. I look forward to spending more time with all of you now that this thesis is finally done!

Espoo, Finland, 15 November 2019
Hannu Karvonen

GLOSSARY

ACD	Activity-centred design
Co-design	Co-operative design
ConOps	Concept of operations
CFA	Confirmatory factor analysis
CTA	Core-Task Analysis
DDW	Driver distraction warnings
EFA	Exploratory factor analysis
FSB	Future ship bridge
HCD	Human-centred design
HF	Human factors
HF/E	Human factors/ergonomics
HFE	Human factors engineering
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
JCS	Joint cognitive systems
LOA	Level of automation
HAZOP	Hazard and operability
HCI	Human-computer interaction
MTD	Metro train driving
PSV	Platform supply vessel
ROS	Remote operator station
SU	Systems usability
RE	Resilience engineering
R&D	Research and development
RQ	Research question
UC	Usability Case
XD	Experience design
UI	User interface
UX	User experience
VTT	VTT Technical Research Centre of Finland Ltd

FIGURES

TABLES

CONTENTS

LIST OF INCLUDED ARTICLES

I Karvonen, H., Aaltonen, I., Wahlström, M., Salo, L., Savioja, P., & Norros, L. (2011). Hidden roles of the train driver: A challenge for metro automation. *Interacting with Computers*, *23*(4), 289–298. https://doi.org/10.1016/j.intcom.2011.04.008

II Karvonen, H., Koskinen, H., & Haggrén, J. (2012). Enhancing the user experience of the crane operator: comparing work demands in two operational settings. In *Proceedings of the 30th European Conference on Cognitive Ergonomics* (pp. 37–44). ACM.

III Koskinen, H., Karvonen, H., & Tokkonen, H. (2013). User experience targets as design drivers: a case study on the development of a remote crane operator station. In *Proceedings of the 31st European Conference on Cognitive Ergonomics* (Article No. 25). ACM.

IV Kaasinen, E., Roto, V., Hakulinen, J., Heimonen, T., Jokinen, J., Karvonen, H., Keskinen, T., Koskinen, H., Lu, Y., Saariluoma, P., Tokkonen, H., & Turunen, M. (2015). Defining user experience goals to guide the design of industrial systems. *Behaviour & Information Technology*, *34*(10), 976–991. https://doi.org/10.1080/0144929X.2015.1035335

V Wahlström, M., Karvonen, H., Norros, L., Jokinen, J., & Koskinen, H. (2016). Radical innovation by theoretical abstraction – A challenge for the user-centred designer. *The Design Journal*, *19*(6), 857–877. https://doi.org/10.1080/14606925.2016.1216210

VI Wahlström, M., Karvonen, H., Kaasinen, E., & Mannonen, P. (2016). Designing user-oriented future ship bridges – An approach for radical concept design. In M. Soares & F. Rebelo (Eds.): *Ergonomics in design: Methods and techniques* (pp. 217–231). Boca Raton, FL: CRC Press.

VII Roto, V., Kaasinen, E., Heimonen, T., Karvonen, H., Jokinen, J., Mannonen, P., Nousu, H., Hakulinen, J., Lu, Y., Saariluoma, P., Kymäläinen, T., Keskinen, T., Turunen M., & Koskinen, H. (2017). Utilizing experience goals in design of industrial systems. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 6993–7004). ACM.

VIII Karvonen, H., Koskinen, H., Tokkonen, H., & Hakulinen, J. (2014). Evaluation of user experience goal fulfillment: Case remote operator station. In *Proceedings of the International Conference on Virtual, Augmented and Mixed Reality,* (pp. 366–377). Springer.

IX Kujala, T., Karvonen, H., & Mäkelä J. (2016). Context-sensitive distraction warnings – Effects on drivers' visual behavior and acceptance. *International Journal of Human-Computer Studies*, *90*, 39–52. https://doi.org/10.1016/j.ijhcs.2016.03.003

AUTHOR'S CONTRIBUTIONS

Article I 'Hidden roles of the train driver: A challenge for metro automation' presents a study of Helsinki metro train drivers' work activity. The analysis of this work activity is conducted with the Core-Task Analysis method developed in the human factors research team at VTT Technical Research Centre of Finland Ltd. (VTT). Based on this analysis, insights into the potential challenges of driverless metro operation are elaborated. These challenges also include user experience-related factors that should be accounted for in the design of the automated metro environment. The empirical research presented in this article was planned and conducted collaboratively with colleagues from the VTT human factors research team and Helsinki Institute for Information Technology. Karvonen took on a central role in the analysis of the results of the studies and in writing of the article. Insight-wise, he identified, thematised, and communicated the results regarding the planned metro automation concept's safety challenges related to operations conducted currently by the drivers. The themes these challenges considered were 'driving the train on the track', 'stopping at a station', 'passenger care', and 'interactions of the driver with other actors of the metro system' (see Section 3.5 of the article for details). Karvonen is the main co-author of the article.

Article II 'Enhancing the user experience of the crane operator: comparing work demands in two operational settings' describes a study of container crane operation in two operational port environments. As in Article I, the analyses in these studies were conducted with the Core-Task Analysis method. In this case, the ultimate aim was to design a novel remote operator station prototype for container cranes by utilising user experience goals. Key UX goals for the design work were identified and set based on user studies introduced in this article. The empirical research presented in the article was planned and conducted in close collaboration with Karvonen's colleague at VTT and a representative from an industrial partner company who are also co-authors of the paper. Karvonen took a central role in analysing the results of the field studies, in the analysis of other results, and in writing the article. He derived particularly the following insights from the empirical research analysis (presented here in a condensed format): 1) remote operation brings more uncertainty factors to the operator's work, 2) in conventional cabin operation, there is a stronger emphasis on dynamism factors, 3) sense of control and feeling of presence as UX goals, 4) the design implications of these UX goals for the development of a new remote operator station (ROS), 5) special attention should be paid to the development of a rich and realistic feel of operation in the design of the ROS. Karvonen is the co-first author of the article.

Article III 'User experience targets as design drivers: a case study on the development of a remote crane operator station' is a sequel to Article II. In Article III, the specification, utilisation, and evaluation of the UX goals identified in the analysis of Article II are presented. The design work described in Article III was done in intensive collaboration with partner company representatives. The writing process of the article was conducted collaboratively with the co-authors of the article from VTT and the University of Jyväskylä. Karvonen contributed significantly to the definition of the chosen UX goals, their design implications, and on the user study evaluations. Karvonen is the second co-author of the article.

Article IV 'Defining user experience goals to guide the design of industrial systems' is an article that describes different design cases where UX goal identification and setting took place in a partly government-funded FIMECC UXUS program in Finland. One of the four cases in the article is the remote operator station design case, whose results regarding the specification of UX goals are analysed from a bit different perspective than in Articles II and III. Specifically, the article looks at different approaches and sources for identifying UX goals based on the design cases and related literature. Along with writing the case description of the remote operator station design case, Karvonen also took part in the general feedback provision process of the article and wrote contributions especially to Sections 1 and 2 of the article. Specifically, he introduced the idea of UX goals as guides in the design process. Karvonen is the sixth co-author of the article: after the two main authors of the article, the rest of the authors were agreed to be listed in alphabetical order by their surname.

Articles V 'Radical innovation by theoretical abstraction – A challenge for the user-centred designer' and **VI 'Designing user-oriented future ship bridges – An approach for radical concept design'** both focus on the same empirical research about a future ship bridge concept design case, but from different perspectives. While Article V focuses more on the theoretical design issues of radical design conducted research, Article VI discusses the practical process of user experience design with an emphasis on radical solutions in the case. The empirical research presented in the articles was planned and conducted in collaboration with Karvonen's colleagues at VTT and Aalto University. Jussi Jokinen, University of Jyväskylä, helped in finalizing the Article V by providing somewhat valuable contributions with the identification of related studies and theory. Karvonen took a central role in the empirical case in conducting the field studies and, together with the first author, was central in analysing their results, in the specification and utilisation of the UX goals, in doing the actual concept design work, and in finalising the visualisations with the help of an external partner. In addition, Karvonen was the responsible research in the case on the research side and collaborated tightly with the industry side's main contact and other partners in the project. Karvonen provided insights for the results of the core-task analyses, for the formulation of the concept design approach for radical designs, and for the produced concept solutions in this case. Karvonen is the second co-author in both of these articles.

Article VII 'Utilizing experience goals in design of industrial systems' is a continuation of Article IV. Instead of focusing only on the UX goals identification and setting stage (as in Article IV), a broader perspective is taken in this Article VII. Specifically, the utilisation of the UX goals throughout the design and evaluation processes of different industrial systems design cases is presented. One of the main cases in the article is the remote operator station design case, which is also described in detail in Articles II, III, IV, and VIII of this thesis. Karvonen took the primary role in the process of describing the ROS case for the article. He also contributed to the design process considerations and conclusions of the paper. Karvonen is the fourth co-author of the article: after the two main authors of the article, the rest of the authors were agreed to be listed in the order of their amount of contribution to the article writing.

Article VIII 'Evaluation of user experience goal fulfillment: Case remote operator station' presents a detailed version of the evaluation study of the developed remote operator station prototype, whose design process was described in Articles II, III, IV and VII. The empirical research conducted in this study was planned in close collaboration with the co-authors of the article. Karvonen was one of the three test moderators in the conducted evaluations. In addition, Karvonen took a central role in analysing the results of the experiments and writing of the different parts of the article. Naturally, the overall writing process of the article was conducted collaboratively with the co-authors. Karvonen played a key role in analysing the fulfilment of the UX goals and the found UX and usability problems of the ROS prototype system. The information provided by this evaluation study was utilised in designing the final version of the ROS that was sold to Indonesia by the partner company. Karvonen is the main co-author of the article.

Article IX 'Context-sensitive distraction warnings – Effects on drivers' visual behavior and acceptance' describes an evaluation study of a smartphone application called VisGuard, which gives context-sensitive distraction warnings to car drivers. Karvonen was responsible for the survey part of the conducted study. Specifically, Karvonen took care of the design of the online questionnaires utilised in the study and of the statistical analysis of the questionnaire results. The questionnaire included questions related to user experiences of the study participants regarding the developed application. The implicit user experience goals in this study that were measured included factors, like 'experienced trust in the application', 'experienced usefulness of the application', and 'harmfulness or annoyance of the application'. In general, the goals were seen to contribute to the acceptance of the application. The writing process of the article was conducted collaboratively with the co-authors. In the writing process, Karvonen focused mostly on the parts of the article that were related to the survey. Specifically, he suggested the theoretical underpinnings of the survey questions related to the application's trustworthiness, perceived consistency (i.e., reliability), and timeliness. Additionally, he also utilised user acceptance (e.g., TAM), validity, harmfulness, and satisfaction theories in generating the key items in the questionnaire. Finally, he identified the presented factors from the exploratory factor analysis. Karvonen is the second co-author of the article.

1 INTRODUCTION

The concept of user experience (UX) began to gain momentum in Western human-computer interaction (HCI) research and interaction design in the second half of the 1990s (Robert & Lesage, 2017). Research at that time studied UX particularly in the design of technological consumer products and services for mass markets (see, e.g., Alben, 1996; Cain, 1998; Fleming & Koman, 1998; Maguire, 1999; Segal & Suri, 1997; Shneiderman, Byrd, & Croft, 1998). In the 2000s, groundbreaking UX research was conducted that focused, for example, on the definition of the concept, creation of research and development (R&D) frameworks, and methodological development (e.g., Battarbee, 2004; Bødker, 2006; Desmet & Hekkert, 2007; Forlizzi & Ford, 2000; Hassenzahl, 2008; Hassenzahl & Tractinsky, 2006; Wright, McCarthy, & Meekison, 2003; Wright, Wallace, & McCarthy, 2008). In particular, the development of UX evaluation methods and models was the focus of attention in research at that time (e.g., Obrist, Roto, & Väänänen-Vainio-Mattila, 2009; Roto, Obrist, & Väänänen-Vainio-Mattila, 2009; Väänänen-Vainio-Mattila, Roto, & Hassenzahl, 2008b, 2008a). However, the R&D focus of user experience work typically remained in technologies for consumers or standard office workers.

In the 2010s, UX has quickly developed as a field of research and design and expanded into new application domains. One such domain is safety-critical systems, where the relevance of users' experiences has been recognised in the R&D of advanced technological solutions for complex environments (see, e.g., Kim, Cooper, Carroll, & Murugappan, 2017; Norros, 2014; Rödel, Stadler, Meschtscherjakov, & Tscheligi, 2014; Savioja, Liinasuo, & Koskinen, 2014; Tasoudis & Perry, 2018; Wurhofer, Fuchsberger, Meneweger, Moser, & Tscheligi, 2015). A key aspect of this research has been in recognising that experiences, which are desirable with mass-market consumer products, are not be compatible with complex safety-critical systems that are utilised, for instance, in ships, hospitals, or nuclear power plants.

Instead of hedonic experiences, like fun and enjoyment, very different kinds of experiences are typically relevant in the usage activity of safety-critical systems. With these systems, a positive UX may mean, for example, that the design solution supports the competencies of the users (Saariluoma & Jokinen, 2014) or the development of professional pride (Lu & Roto, 2016). Additionally, the appropriate design and usage of a safety-critical system can enable experiences of success in maintaining a high level of safety in the object environment. In this way, a good user experience can arise as the users feel that the technical system supports their activities and they are in control of the system and the entire (or at least their own part of the) safety-critical environment where the activity is taking place. This experience of control may also allow the users to feel that they are playing an active and relevant part in the overall activity, and not just being passive bystanders.

The best possible user experience is as such a noble goal worth striving for in the design of technology. When users interact with a technology, a good UX may allow them to feel meaningful, positive emotions that also contribute to their general well-being. In the development of technology, appropriate experience design practices can support the emergence of these experiences in the usage of the technology.

However, 'a good UX' is an abstract aim that can be difficult to be taken into account in the production of design solutions in practice. Therefore, if the notion of good UX in some particular context can be elaborated, it can provide more meaningful information for the design work. To reflect on what good UX can mean in a particular case, designers may ponder questions such as 'How should the usage of the system profoundly feel?' or 'What kinds of experiences are important in this particular context?'

To help in answering these questions, the users' goals, challenges, needs, and system usage activity in the target environment have to be understood on an in-depth level. It is possible to gain this understanding through various R&D approaches. One of these approaches, which is also in the core focus of this thesis, is human-centred design (HCD) that 'aims to make interactive systems more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques' (ISO 9241-210, 2010, p. 2). In addition, different task analysis (e.g., cognitive task analysis; Salmon, Stanton, Gibbon, Jenkins, & Walker, 2009) or work analysis (e.g., Cognitive Work Analysis; Vicente, 1999) approaches can be utilised to look into the users' or workers' tasks, aims, and activity with technologies on a detailed level. In this dissertation, the Core-Task Analysis (CTA, see, e.g., Norros, 2004) method is utilised in particular. These types of analysis and design approaches may help in reflecting on points such as 'What the users are trying to accomplish?', 'What kinds of challenges do the users face?', and 'What factors create experiences of success for the users?'

As can be seen from these questions, the empathetic setting of oneself into the users' position also helps the designer. This empathy allows the designer to comprehend what the users' activity is deeply about, what experiences affect the users' activity, and what positive experiences are facilitated by the activity. In this way, it is possible to consider ways to support these experiences facilitated by the activity, for instance, by means of technological solutions. Nevertheless, to proceed from these types of general-level descriptions to actual design solutions is often a challenging task. Producing concrete designs is commonly based on the intuitions of the designer, instead of a disciplined design process targeting specific user experiences in a systematic manner.

This dissertation investigates the use of UX practices in the human-centred design of safety-critical systems. As an approach to aim at specific user experiences, the thesis focuses on a technique called user experience goals. UX goals define the positive experiences that the usage of the designed system aims to awaken in the users. In a design process, the purpose of UX goals is to describe the experiential aspects to focus on in design. UX goals can be utilised during the design process to systematically aim for specific experiences with the produced design solutions. This thesis focuses on UX goals particularly in the early-stage human-centred design of safety-critical systems. The early-stage design phase of these systems in this context refers to the 1) 'fuzzy front-end' of design (Kaasinen, Karvonen, Lu, Varsaluoma, & Väätäjä, 2015; Varsaluoma, Väätäjä, Kaasinen, Karvonen, & Lu, 2015b), 2) concept/prototype design, and 3) early evaluations of concepts/prototypes.

1.1 Research context and motivation

The topic of this dissertation is motivated by the increasing need for good UX research and design practices with safety-critical systems (see, e.g., de Mol, 2007; Savioja, 2014; Savioja, Liinasuo, et al., 2014; Savioja & Norros, 2013). Many modern safety-critical technologies are becoming a common part of people's everyday lives with highly automated solutions such as advanced driver-assistance systems in cars. Moreover, the amount of complex digital systems in professional work in safety-critical environments has increased rapidly (Rajkumar, Lee, Sha, & Stankovic, 2010). These systems are naturally expected to perform very reliably from the technical performance perspective. Additionally, today's users presume good usability and user experience with the systems. These presumptions are heightened by positive experiences with technical mass-market consumer products like, for example, modern smartphones. To answer the increased expectations, proper human activity-centred experience design practices are essential.

Traditionally, the 'human and organisational factors' in safety-critical systems design and evaluation have been taken into account with human factors engineering (HFE, see, e.g., Lee, Wickens, Liu, & Boyle, 2017) and organisational safety culture analysis methods (e.g., Glendon & Stanton, 2000), which do not traditionally consider users' experiences. However, recently UX-considering approaches have also gained a foothold in safety-critical systems engineering – for instance, in the development of control room solutions for nuclear power plants (Norros, Savioja, & Koskinen, 2015; Savioja, 2014; Savioja, Liinasuo, et al., 2014).

In UX-oriented design, experiential factors are deeply considered and taken into account in the design of the technology. Nevertheless, this does not mean that the more traditional safety, performance, and usability aspects, such as reliability, efficacy, and efficiency, would be forgotten in the development. Instead, it can be seen that a good UX with the system also supports these other relevant aspects of safety-critical systems design and use.

This dissertation aims to combine the systematic nature of systems engineering and HFE methods to human-centred and creative UX design practices in developing safety-critical systems. As one key technique, UX goals (e.g., Hartson & Pyla, 2012) are presented and discussed. UX goals are a rather recent technique, which has already been utilised in the design of a few public cases (see, e.g., Kymäläinen et al., 2017; Varsaluoma et al., 2015b) of safety-critical systems in addition to the cases reported in this dissertation. The general-level motivation of this research is to continue on this path and develop the HCD of safety-critical systems forward from the user experience perspective.

1.2 Objectives and scope

This research focuses on the human-centred design of safety-critical systems from the UX perspective. Therefore, the object of research is the human-centred design of safety-critical systems. Specifically, the objective is to study the use of user experience goals in the early-stage human-centred design process activities of safety-critical systems.

The term 'safety-critical systems' refers here to technical systems where a failure can result in threat to human life, significant property damage, considerable financial losses, or threat to the environment (Bozzano & Villafiorita, 2010; Knight, 2002; Vicente, 1999). However, for example, security breaches (e.g., identity theft or cyber-security issues) with similar consequences are not in the scope of this work. Nevertheless, the results of this research may be relevant from a security perspective as well. With safety-critical systems, conceptualising the specific characteristics of the intended user experience is particularly important in their experience design. The choice of appropriate experiences to aim at is essential for the design process to support the overall goal of safety with these systems.

Empirical *cases* from four different safety-critical focus environments are presented in the dissertation. The cases (with corresponding articles of this dissertation in brackets) are related to 1) the analysis of rapid transit system (metro) operations (Article I), 2) the design of a prototype remote crane operator station (Articles II, III, IV, VII, and VIII), 3) the design of future ship bridge concepts (Articles V and VI), and 4) the early evaluation of driver distraction warnings (Article IX). The cases demonstrate the many applied research activities that can be conducted in early stages of the HCD process with safety-critical systems. Here, one case can include many different *studies* (e.g., interview, observation or questionnaire studies), which were conducted as part of the case in the dissertation.

The contributions of each of the included empirical research articles in this dissertation can also be viewed through the main activities of the human-centred design process along the lines of the ISO 9241-210 (2010) standard. For the purposes of this thesis, which emphasises context, UX goals, concept design, and concept evaluation, these activities are relabelled here (with the original ISO-9241-210 phase labels in brackets) accordingly as follows:

1. Understanding the context ('Understanding and specifying the context of use'),
2. Specifying UX goals and other requirements ('Specifying the user requirements'),
3. Producing concept designs ('Producing design solutions'), and
4. Evaluating the concepts ('Evaluating the design').

These listed activities are in practice non-linear, overlapping, and iterative during the early stages of a design process, but are presented for the sake of clarity here as a sequential step-by-step approach where one activity provides input for the next one. The empirical research articles in the List of included articles have been categorised primarily based on their main emphasis on these different activities (and secondly, based on ascending publication year) in the following manner:

1. Understanding the context: Articles I and II
2. Specifying UX goals and other requirements: Articles III and IV
3. Producing the concept designs: Articles V, VI and VII
4. Evaluating the concepts: Articles VIII and IX

The above categorisation of the articles for these activities is approximate, as some articles have overlapping contributions to more than one of the mentioned activities. For example, Articles III and VII have their primary emphasis in the specification and utilisation of UX goals (or 'UX targets' [as in Article III]; or 'Xgoals' [as in Article VII]) in concept and prototype design, but they also discuss the other activities that were included in the case, such as work analysis and concept/prototype evaluations. This overlapping is gone through in detail in Chapter 4, where an overview of the empirical research methods and other basic information of this thesis' studies are presented.

The proposed categorisation based on the design process activity aims to understand how UX goals could best fit into the HCD process, as is presented and discussed in Sections 5.2–5.5, 6.1, and 6.2. Furthermore, these activities are referred to, for example, in the summary table of the empirical cases in Section 4.3.

The theoretical backgrounds of the articles in this dissertation come from different academic traditions. For instance, the experimental nature of studies in Articles VIII and IX has a different scientific foundation compared to the exploratory and naturalistic study approach applied in the other articles. In general, these approaches can be seen as complimentary to each other and the aim here is to bridge these traditions. There are situations where a more experimental approach is better than an exploratory one and vice versa. For example, if the context of the object of study is very complex and not clearly defined in the beginning of the design work, an exploratory research approach can be a more suitable choice. The demonstration of these different approaches in this thesis aims to demonstrate the multitude of possibilities to study UX-related issues. In addition, the aim here is to investigate the different methodological aspects of the specification, utilisation, and evaluation of UX goals through the various approaches.

The data gathering methods utilised in the empirical studies of this dissertation are both qualitative (e.g., interviews, observations, and thinking aloud) and quantitative (e.g., use of surveys with Likert-scale questions, gathering of user performance measures, and acquirement of log data). However, the emphasis in the empirical studies was strong in qualitative methods, as most of the studies were exploratory in nature. Analysis methods of the gathered data included, for example, interview theme analysis, statistical analysis, and Core-Task Analysis methods.

On a general level, the main goal of this dissertation is to study how user experience research and design can systematically be applied to early-stage human-centred design of safety-critical systems through the utilisation of UX goals. In detail, the thesis also produces empirical evidence from actual research and design cases to advance knowledge about the specification, utilisation, and evaluation of relevant UX factors and goals with safety-critical systems.

1.3 Research gap, questions and contributions

The identified research gap in this dissertation is the separation between human factors (HF) methods and experience design. The fact that in some of the handbooks of human factors methods and engineering (e.g., in Salmon, Stanton, et al., 2012; or Stanton et al., 2017), the term 'user experience' is not mentioned at all as a concept is a demonstration of this gap in practice. Therefore, one research aim in the background of this thesis is in bridging the gap between HF methods and creative UX design work. Here, the way to start this endeavour is to focus on human-centred design, which has embraced the concept of user experience, but also considered the HF aspects.

The general-level research question of this dissertation is '*how can user experience goals be systematically used in the human-centred design of safety-critical systems?*' Merriam-Webster's online dictionary[1] defines systematic as 'using a careful system or method' and refers to it being 'marked by thoroughness and regularity'. In the methodical process context of this thesis, systematic also refers here to the particularly analytical and structured design and evaluation process, which emphasises traceability and reflectivity (see, e.g., Savioja, 2014).

The answer to the general-level research question presented above is provided in Chapter 6. In practice, all the articles included in the thesis contribute to answering this general-level research question.

The specific research questions (RQs) of this dissertation are the following:

1 https://www.merriam-webster.com/ (accessed 5th of May, 2019)

RQ1: What is the value of experience design as part of the development of safety-critical systems?
RQ2: How can analysis methods of human activity in safety-critical environments contribute to the identification of UX goals?
RQ3: How can UX goals be systematically set and defined for the human-centred design of safety-critical systems?
RQ4: How can UX goals be systematically utilised in the production of concepts and prototypes for safety-critical systems?
RQ5: How can the experiences of users be evaluated for systematic analysis of the fulfilment of user experience goals in human-centred design of safety-critical systems?

These research questions, their rationale, and the corresponding results are discussed in detail in Chapter 3. The different articles included in this dissertation provide empirical results related to these research questions in the following manner:

RQ1: Articles II, III, IV, V, VI, VII, and VIII
RQ2: Articles I, II, III, IV, V, and VI
RQ3: Articles II, III, IV, V, VI, and VII
RQ4: Articles II, III, V, VI, VII, and VIII
RQ5: Articles III, VII, VIII, and IX

The contribution of this research is in demonstrating how and why experience design is relevant for safety-critical systems development. The dissertation recognises user experience goals as a key technique in taking the experience aspect with a clear focus and systematicity into account throughout the design process.

1.4 Primary fields of research and theoretical background

Firstly, a central field of research of this dissertation is cognitive science. According to Thagard (2005, p. ix), 'cognitive science is the interdisciplinary study of mind and intelligence, embracing philosophy, psychology, artificial intelligence, neuroscience, linguistics, and anthropology'. For cognitive science, this dissertation provides understanding of the elements of user experience and interaction with technology in safety-critical environments. Furthermore, the thesis offers input on how to make cognitive science methods such that they can be more applicable in the design of technology. In line with this need, for example, Flach, Stappers, and Voorhorst (2017, p. 76) mention that 'designers are challenged to apply the discoveries and theories from cognitive science to improve their product designs, and cognitive scientists are challenged to learn from the successes and failures of design innovations'. As a solution, Flach, Stappers, and Voorhorst suggest 'experience-centric thinking', which emphasises the holistic properties of experiences.

Regarding cognitive science, it has also been argued elsewhere (Flach & Voorhorst, 2016, p. 4) that 'conventional assumptions about mind and matter have become obstacles to the development of cognitive science and to the practical application of cognitive science to the design of technology'. Furthermore, for example, Flach (2009) has stated that there still seems to be a wide gap between cognitive psychology and everyday life experiences, and therefore, those who look to cognitive science for inspiration for technology design that enhances the qualities of everyday life and work may be disappointed. Based on these arguments, it can be said that the so-called 'applied cognitive science', mentioned as early as the mid-1980s by Norman (1986), has still a lot of foundational work to be conducted in order to bridge the gap between cognitive science and the design of products, systems, and services.

Secondly, this dissertation contributes to the fields of human-centred design, systems engineering, and experience design by presenting a systematic approach for the utilisation of UX goals in the early phases of a design process. These fields are discussed in more detail in Chapter 2. In addition to the researchers in these fields, engineers, designers, and other professional practitioners can also gain insights for their work from the results of this thesis.

Thirdly, the theoretical background for the empirical cases with the analysis phase included in this dissertation's articles is in activity theory (AT). Activity theory is derived originally from the work of Lev Vygotsky (1978) and it considers activities as being produced by people within a cultural context and developing over time (Engeström, 2014). According to Carroll (1997), the object of description in AT is the 'activity system', which includes the technological and social factors, individual's attitudes, experiences, and actions, as well as the community's practices, traditions, and values. These are also described in the so-called 'Engeström's basic activity triangle' (Engeström, 2014), which adds, for example, rules and norms, and division of labour to Vygotsky's original theory.

From the human factors perspective, the activity-centred approach widens the unit of analysis to consider the entire multi-layered activity system in technical environments instead of only single user-technology interactions (Savioja, 2014). Consequently, it also considers the socio-cultural context of research, design, use, and human experiences of technology. Based on AT, several approaches for studying human-technology interaction have been developed. One example is Core-Task Analysis (CTA) by Norros (2004), which is also used in the analysis phase studies of this dissertation.

In addition to AT, other approaches that also consider human-technology interaction from a holistic perspective are also utilised as the basis of CTA, such as ecological psychology and pragmatism. Ecological psychology (or environmental psychology) emphasises the natural setting of human as a determinant of behaviour. Gibson (1977) elaborated the situatedness (i.e., context-dependence) of actions with the concept of affordance. Affordances are the perceived opportunities for action in the environment. Environments offer affordances for specific actions, therefore favouring particular actions over other ones. In line with AT, ecological psychology sees, that knowing is inseparable from doing.

Pragmatism, on the other hand, has considered the basis of human habits. In Peirce's (1905) triadic structure (i.e., 'semiotic triangle'), a sign/representation stands for something, an object is what is referred to by the sign, and an interpretant is the individual's comprehension of and reaction to the sign/referent association. A 'habit' is formed in the interaction of these three components. This approach – as most of the approaches already mentioned – also emphasises information in the world, instead of inside the head of individuals. The pragmatist approach has been successfully utilised in several human-technology interaction research studies, for example, by Klemola and Norros (1997), Norros (2014), Savioja, Norros, Salo, and Aaltonen (2014) and Wahlström, Karvonen, and Norros (2013).

CTA also has its basis in cognitive systems engineering (see, e.g., Rasmussen, 1986; Rasmussen, Pejtersen, & Goodstein, 1994) and cognitive work analysis (Vicente, 1999). The former takes modelling concepts from engineering, psychology, cognitive science, information science, and computer science as its basis (Rasmussen et al., 1994). The latter views analysis of constraints of the work domain as a relevant part of the development of work environments (Vicente, 1999). The CTA method has previously been utilised in various application domains and work environments, such as in anaesthetist's clinical practice (see, e.g., Norros, 2005), nuclear power plant operation (see, e.g., Norros, 2004), and maritime piloting (see, e.g., Nuutinen & Norros, 2009).

Finally, one approach that affected especially the empirical ship bridge studies in this dissertation is joint cognitive systems (JCS). JCS is a theoretical approach in which the central idea is that humans and technology form a functional unit. Therefore, in JCS, the human-technology system is taken as the unit of analysis instead of single tasks (Norros & Salo, 2009; Woods & Hollnagel, 2006).

Based on the above contemplations, the research in user experience (see, e.g., Law, Roto, Hassenzahl, Vermeeren, & Kort, 2009) should also consider more deeply the way a person's experience forms in interaction with one's environment. Experience is embodied, situated, and social along the lines of the theories presented above and in Chapter 2. Therefore, a key premise in this dissertation is that experience occurs as a result of the usage of external objects, which have affordances, and is constructed in an activity system (Engeström, 2014).

All of these approaches with their theory backgrounds contributed to the researcher assumptions and methods in the empirical cases, and also functioned as a basis for conducting the Core-Task Analysis of work activity in the analysis phase studies of this dissertation. Consequently, the unit of analysis being studied in this research's work analysis cases is the activity of the associated workers (i.e., metro train drivers, container crane operators, and ship bridge personnel) in the corresponding activity system. The detailed descriptions of the units of analysis in the empirical cases of this thesis in general are provided in Section 4.3.

1.5 Structure of the dissertation

The structure of this dissertation chapter by chapter is the following. Chapter 1 introduces the research context briefly and the motivation, objectives and scope, as well as the research gap, questions, and contributions of this dissertation. In addition, the primary fields of research and the theoretical background of the empirical studies are shortly discussed.

Chapter 2 presents a review of the previous related research work by defining the key concepts of this dissertation and considering the different theories, approaches, methods, and techniques relevant to the topic of the dissertation. Chapter 3 outlines the research questions of this dissertation and the rationale behind the research questions. Chapter 4 presents a background summary of the conducted empirical research of the thesis. For instance, the research methods, settings, and other key background factors of each included study are considered in this chapter.

Chapter 5 presents the results and contributions of the empirical research work of the thesis. The articles as part of this dissertation show the empirical cases conducted in the thesis in detail. Based on the empirical cases, the answers to the five research questions (see Chapter 3) are presented as separate in Sections 5.1–5.5.

Chapter 6 analyses the implications of the results of this dissertation. Specifically, the discussion is focused around methodological and practical implications and the lessons learnt from using UX goals in human-centred design of safety-critical systems. Based on the empirical cases, an answer to the general-level research question of this dissertation is addressed in Sections 6.1 and 6.2. Finally, the dissertation's research validity, limitations, and recommendations for future research are presented.

2 REVIEW OF RELATED WORK

This chapter is divided into three sections in the following manner. First, in Section 2.1, different aspects related to safety-critical systems and their development especially from the human factors point of view are presented. Second, in Section 2.2, definitions of user experience and previous work regarding UX's relevance to safety-critical technology are considered. Third, in Section 2.3, human-centred and experience design, user experience goals, and their utilisation as part of design are discussed based on earlier research.

The reason for this kind of structure of this chapter is to first define the basic concepts and their related research. For example, safety-critical systems and the related issues (Section 2.1) are defined before it is considered what UX or experience design means (in Section 2.2) in those contexts. Similarly, human-centred and experience design are considered (in the introduction to Section 2.3 and in Section 2.3.1) before presenting UX goals (Section 2.3.2) and their role as part of design processes (Section 2.3.3) based on earlier literature.

2.1 Safety-critical systems

According to Besnard and Hollnagel (2014, p. 14), 'the definition of safety usually refers to the absence of unwanted outcomes'. For example, the IEC 61508 (2010) standard defines safety as 'freedom from unacceptable risk' that refers to the absence of financial, physical, or social hazards or risks. A broader view is that 'safety is the system property that is necessary and sufficient to ensure that the number of events that could be harmful to workers, the public, or the environment is acceptably low' (Besnard & Hollnagel, 2014, p. 14). Consequently, Besnard and Hollnagel (2014, p. 14) state that 'safe systems produce acceptably low numbers of unwanted events'. In this 'extended view of safety', the concept of safety is considered much more broadly than freedom from accidents and interpreted as a systemic phenomenon instead of a linear one (Savioja, 2014). In practice, this means that the focus on improving safety should be on the socio-technical system as a whole, instead of, for instance, single human errors. 'Resilience', i.e., the socio-technical system's ability to continuously adjust its functioning to meet the operative objectives, can be seen as the operational manifestation of safety in this approach (see, e.g., Hollnagel, Paries, Woods, & Wreathall, 2011; Leveson et al., 2006; Savioja, 2014).

2.1.1 Definitions and characteristics of safety-critical systems

According to the IEC 61508 (2010, sec. 3.4.1) standard, a safety-critical system 'implements the required safety functions necessary to achieve or maintain a safe state for the equipment under control'. The term safety-critical *system* is used in this dissertation to refer to applications of information technology in safety-critical environments. The concept is used in the thesis interchangeably with various other similar terms such as 'safety-critical technology' or 'safety-critical technical system'. Furthermore, different kinds of other specifying terms than 'safety-critical' referring to similar technical systems can be found from the literature, such as 'high-criticality systems' (e.g., Parasuraman & Miller, 2004), 'high-integrity systems' (e.g., Bowen & Hinchey, 1999), 'high-reliability systems' (e.g., Hofmann, Jacobs, & Landy, 1995), 'high-risk systems' (e.g., Kirwan, 1998), 'life-critical systems' (e.g., Boy, 2013), 'mission-critical systems' (e.g., Ponsard et al., 2007), or just 'critical systems' (e.g., Bozzano & Villafiorita, 2010). The 'critical' part in these concepts refers to it being of great importance to the way things might happen, instead of expressing adverse or disapproving comments or judgements (i.e., critique).

In some publications (e.g., Leveson et al., 2006), the concept of safety-critical system is used to refer to the socio-technical environment, which includes the involved people, technology, organisations, and possible other stakeholders, such as regulators. In a socio-technical system, these different parts have to work together in a uniform manner to produce the desired outcome (Savioja & Norros, 2008; Vicente, 1999). When referring to this broader notion of the socio-technical safety-critical system, the terms 'socio-technical system' or 'socio-technical environment' are used here. In this connection, it has yet to be clarified that the empirical research conducted in this thesis focuses mostly on human activity with safety-critical technology, instead of in the context of broader organisational factors, like safety culture (e.g., Glendon & Stanton, 2000).

With safety-critical socio-technical environments, a failure can result in a threat to human life, significant property damage, considerable financial losses, or threat to the environment (Bozzano & Villafiorita, 2010; Knight, 2002; Vicente, 1999). Application domains or industries that can be described as safety-critical include, for example, nuclear power, marine operations, commercial aviation, mining, healthcare, port operations, and land transportation (e.g., Vicente, 1999). Safety-critical domains typically include potential risks related to the above-mentioned threats. A factor often increasing these risks is the high level of complexity of the socio-technical environment. This complexity derives partly from a large number of elements and some level of diversity amongst interconnected and dynamic parts typical for socio-technical systems (Norros, 2004; Savioja, 2014).

Therefore, it is challenging to both design and use the implemented complex safety-critical technologies.

2.1.2 Safety-oriented methods for systems engineering

There is a considerable amount of published research about different approaches and methods that can be used in the software and hardware R&D of safety-critical systems. Complex safety-critical technical systems are commonly developed and studied with robust systems engineering tools that focus on ensuring, for instance, the safety, reliability, and maintainability of the system (Blanchard & Fabrycky, 2006). In addition, the engineering processes of safety-critical systems typically emphasise the systematicity of the process (see, e.g., Pahl & Beitz, 2013).

The development of safety-critical systems has traditionally been conducted according to the V-model. The original V-model depicted by Clark (2009) based on the work of Forsberg & Mooz (1992) and Forsberg, Mooz, & Cotterman (2000) is presented in Figure 1. Today, there are also many variations of this model. Out of the activities mentioned in Figure 1, the empirical research as part of this thesis is related to the top level of the V-model, which includes the 'user requirements', 'system concept', 'validation plan', and 'validate system to user requirements' activities.

In addition to rigorous systems engineering practices, several standards, laws, guidelines, and safety acts govern the development of safety-critical systems. Often, these documents are also application domain-specific (in other words, specifically customised for a certain industry). These documents typically provide the basic safety requirements for the design of some particular safety-critical system. For example, standards for safety requirements exist for the automotive (ISO 26262, 2011) and nuclear (IEC 61513, 2001) industries.

Some safety standards may require the use of a 'safety case'. A safety case is defined as 'a documented body of evidence that provides a convincing and valid argument that a system is adequately safe for a given application in a given environment' (Bishop & Bloomfield, 2000, p. 34). It typically consists of elements, such as claims, evidence, arguments, and inference. These elements can be utilised in the development of a documented safety case concerning a certain safety-critical system. The Bishop's and Bloomfield's safety case approach emphasises traceability between the different levels of the system and subsystems.

In general, traceability is also a key aspect that is highlighted in the R&D of safety-critical technologies. For example, a risk estimated during development should have corresponding documented requirements, design solutions, and validations that are all traceable between each other. Therefore, the design rationale should also be documented and explicitly stated in case of a need to see the decisions during the design process.

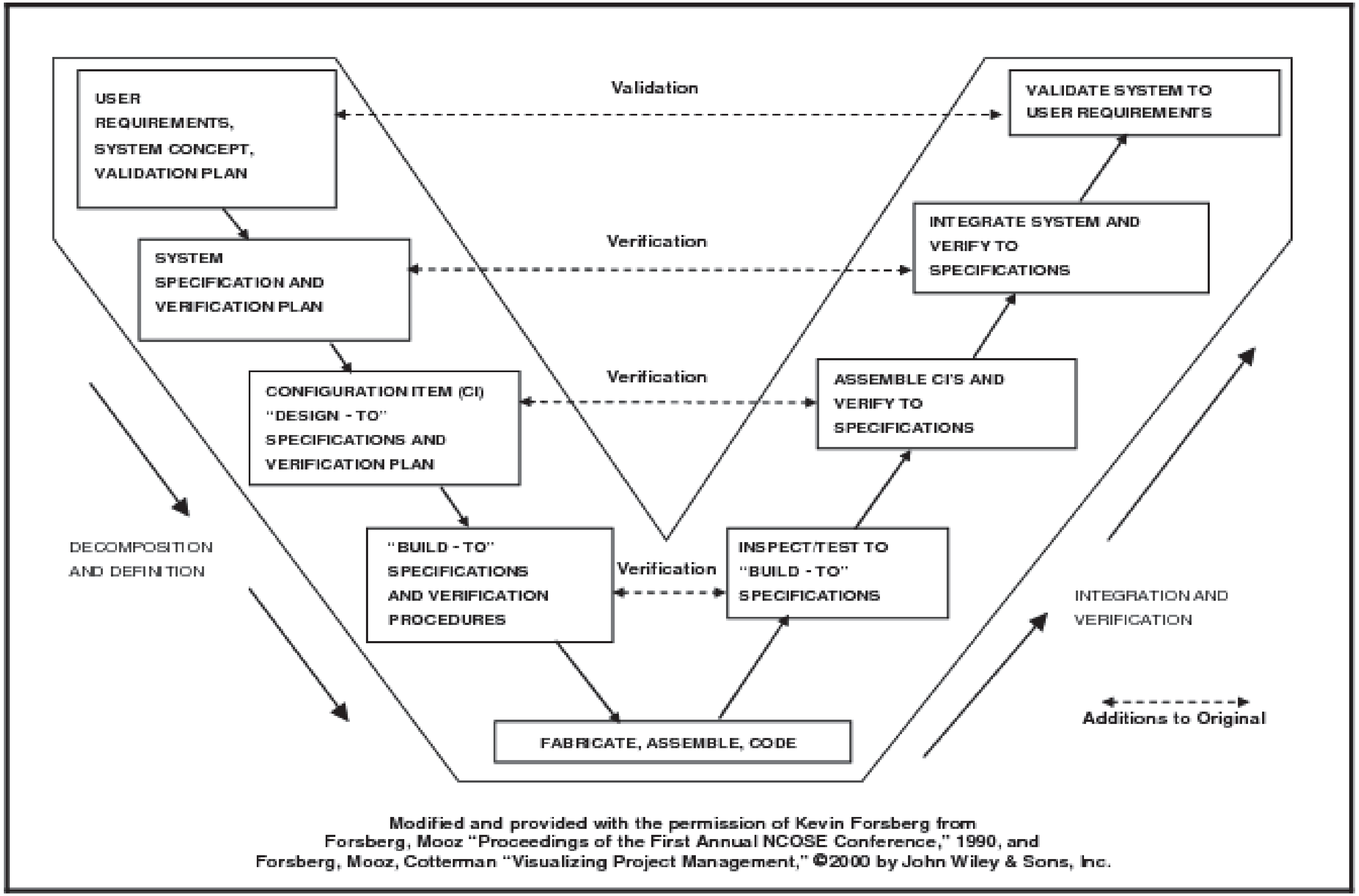


FIGURE 1 Original V-model (from Clark, 2009, p. 384)

In addition, different regulatory bodies may want actively themselves to gather evidence about the fulfilment of the safety functions of some system. This evidence can be gathered by conducting regulatory safety inspections and gathering operative data for evaluation purposes. Additionally, external auditors can evaluate the quality of a system development process of the corresponding technology organisation. However, presenting the details of various regulatory approaches and standards is out of the scope of this work, which focuses more on the human-centred design side of early-phase safety-critical system development. Nevertheless, the value of safety methods in general in the development of high-criticality systems is in no way subject to attempted undermining here.

In addition to regulatory approaches and standards, there are certain essential methods (e.g., from safety and reliability engineering) that are used in industry and academia to analyse and improve the safety of high-reliability systems. For example, the following methods have been suggested to support the development of safe high-reliability systems: 1) risk assessments studies, 2) simulations, 3) verifications and validations, 4) and testing. Below, each of these approaches are briefly presented in their own paragraphs.

Firstly, risk assessment studies are analytical examinations of risks and problems, which may prevent safe and efficient operations (Kletz, 1999). For instance, qualitative hazard and operability (HAZOP) studies can be conducted with a team of safety experts. The team-approach allows the members of the team to build on comments from each other in meetings. The idea with HAZOP is to find possible unsought risks and evaluate their severity. Today, there are also software systems to support HAZOPs, and the method has been utilised in various safety-critical industries.

Secondly, simulations can be described as computer-based dynamic models, which allow the reproduction of how some technology operates in different conditions over time. The aim of simulations is 'to capture the system behavior in a system model and verify the correctness of the system by simulating different scenarios one by one using this model' (Lahtinen et al., 2012, p. 104). The benefit of simulations is that utilising them in some safety-critical technology's design phase is much safer, faster, and cost-effective than, for example, learning by trial and error in real-life conditions with the implemented solution. Additionally, reliability studies by software-based examination of system dynamics (e.g., Coyle, 1997) and statistics-based probabilistic safety/risk assessments (e.g., Montewka, Ehlers, et al., 2014) are approaches which can be utilised to identify and understand some complex socio-technical systems' safety-critical behaviour dynamics and vulnerabilities.

Thirdly, novel formal software verification and validation methods include, for instance, model checking (see, e.g., Baier & Katoen, 2008). Software-based 'model checking is a computer-aided formal method for verifying the correctness of a system design model' (Lahtinen et al., 2012, p. 104). Furthermore, verification and validation studies can also be conducted from the human factors point of view (see, e.g., Laarni, Savioja, Karvonen, & Norros, 2011; Laarni et al., 2014). For example, these types of HF verification and validation studies may include operators conducting certain safety-critical tasks in a simulator environment while their activity with the system is studied. For example, for HF verification and validation in the nuclear industry, the NUREG-0711 (see, e.g., O'Hara, Higgins, & Fleger, 2012) standard offers a reference point.

Fourthly, in testing, the basic idea is to experiment with how a technical system copes with different scenarios and assesses the correctness of its functioning by using a collection of test cases (Lahtinen et al., 2012). For instance, periodic tests are conducted in nuclear power plants on the plant's different safety-critical equipment. These tests are scheduled between frequent intervals (e.g., once a week) to test the correct functioning of the technical devices in different scenarios in the plant. Testing can also be conducted from the human factors point of view. For instance, 'systems usability' evaluations provide metrics and insight about how well a certain tool or environment works in actual usage activity (Savioja, 2014; Savioja & Norros, 2013).

In addition to these approaches, there are several other methods for analysing the reliability, availability, maintainability and safety (RAMS) concerns related to safety-critical systems. These include, for instance, Preliminary Hazard Analysis (PHA, Ericson, 2015), Operating and Support Hazard Analysis (O&SHA, Vincoli, 2006), Probabilistic Risk Assessment (PRA, Mohaghegh, Kazemi, & Mosleh, 2009) Human Reliability Assessment (HRA, Kirwan, 2017), Failure Mode and Effects Analysis (FMEA, Stamatis, 2003), Formal Safety Assessment (FSA, Montewka, Goerlandt, & Kujala, 2014), Fault Tree Analysis (FTA, Ericson, 1999), determination of safety integrity levels (SILs, Gulland, 2004), and qualification and certification (e.g., Kornecki & Zalewski, 2009) processes and methods. With some of these methods, for instance, the probabilities of the risks associated with the safety-critical system can be the results of analysis. Naturally, application domain-specific methods and techniques also exist.

Modern incident and accident investigation methods also provide valuable input for the R&D of safety-critical technologies in different application domains. There are multiple investigation methods of incidents and accidents, such as AcciMaps, Human Factors Analysis and Classification System (HFACS), and Systems Theoretic Accident Modelling and Processes model (STAMP) (for details, see, e.g., Salmon, Cornelissen, & Trotter, 2012), which have influences from human factors research. STAMP specifically emphasises systems thinking and theory, which have also been integrated into systems and safety engineering (see, e.g., Leveson, 2011). Additionally, several other accident modelling methods for complex socio-technical systems exist that come from different disciplines than human factors (see details, e.g., in Qureshi, 2007). However, the specific presentation of the different approaches and methods of safety and risk management or accident and incident investigation is beyond the scope of this thesis, as here the focus is more on the early-phase human activity-centred analysis and design of these systems. The safety requirements related to the methods above can be described as primary in safety-critical systems development while, for example, user experience aspects aim to support these requirements.

In addition to methodical approaches, there are some key technical principles to decrease risks in complex safety-critical environments and they relate to, for instance, redundancy and diversity requirements (Radlinski, Bennett, Carterette, & Joachims, 2009) of the utilised technology. In short, these mean that there needs to be physically separate backup systems (for redundancy), preferably implemented with different technology platforms (for diversity) to prepare for a situation where a failure occurs. Furthermore, different kinds of preventive condition-based maintenance approaches are modern proactive ways of making sure that safety-critical technical systems will not have large technical faults and thereby cause safety threats.

2.1.3 Highly automated safety-critical systems and their human factors challenges

The responsibility of operative actions in complex safety-critical environments is increasingly being handed over from humans to digital automation systems (Bainbridge, 1983). For example, cars are today driven with the assistance of computers, ships are steered by digital autopilots, metro trains in some cities can drive by themselves with intelligent automation, and nuclear power plants may utilise fully digitised automation and control systems.

Automation has been defined in previous human factors literature as full or partial technological replacement of some function which has been previously performed by a human operator (Parasuraman, Sheridan, & Wickens, 2000). In detail, 'automation is technology that actively selects data, transforms information, makes decisions, or controls processes' (Lee & See, 2004, p. 50). Therefore, automation systems can be seen as a modern-day example of 'intelligent systems' (Bainbridge, 1983). In this dissertation, 'intelligent systems' or 'automation' refer to complex contemporary technologies that incorporate advanced digital electronics with (at least semi-) automated capabilities. The terms 'system', 'technology', or 'automation' are here used interchangeably with various possible prefixes such as 'intelligent', 'modern', 'new', and 'novel'.

Intelligent automation is today used in nearly every advanced technological environment. Often, modern automation is utilised with safety-critical systems when aiming to make the operations more precise, efficient, productive, or reliable. A highly automated system can operate without direct manual control especially in structured environments over extended periods of time, and in the process provide its intelligent action and decision-making capabilities (Goldberg, 2012; Shattuck & Woods, 1997). To illustrate different levels of automation (LOA), Parasuraman, Sheridan, and Wickens (2000) have presented a ten-step framework for the LOAs related to decision and action selection where 1 is 'entirely manual' and 10 is 'fully autonomous' (see details in Parasuraman et al., 2000). In levels 2–9, the system still needs humans to monitor and possibly intervene if needed, for example, in exceptional situations that are out of the scope of consideration of the system's designers.

Even though the level of automation with different systems seems to be increasing, humans' role in the operation of these systems will still be crucial in the future. Therefore, research and design approaches that consider the human aspects in the operation, design, and evaluation of these systems are needed to avoid, for example, the so-called 'ironies of automation'(Bainbridge, 1983). Typically, the human operators of highly automated systems become supervisors of the intelligent technology. Additionally, the complexity level of these systems is increasing and it is becoming difficult for one person to comprehend everything about what some complex technical system is doing in different situations. The user's role with these systems is commonly to monitor the system and possibly accept the actions that the automated system is planning to conduct. In some exceptional situations, the user can also possibly intervene and conduct the action manually (Bainbridge, 1983).

When monitoring and deciding to intervene in the functioning of an automated system, the user needs to have an appropriate level of awareness of the current situation (Endsley, 1995, 1996). Situational awareness has been defined by Endsley (1995, p. 36) as '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'. Additionally, with highly automated systems, the understanding of the current (and also possibly predicting the future) state of automation is essential (see automation awareness, e.g., in Karvonen, Lappalainen, & Liinasuo, 2014).

If the human operator does not understand what the system is doing, trust-related problems may appear. In a distrust situation, the user does not trust the technology and he/she can prefer to take manual control even in situations where it slows down the actions considerably. Therefore, this type of disuse of automation may cause performance decreases and inefficiency in the functioning of the joint intelligent system formed by the human and the automation together (Norros & Salo, 2009; Woods & Hollnagel, 2006).

On the other hand, mistrust (i.e., overtrust, see, e.g., Itoh, 2012; Itoh & Tanaka, 2012; Lee & Moray, 1994) can also be a problem. If the user trusts the technology too much, he/she will not monitor its actions on a required level and does not have the confidence to intervene manually, even though that would be needed in some situations. In safety-critical environments, this problem can lead to ill-advised actions that can even have catastrophic consequences. Overtrust of humans in technology has been identified as a contributing factor to some accidents, like the grounding of Royal Majesty (NTSB, 1995) or the mid-air collision at Lake Constance (Sträter, 2016).

Nowadays, a substantial amount of both naturalistic and laboratory studies have shown that trust is an important HF concept to describe human-technology interaction especially with highly automated systems (Lee & Gao, 2005; Salvendy, 2012). In human factors research, trust has been defined as 'the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability' (Lee & See, 2004, p. 51). In addition to distrust or mistrust, the concept of 'appropriate trust' has been suggested. When a user's trust in automation is calibrated to take into account the capability of automation (i.e., 'well-calibrated', Rajaonah, Anceaux, Tricot, & Pacaux-Lemoine, 2006; or 'well-placed' trust, Riegelsberger, Sasse, & McCarthy, 2005), the emergence of appropriate user trust in automation becomes possible (Lee & See, 2004). When the trust in a technology is appropriate, the user utilises automation in situations where it is meant to be used and where it can perform well.

'Appropriate trust' is also included as part of the earlier mentioned systems usability approach's 'user experience: the development potential of use' activity perspective, which is related to the communicative tool function (see, e.g., Savioja & Norros, 2013). According to Savioja (2014, p. 89), in this context, appropriate trust refers to the point that 'in the communicative function, the user feels that they can trust the tool in the same way one can trust another operator . . . '.

Other HF challenges related to the increasing level of automation include (in no specific order) operator out-of-the-loop performance problems (Endsley & Kiris, 1995), complacency (Lee et al., 2017), loss of situation or automation awareness (Endsley, 1995, 1996; Karvonen, Heikkilä, & Wahlström, 2019; Karvonen et al., 2014), operator boredom (Bainbridge, 1983; Cummings, Mastracchio, Thornburg, & Mkrtchyan, 2013), inappropriate reliance on automation (e.g., misuse or disuse, see Lee & See, 2004), and excessive mental workload (Parasuraman, Sheridan, & Wickens, 2008). For these challenges, alarm systems have been developed as one solution. However, when the alarm system is not working correctly, it can worsen the above problems.

When automation is doing most of the tasks, a large issue is that the user may not notice a problem situation as easily as compared to manual operation (Bainbridge, 1983). Furthermore, the monitoring work of automation done by the operators requires patience and deep expertise to spot potential threats. Especially in exceptional situations, professional skills and interpretations are needed according to Bainbridge (1983). It can therefore be difficult to recruit professionals with the required expertise and characteristics needed for this type of supervisory work.

To exacerbate these issues, modern safety-critical technology is becoming more and more autonomous. As mentioned, for example, in Karvonen and Aaltonen (2017), from the human factors perspective, highly automated intelligent systems with complex hidden functioning logic may end up being opaque. In practice, the humans monitoring and operating these systems may not have the possibility to follow what the system is doing and why. Therefore, an appropriate level of situation and automation awareness can be difficult to achieve. Additionally, when systems utilise artificial intelligence for decision-making, also 'artificial intelligence awareness' becomes relevant (Karvonen et al., 2019).

In addition to the above-mentioned problems, some other related human factors challenges with highly automated safety-critical systems include (in no specific order), for example, automation surprises (see, e.g., Sarter, Woods, & Billings, 1997), human vigilance (see, e.g., Cabon, Coblentz, Mollard, & Fouillot, 1993), development of an appropriate concept of operations (i.e., ConOps, including, human roles/responsibilities, tools, procedures, etc., see, e.g., Bilimoria, Johnson, & Schutte, 2014; Fairley & Thayer, 1997), automation reliability (see, e.g., Lee et al., 2017), operator deskilling (see, e.g., Bainbridge, 1983), automation mode confusion (see, e.g., Lee et al., 2017), and human-automation function allocation (i.e., task division, see, e.g., Pritchett, Kim, & Feigh, 2014). However, the inclusive description of these and all other identified human factors issues and challenges with highly automated systems that have been considered in conducting the empirical cases in the thesis is far beyond the scope of this dissertation.

2.1.4 Approaches to human factors challenges with safety-critical systems

To study and solve issues and challenges mentioned in the previous section, several different disciplines, methods, and techniques have been established. For example, human factors/ergonomics (HF/E), usability engineering, and resilience engineering are briefly presented here.

The ISO 9241-11 (2018, p. 25) standard states that 'ergonomics/human factors is a scientific discipline concerned with the understanding of interactions among human and other elements of a system'. Additionally, according to Stramler (1993, p. 148), human factors is:

> That field which is involved in conducting research regarding human psychological, social, physical, and biological characteristics, maintaining the information obtained from that research, and working to apply that information with respect to the design, operation, or use of products or systems for optimizing human performance, health, safety, and/or habitability.

Moreover, human factors engineering has traditionally been utilised in system design work to attain these qualities. Particularly with safety-critical systems, also, for example, personnel selection, team design, and training can also be seen as part of HFE (Lee et al., 2017).

Originally, HF/E research started to gain more interest during the Second World War (see, e.g., Hawkins, 2017) when, for example, military plane cockpit's information ergonomics (e.g., gauge design) was a matter of life and death. Today, the basic goals of human factors work are typically related to human-technology interaction efficiency, well-being of humans, reliability of the technology from the usage point of view, and effectiveness of the technology in question (Beuscart-Zéphir, Borycki, Carayon, Jaspers, & Pelayo, 2013; Savioja, 2014).

Some of these same goals are also mentioned in the definition of usability in the ISO 9241-11 (2018, p. 11) standard, which states usability to be the 'extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use'. In this definition,

- *effectiveness* refers to the 'accuracy and completeness with which users achieve specified goals';
- *efficiency* to 'the resources used in relation to the results achieved';

- *satisfaction* to 'the extent to which the user's physical, cognitive and emotional responses that result from use of a system, product or service meet the user's needs and expectations'; and
- *the context of use* to the 'combination of users, goals and tasks, resources, and environment' (ISO 9241-11, 2018, p. 8).

Good usability can be seen as one of the design-related premises of safety-critical technical systems, for example, because a poor level of usability of some system can lead to life-threating situations. Usability is also one of the key concepts in HCD (ISO 9241-11, 2018), which is discussed in Section 2.3 in more detail. As a specific approach to take usability into account in system development, for instance, the Usability Engineering Lifecycle has been developed (Mayhew & Mayhew, 1999).

With safety-critical systems, *safety* is obviously also a factor that needs crucial attention from HF/E. The contemporary view of safety in HF/E is that it is more than the absence of operative (human) errors. As mentioned in the introduction of Section 2.1, the modern resilience engineering (RE) approach sees that safety is based on resilience. According to Hollnagel et al. (2011, p. 16), resilience can be defined in detail as 'the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions'. In RE, humans are seen to be in a crucial role in conducting these adjustments proactively. Therefore, the positive role of people in socio-technical systems and human problem-solving skills in bringing resilience to recover from unexpected conditions is emphasised in RE (Hollnagel, Woods, & Leveson, 2012).

In addition to these above-mentioned approaches, there is a number of other related fields to human factors engineering, ranging from engineering psychology and human-systems integration to macroergonomics and cognitive engineering (see, e.g, Lee et al., 2017). However, the discussion of these approaches can be seen as beyond the scope of this thesis.

More relevant to the topic of this dissertation is the fact that the experiential factors of users have also recently started to gain more attention in the design of safety-critical systems (see, e.g., Norros et al., 2015; Savioja, 2014; Savioja, Liinasuo, et al., 2014). User experience has traditionally been studied in connection to non-safety-critical consumer products for pleasure in the field of human-computer interaction (Lee et al., 2017). However, user experience aspects can nowadays also be seen to contribute to the overall performance and safety aspects of the operation of safety-critical technologies. The concept of user experience and what it means in the context of safety-critical systems in this thesis is discussed in the next section.

2.2 User experience

In the mass-market consumer business, user experience has long been an important factor in the success of products and services. In this context, the concept of UX is said (e.g., by Hassenzahl & Tractinsky, 2006) to refer to hedonic experiences (like fun or pleasure) with the product in contrast to pragmatic or instrumental aspects (like usability). However, with complex safety-critical systems meant especially for professional use, UX may mean very different types of experiences than hedonic ones.

2.2.1 Definitions of user experience

There is a plethora of definitions for user experience, from which only a few of the most referred ones are presented here. According to the ISO 9241-11 (2018) standard, user experience is the 'user's perceptions and responses that result from the use and/or anticipated use of a system, product or service' (p. 17). In the ISO standard (2018), users' perceptions and responses are seen to include their 'emotions, beliefs, preferences, perceptions, comfort, behaviours, and accomplishments that occur before, during and after use' (p. 27). This definition can be seen as the industry's accepted definition of UX. The standard (ISO 9241-11, 2018, p. 27) also states that UX 'focuses on the user's preferences, attitudes, emotions and physical and psychological responses that occur before, during and after use (including perception of trust, safety, security, and privacy)'. Finally, the standard says that 'human-centred design can only manage those aspects of user experience that result from designed aspects of the interactive system' (ISO 9241-11, 2018, p. 9).

In the academic literature, one of the most cited definitions of user experience is from Hassenzahl and Tractinsky (2006, p. 95), who state that 'UX is a consequence of a user's internal state . . . , the characteristics of the designed system . . . and the context (or the environment) within which the interaction occurs . . . '. Additionally, while user experience is viewed to focus on the users' feelings, values, and responses, usability may be seen as instrumental and focused on task-related aspects (Hassenzahl & Tractinsky, 2006).

Secondly, Hekkert (2006, p. 160) has defined UX as

> The entire set of effects that is elicited by the interaction between a user and a product, including the degree to which all our senses are gratified (aesthetic experience), the meanings we attach to the product (experience of meaning), and the feelings and emotions that are elicited (emotional experience).

Thirdly, a definition from Hassenzhal (2008, p. 12) sees UX as a 'momentary, primarily evaluative feeling (good-bad) while interacting with a product or service'. Finally, a white paper by Roto et al. (2011, p. 7) states that 'the noun "user experience" refers to an encounter with a system that has a beginning and an end. It refers to an overall designation of how people have experienced (verb) a period of encountering a system'. However, despite the high amount of scientific references to these definitions, none of them works as such for the purposes of this thesis.

In activity theory (AT), the environment is seen to include the technical artefacts that humans utilise in their activity. Savioja (2014, p. 64) states that in AT, 'human-environment interaction is considered to take place as continuous embodied action-perception cycles'. In practice, this conveys that a repeated cycle of assumption, perception and action (see Neisser, 1976) takes place in human activity in an environment.

Activity theory takes a broad perspective of what an activity system includes: for example, the tools, subject, object, rules, community, and division of labour are considered (Engeström, 1999). These factors and their relationships are also depicted in the activity system model that has been created originally by Yrjö Engeström (see details in Engeström, 1999). It can be argued that with the help of AT, user experience has increased its relevance as a human-technology interaction research concept (Savioja, Liinasuo, et al., 2014).

Additionally in HCI research, an emphasis on the general role of socially construed meanings transmitted by technologies in people's everyday life has emerged (McCarthy & Wright, 2004). In this research, an increased interest in aesthetics and experience can be seen (e.g., Wright et al., 2008). In this context, Wright et al. (2008, p. 3) state that 'etymologically, "experience" stands for an orientation toward life as lived and felt in all its particulars. It tries to accommodate both the intensity of a moment of awe and the journey that is a lifetime'. Here, Wright et al. also refer to the aesthetic potential of all experiences based on the work of John Dewey (1925).

It is clear that the term 'user experience' is easily used without further contemplations of what is actually an experience. User experience may mean various phenomena and the context strongly affects the contents of user experience. A positive experience in one context can turn out to be a very bad experience in another. Consequently, there is a need to consider the concept in this thesis more specifically for the application domain in question. Therefore, in the next section, it is briefly presented what UX is here to seen to mean especially in the context of safety-critical technical systems.

2.2.2 User experience with safety-critical systems

Today, the users' expectations of how information systems should look, feel, function, and respond can be set by experiences with mass-market consumer products, such as smartphones. Consequently, consumer products' UX design guidelines have also been implemented into application areas with safety-critical features (see, e.g., Rice et al., 2016). However, there are many differences in what affects to good user experience when comparing consumer products for entertainment and complex safety-critical systems.

With safety-critical technologies, different kinds of experiences become relevant in usage activity than with consumer products. Considering good UX as a hedonistic experience along the lines of Hassenzahl and Tractinsky's (2006) definition, as mentioned herein earlier in the beginning of Section 2.2, does not apply as such with safety-critical systems. Euphoric peak experiences (e.g., Maslow, 1971) in a safety-critical environment may even distract the user from the primary task and risky situations can occur.

According to Savioja (2014), user experience with safety-critical systems can be viewed through the fulfilment of the instrumental function of the system: this fulfilment allows positive emotions to arise, because a feeling of accomplishment with the system can be achieved. Additionally, the user experiences in particularly work activity with safety-critical technologies can 'be interpreted to concern the appropriateness of the work practices and tools from the point of view of promoting the general goals of activity of which safety and effectiveness form an important part of' (Savioja, Liinasuo, et al., 2014, p. 431).

In line with these contemplations, Norros, Savioja, and Koskinen (2015) state that from the pragmatist view, the experience of a tool emerges from connecting it to activity and identifying the possibilities that it may provide to achieve intended outcomes. As a reference for good experience, they suggest the following idea, which originally comes from Vygotsky's (1978) activity theory: positive emotions emerge in activity when it creates new mediation that promises the development of the activity (Koski-Jännes, 1999). In other words, a positive experience in tool usage is an indicator of the successful activity that could be achieved with the proposed tool. These thoughts also conform to the EN 16170-2 (2016, p. 6) standard, which mentions that especially for work analysis 'it is essential to understand the motivations induced by the work system and the experiences of the user, which determines their observed activity'.

Based on the above ideas, in this dissertation an adapted version of activity theory is taken as a theoretical background to approach the concept of user experience. Related to user experience in complex work systems from the activity theoretical point of view, Savioja, Liinasuo, and Koskinen (2014, p. 429) have defined UX to be 'an indicator of the users' subjective feeling of the appropriateness of the proposed tool for the activity'. Specifically, this means that the users' positive experience of some (even early-stage representation) of the proposed technology indicates 'the potential to develop into a meaningful tool for the activity . . . , and benefits the interaction with the object of activity' (Savioja, 2014, p. 86).

These ideas, which have their origins in activity theory, are also central to considering the user experience of safety-critical technologies in this dissertation. Especially the approaches (e.g., systems usability) that are used in some of the empirical cases lean on these ideas. On a general level, the aim of this thesis is to bridge the gap between safety-critical systems development and user experience research and design. The backgrounds for human-centred design, experience design, and UX goals will be presented in detail in the next section.

2.3 Human-centred design, experience design and user experience goals

As presented earlier in this thesis, human-centred design is an approach to the design and development of systems in which the aim is in 'making interactive systems more usable' (ISO 9241-11, 2018, p. 10). According to the standard (2018), this aim can be achieved by focussing on the use of the system; by applying human factors, ergonomics and usability knowledge and techniques. Human-centred design can be seen as the most often used umbrella term when discussing the design of interactive systems by using UX/usability/human factors methods.

HCD's principles have received a lot of criticism, for instance, from Donald Norman (see, e.g., Norman, 2005). Norman suggests that, at times, the HCD principles might even be harmful (details in Norman, 2005). For example, Norman (2005) states that 'one concern is that the focus upon individual people (or groups) might improve things for them at the cost of making it worse for others'. The more something is tailored for the particular likes, dislikes, skills, and needs of a particular target population, the less likely it will be appropriate for others'. Instead of HCD, Norman (2005) suggests activity-centred design to be a better approach. ACD demands an in-depth understanding of the technology, the tools, and the reasons for the activities (Norman, 2005). Widely, ACD is a call for designers to avoid engaging with the users too much, and instead focus the design process on the activity of the target users (Cruickshank & Trivedi, 2017). A typical starting point for activity-centred design work is an analytical understanding of the 1) users' activity and the goals of the activity, 2) constraints and possibilities set by the target environment and its related factors, and 3) relevant experiential factors of the users and their activity. Therefore, instead of listening blindly to the users' wishes, the broad activity of the users is emphasised in ACD, as suggested by Norman (2005). Nevertheless, for the purposes of compatibility with the ISO 9241-11 (2018) standard, in this thesis, the term *human-centred design* is used. The tradition of ACD is seen to be included in this thesis' understanding of what should be considered in human-centred design in addition to what is said in the ISO 9241-11 (2018) standard.

'User experience design' or 'experience design' as concepts are not mentioned in the ISO 9241-11 standard. These terms will be used in this thesis almost synonymously, although it is acknowledged that experience design can be seen as a broader approach compared to user experience design, which focuses particularly on *users'* experiences. As one specific technique, UX goals are focused upon particularly from the perspective of how to use them in the human-centred design of safety-critical systems. Integrating UX goals to the human-centred design process is a more natural and straightforward choice (as UX is already considered in HCD) than trying to fit them to the complex process and aspects of the human factors engineering life cycle (although they have many similarities). Additionally, this thesis focuses on the *design of the system*, without including, for example, training, personnel selection, or organisational factors that are included in the scope of HFE. Therefore, in this thesis the focus is specifically in the UX goals' integration to the early-stage HCD process. Yet methodically, especially in the studies of this thesis, approaches related to HFE are also utilised.

In the forthcoming sections, some of the definitions of experience design and user experience goals based on previous literature are presented. In addition, ideas regarding UX goals as part of the design process are presented from earlier research.

2.3.1 Experience design

In academic literature, many experience design approaches with various labels have been presented. According to Lu and Roto (2016, p. 2), for example, experience-centred design (Wright & McCarthy, 2010), experience-driven design (Desmet & Schifferstein, 2011), and positive design (Desmet & Pohlmeyer, 2013) are examples of approaches which 'prioritize quality experience goals over material-level requirements'. Therefore, the aim with these approaches is to decide first to what kind of experience to aim for and second design features that evoke the targeted experience (Desmet & Schifferstein, 2011). The basic idea here is that with experience design, it is possible to support the emergence of particular experiences that are aimed for in the design work.

For UX design, the promise users experience in some new suggested solution may inform the carried design work about the correct way to proceed. This subjective evaluation of the users about the tool's potential allows the designers to anticipate the future use and the appropriateness of the tools to serve future activity (Norros et al., 2015). In this same regard, Savioja (2014) has considered the idea of utilising UX from the general design evaluation perspective as an early expression of the success of a design. In this stage, there may not even be interactive features available in the system, but rather, only an idea of a concept that is concretised in the form of scenarios and visualisations. According to Savioja (2014), user experiences of expert users are 'anticipatory indicators of the effects that the new system will have on the usage activity' (p. 91). For example, in the concept and prototype design stage, users may examine the designs and allow emotions regarding the solutions to surface. Therefore, UX is particularly significant in informing the designers about the potentiality of the design in the beginning stages of the development when it is not yet reasonable to use performance measures, such as errors or task times (Savioja, 2014).

Regarding design, Von Stamm (2008, p. 17) states that 'design is the conscious decision-making process by which information (an idea) is transformed into an outcome, be it tangible (product) or intangible (service)'. In addition, design can be seen as problem-solving activity (Falzon, 2008) – the designer strives to create solutions which effectively address the design problem. Simultaneously, the design problem itself is re-iterated, understood more profoundly, and even reframed (Falzon, 2008).

Accordingly, experience design is also a fundamentally creative activity. To support this activity, a human-centred designer's work process may include, for example, collecting user data (through, e.g., user research methods), examining the data, sketching concepts, prototyping, and evaluating the results with users to receive feedback in a structured format. In addition, some analytical design processes can be followed. Many experience design (XD) projects follow the ISO 9241-210 (2010) standard for human-centred design, as described earlier in this thesis. Kansei engineering has been applied especially in Japan for decades to account for user emotions in the design of products (Nagamachi, 1995, 2002). Desmet and Schifferstein (2011) have also proposed a process called Understand–Envision–Create for experience-driven design.

Regardless of these processes, the actual production of designs is often based on previous experience, trial and error, and iteration. Therefore, the production of designs may be thought to be more of a craft or a form of art, rather than an engineering process. Some designers might also experience that a dogmatic structure hinders creativity in design work.

There are some academic studies about implementing UX design into system development processes. Instead of introducing an entirely new approach into the development process, these studies have mostly investigated how UX research and design methods and techniques can be integrated, for example, into agile software development processes (Salah, Paige, & Cairns, 2014) or the approaches that design consultancies use for experience design and comprehension of UX in industry (Rozendaal, 2010). Regarding the latter, Rozendaal (2010) found out that XD in the industry includes features, such as 1) iterative research and design, 2) unpredictability of the process, 3) reliance on user insights, 4) prototypes as physical models, and 5) empathy tools. Some of these factors are also recognised as goals in the general HCD process (ISO 9241-210, 2010).

In the academic literature, there is a clear demand for more design-led approaches to enrichen the traditional human factors engineering (see, e.g., Lee et al., 2017) for safety-critical systems. For example, de Mol (2007) calls for more design-oriented approaches to human factors. UX research and design can have a lot to offer in this regard. Nevertheless, there are not many publicly available empirical case descriptions that investigate the use of experience design as part of the product development processes of safety-critical technologies.

2.3.2 User experience goals

In this dissertation, the utilisation of *UX goals* is suggested to be one central way to address UX in the design of new solutions. In practice, UX goals aim to support the emergence of specific experiences with the product or service under design. In addition to being the goals set for the design work, 'UX goals' can also be considered to be a technique by which the goals can be used in actual human-centred design work. The specification of UX goals as technique implies that there can be a certain orderly process in the use of UX goals throughout the design. Therefore, the technique can also be used as a part of some established more general-level approach or method for design (e.g., systems engineering). For the systematic use of UX goals as a technique, a certain level of structured stages are needed and these are discussed later in Sections 5 and 6.

The first academic workshop related specifically to user experience goals was organised in 2012 as part of the Nordic Conference on Human-Computer Interaction (NordiCHI) conference with the title 'How to Utilise User Experience Goals in Design?' (see Väätäjä, Olsson, Savioja, & Roto, 2012; Väätäjä, Savioja, Roto, Olsson, & Varsaluoma, 2015). A second UX goals workshop was held at the 2014 NordiCHI conference regarding 'The fuzzy front end of experience design' (see Kaasinen et al., 2015; Varsaluoma et al., 2015b). Thereafter, there have been several publications related directly to UX goals (e.g., Kymäläinen et al., 2016, 2017, 2015; Lu, 2018; Lu & Roto, 2014, 2016; Roto, Kaasinen, Nuutinen, & Seppänen, 2016; Varsaluoma, 2018; Varsaluoma et al., 2015b) in addition to the articles of this thesis.

Therefore, UX goals are a rather novel approach and the definitions of UX goals in the academic literature are diverse. For example, Lu and Roto (2014, p. 719) state that 'an experience goal describes the intended momentary emotion or the emotional relationship/bond that a person has towards the designed product or service'. Therefore, user experience goals are seen to reflect the intended user emotions or feelings, and not, for example, the functionality of the system. Varsaluoma et al. (2015b) suggest a bit broader temporal perspective by noting that UX goals 'concretize what the users are intended to experience before, during or after interacting with the product or service' (p. 324). UX goals have also been mentioned 'to describe the kinds of experiences that a product, service or system should evoke in the users' (Roto et al., 2016, p. 836).

Compared to *usability goals* (see, e.g., Maguire, 2001), UX goals are a more holistic construct. While usability goals are instrumentally focused (related, for example, to effectiveness, efficiency, safety, utility, learnability, and memorability of the system) (Maguire, 2001; Sharp, Preece, & Rogers, 2019), UX goals define how the system use is desired to be experienced. Whereas usability goals define 'how useful or productive a system is from its own perspective' (Sharp et al., 2019, sec. 1.7.2), UX goals focus on users' experiences. Therefore, usability goals can be described to aim to be objective while user experience goals are typically subjective to the specific users. Sharp et al. (2019) list, for instance, the following desirable qualities that can be used as user experience goals: satisfying, helpful, fun, enjoyable, motivating, provocative, engaging, challenging, surprising, pleasurable, rewarding, exciting, emotionally fulfilling, entertaining, and cognitively stimulating.

In the academic literature, other terms that are similar to user experience goals have also been presented, such as 'user experience targets' (Hartson & Pyla, 2012), 'experience qualities' (Arvola & Holmlid, 2015) 'experience goals' (Lu & Roto, 2014), or 'Xgoals' (Joutsela & Roto, 2016; Lu, 2018). However, in this dissertation, the term 'user experience goal' (i.e., 'UX goal') is used. The usage of the term 'goal' instead of the term 'target' or 'requirement' for the experiences to work as aims of the design is based here on the grounds of the premise that a designer cannot guarantee a certain strictly defined UX for the users. Instead, for the designer, it is only possible to aim to facilitate the emergence of the experience, for example, through appropriate UX goals.

Finally, a *user* experience goal refers to the experience induced by the expectations and usage of a system and the related activity for the user. The term 'experience goal' can be seen to be a broader concept than a UX goal: an experience goal aims at the experience facilitated by other factors as well (e.g., service quality induced by service personnel) than the technical system in question. Therefore, instead of a user, the experiencer with experience goals may be, for example, an employee, a customer, or some other human stakeholder. For instance, an employee experience consists of many other factors (e.g., workplace atmosphere, amount of work tasks, etc.) than the associated work tools the employee has.

2.3.3 User experience goals in the design process

According to the results of the first academic workshop on UX goals, a good experience goal is something that 1) helps aim the design as a guiding light, 2) is measurable, 3) describes positive emotions, and 4) is a way to communicate the desired experience with other people (Väätäjä, Savioja, Roto, Olsson, & Varsaluoma, 2015; Väätäjä, Olsson, Savioja & Roto, 2012). In addition, Lu and Roto (2014) have considered experience goals to be a 'starting point' and 'driver of design space expansion'. The fuzzy front end of experience design workshop's results (see Kaasinen et al., 2015; Varsaluoma et al., 2015b) indicated that the UX goals should be clear enough so that the design team can share and commit to them. Moreover, for instance, the following ways to concretise UX goals were recognised: co-construction (reciprocal discussion with context) to ensure that the goals reflect users' world and storytelling (including interesting personas) (Kaasinen et al., 2015). Furthermore, it has been suggested that experience goals should be communicated to all stakeholders who participate in creating the user experience (Varsaluoma et al., 2015b). Finally, a model for an experience goals elicitation process has also been presented in the academic literature (see Varsaluoma et al., 2015b).

If considering UX goals from a human-centred design process stage perspective, the first step of the UX goal-driven design can be said to be to *identify* the potential user experience goals. Next, based on the gained knowledge, the UX goals have to be *set* for the design work. Here, this stage is interchangeably referred to as *choosing* of user experience goals, as it is in some academic publications as well (see, e.g., Väätäjä et al., 2015). Finally, the user experience goals need to be *defined*. This entire process is here referred to as the *specification* of user experience goals.

A study by Varsaluoma et al. (2015) suggests that designers prefer to combine multiple sources of information when identifying user experience goals. Although the set UX goals at the early stages of the design project 'should guide the design process, in practice, it is possible that the original goals are iterated later on, as designers learn more about the users and the context' (Varsaluoma et al., 2015b, p. 331) where the system is planned to be used.

In addition to UX goals, a UX vision can also be set. According to Roto et al. (2016, p. 835), 'a UX vision reflects the overall experience that the design team wants to facilitate for the user'. As with UX goals, the UX vision should also be based on empathic knowledge of users and their activity (Roto et al., 2016). Furthermore, disseminating the UX vision appropriately aids in committing the design team and in focusing on the user viewpoint during the different phases of the design project (Roto et al., 2016).

To operationalise the chosen UX goals, it needs to be defined in detail what are the design implications of the goals (see, e.g., Varsaluoma, 2018). Furthermore, requirements for the system to be designed can be defined in this stage and should also be connected to the chosen UX goals to enhance traceability (Varsaluoma et al., 2015b). The corresponding design solutions can therefore be seen as concretisations of the user experience goals (Jokinen, 2015). Consequently, the produced design solutions are also trackable back to the set UX goals (Karvonen, Koskinen, & Haggrén, 2012). This traceability is especially important for complex safety-critical systems design, where traceability and structure is also emphasised in the technical systems engineering processes (Jarke, 1998). If there is a need to argue at some point for a design solution with a safety-critical technology, the reasoning behind the solution also needs to be shown.

In the evaluation phase, the fulfilment of the UX goals can also be evaluated. According to Varsaluoma et al. (2015b), the selection of appropriate (qualitative) metrics for evaluation is important in this phase. Furthermore, the evaluation of the fulfilment of the goals can be conducted iteratively in different phases of the design (Varsaluoma et al., 2015b).

3 RESEARCH QUESTIONS AND THEIR RATIONALE

The general-level research question of this dissertation is 'how can user experience goals be systematically used in the human-centred design of safety-critical systems?' As mentioned in Section 1.3, *systematic* in this context refers to using a careful system or method and to it being marked by thoroughness and regularity. From the methodical process perspective, it also refers to a particularly analytical design and evaluation process, which emphasises traceability and reflectivity (see, e.g., Savioja, 2014). Therefore, traceability and reflectivity are also seen as a key element in the systematicity of the UX goals' usage process.

As the general-level research question focuses on UX goals specifically in the *human-centred design* of safety-critical systems, it means that other design and development approaches are not considered to a wide extent in this thesis. However, it is acknowledged that UX goals are not the only goals that are involved in the development of safety-critical technologies in general. There is a variety of other (e.g., safety and risk-related) approaches, methods, and techniques used typically in safety-critical systems development to ensure, for example, safety, reliability, and maintainability (see, e.g., Blanchard & Fabrycky, 2006). Therefore, the importance of, for example, risk-based approaches to the validation and verification that happens in safety-critical systems development is significant, but not in the specific scope of this thesis.

The answer to the above-mentioned general-level research question is provided in Sections 6.1 and 6.2. Essentially, all the articles of the thesis contribute to answering the general-level RQ.

The specific research questions of this dissertation, their short justifications (i.e., rationale), and the sections of this dissertation addressing the questions along with the articles providing the basis for these answers are presented next.

RQ1: What is the value of experience design as part of the development of safety-critical systems?

Research question 1 asks about the value of experience design '*as part of*' the development of safety-critical systems. In practice, this question acknowledges that experience design may be only a part of safety-critical systems development, as there are many other already established practices for this work (e.g., systems engineering).

Additionally, RQ1 acknowledges that there is a wide gap between development practices of safety-critical systems and experience design approaches, methods and techniques. Therefore, in this thesis, it is crucial to ask first what the value of experience design for safety-critical technical systems development actually is. The value of XD can often be questioned in many contexts, which are not even safety-critical. Additionally, in the development of safety-critical environments, the experience design perspective has not traditionally been taken into account.

Practical case descriptions may demonstrate the real benefits of experience design in development projects. However, public cases related to the usage of experience design as part of safety-critical systems development are scarce. Empirical examples of what the value of experience design is as part of safety-critical technology development are thus needed.

The answer to this research question is derived from the empirical research of this dissertation. The empirical cases provide key points about the value of experience design as part of different phases of safety-critical systems design. Specifically, the focus is on the demonstration of value with systematic experience design techniques, such as UX goals.

RQ1 is addressed in Section 5.1, and Articles II–VIII provide the empirical basis for answering this question.

RQ2: How can analysis methods of human activity in safety-critical environments contribute to the identification of UX goals?

There is a multitude of different analysis methods of user activity, for instance, in the research fields of human factors and human-computer interaction. Examples of human factors analysis methods are provided, for instance, in a book by Stanton et al. (2017). The huge amount and different variants of these methods can be perplexing. For example, there are over 20 different established methods developed exclusively for conducting task analysis (see, e.g., Stanton, 2003).

Often, interview and observation methods are utilised as data gathering methods in studying user activity. These methods may produce a lot of data about the user and the usage activity of technology, which can be difficult to analyse and interpret from the design perspective.

The identification of appropriate UX goals based on the gathered user data (in addition to other potential sources) is also a challenging task. This *identification of UX goals* refers specifically to a phase where different sources of UX goals can be utilised to gather a list of potentially relevant UX goals. Therefore, the sources and methods to identify relevant UX goals may vary (see, e.g., Varsaluoma et al., 2015b for details).

The identification phase of UX goals is very demanding especially with complex safety-critical systems, as there are typically also many other prioritised goals (e.g., reliability or maintainability). These other goals may also either complement or compromise the UX goals' role in early-stage analysis, definition, and implementation of practical design solutions.

This second research question recognises these challenges and asks for ways to analyse user activity in safety-critical environments to extract and identify relevant UX goals from the gathered analysis data. The RQ2 is addressed in Section 5.2, and Articles I–VI provide the empirical basis for answering this question.

RQ3: How can UX goals be systematically set and defined for the human-centred design of safety-critical systems?

Setting UX goals refers to a phase where specific UX goals for the design work are chosen from a larger group of potential goals. In addition, *defining* UX goals is a phase where the meaning of the chosen UX goals is defined in detail. The sources and ways to set UX goals and define their meaning for the human-centred design of particularly safety-critical systems may differ significantly compared to systems meant, for instance, for entertainment. Notably, a certain level of systematicity may be needed in order for the UX goals of a safety-critical technology to be truly relevant for the users and not endanger the critical safety aspects of the system under design.

Therefore, it is relevant to study the ways in which UX goals can be systematically set and defined specifically for the design of complex safety-critical systems. This specification process should target to a result where the finally chosen UX goals are genuinely relevant and support the other goals of the safety-critical technology under design. Systematicity and carefulness is needed in order for the specification process to reach actually this target. However, in industry settings, this process cannot last particularly long, as typically practical product development projects are conducted within a tight schedule and a rapid series of milestones.

This research question continues from where RQ2 left off. In principle, after the identification of potential UX goals, which may be a long list in the beginning, the final UX goals should be chosen and defined for the design work. Particular attention and systematicity needs to be put onto this stage in the case of safety-critical systems design. Otherwise, the chosen and defined goals may not be relevant in the final usage of the system and can even compromise safety.

RQ3 is addressed in Section 5.3, and Articles II–VII provide the empirical basis for answering this question.

RQ4: How can UX goals be systematically utilised in the production of concepts and prototypes for safety-critical systems?

The utilisation of UX goals refers to the phase where they are used in the ideation and creation of new concepts and prototypes. This design phase is a creative process, which requires both space for ideas and structure for them to evolve as concepts. In the production of concept designs and prototypes, on the one hand, intuition is required for novel solutions, and on the other, systematicity for the design outcomes actually to be produced in a clear format. UX goals may bring both inspiration for the design, and, if utilised systematically, focus for the production of concepts and prototypes.

In human-centred concept design, for example, scenario stories, personas, user flows, storyboards, UX journey maps, visualisations (e.g., concept pictures or videos), 3D-printed mock-ups, or prototypes of different fidelity levels can be produced with the support of UX goals. For the production of these outcomes in the concept design stage, for instance, workshops, focus groups, design probes, brainstorming, brainwriting, drama methods, bodystorming, sketching, storyboarding, and low-/high-level prototyping approaches may be utilised. From a broader perspective, also a multitude of methods from scenario-based design (see, e.g., Carroll, 2000; Carroll & Haynes, 2017) and participatory design or cooperative design (i.e., co-design, e.g., Schuler & Namioka, 2017) to contextual design (e.g., Holtzblatt & Beyer, 2016) and Core-task Design (e.g., Norros et al., 2015) exist that may be used individually or in a combined manner to support this process. In addition, for the design of safety-critical interfaces, for example, Ecological Interface Design (EID, e.g., Burns, 2004) or Information Rich display Design (IRD, e.g., Braseth & Øritsland, 2013) are examples of potential methods.

Moreover, the development of safety-critical systems is typically conducted with structured systems engineering practices that emphasise traceability (see Section 2.1.2). Therefore, if UX goals are to be utilised in the human-centred design of safety-critical technical systems, the produced solutions should also be traceable back to the requirements and the set UX goals. Consequently, it is also relevant to ask how UX goals can be systematically a part of the production of concepts and prototypes for safety-critical technologies. This research question can be seen as a continuation of RQ3; after setting and defining the UX goals, they should be systematically utilised in the concept design.

RQ4 is addressed in Section 5.4, and Articles II–VIII provide the empirical basis for answering this question.

RQ5: How can the experiences of users be evaluated for systematic analysis of the fulfilment of user experience goals in human-centred design of safety-critical systems?

The outcomes from the concept design stage may be evaluated with various methods, depending on the aim of the evaluation. If the aim is to assess the experiences of users, there is a multitude of potential UX evaluation methods from which to choose. However, to evaluate specifically the fulfilment of UX goals, publicly established methods do not exist.

Therefore, it is relevant to ask what the key methods and techniques are to enable a systematic evaluation of experiences of users for analysing the fulfilment of UX goals in the early stages of the design of safety-critical technologies. This dissertation's empirical research focuses particularly on the beginning stage evaluation activities of safety-critical system design. The potential evaluation artefacts in this connection were mentioned in the RQ4 description above (e.g., prototypes). This RQ5 can be seen as a sequel to RQ4. After the production of the concepts or prototypes, they need to be evaluated with users. In this stage, their user experiences can also be analysed in order to evaluate the fulfilment of the UX goals. Naturally, this can be an iterative process where design and evaluation take turns in order for the final concept or prototype system to be achieved.

RQ5 is addressed in Section 5.5, and Articles III, VII, VIII, and IX provide the empirical basis for answering this question.

4 OVERVIEW OF THE EMPIRICAL RESEARCH

In this chapter, an overview of the empirical research cases of this dissertation is provided. For example, the research methods, settings, and other key background factors of each included study are presented. More information about the empirical research can be found in the thesis' articles, which present the different studies in detail.

For clarification, in the empirical context of this dissertation the term *study* refers to a single user study conducted as part of a case. The term *case*, on the other hand, refers here to the larger project case, in which the different studies (e.g., work analysis or evaluation studies) were conducted. In this article-based dissertation, some of the included articles include the description of different studies from a single case. In conducting the studies of these cases, principles of research ethics were followed. The aims and topics of the detailed empirical cases of research in the thesis are the following (with the acronym of the case used in this thesis and the related articles of the dissertation presented in brackets):

1. Analysis of the safety-related operational human factors challenges of automated metro train driving (MTD, Article I)
2. Design of a remote container crane operator station prototype (ROS, Articles II, III, IV, VII and VIII)
3. Future ship bridge concept design (FSB, Articles V and VI)
4. Evaluation of driver distraction warnings (DDW, Article IX)

When considering these cases through the research questions presented in Chapter 3, the first empirical case (which included only an analysis study) provides insights from the analysis methods perspective into RQ2, the second case to RQ1–RQ5, the third case to RQ1–RQ4, and the fourth case (which included only an evaluation study) to RQ5. In Table 1, these relations are presented in a table format for clarification purposes.

TABLE 1 Contribution of the cases to the research questions

Case	Contributions to the research questions
1. Analysis of the safety-related operational human factors challenges of an automated metro train driving (MTD)	RQ2: How can analysis methods of human activity in safety-critical environments contribute to the identification of UX goals?
2. Design of a remote container crane operator station prototype (ROS)	RQ1: What is the value of experience design as part of the development of safety-critical systems? RQ2: How can analysis methods of human activity in safety-critical environments contribute to the identification of UX goals? RQ3: How can UX goals be systematically set and defined for the human-centred design of safety-critical systems? RQ4: How can UX goals be systematically utilised in the production of concepts and prototypes for safety-critical systems? RQ5: How can the experiences of users be evaluated for systematic analysis of the fulfilment of user experience goals in human-centred design of safety-critical systems?
3. Future ship bridge concept design (FSB)	RQ1: What is the value of experience design as part of the development of safety-critical systems? RQ2: How can analysis methods of human activity in safety-critical environments contribute to the identification of UX goals? RQ3: How can UX goals be systematically set and defined for the human-centred design of safety-critical systems? RQ4: How can UX goals be systematically utilised in the production of concepts and prototypes for safety-critical systems?
4. Evaluation of driver distraction warnings (DDW)	RQ5: How can the experiences of users be evaluated for systematic analysis of the fulfilment of user experience goals in human-centred design of safety-critical systems?

Next, in Section 4.1, a key method of the empirical research is presented, the Core-Task Analysis. In Section 4.2, the research methods, settings and other key background factors of the different empirical research studies of the dissertation are examined. Finally, in Section 4.3, a summary of each case's background factors in their different stages is presented for clarity purposes in tables.

4.1 Core-Task Analysis as a method in the empirical studies

As an analysis method in this dissertation's empirical studies, Core-Task Analysis (CTA, see, e.g., Norros, 2004) was utilised in most of the cases (i.e., in the metro operation, container crane operation, and ship bridge work cases). CTA is a work analysis approach, which aims to identify the core task of a specific work. A core task can be defined as the 'main result-oriented content of work that can be derived by analysing the objective of work and the demands that the objective lays on workers both in general and in specific situations' (Article II of this thesis). In other words, it describes the aims and purposes of the work activity in connection to the environmental possibilities and constraints (Norros et al., 2015).

According to the CTA approach, characteristics related to dynamism, complexity, and uncertainty are common for safety-critical environments (Norros, 2004). Firstly, dynamism factors are connected to the environment's temporal demands, such as timing, duration, and delays (Article I). For example, a need to make quick decisions or actions indicates a high level of dynamism. All of the application domains of the four cases presented in this thesis include requirements for fast action in some specific situations. If comparing the application domains in this respect, the fastest action requirements are in car driving where traffic situations can change in a matter of seconds. In contrast, with large ships (e.g., platform supply vessels, as in this thesis) the unfolding of events can be seen to be the slowest out of these empirical case environments.

Secondly, complexity factors may arise from a vast amount of diverse elements in the system, which have dynamic interactions with each other. Furthermore, self-organised adaptations and unanticipated variability are common phenomena as the complexity level of systems increases (Savioja, 2014). It is obvious that there are differences between the complexity levels of the case environments presented here. The operation of a platform supply vessel may be to some extent seen to be more complex than driving a car. These differences in the target environment's complexity are also dependent on the scope of the examination that takes place. For instance, metro train driving with the provided controls in the cabin of the train may not be seen to be a complex task in itself, but the whole socio-technical system of a metro is a very complex environment (see Article I). Similarly, driving a car may not be seen very complex as such, but with the increasing amount of intelligent driver-assistance systems and complicated traffic environments, the level of task complexity for the driver can also increase.

Thirdly, uncertainty factors are related to the unexpectedness of events and to possible exceptional situations (Norros, 2004). With high-uncertainty safety-critical technologies, a technical failure or an unexpected combination of events can trigger even catastrophic consequences. What is typical of high-uncertainty systems is that human decisions and actions need to be made with insufficient information. In all of the case contexts of the dissertation, a high level of uncertainty can be seen to be present, particularly in exceptional situations, such as with technical faults of the system. In these situations, especially when aiming to understand and fix the fault, uncertainty in decision-making and actions is typically present.

In work environments, when the characteristics of a certain safety-critical work domain activity have been analysed, it is possible to interpret the work's core-task demands. Core-task demands are connected to the environment's critical functions, which the workers must support and fulfil in all circumstances (Article II). The demands can be seen as the environment's and the work's requirements that manifest themselves in specific forms in different work situations and set possibilities and constraints for interactions in the usage of the associated tools (Norros, 2004; Article II). Additionally, Savioja (2014, p. 84) has stated that:

> The elaboration of the core-task demands in a particular context provides knowledge about what specifically are, in detail, the features of the object that the tool is supposed to mediate in order for the workers to fulfil their core-task demands.

According to the CTA approach, resources or means of managing these demands are related to the actors' skills, knowledge, and collaboration (Norros, 2004; Norros et al., 2015). Consequently, the conducted activity in some work can be analysed by exploring how these demands and resources or means connect with each other (Wahlström, Seppänen, Norros, Aaltonen, & Riikonen, 2018).

CTA was utilised in the empirical analysis cases that included the analysis phase in this dissertation. In detail, the results of the conducted core-task analyses can be found from Articles I (metro train driving case), II (remote operator station for container cranes case), and VI (platform supply vessel type of the future ship bridge design case).

In the next section, the methods and materials of the different studies in the cases of this dissertation will be discussed more specifically.

4.2 Empirical research studies

The empirical research studies presented in this dissertation vary in their application domains, focuses, and aims. For example, in the metro train driver work analysis study (Article I), the focus of the research was in the analysis of the work domain, while in the remote container crane operator station case (Articles II, III, IV, VII and VIII), a concrete prototype to be studied with users was the aim. Nevertheless, one common denominator for all the empirical research cases of this dissertation is their application domains' (metros, ports, ships, and cars) safety-critical nature. In practice, this safety-criticality means that in an accident situation, human lives or the environment can be threatened. Firstly, in a metro environment, the passengers' lives can be threatened by an accident situation (e.g., a derailing accident or a fire outbreak in a tunnel). Secondly, in a port environment, for example, the truck drivers in the port area are in danger of being hit by a 40-ton container if the crane operator makes a mistake. Thirdly, the crew of a ship or the surrounding sea environment (e.g., with an oil spill) can be in danger in a distress situation at sea. Fourthly, with cars, the driver, the passengers, and the surrounding people may be in danger in the event of an accident.

Another commonality between the focus application domains of the dissertation is mobility and logistics. Ultimately, the aims of the application areas mentioned here focus on the mobility of people (metros or passenger cars) or on the logistics of goods (container terminals and non-passenger ships). With metros and passenger cars, the user experiences of both the drivers and passengers can be considered. With container terminals and non-passenger ships, mostly the user experiences of the related workers may be analysed and designed for in system development for those environments.

4.2.1 Analysis of metro train drivers' hidden roles

In the metro train driving (MTD) case, a primary aim was to investigate the safety-related operational challenges of Helsinki's metro automation in a situation where the drivers would be absent from the trains' cockpit. Therefore, the metro trains would be computer-driven and monitored remotely from a control room (Article I). This case's research was conducted to study both the hidden and explicit ways by which the train drivers (who operate the metros manually) drive the metro trains and support the operation of the entire Helsinki metro system as a socio-technical environment. Three separate, but interrelated empirical studies were conducted that had their basis in the CTA method (Norros, 2004).

Before these empirical studies, several meetings were organised with the metro operation organisation to understand the context and the different work roles in the metro environment. These meetings included free-form discussion sessions with metro traffic control room operators (i.e., traffic controllers) and the management representatives about the manual operation of the metro line in Helsinki. All of these discussions were important for the researchers conducting the studies to understand better the application domain and its workers, the terminology, and the utilised tools. To learn about the metro drivers' work tools, the researchers also visited the trains' cockpits in the depot area. A picture for illustration purposes of a metro train in the Helsinki metro line is provided with a driver's cockpit in the front in Figure 2. Additionally, a scientific literature review on the topic of metro automation and a benchmarking of existing automated metro solutions from around the world were also conducted based on publicly available literature.

FIGURE 2 A picture of a metro train on the Helsinki metro line © Leena Salo

The empirical studies included three interrelated qualitative investigations in the following chronological order: 1) an interview study of 12 metro drivers and traffic controllers, 2) an observational study of four train drivers, and 3) a mirror data workshop with the participants of both the interview and observation studies. The mirror data approach used here has its roots in Vygotsky's (1978) activity-theoretical principle of double stimulation (see also Engeström, 2011) and in retrospective think-aloud protocols (van den Haak, De Jong, & Schellens, 2003). In practice, this double stimulation in the metro case was provided with video materials and pictures collected from the user studies and the workshop participants were asked to comment freely, for example, on the work activity depicted in them. Therefore, the method is a form of self-confrontation from the video and picture material.

After each qualitative study, the results were analysed with the CTA approach, which helped the researchers to refine the goals for the next study. In addition, a card-sorting method (Wood & Wood, 2008) was utilised to categorise the mirror data workshop results (e.g., potential problems with a fully automated metro) under identified themes (Article I). Finally, a communication/interaction analysis was conducted about the drivers' and the metro environment's other primary human stakeholders' communication/interaction relationships.

In addition to these qualitative studies, a small quantitative analysis of the log data of the previous exceptional situations in the Helsinki metro was conducted. Thus, the potentially most frequently occurring exceptional situations in the automated metro environment could also be estimated to some extent. Details of the background of all of the conducted studies in this case can be found in Article I.

Although Article I does not mention the identified UX goals of this case explicitly, there were still many UX-related factors that were analysed (see Section 5.2.1). The primary meaning of this case in the dissertation is to demonstrate ways to analyse these UX-related factors with semi-structured interview, observation, and mirror data workshop methods. From the analysis results, it could also be possible to specify UX goals for further design activities.

The methods, analysis, and results of the MTD study are reported in detail in Article I of this dissertation. Article I and its further contemplations in Section 5.2.1 provide results related especially to RQ2.

4.2.2 Design of a remote operator station prototype for container cranes

The main aim of the ROS case of this dissertation was to design the final concept and an interactive prototype of a new remote operator station for container cranes by particularly emphasising the user experience perspective in the design process. As a critical part of this endeavour, user experience goals were studied and used throughout the design project's different phases. This case is a demonstration of how UX goals can be systematically taken into account especially in the early stages of the design process. The term 'early stages' here refers to all the stages leading up to the developed functional interactive prototype and its user evaluations.

The ROS case was conducted in collaboration with an industrial partner company (Konecranes), who also gave the design brief for the case. In the next sections, the basic facts about the empirical user studies of the case are summarised. First, the field studies that were conducted to understand the context and to specify the UX goals and their design implications are presented. Second, the basic facts about the ROS prototype evaluation study to investigate user experiences and the fulfilment of UX goals are reported.

4.2.2.1 Field studies for understanding the context and for the definition of UX goals

The purpose of the container crane operator field studies was to find out the core-task demands of both conventional and remote container crane operation by studying the crane operators' work in these contexts (Article II). In this way, the aim was also to identify suitable UX goals, define their meaning in the object context, and draw implications for the design of a novel remote crane operator workstation.

Before the field studies, two application-domain experts of the partner company, who were not familiar with the case, were interviewed to understand the conducted work with the container cranes and the related UXs better. The actual field studies consisted of interviews and observations of work activity in two international container terminals. In the first terminal, the crane operations were carried out manually on the spot of the operation from a conventional cabin in the crane structure. In this terminal, altogether six crane operators were interviewed and their work was observed individually in morning and evening shifts.

In the second terminal, most of the crane operations in the container yard's seaside and container stack sections were automated. Therefore, typically only the landside loading and unloading of container trucks was carried out by crane operators through remote operating stations with no direct sightline to the crane. Undoubtedly, the work in a remote control room is more pleasant, ergonomic, and social than being alone in a trembling cabin in the crane structure looking down most of the time while sitting in a bad posture. In this highly automated terminal, altogether five remote operators and one maintenance operator were interviewed. Furthermore, the operators' daily work together in a centralised remote control room located in an office building was observed (see Article II for details).

The interviews were conducted as semi-structured theme interviews. Most of the questions in the interviews were based on the Core-Task Analysis (Norros, 2004) and critical decision method (Klein, Calderwood, & MacGregor, 1989; Wong, 2004). Furthermore, the systems usability framework (Savioja & Norros, 2013) provided a basis especially for the interviews' UX-related questions. The interviews were audio recorded in the first terminal, and in the second terminal, notes were written down with laptops during the interviews for purposes of later analysis.

The process of the conducted interview studies was also to some extent similar to the Contextual Inquiry method (see, e.g., Holtzblatt & Beyer, 2016). The researchers took the role of an apprentice in the interviews with the crane operators (i.e., the masters). In other words, the operators taught the researchers the basics of the operations and answered the detailed questions that the researchers had. In addition, the researchers followed and observed the crane operators to understand better the work activity of the operators. During the observations of the work activity, the researcher did not distract the operators with questions and also otherwise remained quiet. This was due to the fact that some verbal interventions by the researchers during the critical phases of operation might have caused safety problems in the operation.

During the workplace observations, the operators were encouraged to think aloud while doing their work and explain the controls they were using in more detail (see Article II). The observations were video recorded in the first container terminal for analysis of the work in detail. In the second terminal, notes during the observations were written down with laptops.

In the analysis phase, the interviews were transcribed and relevant data from the video and notes material was sought through (Article II). The main aim of the analysis was to identify the crane operators' control and core-task demands related to dynamism, complexity, and uncertainty (Norros, 2004). The recognised demands from the CTA also worked partly as requirements for design. Based on these results, suitable UX goals and their design implications were identified and defined for the novel ROS system to be designed. Furthermore, concept and user interface (UI) requirements were defined and connected to the set UX goals (Article VIII).

As described in Article III, in addition to the field studies, a benchmarking study was conducted in order to understand specifically what kinds of remote crane operation solutions already exist and have been implemented in different international ports. Details about this benchmarking are provided in Articles II and III.

The field study and benchmarking results along with the specified UX goals and other requirements worked as a basis in co-design workshops where the actual concept brainstorming and design took place. Tens of different concept ideas were produced in the design workshops based on the defined requirements and other analysis results. The participants also evaluated the concepts in the workshops. The end-result was a one final concept of operations with two alternative UI designs that also answered the specified requirements. Based on these results, the iterative development of low-fidelity mock-ups and prototypes, such as a virtual camera view-based prototype of the ROS, was embarked upon. The end-results also included a description of the ConOps of the ROS system.

As mentioned in Article IV, the design activities of the case were conducted in a similar manner as many other concept design processes (cf., e.g., Keinonen & Takala, 2010; Takala, Keinonen, & Mantere, 2006), but with a particular focus on UX-related matters, as is, for example, in the Understand–Envision–Create process by Desmet and Schifferstein (2011).

To some extent, the design approach also resembled Contextual Design suggested by Holtzblatt and Beyer (2016). Contextual Design is a process where the system design is conducted based on actual information about the user activity and the context of this activity. The Contextual Design process includes different phases, the description of which is far beyond the scope of this thesis due to their vast extent (see details in Holtzblatt & Beyer, 2016). This process is claimed to ease the requirements change management during the design process and bring clarity to following the reasons and effects of the changes.

The methods, analysis, and results of the crane operator studies in both operating environments are reported in detail in Article II of this dissertation. In Articles III, IV, and VII the specification and utilisation of the UX goals in this design case are described from different perspectives. These articles and their further contemplations in Chapter 5 provide results related especially to the research questions RQ1–RQ4. Although the term 'UX target' is used in Articles II and III, it is meant to refer in this context to the same concept as a UX goal.

4.2.2.2 ROS prototype evaluation study for evaluating the experiences of users and fulfilment of UX goals

Before the ROS prototype evaluation study described in this dissertation, a preliminary evaluation of the ROS system with 20 university students was conducted. This first-stage evaluation investigated how, and whether, the user experience of the ROS simulator interface could be enhanced with either force feedback or visual augmentations (see Article VII). As the author of this thesis was not directly involved in this preliminary evaluation, the details of this study are not reported here.

In the evaluation study described in this thesis, there were six work-domain experts as users. This low number of participants can be justified with convenience sampling and the early stage of the design process (see Section 6.3.2 for more discussion on this topic). The objectives of the evaluation study were to 1) compare the user experiences of two optional ROS UI concepts and to 2) receive data on how well the UX goals 'experience of safe operation', 'sense of control', and 'feeling of presence' are fulfilled with the ROS simulator prototype (Article VIII). These types of simulators are common in human factors studies to gain proof about the usage safety and usability of the proposed system. Compared to the previous studies in the different phases of the case, this evaluation study was a more controlled one and conducted with an experimental approach where certain factors were measured. The details about the technical setup and the participants of the evaluation study are presented in Article VIII. Methodically, in this evaluation, a combination of different interview, observation, thinking aloud, and questionnaire methods was used for triangulation purposes. The procedure of a single evaluation session is presented in detail in Article VIII.

As stated in Article VIII, two different UI concepts (the four- and the two-camera view concept) were tested one at a time and the presentation order of these concepts was counterbalanced across participants. After each operational task, a brief, semi-structured interview was conducted. In addition, a general user experience survey and a systems usability survey was utilised (see details in Article VIII).

In the data analysis phase, the Usability Case (UC; see, e.g., Liinasuo & Norros, 2007) method was utilised. As mentioned in Article VIII, the Usability Case provides a systematic reasoning tool and reference for gathering of data regarding the technology under design and for testing its usability and UX in the targeted work (Norros, Liinasuo, Savioja, & Aaltonen, 2010). UC has its roots in the safety case methodology (Bishop & Bloomfield, 2000) and applies a similar case-based reasoning approach. In short, the UC method creates an accumulated body of evidence that provides arguments about the degree of usability of a system (Liinasuo & Norros, 2007). Further contemplations about the UC methods are provided by Laarni et al. (2014) and Norros et al. (2015).

As mentioned in Article VIII, it was possible to determine whether a certain UX goal (i.e., a claim in UC) was fulfilled based on the evidence provided by the user evaluation study. This evidence was first evaluated in connection with the fulfilment of the defined user requirements, which were connected to the UX goals. If most (i.e., over half) of the user requirements connected to a certain UX goal were met, then that UX goal could also be said to be fulfilled (Article VIII).

The methods, analysis, and results of the ROS evaluation study are reported in detail in Articles III, VII, and VIII of this dissertation. This evaluation study provides results (in Section 5.5.1) related especially to RQ5.

4.2.3 Design of future ship bridge concepts

In the future ship bridge concept design case (FSB), the main aim was to design novel future bridge concepts, which would be user-oriented in such a way that they are accepted and appreciated by professional mariners. The design brief from the industrial partner organisation (Rolls-Royce) in the case was that the concepts would represent the ship bridges of tugboats, cargo ships, and platform supply vessels (PSVs) in the year 2025. In this thesis and the related articles, the focus is on the tugboat and PSV designs, as the involved researchers of the case were able to conduct user studies related only to these vessels. Undoubtedly, the user studies conducted in these environments also helped foster understanding of the work at a cargo ship bridge, along with the discussions about cargo ship operation with the partner company representatives, and in this way contributed to the design of also the future cargo ship bridge concepts.

Before the empirical studies, expert workshops were arranged and a literature review on the relevant topics were conducted to familiarise the researchers with the application domain. In this case, two empirical studies were conducted. First, mariners and other maritime experts were interviewed. In the interviews, there were altogether 12 participants, who worked as shipping company directors, trainers, officers, sea captains, designers, and researchers. The interviews were semi-structured theme interviews focusing on CTA- and UX-related issues. Second, observations of work activity were conducted both on a real tugboat bridge and in a realistic PSV simulator. Both the interview and observation sessions were video recorded. A photo from the tugboat bridge observation is provided for illustration purposes in Figure 3. As in the ROS case, the basis and procedure of the interviews conducted in the FSB case bore some resemblance to the Contextual Inquiry (Holtzblatt & Beyer, 2016) approach.

In the analysis phase, the work on both the tug and PSV bridges was analysed with the CTA approach (Norros, 2004, 2014; Norros et al., 2015; Nuutinen & Norros, 2009). Results of these analyses are provided in Articles V and VI. The identified CTA demands from the results functioned also partly as requirements for the design work. Based on these results, the following types of goals were set: 1) general design goals, 2) an inspirational 'UX theme', 3) user experience goals, and 4) systems usability goals for the design work to be conducted.

After the analysis, a series of co-design workshops was organised where the results of these studies and the defined design requirements and goals were presented along with future maritime, interaction technology, and societal trend analyses to create a basis for the design work. With the trend analyses, the main aim was to facilitate the creation of far-reaching, futuristic, and radical concept ideas.

FIGURE 3 A photo from the tugboat bridge observation session © Daoxiang Zhang

The outcomes of the workshops were several different concept design alternatives presented as scenario stories, personas, and visualisations. The design method resembled also the Contextual Design approach by Holtzblatt and Beyer (2016). The design outcomes described the ConOps of the future ship bridge work and tools. These outcomes were reported in a confidential concept design report for the industrial partner company. With the help of the report and other provided material, experts from the partner company evaluated the concepts. Furthermore, based on this feedback, the developed final concepts were evaluated by maritime experts from the acceptance and experienced usefulness perspectives. However, the results of these evaluations are not reported here in detail as the confidential evaluation was planned and conducted by an external UX partner company. During the project, the final concepts were visualised as concept pictures and concept videos.

The design approach, the conducted studies, and the results of the FSB case are reported in detail in Articles V and VI. These articles and their further contemplations in Chapter 5 provide results related especially to RQ1–RQ4.

4.2.4 Evaluation of driver distraction warnings

In the DDW case, an evaluation study of context-sensitive driver distraction warnings was conducted (see Article IX). Although the study did not include explicitly mentioned UX goals, there were underlying implicit UX goals in the case that were identified from the literature and evaluated. The main meaning of this case in the dissertation is to demonstrate ways to evaluate these UX-related factors with survey and statistical analysis methods. Similar approaches could also be applied to the detail-level evaluation of the fulfilment of UX goals.

Compared to the other more naturalistic and exploratory studies in this dissertation (with the exception of the evaluation study of the ROS case), this DDW case included a more controlled study that was conducted with an experimental setup. In the conducted evaluation study, a test track experiment and a survey study was organised with 31 participants of which 10 were clearly novice (in or just out of driving school with less than 2,000 km of lifetime driving experience) and 21 experienced drivers (over 50,000 km of lifetime driving experience) (see Article IX for more details). In addition, drivers from different age groups were included. The number of participants can be justified with convenience sampling and the early stage of the design process (see Section 6.3.4 for more discussion related to this topic).

The evaluation study's aim was to explore the behavioural effects of a smartphone application called VisGuard[2] giving context-sensitive distraction warnings related to drivers' in-car glance behaviours and investigate how the drivers subjectively experience the application. In the experiment, the drivers conducted different kinds of tasks with a smartphone while driving and received context-aware distraction-warning messages. The purpose of these messages was to warn the drivers if they looked away from the road in front of them (mostly at the smartphone) too long when they were driving, or if there was an upcoming situation, which required their attention. The experiment was conducted on a real driving practice test track, which was closed off from other vehicles. A photo of the VisGuard application on a smartphone screen in a driving situation can be seen in Figure 4.

Video recording captured the driver's gaze and the screen of the smartphone where the application giving the warnings was installed. The experimenter was also present in the car, observing the driving activity of the participant. In addition, a laptop with an internet connection was used for filling in the two (pre- and post-experiment) online questionnaires before and after the driving activity. As the author of this thesis focused in this study mostly on designing the web-based questionnaire and analysing its results, only the basic facts about the questionnaire items and their analysis are presented here. Further details about the study are provided in Article IX.

2 More information online at http://www.visguard.com/ (accessed 10th of August, 2019)

FIGURE 4 Photo of the VisGuard application on a smartphone screen © Tuomo Kujala

Most of the questionnaire items in the post-study survey were related to the participants' experiences about the application, such as the experienced usefulness of the application, trust in the application, and acceptance of the application. These questions were formulated based on previous literature in the corresponding topics (for details, see Article IX). In the analysis phase, the gathered survey data was analysed both qualitatively and quantitatively. Regarding the latter, an exploratory factor analysis (EFA) of the survey results was conducted.

In the study, the means of the constructed experience factors was hypothesised to lean towards a positive direction from the midpoint of the scale, indicating positive general experiences regarding the application (Article IX). In this way, it was possible to measure and evaluate the fulfilment of these constructed experience factors, which can be interpreted to be also UX goals in this connection, although they were not explicitly called as such in Article IX.

The methods, analysis, and results of this driver distraction warnings study are reported in detail in Article IX. This article and its further contemplations in Section 5.5.1 provide results that are related especially to RQ5.

4.3 Methodical details of the empirical research

To describe the empirical research cases in a condensed format for clarification purposes, this section presents their background facts and methodical details in a table format. In Table 2, the basic facts of the different empirical research cases of this thesis are presented. In the next tables (Tables 3–6), the methods and other relevant factors (e.g., aim of the study, included early-stage design phases, study participants, defined requirements, and possible design outcomes) are summarised from the empirical studies of this dissertation. The first columns of these tables are a title column for the rest of the corresponding rows describing the facts of the different empirical studies in the cases (which are mentioned in the top rows).

Tables 3, 4, 5, and 6 are organised according to the modified labels (for the purposes of this dissertation; see Section 1.2) of the ISO 9241-210 (2010) human-centred design activities. This division of the tables is done to clarify which design activity was in question in the described studies. Table 3 presents the basic facts of the 'Understanding the context' activity of the studies of the cases. Table 4 presents the basic facts of the 'Specifying UX goals and other requirements' activity of the studies of the cases. Table 5 presents the basic facts of the 'Producing concept designs' activity of the studies of the cases. Finally, Table 6 presents the basic facts of the 'Evaluating the concepts' (and UX goal fulfilment) activity of the studies of the cases.

'N/A' in a cell means that this factor (or activity) was not included in the case's scope. In this dissertation, there are at least two studies described from each of the above-mentioned design activities. In addition, the dissertation's article numbers from which more details can be found are mentioned in each corresponding cell.

TABLE 2 Basic facts of the empirical cases

	Analysis of metro train drivers' hidden roles (MTD case)	**Design of a remote operator station prototype for container cranes (ROS case)**	**Design of future ship bridge concepts (FSB case)**	**Evaluation of driver distraction warnings (DDW case)**
Main aim of the cases:	Studying the safety-related operational challenges of metro automation in a driverless mode of operation (i.e., the driver would be absent from the train's cockpit and computers would drive the trains)	Concept and functional prototype design of a novel remote operator station for container cranes with an emphasis on the user experience perspective	Concept designs for future ship bridges of tugboats, cargo ships, and platform supply vessels for the year 2025 with an emphasis on the user experience perspective	Evaluation of context-sensitive driver distraction warning messages' effects on drivers' user experience and system acceptance
Included design activities of the case described as part of this thesis (modified labels from ISO 9241-210, 2010):	- Understanding the context	- Understanding the context - Specifying UX goals and other requirements - Producing concept designs - Evaluating the concepts	- Understanding the context - Specifying UX goals and other requirements - Producing concept designs	- Evaluating the concepts
Unit of analysis in the case:	Metro train driver's work activity with the provided tools	User experiences of the container crane operators in their work activity with the proposed tools	User experiences of the ship bridge workers' activity with the proposed tools	User experiences of car drivers' activity with the proposed driver distraction warning application

TABLE 3 Summary of the empirical studies conducted in the 'Understanding the context' activity

	Analysis of metro train drivers' hidden roles (MTD case)	**Design of a remote operator station prototype for container cranes (ROS case)**	**Design of future ship bridge concepts (FSB case)**	**Evaluation of driver distraction warnings (DDW case)**
Data gathering methods:	Article I - Scientific literature review - Benchmarking of existing solutions from public literature - Semi-structured theme interviews - Observations - Mirror data workshop - Video/audio recording	Articles II, III, IV, and VII - Benchmarking of existing solutions from public literature - Semi-structured theme interviews - Observations of actual work (including the usage of a think-aloud protocol)	Article VI - Literature review - Expert workshops - Semi-structured theme interviews - Observations of actual work	N/A
Data analysis methods:	Article I - Core-Task Analysis - Exceptional situation log analysis - Card sorting - Communication/Interaction analysis - Identification of possible challenges related to automation	Articles II, III, IV, and VII - Transcribing of the interviews - Core-Task Analysis - Identification of possible UX goals	Articles V and VI - Future maritime, interaction technology, and societal trend analyses - Core-Task Analysis - Identification of possible UX goals	N/A

TABLE 4 Summary of the empirical studies conducted in the 'Specifying UX goals and other requirements' activity

	Analysis of metro train drivers' hidden roles (MTD case)	**Design of a remote operator station prototype for container cranes (ROS case)**	**Design of future ship bridge concepts (FSB case)**	**Evaluation of driver distraction warnings (DDW case)**
Types of requirements defined:	Article I - Core-task demands from CTA - Requirements for the automation of the Helsinki metro from the human factors perspective - Implicit UX goals	Articles II, III, IV, and VII - Core-task demands from CTA worked partly as requirements for design - Concept and UI requirements - UX goals - Initial design implications of the UX goals	Articles V and VI - Core-task demands from CTA worked partly as requirements for design - General design goals - Inspirational 'UX theme' - UX goals	N/A

TABLE 5 Summary of the empirical studies conducted in the ’Producing concept designs’ activity

	Analysis of metro train drivers’ hidden roles (MTD case)	**Design of a remote operator station prototype for container cranes (ROS case)**	**Design of future ship bridge concepts (FSB case)**	**Evaluation of driver distraction warnings (DDW case)**
Practices utilised in concept design:	N/A	Articles III and VII - UX goals and other requirements presented in the beginning of co-design workshops - Concept brainstorming - Iterative development of low fidelity mock-ups and prototypes	Articles V and VI - User and future trend study results presented at the beginning of co-design workshops - Thematisation, conceptual reconfiguration, and theory visualisations - Abstraction and reformulation of user study findings and UX goals into themes, scenario stories, personas, storyboards, and concept visualisations	N/A
Design outcome:	N/A	Articles III and VII - The final concept of operations and a functional virtual camera view-based container crane remote operator station prototype (i.e., a simulator) with two alternative UI designs	Articles V and VI - Scenario stories, personas, storyboards, concept pictures, and concept videos	N/A

TABLE 6 Summary of the empirical studies conducted in the 'Evaluating the concepts' (and UX goal fulfilment) activity

	Analysis of metro train drivers' hidden roles (MTD case)	**Design of a remote operator station prototype for container cranes (ROS case)**	**Design of future ship bridge concepts (FSB case)**	**Evaluation of driver distraction warnings (DDW case)**
Number and types of users in the evaluation study:	N/A	Articles III, VII, and VIII Six work-domain experts who all had operated different types of manual industrial cranes: - Three had 1-5 years of previous experience - One had 6-10 years of previous experience - Two had over 10 years of previous experience In addition, three of the participants had previous remote crane operation experience.	N/A	Article IX 31 participants owning a driver's licence: - 10 clearly novice (in or just out of driving school with less than 2,000 km of lifetime driving experience) - 21 experienced drivers (over 50,000 km of lifetime driving experience)

	Analysis of metro train drivers' hidden roles (MTD case)	**Design of a remote operator station prototype for container cranes (ROS case)**	**Design of future ship bridge concepts (FSB case)**	**Evaluation of driver distraction warnings (DDW case)**
Evaluation data gathering and analysis methods:	N/A	Articles III, VII, and VIII - Interviews - Thinking aloud - Task performance indicators - Surveys - Observations - Video recording - Usability Case analysis	N/A	Article IX - Surveys - Observations - Video recording - Log data analysis - Exploratory factor analysis
Evaluated UX-related factors:	N/A	Articles III, VII, and VIII - Sense of control - Feeling of presence - Experience of safe operation	N/A	Article IX Based on earlier literature: - Trust in technology - Experienced usefulness - Harmfulness/Annoyance of the application - Acceptance

5 RESULTS AND CONTRIBUTIONS

This chapter presents the empirical results and research contributions of the dissertation divided into sections based on the specific research questions given in Chapter 3. Table 7 lists the specific research questions of the dissertation and the related articles, which provide the basis for answering the questions. In addition, the corresponding sections in this chapter that address the research questions are referred to in Table 7.

TABLE 7 The research questions, related articles, and the corresponding sections

Research question	**Related articles**	**Section addressing the RQ**
RQ1: What is the value of experience design as part of the development of safety-critical systems?	II, III, IV, V, VI, VII, and VIII	5.1 Value of experience design as part of the development of safety-critical systems
RQ2: How can analysis methods of human activity in safety-critical environments contribute to the identification of UX goals?	I, II, III, IV, V, and VI	5.2 Analysis of human activity in safety-critical environments for the identification of UX goals
RQ3: How can UX goals be systematically set and defined for the human-centred design of safety-critical systems?	II, III, IV, V, VI, and VII	5.3 Setting and defining UX goals for the human-centred design of safety-critical systems
RQ4: How can UX goals be systematically utilised in the production of concepts and prototypes for safety-critical systems?	II, III, V, VI, VII, and VIII	5.4 Utilisation of UX goals in the production of concepts and prototypes for safety-critical systems
RQ5: How can the experiences of users be evaluated for systematic analysis of the fulfilment of user experience goals in human-centred design of safety-critical systems?	III, VII, VIII, and IX	5.5 Evaluation of the experiences of users for analysing the fulfilment of UX goals in human-centred design of safety-critical systems

5.1 Value of experience design as part of the development of safety-critical systems

As presented in Section 2.3, experience design emphasises the meaning of positive experiences with a product or service. From the activity theoretical perspective, positive emotions may emerge in an activity when it creates new mediation that promises the development of that activity. In other words, a positive experience in technology usage can be seen as an early indication of the success that could be achieved with the proposed technology.

In this dissertation, two detailed practical example cases regarding the different phases of UX-driven early-stage design of safety-critical systems are provided, namely the ROS and the FSB cases. In the next two sections, both of these cases are considered from the perspective of how they contribute to understanding the value of experience design with safety-critical technologies. In addition, other experience design value implications from the included articles of this dissertation are elaborated upon in Section 5.1.3.

5.1.1 The value of experience design in the remote operator station case

In practice, the ROS case included the hands-on experience design and user evaluation of a remote operator container crane prototype solution. The case demonstrates aspects that can be considered in UX research and design if the object of design is a complex safety-critical system. For example, interviews and observations of professional users in their authentic system usage contexts were conducted. Based on these results, a Core-Task Analysis was undertaken. Some of the identified core-task demands from the CTA helped to specify the relevant UX goals for the concept and prototype design work. The actual design work was done in co-design workshops with brainstorming approaches, where the UX goals were taken as centrepieces. Finally, an interactive virtual camera view-based prototype was developed that was tested with users. In the evaluation tests, the fulfilment of the UX goals and other experiential aspects were evaluated. The ROS case's results are presented in detail in Articles II, III, and VIII.

Several value-related aspects regarding experience design of safety-critical systems arise from the ROS design case. Firstly, the position of UX in general with systems for safety-critical environments can be considered. For instance, in Article III, it is emphasised that the functionality of these types of systems is strongly connected to the experience-related aspects of activity. Therefore, isolating functionality and experience aspects from each other cannot be seen to benefit the design of safety-critical systems. Instead, experience aspects and functionality should be seen as tightly intertwined in design and use.

Secondly, one key result from the ROS case and Article III was that in a safety-critical context, different types of UXs become important compared to mass-market consumer products. For instance, carrying out safety-critical activity successfully with professional-quality tools will presumably contribute to positive experiences, like sense of control, experience of competence, and professional pride. These factors, on the other hand, contribute to the general-level meaningfulness, efficiency, and safety of the work activity. From this perspective, the consideration of appropriate UXs can be seen as just as important in the design of safety-critical systems as it is in the design of commercial mass-market products (Article III).

Thirdly, in Article III, it is also mentioned that professional users may hold certain intuitive abilities about the appropriateness assessment of the technologies that are suggested to aid them in their daily activity. This view highlights the conception that the user experiences of professional users should be considered in different phases of the design process (Article III). In the ROS case, it was noticed that even low-fidelity concept visualisations and interactive prototypes help experts to intuitively evaluate the value of the proposed design.

The systems usability framework discussed earlier in this dissertation, for example, in Section 2.2.2, also supports this line of reasoning: the framework bases its comprehension of UX on AT, which emphasises the experience of a user regarding the development potential of the technology to be used in the activity. This phenomenon was also realised in the ROS case: the users' feedback on the produced virtual camera view-based prototype's design solutions, which the users saw as promising, encouraged further design work on the ROS (Article III).

Fourthly, in line with activity theory, in Article III, it is also stated that 'UX is a subjective phenomenon that concerns the overall status of the activity and its objectives'. Initial UXs of professional users regarding concepts and prototypes can provide information about the ways a novel solution can support the emergence of relevant UXs in later development stage versions. This on the other hand can inform the designers in a substantial manner about the appropriateness of the UX goals that are aimed at during the design work of the safety-critical technology. Consequently, involving the potential users early on in safety-critical system design can help make the proposed system understandable by addressing both the functional (e.g., safety-related) and emotional (e.g., UX-related) aspects of the use activity.

Finally, the design results of the ROS case were utilised in the development and implementation of the final remote operator station system by the industrial partner company (Konecranes) of the case. Furthermore, the collected information, analyses, and experiences provided by the studies guided the partner company to successful new product launches. For the first time, the ROS system was sold to Lamong Bay Terminal in Surabaya, Indonesia[3] as part of an order of more than EUR 100 million[4]. Additionally, the author of this thesis later conducted a weeklong field study in Lamong Bay and the resulted final system was successfully used in real operations by the satisfied local users. A picture of the

---

3 See Karvonen, Koskinen, Wahlström, Perä, and Hakulinen (2015)

4 https://www.konecranes.com/press/releases/2013/konecranes-wins-record-order-for-container-handling-equipment-from-indonesian-terminal-operator and https://www.konecranes.com/press/releases/2014/konecranes-wins-historic-order-for-automated-rtg-system-from-indonesian-container-terminal-operator (accessed 30th of May, 2019)

final system installed with the main display showing some of the design solutions is provided in Figure 5. This further developed ROS system also became an integral part of the partner company's product portfolio. A webpage, which briefly describes the final ROS, is available online[5]. These above achievements highlight the value of experience design as part of the development of the safety-critical system in this case.

FIGURE 5 One of the remote operator stations installed in the Lamong Bay Terminal with the main display showing some of the final UI design solutions © Konecranes

5.1.2 Value of experience design in the future ship bridge case

The FSB case's main aim was to produce concept designs for future ship bridges with a focus on the UX perspective. As in the ROS case, in this case interviews and observations of users in realistic system use contexts were also conducted. One key differentiating factor of this case was that trends, such as human-technology interaction and societal trends, were studied from literature sources to inform the design work. Based on the user study results, also a Core-Task Analysis was conducted. In this case specifically, the identified core-task demands from the CTA worked partly as requirements and as a basis for the specification

5 https://www.konecranes.com/equipment/container-handling-equipment/remote-container-handling/remote-operating-station-ros (accessed 30th of May, 2019)

of the general UX theme and detailed UX goals for the concept design. The future ship bridge case is presented in detail in Articles V and VI.

As described in Article V, the developed inspirational UX theme (or vision) for the design work in the case was titled 'being one with the ship and the sea', which received its inspiration from the joint cognitive systems approach mentioned in Section 1.4. In JCS, the joint human-technology activity is viewed holistically rather than with atomism; it is inferred by considering interrelated elements, such as physical settings, learned meanings and notions, communication practices, and usability issues in the used tools (Article V). Furthermore, this theme was also relevant considering the actual maritime context — anything can happen in the open sea, so 'togetherness' and unity are important for the mariners at embodied, cognitive, and social levels (Article V).

The initial set of UX goals in this case included a wide range of different goals (for the ship bridge workers) to be supported with the developed design solutions. These UX goals were, for example, 'Sense of control' (that meant 'feeling confident, in command, and one step ahead' in this case), 'Feeling of community' (among the ship crew), 'Feeling of trust towards self and peers' (among the ship crew and external actors), 'Feeling of efficiency' (in the operative work), and 'Feeling of ownership' (for the ship and its equipment). Out of this initial set of UX goals, the suitable ones for different contexts were then utilised in the creation of the concept designs.

The actual UX goal and trend-driven concept design work was done in co-design workshops. In the workshops, UX goals, user and trend study results, and visualisations from the CTA guided the work – they were presented in the beginning of the workshop and held visible all the time in the co-design work. Based on the workshop results, personas and scenario stories were produced with various concept-level solutions described in them. These design outcomes were then evaluated by the industrial partner experts and based on these results the concepts for the real user evaluations were chosen. An external partner conducted the real user evaluations in this case. As these results are confidential, they are not included in this thesis. Similarly to the ROS case, it was noticed that even from low-fidelity concept visualisations and mock-ups, expert users can intuitively evaluate the value of the design based on their vast prior experience.

The results of these evaluations helped in choosing the final concepts of operations (ConOps). Based on the final ConOps, concept-level pictures and videos were produced with the help of a partner company. An example of one of the produced platform supply vessel concept pictures is provided in Figure 6. Some of the other concept pictures with the description of the produced commercial product (InnoLeap) from the case are also available online[6]. Furthermore, videos

[6] https://www.vttresearch.com/Documents/Augmented_Reality_Industrial_Reality_22March2018/Radical%20concept%20design%20with%20InnoLeap%20%E2%80%93%20Case%20Rolls-Royce%20Marine%20-%20Hannu%20Karvonen%20VTT.pdf_ (accessed 30th of May, 2019)

of the tug[7], cargo vessel[8], and PSV[9] concepts are also available online on YouTube. The concepts got considerable media attention after publication with press releases: for example, over 500 separate news articles were published about them online. After this project and publication of the concepts, Rolls-Royce has also been working to make the solutions presented in the concepts a reality in a UX-driven manner. These achievements also highlight to some extent the value of experience design as part of the development of the safety-critical system.

FIGURE 6 Concept picture of the produced PSV concept in the FSB case © Rolls-Royce

5.1.3 Other experience design value implications from the included articles

In Article IV, the main focus was in 'how to get insight and inspiration to define UX goals that concretise the intended experience' for design work. A basic requirement for a successful experience design approach is to consider the relevant experience(s) to aim for (Article IV). Therefore, it was suggested in Article IV that at the heart of experience design work should be the identification, setting, definition, and utilisation of UX goals. Setting the UX goals requires an understanding of the potential different approaches to UX goal identification. This understanding also makes the evaluation and reporting of the studies under work more systematic by bringing traceability and reflectivity to the work with UX goals. Moreover, on a general level, understanding of the potential sources for UX goal identification also brings structure and clarity to this fragmented field of research (Article IV).

---

7 https://youtu.be/27uCL90s20o (accessed 30th of May, 2019)
8 https://youtu.be/_nApv-C7qSg (accessed 30th of May, 2019)
9 https://youtu.be/_kv1hQLKOB0 (accessed 30th of May, 2019)

In Article VII, it was stated that the main idea of experience-driven design is to define the intended experience before the actual functionality and the technology. Therefore, the initial design space is determined by the intended experience rather than by the available technology (Article VII).

As an example, if the design problem is about safety, then designing *beyond* preventing safety problems can help. In this way, experience design aims to raise the bar of design by aiming for positive experiences, instead of going directly to answering the problem. Therefore, the positive effect created with the design solutions may also solve the original problem as a by-product. This line of reasoning also resonates with the ideology behind resilience engineering where the main ideas are in the promotion of recovery actions – and in humans being keys to mitigating failures. Thus, instead of focusing on the prevention of human errors, in solving safety challenges, it can help to focus on the positive impact humans have in complex safety-critical systems.

The design processes of the different cases presented in Article VII varied from scenario-based design to iterative agile development. However, in all the cases, 1) Investigation, 2) Design, and 3) Evaluation activities could be identified. These activities are similar to the ISO 9241-210:2010 human-centred design standard's activities with the exception of the 'specifying user requirements' activity. Firstly, 'Investigation' is seen to include all activities (e.g., background research and specification of design requirements [also UX goals]) aiming at improving the design team's understanding of the task at hand. Second, 'Design' activities include the generative activities of ideation, concept design, and prototyping the actual product concept. Finally, 'Evaluation' includes the assessment of the concepts to identify whether they evoke the intended experiences and evaluation of the relevance of the chosen UX goals. Article VII describes these above-mentioned activities in more detail.

Furthermore, as mentioned in Article IV, 'concrete UX goals may be most useful in experience-driven design in an industry context where various stakeholder groups need to agree on what to design'. This is because UX goals may help in keeping UX in focus through the multidisciplinary product design process (Article IV). In this way, UX goals may work as boundary objects to communicate to different stakeholder groups what is aimed at in the design work. This notion also applies to UX design of safety-critical systems.

From a design process perspective, UX research and design activities have been considered in this dissertation along the lines of the ISO 9241-210:2010 standard's activities for human-centred design. As complex safety-critical systems are typically developed with structured and systematic engineering practices (e.g., from systems engineering), it is also beneficial to present a stage-by-stage process for UX goals in human-centred design activities to be considered as a potential approach. In detail, the following numbered activities specific to this dissertation have been identified (the original ISO-9241-210 standard's activity labels are in brackets):

1. Understanding the context ('Understanding and specifying the context of use'),
2. Specifying UX goals and other requirements ('Specifying the user requirements'),
3. Producing concept designs ('Producing design solutions'), and
4. Evaluating the concepts ('Evaluating the design').

In line with these activities, in Article III ('UX targets' mean the same as 'UX goals' in this article; see the corresponding headings in the article) the following activities are described and included in the utilisation of UX goals in the ROS case (see Section 5.4.1 for details):

1. 'From contextual inquiry of crane operation work to UX targets' (i.e., UX goals)
   a. Discovering relevant UX targets
   b. Defining the guiding UX targets
2. 'UX targets guiding the concept design and the design of innovative features'
   a. Articulating the design implications of the UX targets
   b. Defining user requirements
   c. Design and implementation of solutions supporting the UX targets
3. 'Bringing the new system into use: concept evaluation'
   a. Concept and UX target evaluation

In addition to answering RQ2–RQ5, the next sections are also highly relevant and fit appropriately into the numbered activities that were modified based on the ISO 9241-210:2010 standard's activities. This fit can be paired in the following manner: Section 5.2 to activity 1) Understanding the context, Section 5.3 to activity 2) Specifying UX goals and other requirements, Section 5.4 to activity 3) Producing concept designs, and Section 5.5 to activity 4) Evaluating the concepts. Based on the empirical research, these activities are also gone through from the UX goals perspective in more detail in the next sections (Sections 5.2–5.5).

5.2 Analysis of human activity in safety-critical environments for the identification of UX goals

The first step in any human-centred design process is the analysis of the related human activity. This analysis can include, for example, literature reviews, interviews, or field observations related to the users. The main aim in this phase of a design-oriented project is to get the authentic context of the current technology usage activity understood thoroughly.

Traditionally, in UX research, analysis approaches such as contextual inquiries, context maps, workflows, personas, stakeholder maps, and affinity diagramming can be utilised in making sense of users, their activity, and their target context. However, for complex safety-critical systems and environments, more structured and robust approaches are often needed. Typically, human factors engineering approaches are utilised that use, for instance, task analyses to make sense of what the tasks and objectives of the users are. One activity-oriented task-analysis approach, the Core-Task Analysis, was also utilised in all of this dissertation's analysis phase studies. The results of these analysis studies are described next, particularly from the UX factors and UX goals identification perspective.

5.2.1 Analysis of metro train drivers' hidden roles

While conducting the MTD case, the plan of the metro organisation was to automate the Helsinki metro by the year 2014 to a driverless mode of operation. Therefore, in the MTD case the current tasks of the train drivers and the metro system as a whole were studied. As a key approach, the Core-Task Analysis was employed in the case. The results of the Core-Task Analysis are presented in Article I of this dissertation.

As mentioned in Article I, one key result from the analysis was that if the tasks, roles, and experience of the current drivers are not properly understood, it will be very difficult to design, handle, and manage the automated system to take over some of these tasks that are currently done by the drivers. From the analysis, it was also clear that trivial tasks for the train drivers could be difficult tasks for the automated system. Moreover, with safety-critical systems, it is important to emphasise that computers and automatic systems excel only at handling monotonous and repetitive basic tasks, which do not require complex problem-solving skills (that humans have at a much better level than computers; Article I).

From the passengers' perspective, the fact that the train driver is a human, like passengers, means that he or she is in this way a natural 'interface' towards them. Therefore, a driver can also provide much better intelligibility, service, and situational awareness for the passengers than an automated system. If the passengers do not understand what an automated system does in different situations, distrust in and disuse of the system can become a problem (Article I).

These trust-related issues are clearly linked to the experiential point of view of the passengers. Therefore, it is also crucial that in the beginning stages of automated mobility-related safety-critical technology deployment, the technology works well so that it can gain the trust of the passengers. Trust in technology is a fragile phenomenon, as it can be lost with a one single breach. Trust in automation was also discussed in Section 2.1.4, and appropriate trust can be seen as an important UX goal for complex safety-critical systems with high levels of automation.

The CTA results also yielded key core-task demands, which included UX-related factors (i.e., implicit UX goals) that could be identified from the train driver's perspective. From the train driver's viewpoint, such UX-related factors were 'trusting one's own competence in solving problematic situations', 'sharing experiences with colleagues', and 'the ability to work under uncertain information'. All of these factors typically have to do with the tools utilised in the train driver work (e.g., train controls, radio communication system, and the lineside signals) and how the driver experiences their usage. These factors also highlight the means to handle the core-task demands related to the dynamism and complexity of the metro environment from the driver's perspective.

In the analysis results of the MTD case in Article I, it is concluded that if the identified 'hidden roles' of the drivers are not accounted for in an automated system, the safety and quality of service of the automated metro can be affected negatively. This empirical case is an example of how the more traditional human factors approaches can be utilised to also identify and consider the experiential side of safety-critical operations. One key aim in the study was to look beneath the surface of the everyday work of the metro train drivers and identify issues that may not have been considered in the automation project of the metro. Specifically, the Core-Task Analysis and the mirror data workshop provided deep insights into the drivers' work and their potential hidden roles in the socio-technical metro system, which affect the quality of service and safety in everyday operations. Particularly, the mirror data (i.e., photos and video clips from the work) workshop approach allowed the participants to reflect on the presented issues on the grounds of previous experience and go deeper into the meaning of the factors in various situations from the experiential perspective.

These kinds of results provide information about the profound meanings behind different conducted tasks, situations, utilised tools, and certain behaviour. In this way, they also provide insightful subject matter regarding the UX factors that are relevant from the metro operation work and the passengers' safe and enjoyable mobility perspective in the usage of the metro service.

If UX matters had been in the focus of the development project of the Helsinki automated metro, the next activity in this work would have been in specifying UX-related design goals for the automated metro system. However, the research project ended in the analysis of the results provided in Article I.

Further phases in the metro case would have allowed the development of a new concept of operations for the automated environment based on the gained results. The starting point for this ConOps would have been the main aims of the current activity (i.e., the core task) and the desired user experiences, if UX had been emphasised. From the metro system's point of view, this would have meant in particular the consideration of the general goals of the metro activity (e.g., safety and good quality of service), but also both the passenger and worker experience perspectives with the new automated system. Therefore, in addition to material benefits, experiences related to, for example, quality, style, aesthetics, and experience of safety would also have been essential.

5.2.2 Analysis of container crane operators' work

In the ROS case's analysis phase (see Articles II, III, and IV), the main aim was to understand the different crane operation experiences by analysing the work demands in both conventional and remote operation settings. The results of the field study analysis included the identified core-task demands for both conventional and remote crane operation (presented in detail in Article II) and confirmatory results for the identified UX goals like sense of control or feeling of presence. Furthermore, initial design implications could be elaborated based on the results.

The field studies conducted in the ROS case were the first time the CTA method was utilised for this kind of comparative research of conventional and remote operational settings to facilitate the design of a new system. Through the case, it became evident that the CTA results were also suitable as a basis for the identification of UX goals and for the definition of their meaning. In addition, user requirements for the design of the remote operator station could be specified based on the CTA results and the identified UX goals (Article II).

In Article III, it is mentioned that the design task at hand often also determines to some extent the kinds of experiential goals that should guide the development. Therefore, if the design task is focused only on the concept of operations, the UX goals may be different compared to a situation where a fully implemented system is the focus of the design project. In the ROS case, a functional virtual camera view-based prototype was the ultimate design task for the project, and that also undoubtedly affected the choosing of the UX goals.

As a result of the container terminal benchmarking and other analysis activities, an initial broad list of relevant UX goals was identified (see details in Article III), which included UX goals such as 'feeling of presence', 'sense of control', 'experience of fluent cooperation', 'feeling of safe operation', 'experience of appropriate functioning of the tool', 'experience of fit for one's own use', 'experience of one's own competence, 'experience of interesting and rewarding work tasks', 'feeling of having a professional tool to operate', 'professional pride and motivation', 'feeling of being an essential part of the work community', and 'appropriate trust in technology'. In the next stage, which was the setting and defining of the final UX goals for the ROS case (see Section 5.3.2), this list needed to be limited to a shorter version for the actual task of the ROS concept and prototype design. In this stage, the analysed field study results also worked as a basis and were ultimately emphasised in choosing the 'feeling of safe operation' and 'experience of fluent co-operation' as the final UX goals (along with 'sense of control' and 'feeling of presence'). Choosing these two goals based on the field study results was affected by the fact that the crane operators constantly emphasised the safety aspects of the operations in their work and the daily work activity was very social and included a lot of communication with different professionals. The latter result was surprising for the case's researchers, as they initially thought the container crane operation to be rather lonely and quiet work.

5.2.3 Analysis of ship bridge operation work and relevant trends

In the future ship bridge case's (Articles V and VI) analysis phase, the aim was that the users' target context, tasks, and needs are methodically taken into consideration as well as possible. Therefore, a range of user studies was conducted. For example, 1) several mariners, partner company employees, and other maritime experts were interviewed; 2) some of the mariners' command bridge work was observed in detail; and 3) the work domain and the mariners' tasks were analysed with the Core-Task Analysis method.

In this case, the CTA results (i.e., the identified core-task demands) provided an 'analytical grid' (see Articles V and VI) through which the users' work tasks and needs could be viewed in the design work. The core-task demand findings represent both the actual methods in which the control demands are addressed (i.e., as expressed in the interviews or observed by the researchers) and also potential methods for this addressing (i.e., as inferred or suggested by the researchers or interviewees). These interrelations can be used as indications of the instrumental UX and systems usability goals and they form the basis of the analytical grid (see Article VI). In this way, the CTA model helped to analyse the field study data, as the findings on work activities could be arranged and viewed through the model (Article V).

As a difference to the earlier presented analyses of the MTD and ROS cases where CTA was also utilised, the analysis method worked in this case more as a tool for fostering creativity in design. Questions such as 'Perhaps some demands related to uncertainty could be addressed better by enhancing collaboration?' could be asked within the results after the analysis phase.

From a design theoretical perspective, in the analysis of the case, it was noticed that psychological, sociological, and anthropological theories and models can provide inspiration for UX goal specification and concept design as 'design frames' and principles. More specifically, goal setting, visualisation, thematisation, and conceptual reconfiguration were discovered as general mechanisms through which theories may translate into design contributions (see details in Article V). Particularly for design aiming at radical solutions, relevant human scientific theory and model interpretations with these mechanisms may provide a powerful vehicle for abstraction, which allows the necessary 'distance' to the gathered user data. In this way, a departure from the existing design and user paradigms towards what has not yet been imagined is made possible (Article V).

In addition to the analysis of the domain-specific work activity, this case included the analysis of trend foresight. Future maritime, interaction technology, and societal trend analyses were conducted. In detail, the trend studies helped in identifying, for example, relevant human-technology interaction trends such as the use of augmented reality and its potential in future ship work activity. Furthermore, societal challenges, such as environmental concerns, were considered in the analysis and, for instance, partly led to the consideration of eco-efficiency in the suggested ship bridge concepts. The trend analysis results were presented in the beginning of the co-design workshops where the actual concept ideation and design work was conducted: the main aim was to facilitate and provide inspiration for the creation of radical concept ideas with the trend analysis results.

5.3 Setting and defining UX goals for the human-centred design of safety-critical systems

As mentioned in Article VII, UX goals can be set and defined from the outcomes of the analysis phase. Furthermore, the gained knowledge about the context and users can enable the specification of user and other requirements. Compared to functional requirements that consider *what* to design, UX goals define *how* the design outcome should feel and what kind of experiences it should facilitate (Article VII).

The cases where UX goals were explicitly set for the design work in this dissertation were the ROS and FSB cases, which both focused on user experiences of workers in safety-critical environments. In Article VII, user experience in work environments was defined as 'the way a person feels about using a product, service, or system, in a work context, and how this shapes the image of oneself as a professional'. Moreover, based on the definitions of UX goals in Articles II and III, the purpose of UX goals in design is to describe the experiential qualities to focus on in the design work. In detail, UX goals define the positive experiences the designed system should aim to awaken in the users.

In this thesis, one of the main arguments is that the specification of UX goals should be based on a thorough understanding (from the analysis phase) of what the users want to achieve in their activity. In later design stages, this understanding also allows a deep comprehension of how the desired experiences could best be supported with design solutions.

The results of the articles of this dissertation contribute to different aspects of specifying UX goals. Article IV describes what can be the general sources for the identification of UX goals (see next Section 5.3.1). On a more practical level, Articles II, III, and IV describe the process of specifying UX goals in the ROS case (Section 5.3.2) and Articles V and VI the specification of the general UX vision and detailed goals in the FSB case (Section 5.3.3).

5.3.1 Sources for the setting of UX goals

In Article IV, the focus is on the information sources for the 'setting' of UX goals. In the context of this thesis, this setting is actually related more to the UX goal *identification* phase. In Article IV, five different sources to acquire insight and inspiration for this UX goal 'setting' were identified. Each source brings in a different viewpoint, which also reflects the multidisciplinary nature of UX in general (the following list is adapted from the text in Article IV):

1. The *Brand* approach ensures that the UX goals are in line with the company's brand promise.
2. The *Theory* approach utilises the available scientific knowledge of human behaviour.
3. The *Empathy* approach focuses on knowing the actual users and stepping into their shoes when specifying UX goals.

4. The *Technology* approach considers the new technologies that are being introduced and their positive or negative influence on UX.
5. The *Vision* approach focuses on renewal – i.e., introduction of a new UX.

Combining these approaches can bring in the viewpoints of different relevant stakeholders (Article IV) to UX goals. A clear understanding of these approaches helps identify the relevant UX goals more systematically. The content of Table 8 describes a summary adapted from Article IV to present the sources, their contributions, benefits, and challenges related to UX goal setting.

TABLE 8 Sources for 'setting' UX goals

Source/ Approach	**Short definition**	**Contribution**	**Benefit**	**Challenge**
Brand	UX goals derived from company brand image	A high-level UX vision can be defined to unite products under the same brand	Pre-defined, focused, and easy-to-share UX vision may be created	Interpretation of the vision of UX goals for different products under the brand
Theory	Deriving UX goals from scientific understanding of human beings	A collection of possible UX goals to choose from can be provided	Science-based evidence supports the set UX goals	Choosing the goals to focus on from a wide set of possible UX goals
Empathy	Inspiration from designer's empathic understanding of the users' world	Good ecological validity of the goals and a mindset focusing on the users' world	Mindset-focus supports decision-making beyond the goal-setting phase	Gaining insights into the deep emotional aspects of differing users' worlds
Technology	UX goals identified based on possibilities and challenges of new technology	UX possibilities and UX challenges raised by technical enablers may be considered	UX goals support the successful adoption of new technologies	Focusing on a certain technology may not cover all aspects of use
Vision	Inspiration from investigating the deep reasons for product existence and envisioning renewal	Inspiration from other domains can be gained and fixations on familiar solutions may be eliminated	Support for a holistic renewal with UX goals	User acceptance of the visionary goals

5.3.2 User experience goals specification in the ROS case

The specification of UX goals in the ROS case is described in detail in Articles II, III, and IV. Here, only a condensed version of the process is described with a focus on the generalisability of the results.

In the identification of UX goals for the ROS case, firstly, the utilisation of appropriate theoretical underpinnings was of the essence. For example, the systems usability framework, which recognises UX as an important element in the design of new tools for work, was utilised as a theoretical basis when identifying the UX goals. Particularly, the 'User Experience: The development potential of use' part of the framework was taken as one component which fostered the identification of the first vast set of UX goals for this case to be worked on further.

Secondly, the familiarisation of the case's researchers and designers with the application domain and the work in question was embarked upon. A benchmarking of international ports with different remote operation solutions was conducted. This analysis supported the identification of some of the domain-specific UX goals.

Thirdly, based on both the theoretical SU framework and the results of the benchmarking, a workshop was held with the case's key people about what the relevant UX goals could be in the design of the remote operation user interface. As a result, a list of 13 possible UX goals was created (see Article III). To validate these goals, two application-domain experts of the partner company evaluated each of them. After these interviews, the list of UX goals was refined based on the results.

Fourthly, the field studies of the case were conducted and operating experiences of expert users were collected. The refined list of UX goals, the Core-Task Analysis approach, and the critical decision method worked as a basis for the definition of the interview questions for these field studies. The final questions were related to UX goals such as 'sense of control', 'feeling of presence', and 'experience of fluent co-operation'. Through the field studies, an evaluation of the chosen UX goals could be done with application-domain experts and with the results of the work activity and domain analysis. A more detailed description of these studies is provided in Article II.

Finally, after the analysis of the results from the field studies was conducted, the most relevant UX goals for the design of the remote operation user interface in question were chosen (altogether four final goals). In addition to setting these goals, it was defined in detail what the goals actually meant in the context of the remote operator station prototype to be designed. Therefore, it was elaborated what the meanings were of each of the chosen UX goals in the case of remote crane operation and what their high-level design implications in this context were. To see the list of the final chosen UX goals and their importance in the

ROS case (based on Article IV), see Table 9[10]. A detailed description and some of the design implications of these goals are found in Article III.

In addition, an inspirational UX vision was defined for the design work in the ROS case, called 'hands-on remote operation experience'. In practice, this vision meant that although the operations were carried out remotely, the operators should still maintain a hands-on capability to guide the conducted operations through the remote user interface. In detail, the remote operation with the new UI should 'feel as vivid and safe as it would be carried out on-site where the crane is located' (Article IV).

TABLE 9 The final chosen UX goals and their importance in the ROS case

UX goal	Importance in ROS case (i.e., why this goal was chosen)
Sense of control	An accurate sense of control with the ROS is crucial, as the operator is not directly in touch with the operated crane and does not have a direct sightline to it.
Feeling of presence	Although the remote operator is not physically present at the site, she or he still has to perceive the prevailing conditions in the target environment vividly and at a sufficient level of realism. Therefore, feeling 'present' is a needed experience for the operator and an important design goal for the ROS.
Experience of fluent co-operation	The crane operation work is – contrary to the researchers' initial conceptions – a very social activity with a great deal of communication between different professionals. Thus, fluent collaboration tools are also needed. Especially in the remote operation setting, communication with other experts in the remote control room is essential.
Experience of safe operation	The cranes are lifting heavy loads, and human lives can be in danger if something goes wrong. The operators are experts who know when the operation is safe, and their experience or feeling about safety indicates to a good extent the real safety level of the operation. The experience of safe operation can be seen as a basis for trust in the utilised technology.

5.3.3 User experience goals specification in the future ship bridge case

The specification of the appropriate UX goals for the FSB case began from interaction with the industrial partner organisation – Rolls-Royce. Rolls-Royce's aims and hopes were the starting point for the UX goal specification, but also the personal values and research background of the researchers and designers in the case influenced them.

For exploring the creation of radical designs, the lenses of social psychology, cognitive science, and the arts were utilised in the case. An additional 'lens',

10 TABLE 9 is originally adapted from and further modified based on the project's final report available at https://issuu.com/vttfinland/docs/fimecc_115_uxus_ (accessed 15th of September, 2019)

through which the actual concept design was partly conducted was the idea of considering possible futures. Therefore, future scenarios were considered early on in the project; the produced concepts were not to serve only the needs of the present, but also those of the future. Major global trends that were taken into account were 1) the globalisation of markets, 2) environmental concerns, 3) exhaustion of natural resources, and 4) future predictions and weak signals regarding technological development.

The theoretical themes from which inspiration was drawn in the case were, for instance, the idea of emphasising the importance of positive user experiences. However, given the multitude of possible interpretations, 'a positive user experience' provides a very abstract aim for the design work. In other words, specifying detailed frames for positive experiences is necessary to bring more precise guidance to design (Article V). Therefore, clear user experience goals needed to be established in the case: for example, it was defined that the UX in the ship work should be comfortable both psychologically and physically and that the users would experience a sense of control in their operative work.

In this case, the theoretical backgrounds of CTA and joint cognitive systems approaches worked as a basis for understanding the preconditions for positive user experience in the ship work domain (Article V). In this way, as was previously mentioned in Section 5.1.2, the UX vision or theme, the feeling of 'being one with the ship and the sea', which emphasises 'togetherness' with the vessel and the environment, could also be created for the concept design work.

In detail, general-level design goals for the tugboat study included, for example, 'enhanced communication', 'enhanced anticipation of the escorted boat', and 'enhanced interpretation of the environment'. Finally, although the set UX goals were not explicitly stated as UX goals in the FSB case's Articles V and VI, they were still clearly defined as such and worked as an underlying driver for the design. Specifically, UX goals such as 'sense of control', 'feeling of comfort with using technology', 'feeling of ownership', 'feeling of efficiency', and 'feeling of community' were defined for the design. To make these goals more vivid for evaluation purposes, scenario stories, personas, and low-fidelity visualisations were produced in later phases of the case.

5.4 Utilisation of UX goals in the production of concepts and prototypes for safety-critical systems

The process of producing concepts, mock-ups, and prototypes in the beginning stages of technology design is a creative one and therefore challenging to cover in a formal manner. Nonetheless, it is not an arbitrary process, but features certain distinctive elements. The basic assumption here is that new design ideas do not come out of thin air – they always reflect and draw on some existing ideas. A novel idea is typically created by combining existing ideas in a way never conceived of before.

Traditional tools and approaches in UX research and design include, for example, personas, scenario stories, storyboards, probes, wireframes, and mock-ups. These may help in the production of concepts and prototypes, for example, for consumer products. Furthermore, participatory design is often emphasised, which essentially means that the users are taken to be involved as part of doing the design work.

The original purpose of setting and defining UX goals (see Section 5.3) is for the goals to guide the design work. However, in practice, the UX goals are not utilised in a vacuum, but many other forces also affect the design decisions. Acknowledging this in practical design work is essential. For example, safety standards, organisational factors, and resource availability may affect the utilisation of UX goals with safety-critical systems. Nevertheless, for the purposes of this dissertation, the utilisation of user experience goals in an ideal manner in design is focused on here without a vast consideration of other affecting factors.

5.4.1 UX goal-driven concept design of a remote operator station for container cranes

In this section, a connection is made between the final chosen UX goals and the design solutions developed in the ROS case into the new concept. In this case, first lots of design solutions were produced that were considered to be in line with the specified UX goals (Article VII). In Article III, the four selected UX goals in the ROS case are discussed in detail and their design implications in the design of the novel remote operator station are presented. The actual concept design activity included several co-design workshops with the researchers and designers of the ROS team. In order for the workshops to focus on the right issues, background insights about the specified UX goals, user requirements, and design implications were all introduced at the start of each workshop session (Article VII).

After the workshops, the most promising ideas were collaboratively materialised into low-fidelity concepts, mock-ups, and prototypes. During this process, a vast array of different design solutions was produced, from which the best ones were chosen for further development. These final solutions were chosen and refined based on an interplay of the gathered analysis results and the produced concepts. The used iterative approach for the design produced one basic ROS concept of operations in which two different camera-view setups could be modified (i.e., with a four-view setup and a two-view setup). Finally, an interactive virtual camera view-based prototype 'simulator' was developed with realistic physical controls and displays to support the evaluations of the designs and fulfilment of the UX goals with user tests.

Table 10 summarises the specification stages of the UX goals in this case that are partly based on Article III[11]. As a specific example, some design implications and high-level requirements related to the chosen UX goal 'the feeling of presence' are presented.

[11] TABLE 10 is also partly based on a visualisation about the UX goals in the design process of the ROS case in a booklet published about the project (see Roto et al., 2014).

TABLE 10 UX goals' specification in the concept design stage of the ROS case (see Articles II and III for details)

Early design process tasks in the ROS case	**UX goals' evolution stage based on the conducted tasks**	**Example UX goals from the ROS case (and their source from the different early design process tasks, see also Section 5.3.2) and design implications and requirements examples from the feeling of presence UX goal**
- UX goals derived from the systems usability (SU) framework - Benchmarking of current solutions - Field studies to get to know the domain, work activity and users' future expectations	1. Initial broad set of UX goals	- Feeling of presence (Benchmarking) - Sense of control (SU framework) - Professional pride and motivation (SU framework) - Experience of one's own competence (SU framework) - Appropriate trust in technology (SU framework) - Feeling of having a professional tool to operate (SU framework and benchmarking) - Feeling of being an essential part of the work community (SU framework) - Experience of fitness for one's own use (SU framework) - Experience of interesting and rewarding work tasks (SU framework) - Experience of appropriate functioning of the tool (SU framework)
- Core-Task Analysis of field data	2. Refined and focused set of UX goals	- Feeling of presence (Benchmarking) - Sense of control (SU framework) - Experience of safe operation (Field studies) - Experience of fluent co-operation (Field studies)
- Interpretation of the UX goals' meaning in the target environment	3. Design implications of the chosen UX goals	**Examples of design implications of the 'feeling of presence' UX goal:** - Quality of interaction (including, e.g., the feel of operation and clarity of the operating view) - Support for the comprehension of the physical dimensions in the target environment - Availability of rich data without disturbing delays
- Requirements definition for interaction and creating corresponding concept solutions for the ROS	4. Requirements and concept solutions connected to the chosen UX goals	**Examples of high-level requirements of the 'feeling of presence' UX goal:** - Representation of the operated entity should be primarily based on the video feeds from the truck-loading area or the stack. - The operation views present the loading zone in integrated and consistent manner. - The operation view should compensate for the loss of visual stereoscopic effect

In this connection, it has to be emphasised that in practice, this process was not as straightforward as described in Table 10. For example, some stages were done in parallel. As the stages 1 and 2 in Table 10 had already been focused on in detail in Section 5.3.2, the details of the stages 3 and 4 are subsequently discussed based specifically on Article III.

As can be seen from Table 10, the meaning of the chosen UX goals needed to be understood in the container crane remote operation context before the goals could be utilised in design. Therefore, an investigation of their conceptual meaning was conducted. The produced design implications described the specific ways to fulfil each UX goal with the new system. This operationalisation of the UX goals into the context was essential in defining what the UX goals actually meant in this context and how to measure their fulfilment. The defined user requirements were also each connected to the appropriate UX goals at this stage. Therefore, all the produced design solutions were based on these requirements, and in this way, were also trackable back to the set UX goals (Article VII).

A concept-level picture from Article VIII about one of the produced concept solutions is provided in Figure 7. As described in Article VIII, in the evaluation version of the ROS, the main display's user interface consisted of virtual world camera screens and simulated operational data. In this display, two different camera-view user interface setups were implemented into the prototype: a four-view (see Figure 7, which is adapted from Article VIII and presents a simplified concept illustration version in the main display) and a two-view setup (see Figure 5 for the real, finally implemented version in the main display). Wireframes of the layouts of these two alternative display setups can be seen in Figure 8, which is adapted from Article VIII. A detailed description of the developed final setup of the virtual camera view-based prototype system is provided in Article VIII.

FIGURE 7 Concept illustration of the ROS system with the four-view setup in the main display (from Article VIII) © Konecranes

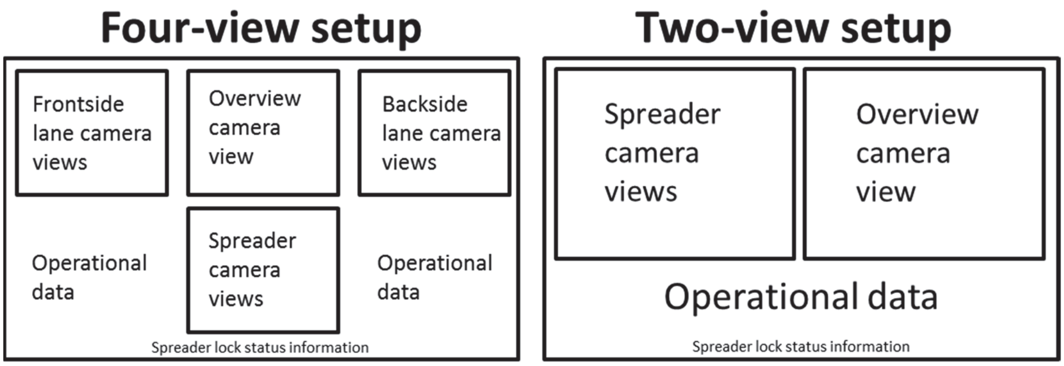


FIGURE 8 Wireframe versions of the two alternative main display setups of the concepts (from Article VIII)

5.4.2 UX goal-driven concept design of future ship bridges

The main aim in the FSB case was to study and innovate radical improvements to command bridge operations for future ships. The developed new bridge concepts included renewals to the bridge as a work environment, to the work processes, and to the ways of human-technology interaction. The creation of the renewals was based on a multidisciplinary co-creation process with research and design groups that had experience in design, work processes' renewal, collaborative work research, new interaction tools, and the application field of ship operations.

In this case, different UX goals were utilised in the brainstorming of the new bridge concepts of operations in co-design workshops. The overall UX vision for the design work was entitled 'being one with the ship and the sea'. This vision consisted of many UX goals, such as 'awareness of one's role in the system', 'feeling of trust towards self and peers', and 'feeling of human-environment system harmony'. In Article V, four design implications are discussed: one related to overall activity and three related to uncertainty aspects of maritime activity specifically. Scenario stories and personas were produced based on the UX vision, goals and the design implications. The scenario stories and personas were then used when the actual concept design of future bridges for different vessels was done.

In practice, the new design ideas were produced jointly in co-design workshops with the design team after the field research and gathering of the trend material. The workshops produced different descriptions of problems to be solved and possible solutions. The results were analysed together and most potential ideas were chosen (based on a vote by the participants) to be worked further into usage scenario stories, interaction concepts, and look and feel illustrations of the concepts. After the workshops, the researchers refined the ideas and the concepts of operations' descriptions, which were first presented in a report with scenario stories and low-fidelity visualisations.

After the creation of different concepts and the report describing them, a questionnaire study for the partner company personnel regarding the concepts was conducted. Based on the questionnaire results, the final concepts of operations to be visualised and evaluated with the partner company's customers were chosen. In the evaluations, scenario stories and cartoon-like concept pictures functioned as artefacts that aimed to concretise the UX goals. In the evaluations, the users were encouraged to put themselves into the proposed roles in the scenarios and consider the concept solutions from the perspective of actual activity on-board a ship. As the evaluations were confidentially planned, conducted, and reported by a partner UX company, the results of these evaluations are not reported here in detail.

After the user evaluations, the concepts of operations and the associated human-technology interactions were finalised for the different ship types and concept pictures. Additionally, 3D-animated videos were created with the help of a graphic design partner company. The visualisations were produced in tight collaboration with this partner. In this collaboration, several online meetings were held. The researchers and designers of the case instructed the graphic partner in the online drawing and graphic design sessions about what the developed concept solutions included in detail and how they should be visualised. One of the concept pictures is provided for illustration purposes in Figure 9 and another one was presented earlier in Figure 6. In addition, some other pictures from the case are also available online[12].

As an example, one design goal derived specifically from the CTA's identified control demands of the tugboat captain's bridge work was based on the fact that the escorting of ships requires anticipation of the movements of the escorted vessel. This core-task demand was also included in the chosen UX goals, which were the following: 1) ' sense of control' (i.e., feeling confident, in command, and one step ahead in the operation in this context), 2) 'feeling of trust towards self and peers' (i.e., captain's sense of self-trust and trust in peers [meaning here, e.g., the tugboat bosun or the pilot in the escorted ship]), and 3) 'feeling of efficiency' (in the operative work of tugboat escorting). In specific, the anticipation of the movements of the escorted vessel requires skill and active radio communication between the tug captain and the pilot in the escorted ship (Article V). Therefore, design solutions supporting this type of anticipatory actions and the associated UX goals are essential and were also generated in the concept solutions. More details about this tugboat concept solution (labelled 'Intelligent Towing') and its visualisations can be found in Articles V and VI.

12 https://www.vttresearch.com/Documents/Augmented_Reality_Industrial_Reality_22March2018/Radical%20concept%20design%20with%20InnoLeap%20%E2%80%93%20Case%20Rolls-Royce%20Marine%20-%20Hannu%20Karvonen%20VTT.pdf_ (accessed 30th of September, 2019)

FIGURE 9 A user interface concept illustration of the 'Logbook' system for tugboats from the FSB case © Rolls-Royce

5.5 Evaluation of the experiences of users for analysing the fulfilment of UX goals in human-centred design of safety-critical systems

The evaluation activities in the context of this dissertation cover the user experience evaluations of concept ideas, visualisations, scenario stories, sketches, and mock-ups. In addition, the user test studies of different fidelity-level interactive prototypes are considered. UX goals can also be evaluated from many perspectives during the design process. Firstly, it can be evaluated whether the designed solutions evoke the desired experiences (i.e., validation of the produced design solutions), and if so, how they actually support them. Secondly, it is possible to evaluate whether the intended experiences of the UX goals are actually relevant for and appreciated by the users (i.e., evaluation of the validity of the chosen UX goals).

Thirdly, especially in safety-critical systems development, it can be assessed whether the goals are actually relevant from the safety perspective even though the users would not experience them as suitable for them. This means that even though users might not feel the utilised UX goals to be relevant, it is still possible that the produced system solution, which is based on the UX goals, will support the general safety goals of the activity on a very good level. For example, even though 'experience of trustworthiness' can be rated as below average with some systems, it is still possible that from the safety perspective, an appropriate level of trust (see Section 2.1.4) in the specific system in question is supported. Therefore, if higher levels of trust would be rated by users, there may be some danger entailed in trusting the system excessively (from an objective perspective), which can ultimately result, for example, in accidents with safety-critical systems.

Next, two evaluation studies will be presented. The first study is clearly related to UX goal evaluation activities (i.e., evaluating the fulfilment of the specified UX goals) while the second one demonstrates, on a general level, more objective UX evaluation practices of underlying UX goals with a driver-assistance system.

5.5.1 Evaluation of prototype remote operator station for container cranes

As mentioned in Section 4.2.2.2 and Article VIII, one of the main objectives in the ROS evaluation was to assess whether the UX goals 'experience of safe operation', 'sense of control', and 'feeling of presence' are fulfilled with the developed version of the ROS prototype system. To assess whether these goals were fulfilled, a Usability Case-based reasoning method was utilised. Full details of the used method and the procedure can be found in Article VIII. In addition, the prototype system that was used in the evaluations is briefly described here and illustrated in Article VIII in detail.

Based on the UC method, the data gathered from the user studies (e.g., interview, questionnaire, and test session video and simulator data) was carefully analysed regarding each defined user requirement (i.e., a *subclaim* in UC) on whether positive or negative cumulative *evidence* was found about the fulfilment of each requirement (Article VIII). This analysis of *arguments* was done based on the gathered evidence from the user evaluation studies. Depending on the fulfilment of different user requirements, it was possible to determine whether a certain UX goal (i.e., a *claim* in the UC) was fulfilled or not (Article VIII). If most (i.e., over half) of the requirements connected to a certain UX goal were met, then the goal could also be said to have been fulfilled. The UC method also provided data on the usability and UX of the concepts under evaluation in addition to the evidence-based reasoning arguments for UC. One example of a UC-based reasoning in the ROS evaluations is provided in Figure 10 (adapted from Article VIII). The final Usability Case was presented in a separate confidential report for the partner company. The UC in its entirety cannot be presented here (or in an academic article) due to its large size.

Based on the gained results from the UC, the 'experience of safe operation' and 'feeling of presence' goals were not supported with the used version of the system in the evaluations. Some details of the reasons for this result are provided in Article VIII. In general, it was difficult to assess the fulfilment of these goals with the developed version of the prototype, as the crane operations were conducted in the evaluations in a virtual world where no human lives were in danger and the presented camera views were not real ones (Article VIII).

As mentioned in Article VIII, there was, however, support for the fulfilment of the 'sense of control' UX goal in the results, for example, because the users had a possibility to freely decide when to start and stop operating and to easily adjust the speed of operation with the provided joysticks. Additionally, the joysticks were felt to be sufficiently robust and suitable for controlling the crane with an appropriate feel of the operations (Article VIII).

Based on the results, it can also be said that the originally defined main UX vision of 'hands-on remote operation experience' for the ROS was not yet fulfilled with the prototype system version used in the evaluations. In the future design and development, the requirements that were not met were suggested in Article VI to be taken under careful investigation and answered with sufficient solutions. In this way, the specified UX goals could also be met better.

The experiences gained from this study suggest the Usability Case method also to be also a suitable systematic approach for evaluating the fulfilment of UX goals. Through the method, it was possible to evaluate the user experiences with the prototype system and measure and evaluate the fulfilment of the UX goals systematically in this prototype stage (see Article VIII).

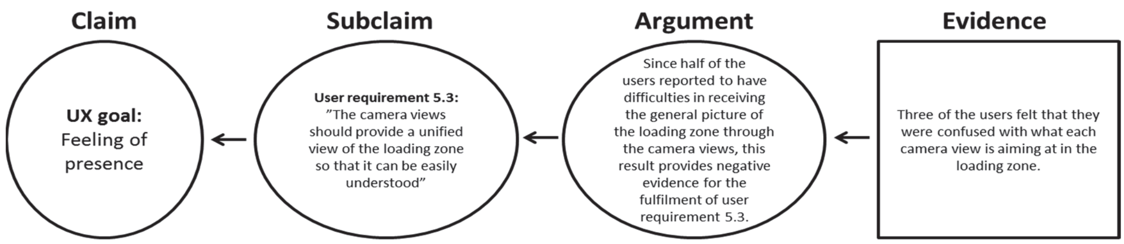


FIGURE 10 Example of UC-based reasoning in ROS evaluation (from Article VIII)

5.5.2 Evaluation of driver distraction warnings

Although in Article IX, it was not explicitly stated that the focus of the DDW case is in evaluating the fulfilment of UX goals, there were implicit UX goals in the case that were evaluated with a survey-based approach. Another reason for the presentation of this study here is that the analysis methods presented in the article may also suit the evaluation of UX goals.

In the study, a survey was designed to assess factors related to the application's acceptance based on previous scientific literature about trust in technology and technology acceptance. The measured experience-related factors included survey items related, for example, to participants' trust in the application and experienced usefulness of the application.

The statistical analysis of the survey results was conducted as exploratory factor analysis (Principal Axis Factoring). As mentioned in Article IX, the purpose of the EFA in this study was not to create a novel general scale, but to reduce the data set by constructing the most suitable factors amongst the items for the purposes of the conducted experiment.

Four different factors were identified in the EFA, namely, 'Trust in the application', 'Usefulness of the application', 'Harmfulness/annoyingness of the application', and 'Functioning of the circle symbol' (see details in Article IX). Out of these factors, especially the three first ones can be considered to be looking more at the experiential aspects. Furthermore, all four of the factors can be interpreted as indicative of an underlying factor contributing to the acceptance of the application.

Consequently, the main aim of the conducted evaluation studies was to assess whether the participants experienced the prototype mobile application's warning messages about potential driver distraction as acceptable. As mentioned above, this acceptance, on the other hand, was noticed in the statistical analysis of the results of the study to be constructed from 'UX goals', such as participants' experienced trust in the application, experienced usefulness of the application, experienced harmfulness/annoyingness of the application, and how the users experienced the functioning of the circle symbol in the application. These goals also formed the constructed experience factors in the results of the statistical analysis.

As mentioned in Section 4.2.4, the means of the four constructed experience factors were hypothesised to differ in a positive direction from the midpoint of the scale, indicating positive general experiences towards the application' (Article IX). With the harmfulness/annoyingness factor, this idea was the other way around (i.e., to differ in a negative direction). The rationale for the analysis of the factors here is that through the fulfilment of these UX goals, the users would accept the application as part of the driving and thus also obey the warnings given by the application. The line of reasoning behind this is that if the drivers felt that the application was, for example, trustworthy, useful, and not annoying, that also has a considerable effect on the observation and compliance of the warning messages by the drivers. Therefore, if the acceptance factor is on a positive level, then the safety of the car-driving activity is also enhanced.

The reason that the developed application gave specifically *context-adaptive* distraction warnings was the reduction aim of drivers' experiences of false positive messages. Consequently, the drivers would be expected to experience the application to work well, if they could understand and validate the reason for the warning from the surrounding driving environment (e.g., an intersection area). Therefore, trustworthiness and experienced usefulness of the application were also expected to be affected positively by the context-awareness of the warnings. In this way, UX goals could be seen in this case seen as central means of enhancing the traffic safety of driving with the application.

After their statistical analysis, the survey results indicated that the users experienced the system as trustworthy and useful. In addition, in general, the application was soundly accepted by the participants. This is somewhat in contrast with the fact that the application did not work appropriately at certain points with objective measures. In other words, the application had technical problems related to the tracking of the head position of the driver and therefore sometimes gave random warnings for particular drivers. This apparently had a negative effect at least on the answers related to these participants' experienced trust in the application.

The following conclusions regarding UX goals may be drawn from the study. In addition to evaluating whether the users achieve a certain user experience with the system, it is also beneficial to evaluate whether the experience is actually appropriate from safety or usefulness perspectives. The issue here is that even though the users would experience the system to be, for example, trustworthy and useful, if the application is not actually safe, the users may be lulled into a false belief of safety and usefulness of the application for different situations. Therefore, the validation of the appropriateness of the defined implicit UX goals is also necessary in the use of UX goals, especially with safety-critical systems.

The case also demonstrates the usage of quantitative and statistical approaches to study UX-related factors. With the defined 'UX goals' of the case, the introduction and adoption of the application may also be enhanced and evaluated in later phases of product development of the application.

6 GENERAL DISCUSSION

The empirical results of this dissertation bring forth various methodological and practical implications regarding experience design and the use of UX goals in the early stages of human-centred design of safety-critical systems, which will be discussed in this chapter in detail. One can also see some theoretical implications behind the presented methodological implications in Section 6.1. Additionally, the limitations and validity of the conducted research (Section 6.3) and possible future research topics (Section 6.4) are discussed in this chapter.

Particularly Sections 6.1 and 6.2 aim to address the general-level research question of this dissertation, which is 'how can user experience goals be systematically used in the human-centred design of safety-critical systems?' The results of the empirical cases and their conducted studies in this thesis provide the evidence and basis for answering the general-level research question in these sections.

6.1 Methodological implications

This thesis includes a wide array of results from a methodological point of view. The implications of these methodological findings for research and design are discussed in the next sections.

6.1.1 Rationale for the consideration of user experience and use of UX goals in human-centred design of safety-critical systems

A 'good user experience' is sometimes stated as an aim for a technology design project. However, for the designer it is an abstract and ambiguous starting point. From the designer's perspective, it is more practical to aim for a good UX if this experience is elaborated in more detail. Based on the empirical cases of this thesis, pondering questions such as 'how should the design outcome's usage feel?' (see Article VII) may help in defining the focus of the intended user experience. To answer these types of questions meaningfully, an understanding of the aims, needs, and challenges of the users is needed. Especially in design cases related to work environments, which were also a majority in this thesis' empirical research cases, an understanding of what kinds of factors produce positive experiences in that particular work activity is crucial.

For the designer, this understanding requires profound empathetic insight about the users and their activity. To empathise with the users, the designer has to strive to understand the technology usage deeply from the users' point of view, i.e., to try to 'step into the users' shoes', as mentioned in Article IV. Setting into the user's role allows the designer to understand what the user is really aiming at and how to support this activity, for example, with different technological solutions. Appropriate design solutions supporting users in their activity also facilitate the emergence of positive experiences for the users.

One key notion in this thesis is that UX goals bring structure and systematicity to the sometimes chaotic and ambiguous design project processes (cf., Leifer & Steinert, 2014). UX goals can be considered a technique, which makes the design work more target-driven, guided, and communicative compared to work without them. This statement is in line with the results of Väätäjä et al. (2012) who state that UX goals affect design by providing a vision, focus, guiding the design process, and helping in communication. Therefore, UX goals can also be seen as a meaningful technique for safety-critical systems development, which typically happens with structured engineering processes. Additionally, the integration of other UX research and design techniques in the same manner may help not only in the adoption of the approaches process-wise, but also in the development work conducted by engineers.

The UX goals also work as boundary objects (see, e.g., Star, 2010) to the wider parties involved in safety-critical systems design to communicate what is aimed at in the human-centred design work. Design is – in many ways – communication between different stakeholders. This communication can be supported with jointly agreed upon goals. By sharing the UX goals and reminding the different involved parties about them at appropriate intervals, the consideration of relevant UX factors in the design work can be enhanced and concretised. UX goals as part of the early stages of safety-critical system design bring additional value, as indicated by the empirical research in this thesis. UX goals help take into account the experiences of users better and more systematically in the early human-centred design and evaluation of safety-critical systems.

Another key point is that as long as the content of the user experiences to be considered in safety-critical system design can be viewed from a standpoint that considers the nature of the safety-critical usage activity, it is then possible to advocate user experience as a meaningful concept in the design of the system. In Article IV, user experience particularly in work environments was defined as 'the way a person feels about using a product, service, or system in a work context, and how this shapes the image of oneself as a professional'. This definition as such does not serve the topic of this dissertation for a number of reasons. Firstly, in addition to professional work systems, this dissertation also covers a non-professional driving-related assistance system. Secondly, the safety-critical aspects of the work where the empirical research of this dissertation has been conducted creates a unique challenge for UX-related issues. Thirdly, the activity-theoretical viewpoint in this thesis presents another difference compared to many other more traditional UX considerations.

For the purposes of safety-critical systems design, specifically in the context of this thesis, the following definition of user experience is suggested to be considered appropriate: 'User experience is an indicator of the user's subjective feeling of the appropriateness of the safety-critical system for activity. This experience shapes the image of the user as being part of the activity and contributes to the overall performance and safety of the human-technology joint entity in the safety-critical environment'. This definition does not aim to be a generalisable definition of what UX is. The definition utilises the activity-theoretical underpinnings and domain-specificity of this research as its basis. From the activity-centred viewpoint, good UX contributes also to the traditional goals with safety-critical systems, such as safety, good performance, reliability (see, e.g., Falzon, 2008), and system security (see, e.g., Knight, 2002). As mentioned in Section 1.4, the activity-centred approach widens the unit of analysis to consider the entire multilayered system of activity instead of only single-user tasks. In a similar way, UX goals do not only concern single tasks or user interactions with the technology, but rather reflect experience in activity in human-technology environments on a more general level (Article VII).

A further key perception here is that based on the results of this thesis, it is suggested that the content of an appropriate UX should always be defined domain-specifically. For example, the general definitions of UX presented in Chapter 2 would not have worked as such in the context of the empirical cases of this dissertation. At least the user experience of the empirical studies' users would not have included the same factors as in mass-market consumer products for entertainment, such as video games. Therefore, it can be concluded that the empirical results support the line of reasoning that there should always be context and domain-specific definitions for both user experiences of the users and the UX goals to be used in design work.

With complex safety-critical environments, a positive UX may mean, for example, that the system supports the user's feeling of competence and therefore allows experiences of success in the activity where the system is utilised. In this way, the user may achieve a sense of control regarding the utilised tools in activity and also of the whole activity (or one's part in it). Therefore, with safety-critical systems, an important part of a good user experience is that the user feels like an essential part of the ongoing activity, and not just a bystander who is left out of the loop in relation to what is happening in the current activity and in the functioning of the entire socio-technical system in question.

In general, UX should be taken into account in safety-critical systems design, because it also supports the achievement of the traditional primary goals of safety, efficacy, and efficiency typically associated with these systems. In the ISO

9241-210 (2010, p. 7) standard, there is a note about UX in safety-critical and mission-critical systems mentioning 'in safety-critical and mission-critical systems, it might be more important to ensure the effectiveness or efficiency of the system than to satisfy user preferences'. However, if user experience is understood as a contributing factor to safety, effectiveness, and efficiency of the whole activity in safety-critical environments (as is presented in this thesis), then this note's point has to be reconsidered. Generally, these types of dichotomies, as presented in the 9241-210:2010 do not necessarily advance the production of good designs. With some safety-critical systems, it can even be considered a precondition that certain UXs are realised in the use of the system. For example, with the driver distraction warning system presented in this thesis, it was a necessity for the activity's effectiveness and efficiency that the driver experiences the system to be useful, trustworthy, and acceptable (see Section 5.5.2 for details).

UX goals related to pleasure and fun, which are associated with mass-market consumer products, are not typically suitable for safety-critical environments where people's lives can be in danger. For example, in the ROS field study of this dissertation, it became evident that one of the dangers in remotely conducted operation work is a phenomenon, which the operators referred to as the 'game effect'. As mentioned in Article II, the game effect meant that because of operating the computerised user interface for a long time, sometimes the operators lost their touch to the real situation in the field and to the safety-critical nature of the remotely conducted task. In other words, this phenomenon means in practice that after doing the operations for several hours, the operator can get numb to the operations and does not remember the power of the forces and seriousness of the situations they are dealing with. This kind of an experience can be exacerbated if the operating environment is not realistically represented or the operating user interfaces have playful or gamified features (Article II).

Furthermore, in the empirical research, it was noticed in practice that at their best, well-designed safety-critical systems can enhance and facilitate the feeling of engagement, motivation, and the emergence of flow (see Article III) in work activity. These factors also support the general-level meaningfulness, safety, and efficiency goals of the conducted activity with the system. From these perspectives, the consideration of appropriate UXs can be seen as relevant in the design of safety-critical systems as it is nowadays for the design of commercial mass-market products (Article III).

6.1.2 Characteristics of an appropriate UX goal for safety-critical systems

On the basis of the empirical cases, this section presents a list of characteristics of an appropriate UX goal for safety-critical systems. UX goals should capture the focused and relevant user experiences for the context in question. This means that the goals should be sufficiently concrete to be aimed for in design work and their relevance should be appropriate, given the object environment. Based on the results of the UX goal specification, utilisation, and evaluation studies presented in the dissertation, the following is a list of general-level characteristics of an appropriate UX goal for a safety-critical system. An appropriate UX goal for a safety-critical system:

- Describes an emotional state of the user, which is enabled by the system usage
- Is sufficiently precise and specific to be understood in the same way by different stakeholders in design, but leaves also room for ideation and innovation
- Evokes design ideas the designers would not come up with otherwise
    - Encourages innovative ideas supporting the intended experience
- Is suitable for the target context in question
    - For safety-critical technologies, the UX goal should be such that it takes into account the potential dangerousness of the operational environment and does not compromise any safety aspects (i.e., the goal should not be in conflict with safety requirements)
- Works systematically as a guiding light throughout the design process by helping to point the human-centred design work in a relevant direction
- Is concretised through design implications, documented, distributed to the design team, made visible, and highlighted at appropriate time intervals in the course of the human-centred design work
    - If designers do not have clear goals, they will work towards their own 'hidden' assumptions
- Is communicated with appropriate systematic means to different stakeholders
- Is assessable
    - It should be possible to evaluate the goal's fulfilment systematically

The last point in the list refers to evaluating the fulfilment of the UX goal as is described in Section 5.5.1. This evaluation should be conducted iteratively and analytically as the design process proceeds and the object of design matures. An iterative evaluation approach allows the estimation of whether or not the goals are fulfilled with the current design version and continue the design-evaluation cycle until they actually are fulfilled. In this way, it is possible to make sure – systematically – that the intended experiences will be supported. The evaluation of UX goals on a general level was discussed in Section 5.5.

The above list is also in line with the results of Väätäjä et al. (2012, 2015), which showed that a 'good UX goal' is measurable, clear and precise, but broad enough to allow space for design ideas (Varsaluoma, 2018; Varsaluoma et al., 2015b). In addition, Väätäjä et al. (2012, 2015) argue that UX goals should guide the design, evoke design ideas, and function in supporting communication between different stakeholders in design.

Furthermore, characteristics for a good UX goal in the results of Varsaluoma et al. (2015b) were that it is possible to evaluate the goal, the goal is grounded in research, it is context-related, it drives creativity, and that it comes from the end-users. On a more general level, it was also found that a good UX goal 1) mediates empathy, 2), guides design work, and 3) is traceable through the design process (Varsaluoma et al., 2015b). All of these characteristics are also in line with the list presented above in this thesis.

6.1.3 Lessons learnt from identifying, setting, and defining user experience goals

Based on the experiences gained from the empirical research, identifying, setting, and defining UX goals can be seen as different phases to be conducted before utilising the goals in design. *Identifying* UX goals refers to the phase where different sources of UX goals can be utilised to gather a list of potentially relevant UX goals. *Setting* (or *choosing*) UX goals refers to the phase where the final UX goals for the design work are chosen based on earlier considerations. *Defining* UX goals is seen to be the phase where the meaning of the chosen UX goals and their design implications are defined before the actual design work starts to take place. Additionally, this entire process including all these phases can be seen as the *specification* of UX goals. This specification activity has also been referred to as the second phase in the UX goal activities presented in Section 1.2 and in later parts of this thesis.

In Article IV (see also Table 8), several sources for identifying relevant UX goals could be recognised. However, one of the main arguments in this thesis is that UX goals should be set based on a thorough understanding of users from the analysis phase. In later design phases, this understanding also allows a deep comprehension of how the aimed experiences could be best supported with design solutions.

Also in the empirical cases of this dissertation (see, e.g., Articles II, III, and VI), empathy for users was noticed to be especially important when identifying UX goals. In practice, this type of empathic UX goal identification required user studies to be conducted. However, if designers in some case already have a lot of information and previous experience from the target user group, a needed level of empathic understanding may also be reached in this way. Nevertheless, in this case, there is still a risk of basing the UX goal setting and the design work on stereotypical views or assumptions (Varsaluoma et al., 2015).

In this dissertation's empirical cases with a work analysis study included, the Core-Task Analysis approach functioned as a useful tool for the systematic analysis of field interview and observational studies (see Articles I–VII). Moreover, especially in the ROS and FSB cases, the CTA (with the support of the other analysis approaches in the studies) helped to generate insight about what the potential UX goals could be for the system to be designed.

Once the context was understood on a sufficient level in these cases, it was possible to start defining the meaning of the chosen UX goals in this context and appropriate requirements for the design of the new concepts. In addition to UX goals, these requirements included in the ROS case user-level, concept-level, and interface-level requirements, and consideration of other design implications from the context study results. These requirements were also linked to the set UX goals. In this way, it was possible to see the connection between the set UX goals and each requirement and the evaluation activities were very systematic in the ROS case.

As mentioned in Article IV, UX goals are not typically the only goals for the design work. With safety-critical systems, there are also other goals related, for example, to the system's safety, reliability, maintenance, price, and compatibility. In practice, UX goals should be consolidated to these other goals to ensure that no conflicts exist (Article IV). The end-result is always a compromise between the different types of goals. However, when considering safety-critical systems, safety-related requirements should be prioritised over other ones.

In the setting of user experience goals, the whole design team should be involved. With goals that the design team has set and defined together, the team members can understand their content and be engaged to aim at their fulfilment with the final system. When the team members are committed to the UX goals, they will also consider them in the produced design solutions. Although in the empirical cases presented in this dissertation the design team members came from very different backgrounds (e.g., engineering, psychology, and industrial design), they could understand the meaning of the UX goals well and commit to their fulfilment during the design and evaluation work.

In sharing and engaging other stakeholders in the set UX goals, an example of a committable approach is to create first a general UX vision for the system. The vision can then be elaborated into a few clear, meaningful, and understandable UX goals. In the ROS case, this vision was 'hands-on experience in remote operation', which consisted of the UX goals 'experience of safe operation', 'sense of control', 'feeling of presence', and 'experience of fluent co-operation' (Article III). In the future ship bridge case, this general UX vision was 'being one with the ship and the sea'. The UX goals that contributed to this experience in each concept solution depended, for instance, on the vessel type and its suggested solutions (see Articles V and VI for details).

When appropriate goals and requirements have been specified, they should be coherently utilised in concept design, for example, in the design of tasks, system features, system user interfaces, and general appearance of the product. The lessons learnt regarding the utilisation of UX goals in the production of concepts, mock-ups, and prototypes are discussed next.

6.1.4 Lessons learnt from UX goal utilisation in the production and evaluation of concepts and prototypes

Article VII lists three distinct challenges that can be identified in UX goals' utilisation in design: 1) 'finding the appropriate abstraction level of UX goals', 2) 'translating UX goals into appropriate guidance for design', and 3) 'directing and keeping focus on experience'. In the ROS case, a specific process was used to generate design solutions from the UX goals. Details of this process are described in Section 5.4.1 and Articles II, III, and VII of this dissertation. The benefit of this type of structured process is that it is possible to trace the produced design solutions to the originally specified UX goals for validation purposes (Article VII).

The potential downside is that the use of a new process that may be unfamiliar to the team involved in the design process can require some learning time when used for the first time (Article VII).

As mentioned in Article VII, in some of the article's cases, the UX goals were originally expected to 'act as evaluation criteria for meeting experiential goals, but evaluations of the design concepts sometimes revealed the need to reconsider the Xgoals themselves'. This evaluation of the relevance and validity of the chosen UX goals (mentioned also in Section 5.5) wound up being a very challenging yet relevant task in this thesis' ROS case as well. In the ROS case, the validity of the chosen UX goals was consistently evaluated and the final chosen UX goals were the result of a long refinement process. If the different phases of the ROS case are considered from the iterative design and evaluation process perspective, it was evident that 'the UX goals worked as powerful means of ensuring that the design process stayed on its correct course throughout the project' (Karvonen, Koskinen, Wahlström, Perä, & Hakulinen, 2015, p. 132). Therefore, UX goal evaluation proved to be useful in ensuring that the goals are focused on the appropriate aspects of the users' experiences.

Through the ROS case, it could be noticed that UX goals brought both structure to UX research and design and opened up novel evaluation possibilities for the targeted user experiences. As recognised in the case, the evaluation of appropriate user experiences requires a realistic setup of the actual system and context. Operating a virtual camera view-based prototype version of the user interface does not necessarily bring about appropriate work-related experiences, such as the experience of competence or feeling of presence (Article VII). Therefore, it can sometimes be very difficult to evaluate experiences realistically in the concept design stage.

In the future ship bridge case, scenario stories, personas, storyboards, concept pictures, 3D prints, and concept videos concretised the concepts and through these artefacts, the users could understand how the concepts worked and empathise themselves to the actual proposed usage situations during design. The UX goals were not as explicitly stated in the FSB case's Articles V and VI as in the ROS case's articles, but their utilisation still guided the design process of the different concepts, which were designed to have radical and futuristic elements. An external UX company did the evaluation of the concepts and the results of these evaluations were confidential. Therefore, the evaluations were not reported as part of this thesis.

In human-centred design, there is often the conception that designs must be carried out according to users' wishes, needs, and desires. However, based on the experiences from the future ship bridge design case, social science theories can contribute positively to design work by offering ways to go beyond the standard approaches to design (Article V). From the UX goal perspective, the designers should, in principle, begin by determining what kind of activity or emotion should be supported by the design; only after this should the product-related design ideas be generated (Article V).

This approach might also allow the designer to ‘think outside the box’, since the design process is not bound to existing products (Wahlström, Karvonen, & Kaasinen, 2014). For example, in the future ship bridge case, the CTA-derived systems usability goals, social science theories, and the feeling-related non-instrumental UX goals drew the design focus from the pre-existing products and related user wishes to potential future activity.

Also, in Article III, it was emphasised that in UX goal-based design, the set UX goals and their design implications should guide the HCD work in its different stages (e.g., for concept design, in 1) understanding the context, 2) specifying UX goals and other requirements, 3) producing concept designs, and 4) evaluating the concepts). The meaning of the UX goals and the amount of work devoted to them is emphasised in the early stages of design (i.e., concept/prototype design and evaluation). In these stages, the goals should be meticulously specified according to the gathered domain and user data and taken as ‘guiding stars’ for the design (Karvonen et al., 2012).

As this dissertation is focused on these early stages of design and evaluation, it does not include the stages where the final system fulfilling the UX goals is being developed. In the case of safety-critical work systems, the work would then continue to the integrated solution design stage, where the evaluation would focus on the usage of the technology and the acceptability of the new tool from the end users' point of view (Norros et al., 2015). For example, the systems usability indicators could in this stage be used in the full-scale assessment of the quality of the final design and to support the fulfilment of the core-task functions in the intended work (Savioja, 2014). Human factors verification and validation (see, e.g., NUREG-0711, O’Hara et al., 2012) of the safety-critical system from the systems usability perspective would also be essential in different later phases of the development work.

6.1.5 Review of the methodological implications and suggestions

A key part of this dissertation has been to investigate the methodological implications of the usage of UX goals in human-centred design of safety-critical systems. The investigation has been based on the set research questions in Chapter 3. Firstly, it has been addressed why experience design is relevant for safety-critical technologies. In short, experience design is important as UX also contributes to the overall safety, efficiency, and meaningfulness of these systems. Second, the contribution of different analysis methods of user activity to the identification of UX goals for safety-critical systems has been presented. The analysis of the field studies with the Core-Task Analysis method was found to be a suitable approach for identifying potential UX goals for safety-critical systems design, particularly in work contexts. Third, empirical examples of how user experience goals could be systematically specified, utilised, and evaluated in different safety-critical technology design cases has been presented. In the empirical cases, the UX goals were seen as an approach to make HCD and experience design more focused, structured, and systematic.

Based on the results of the empirical investigations, setting the right UX goals can provide the designers appropriate empathic understanding of the users of the system. In order for the design team (and potentially the users in user studies) to understand the specified UX goals in-depth, they can be storified to scenarios and visualised. Setting the right UX goals in the beginning stage of design is the basis for a good UX to develop with the system in its actual usage. When the final UX goals are specified, they should be further interpreted with regard to their concrete design implications.

After setting the goals and defining their design implications in the target context, the concept idea generation may take place in different types of co-design workshops where new concept ideas are produced based on the understanding from the previous stages (see the ROS and FSB cases). Users may also be part of this design activity, as suggested, for instance, in participatory design (see, e.g., Schuler & Namioka, 2017) or collaborative design (see, e.g., Détienne, 2006) approaches.

The produced concepts should demonstrate the concrete benefits of the proposed solutions and desired experiences for the users. The concept solutions also need to meet the requirements defined in the earlier stage. In practice, these concept solutions should then be trackable back to the set UX goals. The end result can also be a new concept of operations (see, e.g., Fairley & Thayer, 1997). To demonstrate a ConOps, the outcomes of this stage may include, for example, low-fidelity sketches, scenario stories, and mock-ups of different granularity levels.

After the production of the design solutions, they need to be evaluated. It is in this stage, at the very latest, that the potential users need to be involved. The evaluations can be conducted, for example, with user focus group interviews, which are useful for recognising the weaknesses and strengths of the proposed concepts in order to choose the best concepts for further development. On the other hand, individual in-depth user interviews can provide insights especially on how to enhance some specific concept idea and in ensuring that the proposed solutions work in the planned context of use.

Based on the evaluation results, the final concept may be created. The concept provides the vision of the future activity with the proposed technical system. The final concept can be, for instance, visualised in pictures for different stakeholders. A pictorial visualisation works as a concrete and easy-to-grasp boundary object in sharing the idea behind the concept. In this way, further structured feedback for the development of the actual system can also be gathered from stakeholders. More interactivity in the final concept can be added with an interactive prototype system. For example, in the ROS case, a virtual camera view-based prototype was developed for this purpose.

In general, UX design activities with UX goals should also be integrated into the later stages of system development, for example, by combining them with agile software development methods. In addition, after the implementation of the system, UX factors and feedback should be systematically analysed. However, as the scope of this dissertation is in the early stages of human-centred design, the details of these later stages are not discussed here.

The stages mentioned above are also in line with the UX-driven concept design process that is presented as a simplified version in a booklet (Roto, Nuutinen, & Smedlund, 2014) published from the program where the ROS and FSB cases were conducted. A modified version of the stages of the process (see Roto et al., 2014) for the purposes of this dissertation (especially for early-stage design) is provided below:

1. Gaining insight
    - Conducting, for example, user studies to gain insights into what kind of UX delights the target audience. If the system aims for radical or futuristic usage, studying suitable trends is also required.
2. Identifying and setting UX goals
    - Identifying and setting, for instance, high-level UX vision and goals based on the user studies and other gathered insight.
3. Defining design implications
    - Defining the UX goals' implications for design in the specific context in question.
4. Undergoing concept design
    - Utilising UX goals in the ideation and creation of new concepts and/or prototypes.
5. Evaluating
    - Evaluating both the produced designs and the fulfilment of the UX goals based on user evaluation study results.

The above approach is well in line with the activities of the HCD process defined in the ISO 9241-210 (2010) standard. Stages 4 and 5 should be iterated until a final concept fulfilling user needs and the UX goals is found. An alternative view is that even though the set UX goals would not be reached, they can still be seen as a beneficial approach as they have guided the design process in the right direction. In the evaluation phase, the UX goals themselves can also be validated concerning whether they are the experiences that the users want. This type of evaluation in professional contexts requires that the users in evaluations are as close as possible to the potential real users (e.g., professionals in a certain work field).

Based on the above contemplations and results of this thesis, the UX goal utilisation process can also be analysed from the perspective of integrating it to the human-centred design process of the ISO 9241-210 (2010) standard. Figure 11 depicts this HCD process from the utilisation of UX goals perspective under the scenario that a concept or prototype design would be the focus of the design work, as it was in the ROS and FSB cases of this dissertation. The adapted version presents the modified activities of the ISO 9241-210 (2010) process with sharp-cornered rectangles in imperative mood. Based on the results of the empirical research in this thesis, there are key points added (with round-cornered rectangles) related to the process' different UX goal-related activities.

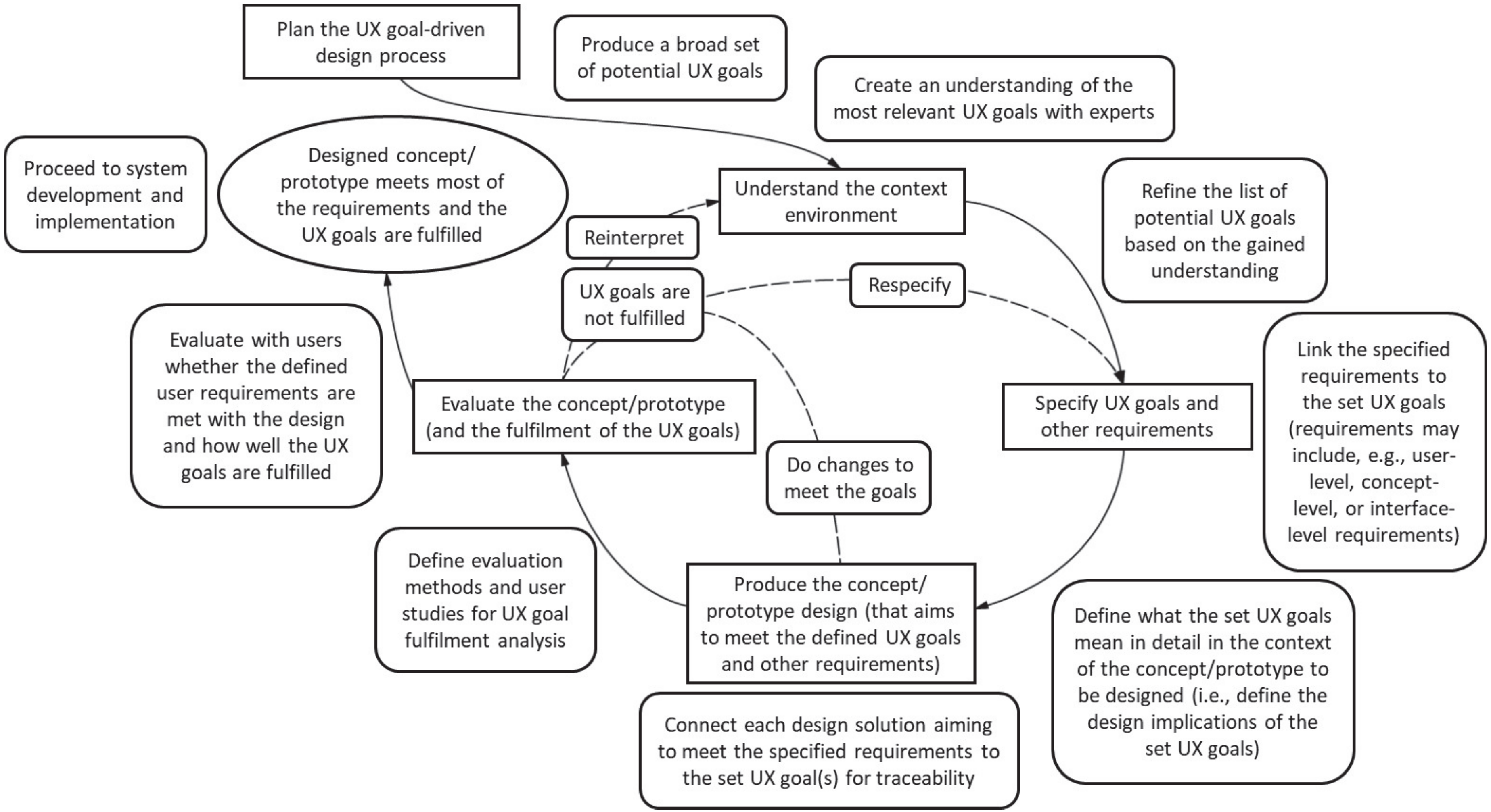


FIGURE 11 The ISO 9241-210 (2010) process modified from the perspective of UX goals' utilisation in concept/prototype design

In addition to these concretisations in Figure 11, guidelines for the usage of UX goals in supporting the process of human-centred design of safety-critical systems are provided in Appendix A. These guidelines are mostly based on the results of the previous academic workshops regarding UX goals (see Kaasinen et al., 2015; Väätäjä et al., 2012), elaborated publications about the results of the second workshop (Varsaluoma et al., 2015b; Varsaluoma, Väätäjä, Kaasinen, Karvonen, & Lu, 2015a), and results gained from the empirical research of this thesis. The list of guidelines in Appendix A is purposefully short, as it is meant for practitioner usage in human-centred design. The guidelines are also aligned with the methodological implications and suggestions presented in this section.

Finally, based on the empirical research cases, it can also be said that one contribution of the UX goals is the advancement of ethical design practices. As UX goals focus on the experiences of human users, they also advance awareness and improved empathy of different stakeholders about the affecting factors (e.g., work conditions) related to the users and aim to improve the factors contributing to this experience. Therefore, real human experiences can be affected positively by using UX goals in the design and development processes of systems, products, and services.

As can be seen from Figure 11, UX goals can metaphorically 'bring flesh' to the general and abstract nature of HCD guidelines, approaches, and processes, such as the ISO 9241-210:2010 standard. Typically, these types of standards are therefore more like a 'skeleton', which provides loose guidance to the actual design work. To increase the practical relevance of the results of this dissertation, in the next section, the work is discussed from the practical implications point of view.

6.2 Practical implications

Compared to entertainment technology for consumers, the requirements for complex safety-critical technical systems are understandably different in many respects. With a mass-market consumer product meant for leisure, a system failure usually only causes nuisance to the user. With a safety-critical system, a failure with the technology can result in the loss of human lives. Therefore, safety-critical technology is typically governed with strict regulations, standards, and other constraints. Moreover, the development of such technology usually requires tightly controlled and documented processes.

Sometimes the multitude of safety requirements and standards may also end up producing cumbersome solutions from the usage point of view. In design, it should be taken into account that the strict requirements do not mean that the users would necessarily appreciate the possible hindering effects these constraints have on their daily lives nor accept poor-quality user experiences in the usage of these systems. On the contrary, if an appropriate level of user engagement with safety-critical systems is aimed at, the UX perspective has to be thoroughly considered in the design activities of this technology.

Users are undoubtedly more engaged with a system that they like using. For the quality and level of user experience of safety-critical systems to keep up with that of mass-market consumer products, novel approaches are required. As a one suitable tool, UX goals have been suggested in this dissertation to also work in the practical human-centred design of safety-critical systems. Nevertheless, the utilisation of different UX approaches in the design of safety-critical technologies is, in practice, always a compromise amongst other demands coming, for example, from safety requirements. Naturally, with safety-critical systems, the safety requirements should override other, less critical requirements.

6.2.1 Implications for the research and design of safety-critical systems

Traditionally, UX research and design has been seen to include a combination of user-centred activities, such as user research, definition of usage scenarios and personas, rapid prototyping, usability testing, and user interface and interaction design. The ultimate aim in these types of activities has been to understand user needs and designing products and services that are both pleasurable and easy to use (Wright & McCarthy, 2005), and that people using them are also highly motivated and willing to use the systems. Therefore, UX research and design has not only been about comprehending *what* people do, but also about *why* they do something. This understanding should include a conception of the users' tasks, goals, motivations, behaviours, and expectations.

Different human factors methods can support the development of this understanding. HF approaches, like task analyses or assessment methods (e.g., situation awareness or mental workload assessments, see Stanton et al., 2017), have traditionally been applied in the evaluation and design of safety-critical systems, such as nuclear power plant control rooms, air traffic control towers, and healthcare systems. These approaches can be seen as overly tedious and costly to be applied in modern-day technology design and development cycles. Conversely, UX approaches, such as rapid iterative user testing or personas, have been applied to web service development and in the design of consumer technology, such as smartphones. The UX methods are often seen as lightweight and innovative approaches with agile, iterative design cycles. Therefore, traditionally a clear division can be said to exist between the HF and UX disciplines and their design and development processes and methods.

UX research and design approaches' ability for quick adaptation is intriguing. For HFE research, it offers grounds to contemplate whether its engineering procedures are overly complex. Fast iterative design should not be utilised only in mass-market consumer products development. In addition, with safety-critical systems and their physical hardware, fast prototyping is nowadays possible and a cost-efficient approach for iterative development. Furthermore, mock-up and simulator environments allow for ideas to be tested and the potential of the proposed safety-critical tools to be evaluated with users without actual system implementations.

Undoubtedly, the technology that is the focus of HF research is designed for and experienced by real people with their various emotions. UX research allows practitioners to gain insights into the users' experiences, emotions, activity, and attitudes. Therefore, these two research disciplines and communities should aim for collaboration – to build bridges and share knowledge between each other.

The results of the research in this dissertation aim to increase understanding of the meaning of user experience in designing safety-critical systems for high-reliability environments in both HF and UX communities. With this understanding, the requirements definition, concept design, and evaluation practices of these systems can also be reflected from the standpoint of UX. This kind of knowledge can be seen to be useful for both academic researchers and HF practitioners or UX designers. For researchers, new insights on the design and evaluation of UX can be provided. For practitioners or designers, guidelines and heuristics can be defined to support the process of designing novel safety-critical technologies from UX point of view. One such list of guidelines for practitioners is provided in Appendix A of this dissertation (see also Väätäjä et al., 2012; Varsaluoma et al., 2015a).

Furthermore, particularly suitable UX goals with safety-critical systems can be identified based on the results of the empirical research in this dissertation and from other previous UX-oriented studies of safety-critical systems design. For example, 'a sense of control' (Article II, III, and VIII), 'an appropriate experience of safe operation' (Articles III and VIII), and 'an appropriate level of trust in technology' (Articles I, IV, and IX; [Kymalainen et al., 2017; Savioja & Norros, 2013]) can be identified as relevant goals that work across many different application domains with safety-critical systems. However, the meaning and design implications of these goals should obviously be defined domain- and application-specifically in order for them to support the design work of some individual system for a particular environment in an appropriate manner.

If these stated goals are considered in detail, regarding the sense of control, it was noted already in Section 5.1.2 that it can mean, for example, 'feeling confident, in command, and one step ahead'. This goal is also mentioned to be important with safety-critical work systems in the systems usability framework: the 'feeling of control' indicator is mentioned in the psychological tool function's 'User experience: the development potential of use' perspective on activity (see, e.g., Savioja, 2014). In detail, Savioja (2014, p. 109) states that 'The feeling of being in control is related to understanding and anticipating the dynamic nature of the controlled process, which is a prerequisite of resilience in the system' while referring to Hollnagel et al. (2011).

The 'appropriate experience of safe operation' UX goal can be partly justified, for example, with the following ISO 9241-11 (2018, p. 27) standard's statement about UX that was also mentioned in Chapter 2: 'User experience focuses on the user's preferences, attitudes, emotions and physical and psychological responses that occur before, during and after use (including perception of trust, safety, security, and privacy)'. Therefore, the perceived degree of safety can be seen to affect (e.g., inhibit) the use of a system. However, if the user has only a subjective experience of the safety of the operation, it does not directly mean that from the objective perspective the operation is on a sufficiently safe level. Typically, in work environments, the users are professionals who have an intuitive feeling about the safety based on their vast previous experience, but even that instinct can be sometimes compromised due to, for example, distractions or fatigue. Therefore, in this section, the UX goal is amended with the word 'appropriate' experience of safe operation. By this word, it is suggested that the feeling needs to be at an appropriate level considering the objective level of safety. The line of reasoning here is similar to the concept of 'appropriate trust' (see, e.g., Lee & See, 2004), which is next discussed as a potential UX goal for safety-critical systems in detail. The appropriate experience of safe operation can also be seen as a precondition for attaining an appropriate level of trust in technology.

As mentioned earlier, a basic human factors issue, especially with highly automated complex safety-critical technology, is whether people should trust them. The concept of trust in technology was discussed in detail in Section 2.1.3. In designing new safety-critical technology, it is essential to understand what trust in technology is and how to design to support the development of appropriate trust in technology (Lee & See, 2004) in order to make the technology in question acceptable and successful. In practice, appropriate trust means that the capabilities of the technology match the trust level of the user. The concept also bears similarities to the concept of 'rational trust in technology' discussed by Saariluoma, Karvonen, and Rousi (2019). To support the calibration of trust to an appropriate level in technology development, experience design can be seen as a suitable approach, as trust can be seen as an attitude that is highly affected by emotional experiences. Therefore, an appropriate level of trust in technology can also be seen as an important UX goal of safety-critical systems, especially as their automation level increases (see, e.g., Lee & See, 2004).

6.2.2 UX goals in practical design work

Regarding UX goals in practical design work, some key points for further consideration emerge from the conducted empirical design work with the different research and design teams. Firstly, in setting UX goals, the whole design team should be involved. In this way, it is easier for the team members to commit to the jointly set goals. Secondly, in specifying and sharing the goals, a general UX vision with a few UX goals works as a good approach for the design team to remember and understand them. Thirdly, as mentioned in Article VII, UX goals may sometimes be too abstract for designers to take as an input for the design work. Conversely, overly concrete interpretations about the goals and their implications may not leave room for ideation (Article VII). Nevertheless, even if there are no established guidelines regarding the level at which the UX goals should be set and in which way they should guide the design, the goals are still beneficial in maintaining an experience mindset in the design team (Article VII). In an optimal case, UX goals form the backbone of the entire human-centred design work in the team. In this scenario, using UX goals may become as important from the outcome perspective as the actual hands-on design work.

According to the empirical results, a suitable design orientation should be developed based on in-depth understanding of users and their activity. This understanding can be gained, for example, with field studies of the system's potential users. A Core-Task Analysis may help to distinguish the basic demands and aims of the work domain and the system's users (Norros, 2004). To accommodate for future usage, the approach should also take into account the application domain's contextual, technological, societal, and theoretical trends (see the FSB case and Articles V and VI).

From the empirical results, it can be interpreted that UX goals can also support the production of radical design ideas. This was noticed especially in the FSB case where the aim was to produce radical concept designs for future ship bridge operations. Notions regarding how UX goals help the emergence of radical designs are discussed subsequently based on the empirical results. Firstly, UX goals help in facilitating designs that would not have been thought of with a product- or technology-centric approach. Secondly, UX goals reflect the intended user emotions and in doing so, they do not inhibit radical design by directly dictating the design idea creation as some other, more dogmatic design approaches can. Thirdly, emotion-related non-instrumental UX goals draw the design focus from pre-existing products to users' potential future activity and in this way provide a view separate from existing and obvious solutions (Article VI).

The design processes of the ROS and FSB case also contributed to the development of a new commercial and practical-level concept design approach called InnoLeap (see, e.g., Wahlström et al., 2014). In summary, InnoLeap is a collection of design principles that facilitates the creation of radical concept design ideas for industrial work activity (Wahlström et al., 2014). Short descriptions of the approach for non-academic purposes can also be found online with the future ship bridge case as an example[13, 14]. In practice, the stages of the InnoLeap approach include similar design process stages as there was in the FSB case: 1) Trend insight, 2) Analysis of user activity, 3) Draft operation concepts, 4) Concept evaluation, 5) Creation of the final concepts, 6) Final concept visualisations, and 7) Concept release and media buzz. The details of these stages can be found in an online brochure about InnoLeap[15].

Ultimately, design is about problem solving and UX goals can help to frame the problem space of design towards good experiences. Therefore, concrete and focused UX goals are needed for good experience design. Without concretising which user experiences to aim at, relevant UXs can easily be forgotten in practical design work. This problem becomes especially evident with safety-critical systems, as the designers receive many requirements from different sources related

---

13 https://www.vttresearch.com/Documents/Augmented_Reality_Industrial_Reality_22March2018/Radical%20concept%20design%20with%20InnoLeap%20%E2%80%93%20Case%20Rolls-Royce%20Marine%20-%20Hannu%20Karvonen%20VTT.pdf (accessed 30th of May, 2019)

14 http://www.vtt.fi/innoleap (accessed 30th of May, 2019)

15 https://www.vttresearch.com/Documents/Business%20Tools/vtt_innoleap_A4.pdf (accessed 30th of May, 2019)

to, for example, technical restrictions, standards, regulations, safety, and marketing. To counterbalance these requirements, the UX goals should be taken as 'guiding stars' throughout the design process in order for the end-result to soundly support the desired user experiences.

6.3 Research validity and limitations

With the research methods and techniques presented in this dissertation, it has been possible to specify relevant UX factors related to the safety-critical systems presented in the empirical cases. For example, appropriate UX goals could be identified, prioritised, set, and defined for the design work and it was possible to describe their meaning and importance in detail in the case target environments and systems under design.

The validity of the approaches used in the empirical cases can be to some extent illustrated through the success of the design cases (see Sections 5.1.1 and 5.1.2). Another factor contributing to the validity of the presented approaches is that the cases were conducted in several different domains, such as port container crane and ship bridge operations. In practice, this diversity provides generalisability for the suggested approaches.

The empirical cases conducted in this dissertation were determined by the projects that were funded during the research work (see Acknowledgements). If the same methods and approaches would have been used in different empirical case contexts, such as the other ones described in Articles IV and VII in addition to the ROS case, the results could have also been very different.

The reliability of the empirical research cases in this thesis can be estimated, for example, through the number of participants in the user studies. As can be interpreted from Tables 3–6, in the MTD case, there were altogether 16 study participants; in the ROS case's work analysis study, altogether 12 participants; in the ROS case's prototype evaluation study, altogether six participants; in the FSB work analysis study, altogether 12 participants; and in the DDW evaluation study, altogether 31 participants.

Therefore, in total, 40 participants took part in this dissertation's work analysis studies. The validity of these exploratory and contextual naturalistic studies can be said to have been increased by the fact that all the participants in these work environment studies were real professional users (and not, for example, novice university students) who did the work that was studied on a daily basis. In addition, altogether 37 users participated in the more experimental type of prototype evaluation studies. In the ROS evaluation study, the participants were domain experts and in the DDW evaluation study, the participants were chosen so that they could be allocated to polarised experience background groups (novice/experts). Next, each case's studies' research validity and limitations are discussed in separate sections in detail.

6.3.1 MTD case

As mentioned in Article I (Section 4), the MTD case included some recognised limitations. Firstly, the participants for the studies in the case were chosen by the metro organisation and might have represented a biased sample of workers. Moreover, in the train driver observation study, there were only four participants. The reason for this low amount of participants is that convenience sampling was utilised. Convenience sampling has its justification especially in qualitative research in many academic publications (see, e.g., Emerson, 2015; Etikan, 2016; Marshall, 1996). Obviously, different kinds of results could have been acquired with a larger amount of participants.

In this connection, however, it is worth considering that the convenience sampling approach could be one way to increase the usage of the suggested design methods in organisations doing system development work. To accept different approaches with a scientific background as part of development work, the suggested methods need to be sufficiently lightweight to be integrated in, for example, agile development processes of modern technology companies. On the other hand, safety-critical systems are also typically developed in the industry with time-taking approaches, which can better allow the integration of in-depth UX goals specification, utilisation, and evaluation with the help of representative stakeholder participant and appropriate user sample sizes.

On a more general level, the attitudes and opinions of the drivers in the MTD case could have also been biased by the fact that the researchers were studying the effects of the planned metro automation. In the plans of the metro automation project at the time of conducting the studies, the automation would have considerably changed the work roles of the drivers in the future. Despite the fact that the metro organisation had promised the drivers that none of them would be let go, the drivers' attitude towards the automation project may have been overly critical. This may have been due to the assumption that the change to an automated system would have rendered the drivers' current occupation obsolete and therefore contradicted with the pride they take in their current work of driving the metro manually. The gained results in the MTD case, however, relied only moderately on the drivers' interviews and observations, and the potential bias would have probably affected only the drivers' comments regarding the metro automation, instead of the tasks and aims of the drivers that they conducted and commented on themselves in the studies.

6.3.2 ROS case

In the ROS case's contextual field studies (i.e., work analysis studies), the number of participants (12) may be considered appropriate when taking into account the exploratory nature of the studies. However, in the ROS evaluation study, the number of participants in the evaluations was small, only six participants. In planning the evaluations, convenience sampling was also utilised (as was in the MTD case). The reasons for this sampling with the ROS case's evaluation study were manifold. For example, the industrial partner company in the given timeframe could recruit to the study only six employees with the required experience. Clearly, with a larger number of participants, the results of the evaluations could have been different.

In the analysis of the evaluation study of the ROS case, the fulfilment of the UX goals was assessed based on the gathered empirical study data. A certain UX goal could be said to have been fulfilled if over half of the specified requirements connected to that goal received support for fulfilment in the evaluation study results (Article VIII). This 50% requirement proved to be reasonably suitable for these kinds of early-stage research evaluations, but in advanced verifications and validations of mature safety-critical systems, more rigorous criteria would be needed. The empirical research conducted as part of this thesis was limited to early-stage evaluations. Therefore, answering how this evaluation analysis should be done in later product development phases is not included in the scope of this thesis, but is a topic for further research (see Section 6.4).

As mentioned in Article VIII, the level of ecological validity of the developed prototype system in the ROS case may have affected the fulfilment of the UX goals in the evaluations. The 'experience of safe operation' and 'feeling of presence' goals were not supported with the prototype system in the evaluations. This fact was partly affected by the context and technical maturity level of the utilised prototype in the evaluations: in the prototype version, the operational tasks were conducted in a virtual world where no human lives were in danger and the presented camera views were not real ones (Article VII). Undoubtedly, the UX of the study participants was affected by the fact that the operations with the prototype were not happening with a real crane and real people in the operational environment, as they would have been with an implemented version of the ROS in a port environment. In the ROS prototype, the virtual camera views provided a virtual-world outlook to the loading and unloading area of the trucks. If the people seen in the operational environment (the truck loading and unloading area) had been real human beings, the participants could have been more cautious and their experiences could have been different with regard to the operations (Article VIII). Therefore, to gain a more realistic picture about the fulfilment of the UX goals, a high-fidelity system with real camera views and an actual remotely operated crane would have been needed.

On the other hand, as discussed in Chapter 2, early-stage human-technology interaction design can benefit if professional users are allowed to experience even a preliminary representation of the proposed technology. Professionals, with their vast prior experience, can evaluate the potential of the tool to develop into a meaningful solution for their daily activity (Savioja, Liinasuo, et al., 2014).

In addition to the studies presented here, the author of this thesis visited the Lamong Bay Terminal in Surabaya, Indonesia where the developed final version of the ROS system by the industrial partner company was for the first time implemented within a real port environment. The aim of this visit was to conduct remote operators' UX evaluations in their control room after the operations in the port had been running for several months. The results of these evaluations are company confidential information of the industrial partner company in the ROS case. It can, however, be said in this connection that the operations were running in the terminal very well and the developed remote operator station was experienced to be of good quality by the local remote crane operators.

6.3.3 FSB case

One clear limitation in the FSB case was that only static scenario stories, personas, storyboards, cartoon-like concept pictures, and animated concept videos were the aim of the research and design task in the associated project instead of, for example, interactive prototypes that the potential users could freely explore. Therefore, the produced design outcomes functioned as artefacts that aimed to concretise the set UX goals in the case. Naturally, these artefacts offered more limited possibilities to study the end-user UXs compared, for example, to the interactive prototype that was developed in the ROS case. Undoubtedly, more in-depth experiences may be experienced when interacting with the functionalities of a system and realistic feedback is provided by the system.

Nevertheless, in this context the participants who took part in the different evaluation sessions (that were conducted by other experts than the researchers and designers of the FSB research and design case) could easily empathise particularly with the descriptions of the personas and scenario stories that were produced in the case. The bridge crew members in the evaluations, who did their work on ships on a daily basis, were especially able (according to their own comments) to step into the described persona's role very well and imagine the suggested scenario situations in their mind based on their previous experience of similar situations in the real world.

6.3.4 DDW case

In the DDW case, the sample size of the participants in the evaluation study was 31. This number is quite low compared to typical factor analysis sample sizes. The justification for this number of participants can be enhanced by the design stage where the evaluation was conducted. In this early stage of prototype development, it was more important to explore what the elements of the evaluation measurements are that will be considered later in product development and evaluations when the design matures.

For example, with the conducted analysis in the case, the correlations of key experiential factors could be examined in this context. The clusters of experience elements (i.e., the identified factors) found in the case could be used later in the design and evaluation of other similar safety-critical solutions. Additionally, this same set of questions and experience factors could also be utilised as a part of later design stages of the driver distraction warning application or some other safety-critical system with a confirmatory factor analysis (CFA) approach.

In this way, it could be affirmed and verified that the design is on the right track and whether the specified UX factors/goals are yet fulfilled. In addition, the relevance of the defined experiential elements in different stages of design could be analysed with this process. On the other hand, this would allow the study of acceptability of the system in later design stages, all the way up to the final human factors validation and verification tests. In the tests with the final version of the system, a larger sample size of participants would be needed. The general conclusion here is that EFA with a relatively low number of participants is reasonably suitable for the analysis of relevant experiential factors in the early-stage design work, which was also the stage of work in the DDW case.

In future evaluations of the VisGuard system, a more representative group of the driver population would also be needed. In the study reported here, there were people of all age groups involved and the experience level of the participant groups varied considerably (clearly novice vs. experienced drivers). However, in reality, there are many people that fall into categories in between these different groups defined in the study.

The use of participants with certain, potentially non-representative backgrounds of the general population in this study can be to some extent justified with the following argument. In pilot testing, the earlier mentioned convenience sampling (Emerson, 2015; Etikan, 2016; Marshall, 1996) is often the case. There can be certain constraints in participant acquisition for studies with humans in real-world cases. In the DDW case, the participant sampling was done in such a way that the procedure was still convenient for the researchers, considering the constraints present in planning and conducting the study.

6.3.5 Review of research validity and limitations

The practical validity, usefulness, and relevance of the design approaches used particularly in the ROS and FSB cases is increased by the fact that they were conducted together with industrial company partners. The developed concepts, solutions, and systems were published, further developed, sold to customers, and brought into use by the companies (see Sections 5.1.1 and 5.1.2). For example, as mentioned in Section 5.1.1, the design results of the ROS case were utilised in the development and implementation of the final remote operator station system by the industrial partner company of the case.

As discussed previously, the results of this dissertation have both methodological and practical implications. From the methodological perspective (see Section 6.1), the results provide new insights into combining user experience and human factors research and design methods in safety-critical systems' early-phase development. From a practitioner point of view (see Section 6.2), a core result is a UX goal-based early-stage design approach and guidelines to be considered in the human-centred design of safety-critical systems.

One clear limitation in the empirical cases of this dissertation was that the ROS case was the only one where all the HCD activities (as suggested in this thesis) from 'Understanding the context' to 'Evaluating the concepts' were examined. The other cases included only one to three of the suggested activities (see Table 2 for details). Although the results of the other cases support the ROS case's results regarding the use of UX goals and consideration of UX factors in safety-critical environments, more cases that go through all the activities of the process would still be needed to increase the validity of the results regarding UX goals' use in human-centred design of safety-critical systems. Furthermore, the use of UX goals in the system development and implementation stages after the concept and prototype have been produced needs to be focused on in further studies.

Finally, in the research conducted for this dissertation, it could not be analysed whether UX goals made the design outcomes ultimately better, since it was not possible to run a controlled experimental study comparing design processes with and without the UX goals (Article VII). In practically oriented research and design projects with industry, it is very rarely possible to do this type of parallel research work to compare the pros and cons of the implemented approaches, methods, or techniques. However, in academia, these types of studies may be possible, at least on the micro scale. Nevertheless, the factors affecting a certain design process can be so complex that it would be nearly impossible to find out clearly what factual effect the UX goals had on the final designs.

6.4 Recommendations for future research

The research in this dissertation contributes to the theory-informed and practice-oriented knowledge about experience design and the utilisation of UX goals in human-centred design of safety-critical systems. In order to inspire researchers and designers analysing and conducting user research and design work, the thesis proposes new ways to incorporate UX goals as part of the design process. Documenting this research in the dissertation helps other researchers and designers to conduct similar studies with UX goals in the future.

As mentioned in Section 2.3.1, the aim with experience design approaches is to 'think experience before product' (Hassenzahl, 2010). Therefore, an ideal would be that the desired user experience is defined before the product design begins, but this is rarely possible in practice due to different development constraints. The specification of the appropriate user experiences to aim for, especially in human-centred design of safety-critical systems, may include different activities (e.g., literature reviews, user studies, workshops, and evaluations) and can take a lot of time, as was the case in the ROS and FSB cases of this thesis.

The research in the cases of this dissertation could be conducted to such a profound extent because the cases were conducted as part of a multi-year applied research project. In practical design work with companies, it is rare to have many years to conduct background research and, for example, specify appropriate UX goals and their meaning. Therefore, more agile and lean ways for this type of specification and utilisation of UX goals are needed to be studied in future research. In addition, the integration of the UX goals technique to modern design thinking (Brown, 2008; Buchanan, 1992) approaches would be beneficial to help practical design work as part of different organisations' product and service development processes.

More research is also needed on what are appropriate UX goals especially for safety-critical technologies. In this dissertation, initial recommendations of potential, more general-level cross-domain UX goals for safety-critical systems based on the empirical studies could be discussed in Section 6.2.1. However, to increase the validity of these recommendations, both conceptual research and practical design work will be needed in the future.

Furthermore, the integration of the UX goals technique to existing safety standards that require risk-based approaches to validation and verification is a topic for future research. For example, it would be relevant to study how UX goals and the presented Usability Case approach could be integrated into IEC61508 (in the industrial control domain) or ISO 26262 (in the automotive domain). In this way, it could be ensured that the UX goals do not lead to interfaces that support some specific positive user experiences in general, but lead to significant risk in certain situations of the domain in question. Integrating UX goals with the systems usability framework with the Systems Usability Case approach (see, e.g., Laarni et al., 2014; Norros et al., 2015) is one potential way forward in this regard as well.

One limitation regarding the design work conducted in the ROS and FSB cases as part of this thesis was that the real potential users were not present in the brainstorming and design workshops where the concepts were produced. Future work could focus on how users could be more involved in the production of designs with the help of UX goals. For instance, combining participatory design (e.g., Schuler & Namioka, 2017) with users for the specification and utilisation of UX goals in design cases would be a one topic for further research.

Furthermore, the academic advancement of the commercial concept design approach (i.e., InnoLeap) that was developed based on the ROS and FSB cases will be important in the future. A good starting point for this work would be another research-oriented study where the theoretical and conceptual basis of the approach could be studied in detail. One possibility would also be to combine the InnoLeap approach to a more academically-oriented design approach, such as Core-Task Design (see, e.g., Norros et al., 2015), and study how concept design could be improved.

The DDW case's main meaning in this dissertation was to demonstrate ways to evaluate the UX-related factors with statistical analyses. Similar approaches could also be applied to the evaluation of the fulfilment of UX goals in future studies. These types of statistically-oriented studies could provide robust evidence for argumentation in a Safety/Usability Case (Bishop & Bloomfield, 2000; Liinasuo & Norros, 2007) type of evaluation procedure of UX goal fulfilment, which was utilised in the ROS case's evaluation study.

One lesson learnt from the ROS case's validity and limitations (see Section 6.3.2) was that a more realistic picture about the fulfilment of the UX goals could be gained with a higher fidelity-level system. Clearly, more applied research is needed in the utilisation of UX goals beyond the early stages of design. In this thesis, the studies were conducted until the end of the concept and prototype stages of the produced designs. Therefore, the empirical research in this thesis clearly focused on the top-level activities of the systems engineering V-model (see Section 2.1.2 and Figure 1). Research on how UX goals could be used in the other activities of safety-critical systems development, for instance, according to the systems engineering V-model process (see Figure 1), are therefore needed.

Connected to this topic is the question of what the appropriate level of fidelity of concept solutions is for early-stage evaluation. As described in the empirical research of this thesis, early-stage evaluations can be conducted with scenario stories, personas, storyboards, concept pictures, and interactive prototypes of different levels of realism. Therefore, the way to conduct analytical evaluation also depends on the level of design solutions that can be produced for the evaluation. In addition, the evaluation should be conducted iteratively throughout the design project. Research on how this iterative evaluation and its results should be systematically taken into account in safety-critical systems development is clearly needed.

Additionally, if UX goals are systematically set and evaluated with particular measures in a certain context, it should be possible to gather a 'user information data bank' about the accumulated knowledge regarding the use and evaluation of the UX goals. In this way, for example, the gathered information could provide indicators about how (i.e., with what kinds of design solutions) to achieve a certain level of fulfilment of a specified UX goal in design. Therefore, predictability could also be brought to the design process. For organisations doing design work with UX goals, this would clearly be one topic for further research.

On a general level, more research is needed in bridging the gap between systems engineering and UX research and design of safety-critical technologies. One key issue would be to integrate UX goals-oriented design practices in a more in-depth manner into the human factors engineering process (see, e.g., Lee et al., 2017) that is already conducted as part of safety-critical systems development. The results and implications of the thesis aim to provide the groundwork for addressing this and the other suggested recommendations for further research in the future.

REFERENCES

Alben, L. (1996). Quality of experience: defining the criteria for effective interaction design. *Interactions*, *3*(3), 11–15. https://doi.org/https://doi.org/10.1145/235008.235010

Arvola, M., & Holmlid, S. (2015). User Experience Qualities and the Use Quality Prism. In E. Kaasinen, H. Karvonen, Y. Lu, J. Varsaluoma, & H. Väätäjä (Eds.), *The Fuzzy Front End of Experience Design - Workshop Proceedings*. Retrieved from http://www.vtt.fi/inf/pdf/technology/2015/T209.pdf

Baier, C., & Katoen, J.-P. (2008). *Principles of Model Checking*. Massachusetts, MA: MIT Press.

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8

Battarbee, K. (2004). *Co-Experience - Understanding User Experiences in Social Interaction* (Doctoral dissertation, University of Art and Design Helsinki). Retrieved from https://aaltodoc.aalto.fi/bitstream/handle/123456789/11462/isbn9515581613.pdf?sequence=1

Besnard, D., & Hollnagel, E. (2014). I want to believe: some myths about the management of industrial safety. *Cognition, Technology & Work*, *16*(1), 13–23. https://doi.org/10.1007/s10111-012-0237-4

Beuscart-Zéphir, M. C., Borycki, E., Carayon, P., Jaspers, M. W. M., & Pelayo, S. (2013). Evolution of Human Factors Research and Studies of Health Information Technologies: The Role of Patient Safety. *Yearbook of Medical Informatics*, *22*(1), 67–77. https://doi.org/10.1055/s-0038-1638835

Bilimoria, K. D., Johnson, W. W., & Schutte, P. C. (2014). Conceptual framework for single pilot operations. In *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI-Aero '14* (pp. 1–8). ACM Press.

Bishop, P., & Bloomfield, R. (2000). A methodology for safety case development. *Safety and Reliability*, *20*(1), 34–42. https://doi.org/https://doi.org/10.1080/09617353.2000.11690698

Blanchard, B. S., & Fabrycky, W. J. (2006). *Systems Engineering and Analysis*. Upper Saddle River, NJ: Pearson Prentice Hall.

Bødker, S. (2006). When second wave HCI meets third wave challenges. In *Proceedings of the 4th Nordic conference on Human-computer interaction - NordiCHI '06* (pp. 1–8). ACM Press.

Bowen, J. P., & Hinchey, M. G. (1999). *High-Integrity System Specification and Design*. London, UK: Springer.

Boy, G. A. (2013). *Orchestrating Human-Centered Design*. London, UK: Springer.

Bozzano, M., & Villafiorita, A. (2010). *Design and Safety Assessment of Critical Systems*. New York, NY: Auerbach Publications.

Braseth, A. O., & Øritsland, T. A. (2013). Visualizing complex processes on large screen displays: Design principles based on the Information Rich Design concept. *Displays*, *34*(3), 215–222. https://doi.org/10.1016/j.displa.2013.05.002

Brown, T. (2008). Design thinking. *Harvard Business Review*, *86*(6), 84–92. Retrieved from https://hbr.org/2008/06/design-thinking

Buchanan, R. (1992). Wicked problems in design thinking. *Design Science*, *8*(2), 5–21. https://doi.org/10.1017/dsj.2015.1

Burns, C. M. (2004). Ecological Interface Design. In *Ecological Interface Design*. Boca Raton, FL: CRC Press.

Cabon, P., Coblentz, A., Mollard, R., & Fouillot, J. (1993). Human vigilance in railway and long-haul flight operation. *Ergonomics*, *36*(9), 1019–1033. https://doi.org/10.1080/00140139308967974

Cain, J. (1998). Experience-Based Design: Toward a Science of Artful Business Innovation. *Design Management Journal (Former Series)*, *9*(4), 10–16. https://doi.org/https://doi.org/10.1111/j.1948-7169.1998.tb00223.x

Carrol, J. (1997). Human–computer interaction: psychology as a science of design. *International Journal of Human-Computer Studies*, *46*(4), 501–522. https://doi.org/10.1006/ijhc.1996.0101

Carroll, J. (2000). Five reasons for scenario-based design. *Interacting with Computers*, *13*(1), 43–60. https://doi.org/10.1016/S0953-5438(00)00023-0

Carroll, J., & Haynes, S. (2017). Scenario-Based Design. In G. A. Boy (Ed.), *The Handbook of Human-Machine Interaction* (pp. 153–164). Boca Raton, FL: CRC Press.

Clark, J. O. (2009). System of systems engineering and family of systems engineering from a standards, V-model, and dual-V model perspective. In *3rd Annual IEEE Systems Conference* (pp. 381–387). IEEE.

Coyle, R. (1997). System Dynamics Modelling: A Practical Approach. *Journal of the Operational Research Society*, *48*(5), 544–544. https://doi.org/10.1038/sj.jors.2600682

Cruickshank, L., & Trivedi, N. (2017). Beyond Human-Centred Design: Supporting a New Materiality in the Internet of Things, or How to Design When a Toaster is One of Your Users. *The Design Journal*, *20*(5), 561–576. https://doi.org/10.1080/14606925.2017.1349381

Cummings, M. L., Mastracchio, C., Thornburg, K. M., & Mkrtchyan, A. (2013). Boredom and Distraction in Multiple Unmanned Vehicle Supervisory Control. *Interacting with Computers*, *25*(1), 34–47. https://doi.org/10.1093/iwc/iws011

de Mol, B. (2007). Designing medical technology. *Safety Science*, *45*(1–2), 283–291. https://doi.org/10.1016/j.ssci.2006.08.003

Desmet, P. (2011). *From Floating Wheelchairs to Mobile Car Parks: Selected Work from TU Delft: a Collection of 35 Experience-driven Design Projects*. The Hague, The Netherlands: Eleven International Publishing.

Desmet, P., & Hekkert, P. (2007). Framework of product experience. *International Journal of Design*, *1*(1), 57–66. Retrieved from http://www.ijdesign.org/index.php/IJDesign/article/view/66

Desmet, P., & Pohlmeyer, A. (2013). Positive design: An introduction to design for subjective well-being. *International Journal of Design*, *7*(3), 5–19. Retrieved from http://www.ijdesign.org/index.php/IJDesign/article/view/1666

Desmet, P., & Schifferstein, R. (2011). *From floating wheelchairs to mobile car parks*. The Hague, The Netherlands: Eleven International Publishing.

Détienne, F. (2006). Collaborative design: Managing task interdependencies and multiple perspectives. *Interacting with Computers*, *18*(1), 1–20. https://doi.org/10.1016/j.intcom.2005.05.001

Dewey, J. (1925). *Experience and Nature*. La Salle, IL.: Open Court.

Emerson, R. (2015). Convenience Sampling, Random Sampling, and Snowball Sampling: How Does Sampling Affect the Validity of Research? *Journal of Visual Impairment and Blindness*, *109*(2), 164–168. https://doi.org/doi.org/10.1177%2F0145482X1510900215

Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems: Situation awareness. *Human Factors*, *37*(1), 32–64. https://doi.org/doi.org/10.1518%2F001872095779049543

Endsley, M. (1996). Automation and Situation Awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance - Theory and Applications* (pp. 163–181). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Endsley, M., & Kiris, E. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(2), 381–394. https://doi.org/10.1518/001872095779064555

Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R.-L. Punamäki (Eds.), *Perspectives on activity theory* (pp. 19–38). New York, NY: Cambridge University Press.

Engeström, Y. (2011). From design experiments to formative interventions. *Theory & Psychology*, *21*(5), 598–628. https://doi.org/10.1177/0959354311419252

Engeström, Y. (2014). *Learning by Expanding*. New York, NY: Cambridge University Press.

Ericson, C. (1999). Fault tree analysis. In *Proceedings of the 17th International System Safety Conference* (pp. 1–9). System Safety Society.

Ericson, C. (2015). *Hazard analysis techniques for system safety*. Hoboken, NJ: John Wiley & Sons.

Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1. https://doi.org/10.11648/j.ajtas.20160501.11

European Committee for Standardization (CEN). (2016). *EN 16170-2, Ergonomics methods - Part 2: A methodology for work analysis to support design*.

Fairley, R., & Thayer, R. (1997). The concept of operations: The bridge from operational requirements to technical specifications. *Annals of Software Engineering*, *3*(1), 417–432. https://doi.org/https://doi.org/10.1023/A:1018985904689

Falzon, P. (2008). Enabling safety: issues in design and continuous design. *Cognition, Technology & Work*, *10*(1), 7–14. https://doi.org/10.1007/s10111-007-0072-1

Flach, J. M. (2009). The dynamics of experience: a search for what matters. In *Proceedings of the European Conference on Cognitive Ergonomics 2009 (ECCE 2009)* (pp. 11–18). VTT.

Flach, J. M., Stappers, P. J., & Voorhorst, F. A. (2017). Beyond Affordances: Closing the Generalization Gap Between Design and Cognitive Science. *Design Issues*, *33*(1), 76–89. https://doi.org/https://doi.org/10.1162/DESI_a_00427

Flach, J. M., & Voorhorst, F. (2016). *What matters?: Putting common sense to work*. Dayton, Ohio: Wright State University Libraries.

Fleming, J., & Koman, R. (1998). *Web navigation: designing the user experience*. Sebastopol, CA: O'Reilly.

Forlizzi, J., & Ford, S. (2000). The building blocks of experience: an early framework for interaction designers. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 419–423). ACM.

Forsberg, K., & Mooz, H. (1992). The Relationship of Systems Engineering to the Project Cycle. *Engineering Management Journal*, *4*(3), 36–43. https://doi.org/https://doi.org/10.1080/10429247.1992.11414684

Forsberg, K., Mooz, H., & Cotterman, H. (2000). *Visualizing project management: a model for business and technical sucess*. New York, NY: John Wiley and Sons.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Glendon, A., & Stanton, N. (2000). Perspectives on safety culture. *Safety Science*, *34*(1–3), 193–214. https://doi.org/10.1016/S0925-7535(00)00013-8

Goldberg, K. (2012, March). What Is Automation? [Editor's Corner]. *IEEE Robotics & Automation Magazine*, *19*(1), 101–102. https://doi.org/10.1109/MRA.2011.2181783

Gulland, W. G. (2004). Methods of determining safety integrity level (sil) requirements-pros and cons. In *Practical Elements of Safety: Proceedings of the Twelfth Safety-critical Systems Symposium* (pp. 105–122). Springer.

Hartson, R., & Pyla, P. (2012). *The UX Book: Process and Guidelines for Ensuring a Quality User Experience*. Waltham, MA: Elsevier.

Hassenzahl, M. (2008). User experience (UX): towards an experiential perspective on product quality. In *Proceedings of the 20th Conference on L'Interaction Homme-Machine - IHM '08* (pp. 11–15). ACM.

Hassenzahl, M. (2010). Experience Design: Technology for All the Right Reasons. In *Synthesis Lectures on Human-Centered Informatics 3(1)*.

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, *25*(2), 91–97. https://doi.org/10.1080/01449290500330331

Hawkins, F. (2017). *Human factors in flight*. New York, NY: Routledge.

Hekkert, P. (2006). Design aesthetics: principles of pleasure in design. *Psychology Science*, *48*(2), 157–172. Retrieved from https://www.persistent-identifier.nl/urn:nbn:nl:ui:24-uuid:4e0b8051-5436-4e12-a24f-43ce8a58d8b2

Hofmann, D. A., Jacobs, R., & Landy, F. (1995). High reliability process industries: Individual, micro, and macro organizational influences on safety performance. *Journal of Safety Research*, *26*(3), 131–149. https://doi.org/10.1016/0022-4375(95)00011-E

Hollnagel, E., Paries, J., Woods, D., & Wreathall, J. (2011). *Resilience Engineering in Practice: A Guidebook*. Farnham, UK: Ashgate.

Hollnagel, E., Woods, D., & Leveson, N. (2012). *Resilience engineering: Concepts and precepts*. Padstow, UK: Ashgate.

Holtzblatt, K., & Beyer, H. (2016). *Contextual design: Design for life*. Cambridge, MA: Morgan Kaufmann.

IEC. (2001). *IEC 61513, Nuclear power plants – Instrumentation and control for systems important to safety – General requirements for systems*. International Electrotechnical Commission.

IEC. (2010). *IEC 61508. Functional safety of electrical/electronic/programmable electronic safety-related systems*. International Electrotechnical Commission.

ISO. (2010). *ISO 9241-210:2010, Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems*. International Organisation for Standardisation.

ISO. (2011). *ISO 26262, Road vehicles — Functional safety*. International Organisation for Standardisation.

ISO. (2018). *ISO 9241-11:2018, Ergonomics of human-system interaction - Part 11: Usability: Definitions and Concepts*. International Organisation for Standardisation.

Itoh, M. (2012). Toward overtrust-free advanced driver assistance systems. *Cognition, Technology & Work*, *14*(1), 51–60. https://doi.org/10.1007/s10111-011-0195-2

Itoh, M., & Tanaka, K. (2000). Mathematical Modeling of Trust in Automation: Trust, Distrust, and Mistrust. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *44*(1), 9–12. https://doi.org/10.1177/154193120004400103

Jarke, M. (1998). Requirements tracing. *Communications of the ACM*, *41*(12), 32–36. https://doi.org/10.1145/290133.290145

Jokinen, J. P. P. (2015). *User psychology of emotional user experience* (Doctoral dissertation, University of Jyväskylä). Retrieved from http://urn.fi/URN:ISBN:978-951-39-6225-8

Joutsela, M., & Roto, V. (2016). Introducing Experience Goals into Packaging Design. In *Proceedings of DRS 2016, Design Research Society 50th Anniversary Conference*. DRS Design Research Society.

Kaasinen, E., Karvonen, H., Lu, Y., Varsaluoma, J., & Väätäjä, H. (2015). *The fuzzy front end of experience design: Workshop proceedings*. Retrieved from http://www.vtt.fi/inf/pdf/technology/2015/T209.pdf

Karvonen, H., & Aaltonen, I. (2017). Intent communication of highly autonomous robots. In *Workshop on "The Role of Intentions in Human-Robot Interaction", The 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2017)*. Retrieved from http://intentions.xyz/wp-content/uploads/2017/01/Karvunen.Aaltonen.Intent-communication-of-highly-autonomous-robots.pdf

Karvonen, H., Heikkilä, E., & Wahlström, M. (2019). Artificial Intelligence Awareness in Work Environments. In B. Barricelli, V. Roto, T. Clemmensen, P. Campos, A. Lopes, F. Gonçalves, & J. Abdelnour-Nocera (Eds.), *Human Work Interaction Design. Designing Engaging Automation. HWID 2018. IFIP Advances in Information and Communication Technology, vol. 544* (pp. 175–185). Springer.

Karvonen, H., Koskinen, H., & Haggrén, J. (2012). Defining user experience goals for future concepts. A case study. In *NordiCHI2012 UX Goals 2012 Workshop Proceedings, Tampere: TUT Publication series* (pp. 14–19). Retrieved from https://tutcris.tut.fi/portal/files/2814790/ux_goals_2012_workshop.pdf

Karvonen, H., Koskinen, H., Wahlström, M., Perä, J., & Hakulinen, J. (2015). User experience in remote operation. In M. Nuutinen & H. Koskinen (Eds.), *User Experience and Usability in Complex Systems – UXUS, Final report 1/2015* (pp. 127–134). Tampere, Finland: FIMECC Oy.

Karvonen, H., Lappalainen, J., & Liinasuo, M. (2014). Automation awareness user interface study: Preliminary results. *Enlarged Halden Programme Group Meeting, EHPG 2014*. The OECD Halden Reactor Project.

Keinonen, T., & Takala, R. (2010). *Product concept design: a review of the conceptual design of products in industry*. Germany: Springer Science & Business Media.

Kim, S. Y., Cooper, J., Carroll, A., & Murugappan, S. (2017). Purple Sky Framework Towards the Flight Deck of the Future Experience: Through Co-design, Rapid UX Prototyping, and User Testing. In N. Stanton, S. Landry, G. Di Bucchianico, & A. Vallicelli (Eds.), *Advances in Intelligent Systems and Computing* (pp. 851–862). Cham, Switzerland: Springer.

Kirwan, B. (1998). Human error identification techniques for risk assessment of high risk systems—Part 1: review and evaluation of techniques. *Applied Ergonomics*, *29*(3), 157–177. https://doi.org/10.1016/S0003-6870(98)00010-6

Kirwan, B. (2017). *A guide to practical human reliability assessment*. London, UK: CRC press.

Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(3), 462–472. https://doi.org/10.1109/21.31053

Klemola, U., & Norros, L. (1997). Analysis of the clinical behaviour of anaesthetists: recognition of uncertainty as a basis for practice. *Medical Education*, *31*(6), 449–456. https://doi.org/10.1046/j.1365-2923.1997.00625.x

Kletz, T. A. (1999). The Origins and History of Loss Prevention. *Process Safety and Environmental Protection*, *77*(3), 109–116. https://doi.org/10.1205/095758299529938

Knight, J. C. (2002). Safety critical systems: challenges and directions. In *Proceedings of the 24th International Conference on Software Engineering. ICSE 2002* (pp. 547–550). ACM.

Kornecki, A., & Zalewski, J. (2009). Certification of software for real-time safety-critical systems: state of the art. *Innovations in Systems and Software Engineering*, *5*(2), 149–161. https://doi.org/10.1007/s11334-009-0088-1

Koski-Jännes, A. (1999). From addiction to self-governance. In Y. Engestrom, R. Miettinen, & R.-L. Punamaki (Eds.), *Perspectives on activity theory* (pp. 435–444). New York, NY: Cambridge University Press.

Kymäläinen, T., Kaasinen, E., Aikala, M., Hakulinen, J., Heimonen, T., Paunonen, H., … Mannonen, P. (2016). Evaluating Future Automation Work in Process Plants with an Experience-Driven Science Fiction Prototype. In *12th International Conference on Intelligent Environments (IE)* (pp. 54–61). IEEE.

Kymäläinen, T., Kaasinen, E., Hakulinen, J., Heimonen, T., Mannonen, P., Aikala, M., … Lehtikunnas, L. (2017). A creative prototype illustrating the ambient user experience of an intelligent future factory. *Journal of Ambient Intelligence and Smart Environments*, *9*(1), 41–57. https://doi.org/10.3233/AIS-160417

Kymäläinen, T., Perälä, P., Hakulinen, J., Heimonen, T., James, J., & Perä, J. (2015). Evaluating a Future Remote Control Environment with an Experience-Driven Science Fiction Prototype. In *11th International Conference on Intelligent Environments (IE)* (pp. 81–88). IEEE.

Laarni, J., Savioja, P., Karvonen, H., & Norros, L. (2011). Pre-validation of Nuclear Power Plant Control Room Design. In D. Harris (Ed.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6781, pp. 404–413). Springer Berlin Heidelberg.

Laarni, J., Savioja, P., Norros, L., Liinasuo, M., Karvonen, H., Wahlström, M., & Salo, L. (2014). Conducting Multistage HFE Validations - Constructing Systems Usability Case. In *Proceedings of ISOFIC/IISSNP 2014*. Korean Nuclear Society.

Lahtinen, J., Valkonen, J., Björkman, K., Frits, J., Niemelä, I., & Heljanko, K. (2012). Model checking of safety-critical software in the nuclear engineering domain. *Reliability Engineering & System Safety*, *105*, 104–113. https://doi.org/10.1016/j.ress.2012.03.021

Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., & Kort, J. (2009). Understanding, Scoping and Defining User eXperience: A Survey Approach. In *Proceedings of CHI 2009* (pp. 719–728). ACM.
Lee, J. D., & Gao, J. (2005). Trust, information technology, and cooperation in supply chains. *Supply Chain Forum: An International Journal*, *6*(2), 82–89. https://doi.org/https://doi.org/10.1080/16258312.2005.11517150
Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, *40*(1), 153–184. https://doi.org/10.1006/ijhc.1994.1007
Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80.
Lee, J. D., Wickens, C. D., Liu, Y., & Boyle, L. N. (2017). *Designing for people: An introduction to human factors engineering*. Charleston, SC: CreateSpace.
Leifer, L. J., & Steinert, M. (2014). Dancing with Ambiguity: Causality Behavior, Design Thinking, and Triple-Loop-Learning. In O. Gassman & F. Schweitzer (Eds.), *Management of the Fuzzy Front End of Innovation* (pp. 141–158). Cham, Switzerland: Springer.
Leveson, N. (2011). *Engineering a safer world: Systems thinking applied to safety*. Cambridge, MA: MIT press.
Leveson, N., Dulac, N., Zipkin, D., Cutcher-Gershenfeld, J., Carroll, J., & Barrett, B. (2006). Engineering Resilience into Safety-Critical Systems. In E. Hollnagel, D. Woods, & N. Leveson (Eds.), *Resilience Engineering: Concepts And Precepts*. Aldershot, UK: Ashgate.
Liinasuo, M., & Norros, L. (2007). Usability Case – integrating usability evaluations in design. *COST294-MAUSE Workshop on Downstream Utility*, 11–13. Institute of Research in Informatics of Toulouse (IRIT).
Lu, Y. (2018). *Experience Goals in Designing Professional Tools: Evoking Meaningful Experiences at Work* (Doctoral dissertation, Aalto University). Retrieved from https://aaltodoc.aalto.fi/bitstream/handle/123456789/34084/isbn9789526080802.pdf?sequence=1&isAllowed=y
Lu, Y., & Roto, V. (2014). Towards meaning change. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction Fun, Fast, Foundational - NordiCHI '14* (pp. 717–726). ACM Press.
Lu, Y., & Roto, V. (2016). Design for Pride in the Workplace. *Psychology of Well-Being*, *6*(1), 6. https://doi.org/10.1186/s13612-016-0041-7
Maguire, M. (1999). A review of user-interface design guidelines for public information kiosk systems. *International Journal of Human-Computer Studies*, *50*(3), 263–286. https://doi.org/10.1006/ijhc.1998.0243
Maguire, M. (2001). Methods to support human-centred design. *International Journal of Human-Computer Studies*, *55*(4), 587–634. https://doi.org/10.1006/ijhc.2001.0503
Marshall, M. N. (1996). Sampling for qualitative research. *Family Practice*, *13*(6), 522–526. https://doi.org/10.1093/fampra/13.6.522

Maslow, A. H. (1971). Peak experiences in education and art. *Theory Into Practice*, *10*(3), 149–153. https://doi.org/10.1080/00405847109542321
Mayhew, D. J., & Mayhew, D. (1999). *The usability engineering lifecycle: a practitioner's handbook for user interface design*. San Diego, CA: Morgan Kaufmann.
McCarthy, J., & Wright, P. (2004, September 1). Technology as experience. *Interactions*, *11*(5), 42. https://doi.org/10.1145/1015530.1015549
Mohaghegh, Z., Kazemi, R., & Mosleh, A. (2009). Incorporating organizational factors into Probabilistic Risk Assessment (PRA) of complex socio-technical systems: A hybrid technique formalization. *Reliability Engineering & System Safety*, *94*(5), 1000–1018. https://doi.org/10.1016/j.ress.2008.11.006
Montewka, J., Ehlers, S., Goerlandt, F., Hinz, T., Tabri, K., & Kujala, P. (2014). A framework for risk assessment for maritime transportation systems—A case study for open sea collisions involving RoPax vessels. *Reliability Engineering & System Safety*, *124*, 142–157. https://doi.org/10.1016/j.ress.2013.11.014
Montewka, J., Goerlandt, F., & Kujala, P. (2014). On a systematic perspective on risk for formal safety assessment (FSA). *Reliability Engineering & System Safety*, *127*, 77–85. https://doi.org/10.1016/j.ress.2014.03.009
Nagamachi, M. (1995). Kansei Engineering: A new ergonomic consumer-oriented technology for product development. *International Journal of Industrial Ergonomics*, *15*(1), 3–11. https://doi.org/https://doi.org/10.1016/0169-8141(94)00052-5
Nagamachi, M. (2002). Kansei Engineering in Consumer Product Design. *Ergonomics in Design: The Quarterly of Human Factors Applications*, *10*(2), 5–9. https://doi.org/10.1177/106480460201000203
Neisser, U. (1976). *Cognition and reality: principles and implications of cognitive psychology*. San Francisco, CA: Freeman and Company.
Norman, D. (1986). Cognitive engineering. In D. Norman & S. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Norman, D. (2005, July 1). Human-centered design considered harmful. *Interactions*, *12*(4), 14–19. https://doi.org/10.1145/1070960.1070976
Norros, L. (2004). *Acting under uncertainty – The core-task analysis in naturalistic study of work and expertise*. Espoo, Finland: VTT.
Norros, L. (2005). The concept of habit in the analysis of situated actions. *Theoretical Issues in Ergonomics Science*, *6*(5), 385–407. https://doi.org/10.1080/14639220500076520
Norros, L. (2014). Developing human factors/ergonomics as a design discipline. *Applied Ergonomics*, *45*(1), 61–71. https://doi.org/10.1016/j.apergo.2013.04.024

Norros, L., Liinasuo, M., Savioja, P., & Aaltonen, I. (2010). Cope Technology enabled capacity for first responder. In *COPE project deliverable D*. Retrieved from http://cope.vtt.fi/COPE Deliverable D2.3 - COPE Technology Enabled Capacity for First Responder.pdf

Norros, L., & Salo, L. (2009). Design of joint systems: a theoretical challenge for cognitive systems engineering. *Cognition, Technology & Work*, *11*(1), 43–56. https://doi.org/10.1007/s10111-008-0122-3

Norros, L., Savioja, P., & Koskinen, H. (2015). Core-Task Design: A Practice-Theory Approach to Human Factors. *Synthesis Lectures on Human-Centered Informatics 8(2)*.

NTSB. (1995). *Grounding of the Panamanian Passenger Ship Royal Majesty on Rose and Crown Shoal Near Nantucket, Massachussetts, June 10, 1995*.

Nuutinen, M., & Norros, L. (2009). Core task analysis in accident investigation: analysis of maritime accidents in piloting situations. *Cognition, Technology & Work*, *11*(2), 129–150. https://doi.org/10.1007/s10111-007-0104-x

O'Hara, J., Higgins, J., & Fleger, S. (2012). Human Factors Engineering Program Review Model (NUREG-0711) Revision 3: Update Methodology and Key Revisions. In *Proceedings of 8th International Topical Meeting on Nuclear Plant Instrumentation and Control and Human-Machine Interface Technologies*.

Obrist, M., Roto, V., & Väänänen-Vainio-Mattila, K. (2009). User experience evaluation. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (pp. 2763–2766). New York, NY: ACM Press.

Pahl, G., & Beitz, W. (2013). *Engineering design: a systematic approach*. London, UK: Springer Science & Business Media.

Parasuraman, R., & Miller, C. A. (2004). Trust and etiquette in high-criticality automated systems. *Communications of the ACM*, *47*(4), 51. https://doi.org/10.1145/975817.975844

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *30*(3), 286–297. https://doi.org/10.1109/3468.844354

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation Awareness, Mental Workload, and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 140–160. https://doi.org/10.1518/155534308X284417

Peirce, C. S. (1905). What pragmatism is. *The Monist*, 161–181.

Ponsard, C., Massonet, P., Molderez, J. F., Rifaut, A., Lamsweerde, A. van, & Van, H. T. (2007). Early verification and validation of mission critical systems. *Formal Methods in System Design*, *30*(3), 233–247. https://doi.org/10.1007/s10703-006-0028-8

Pritchett, A. R., Kim, S. Y., & Feigh, K. M. (2014). Modeling Human–Automation Function Allocation. *Journal of Cognitive Engineering and Decision Making*, *8*(1), 33–51. https://doi.org/10.1177/1555343413490944

Qureshi, Z. (2007). A review of accident modelling approaches for complex socio-technical systems. In *Proceedings of the twelfth Australian workshop on Safety critical systems and software and safety-related programmable systems* (pp. 47–59). Australian Computer Society, Inc.

Radlinski, F., Bennett, P. N., Carterette, B., & Joachims, T. (2009). Redundancy, diversity and interdependent document relevance. *ACM SIGIR Forum*, *43*(2), 46–52. https://doi.org/10.1145/1670564.1670572

Rajaonah, B., Anceaux, F., Tricot, N., & Pacaux-Lemoine, M.-P. (2006). Trust, cognitive control, and control: the case of drivers using an Auto-Adaptive Cruise Control. *Proceedings of the 13th Eurpoean Conference on Cognitive Ergonomics: Trust and Control in Complex Socio-Technical Systems*, 17–24. ACM.

Rajkumar, R. (Raj), Lee, I., Sha, L., & Stankovic, J. (2010). Cyber-physical systems. In *Proceedings of the 47th Design Automation Conference on - DAC '10* (p. 731). New York, NY: ACM Press.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. New York, NY: Elsevier Science.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive systems engineering*. New York, NY: Wiley.

Rice, M., Tay, H. H., Ng, J., Lim, C., Selvaraj, S. K., & Wu, E. (2016). Comparing Three Task Guidance Interfaces for Wire Harness Assembly. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16* (pp. 2279–2284). New York, NY: ACM Press.

Riegelsberger, J., Sasse, M. A., & McCarthy, J. D. (2005). The mechanics of trust: a framework for research and design. *International Journal of Human-Computer Studies*, *62*(3), 381–422.

Robert, J.-M., & Lesage, A. (2017). From Usability to User Experience with Interactive Systems. In G. A. Boy (Ed.), *The Handbook of Human-Machine Interaction* (pp. 303–320). Boca Raton, FL: CRC Press.

Rödel, C., Stadler, S., Meschtscherjakov, A., & Tscheligi, M. (2014). Towards Autonomous Cars. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '14* (pp. 1–8). ACM Press.

Roto, V., Kaasinen, E., Nuutinen, M., & Seppänen, M. (2016). UX Expeditions in Business-to-Business Heavy Industry. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16* (pp. 833–839). New York, NY: ACM Press.

Roto, V., Law, E., Vermeeren, A., & Hoonhout, J. (2011). *User Experience White Paper*. Retrieved from http://www.allaboutux.org/files/UX-WhitePaper.pdf

Roto, V., Nuutinen, M., & Smedlund, A. (2014). *UX HOW? A Glimpse of User And, Industry – Issue 2.* Retrieved from https://issuu.com/vttfinland/docs/ux_booklet_english?e=5313536/10315158

Roto, V., Obrist, M., & Väänänen-Vainio-Mattila, K. (2009). User Experience Evaluation Methods in Academic and Industrial Contexts. In *Workshop on User Expererience Evaluation Methods, in conjunction with Interact'09 conference*.

Rozendaal, M. C. (2010). Investigating experience design in industry. In *Proceedings of the 7th International Design & Emotion Conference* (pp. 4–7).

Saariluoma, P., & Jokinen, J. P. P. (2014). Emotional Dimensions of User Experience: A User Psychological Analysis. *International Journal of Human-Computer Interaction*, *30*(4), 303–320. https://doi.org/10.1080/10447318.2013.858460

Saariluoma, P., Karvonen, H., & Rousi, R. (2019). Techno-Trust and Rational Trust in Technology – A Conceptual Investigation. In B. Barricelli, V. Roto, T. Clemmensen, P. Campos, A. Lopes, F. Gonçalves, & J. Abdelnour-Nocera (Eds.), *Human Work Interaction Design. Designing Engaging Automation. HWID 2018. IFIP Advances in Information and Communication Technology, vol. 544* (pp. 283–293). Springer.

Salah, D., Paige, R. F., & Cairns, P. (2014). A systematic literature review for agile development processes and user centred design integration. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering - EASE '14* (pp. 1–10). ACM Press.

Salmon, P., Cornelissen, M., & Trotter, M. (2012). Systems-based accident analysis methods: A comparison of Accimap, HFACS, and STAMP. *Safety Science*, *50*(4), 1158–1170. https://doi.org/10.1016/j.ssci.2011.11.009

Salmon, P., Stanton, N., Gibbon, A., Jenkins, D., & Walker, G. (2009). Cognitive Task Analysis. In P. Salmon, N. Stanton, A. Gibbon, D. Jenkins, & G. Walker (Eds.), *Human Factors Methods and Sports Science* (pp. 69–107). CRC Press.

Salmon, P., Stanton, N., Lenné, M., Jenkins, D., Rafferty, L., & Walker, G. (2012). *Human Factors Methods and Accident Analysis: Practical Guidance and Case Study Applications*. Farnham, UK: Ashgate Publishing, Ltd.

Salvendy, G. (2012). *Handbook of human factors and ergonomics*. Hoboken, NJ: John Wiley & Sons.

Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (pp. 1926–1943). John Wiley & Sons, Inc.

Savioja, P. (2014). *Evaluating systems usability in complex work – Development of a systemic usability concept to benefit control room design* (Doctoral dissertation, VTT). Retrieved from https://www.vtt.fi/inf/pdf/science/2014/S57.pdf

Savioja, P., Liinasuo, M., & Koskinen, H. (2014). User experience: does it matter in complex systems? *Cognition, Technology & Work*, *16*(4), 429–449. https://doi.org/10.1007/s10111-013-0271-x

Savioja, P., & Norros, L. (2008). Systems Usability — Promoting Core-Task Oriented Work Practices. In E. Law, T. Hvannberg, & G. Cockton (Eds.), *Maturing Usability: Quality in Software, Interaction and Value* (pp. 123–143). London, UK: Springer.

Savioja, P., & Norros, L. (2013). Systems usability framework for evaluating tools in safety–critical work. *Cognition, Technology & Work*, *15*(3), 255–275. https://doi.org/10.1007/s10111-012-0224-9

Savioja, P., Norros, L., Salo, L., & Aaltonen, I. (2014). Identifying resilience in proceduralised accident management activity of NPP operating crews. *Safety Science*, *68*, 258–274. https://doi.org/10.1016/j.ssci.2014.04.008

Schuler, D., & Namioka, A. (2017). *Participatory Design: Principles and Practices*. Boca Raton, FL: CRC Press.

Segal, L., & Suri, J. (1997). Empathic practitioner: Measurement and interpretation of user experience. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *41*(1), 451–45. https://doi.org/https://doi.org/10.1177%2F107118139704100199

Sharp, H., Preece, J., & Rogers, Y. (2019). *Interaction Design: beyond human-computer interaction 5th edition.* Indianapolis, IN: John Wiley & Sons, Ltd.

Shattuck, L. G., & Woods, D. D. (1997). Communication of Intent in Distributed Supervisory Control Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *41*(1), 259–263. https://doi.org/10.1177/107118139704100159

Shneiderman, B., Byrd, D., & Croft, W. B. (1998). Sorting out searching: a user-interface framework for text searches. *Communications of the ACM*, *41*(4), 95–98. https://doi.org/10.1145/273035.273069

Stamatis, D. H. (2003). *Failure mode and effect analysis: FMEA from theory to execution*. Milwaukee, WI: ASQ Quality press.

Stamm Von, B. (2008). *Managing innovation, design and creativity.* Chichester, UK: John Wiley & Sons.

Stanton, N. (2003). The psychology of task analysis today. In D. Diaper & N. Stanton (Eds.), *The handbook of task analysis for human-computer interaction* (pp. 569–584). London, UK: CRC Press.

Stanton, N., Salmon, P., Rafferty, L., Walker, G., Baber, C., & Jenkins, D. (2017). *Human factors methods: a practical guide for engineering and design*. Boca Raton, FL: CRC Press.

Star, S. (2010). This is Not a Boundary Object: Reflections on the Origin of a Concept. *Science, Technology, & Human Values*, *35*(5), 601–617. https://doi.org/10.1177/0162243910377624

Stramler, J. H. (1993). *The dictionary for human factors/ergonomics*. Boca Raton, FL: CRC Press.

Sträter, O. (2016). *Cognition and safety: an integrated approach to systems design and assessment*. New York, NY: Routledge.

Takala, R., Keinonen, T., & Mantere, J. (2006). Processes of Product Concepting. In T. Keinonen & R. Takala (Eds.), *Product Concept Design* (pp. 57–90). London, UK: Springer.

Tasoudis, S., & Perry, M. (2018). Participatory Prototyping to Inform the Development of a Remote UX Design System in the Automotive Domain. *Multimodal Technologies and Interaction*, *2*(4), 74. https://doi.org/10.3390/mti2040074

Thagard, P. (2005). *Mind: Introduction to cognitive science*. Cambridge, MA: MIT press.
Väänänen-Vainio-Mattila, K., Roto, V., & Hassenzahl, M. (2008a). Now Let's Do It in Practice: User Experience Evaluation Methods in Product Development. In *Proceedings of Extended Abstracts on Human Factors in Computing Systems (CHI'08)* (pp. 3961–3964). ACM.
Väänänen-Vainio-Mattila, K., Roto, V., & Hassenzahl, M. (2008b). Towards Practical User Experience Evaluation Methods. *5th COST294-MAUSE Open Workshop on Valid Useful User Experience Measurement - VUUM*, 19–22. Institute of Research in Informatics of Toulouse (IRIT).
Väätäjä, H., Olsson, T., Savioja, P., & Roto, V. (2012). *UX Goals 2012 workshop--how to utilize user experience goals in design?* Tampere, Finland: Tampere University of Technology.
Väätäjä, H., Savioja, P., Roto, V., Olsson, T., & Varsaluoma, J. (2015). User experience goals as a guiding light in design and development-Early findings. In *INTERACT 2015 Adjunct proceedings* (pp. 521–527). University of Bamberg Press.
van den Haak, M., De Jong, M., & Schellens, P. (2003). Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue. *Behaviour & Information Technology*, *22*(5), 339–351. https://doi.org/10.1080/0044929031000
Varsaluoma, J. (2018). *Approaches to Improve User Experience in Product Development: UX Goals , Long-Term Evaluations and Usage Data Logging* (Doctoral dissertation, Tampere University of Technology). Retrieved from http://urn.fi/URN:ISBN:978-952-15-4251-0
Varsaluoma, J., Väätäjä, H., Kaasinen, E., Karvonen, H., & Lu, Y. (2015a). Getting started with the experience design process. In M. Nuutinen & H. Koskinen (Eds.), *User Experience and Usability in Complex Systems – UXUS, Final report 1/2015* (pp. 125–126). Tampere, Finland: FIMECC Oy.
Varsaluoma, J., Väätäjä, H., Kaasinen, E., Karvonen, H., & Lu, Y. (2015b). The Fuzzy Front End of Experience Design. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction on - OzCHI '15* (pp. 324–332). ACM Press.
Vicente, K. J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. Mahwah, NJ: CRC Press.
Vincoli, J. W. (2006). *Basic guide to system safety*. Hoboken, NJ: John Wiley & Sons.
Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. *Cambridge, MA: Harvard University Press.*
Wahlström, M., Karvonen, H., & Kaasinen, E. (2014). InnoLeap - Creating Radical Concept Designs for Industrial Work Activity. *NordiCHI 2014 The Fuzzy Front End of Experience Design Workshop*, 63–68. VTT.
Wahlström, M., Karvonen, H., & Norros, L. (2013). Rehearsing for a major accident in a metro control centre: a naturalistic analysis of situation

awareness. In *Proceedings of the 11th International Conference on Naturalistic Decision Making (NDM 2013)* (pp. 21–24). Arpege Science Publishing.

Wahlström, M., Seppänen, L., Norros, L., Aaltonen, I., & Riikonen, J. (2018). Resilience through interpretive practice – A study of robotic surgery. *Safety Science*, *108*, 113–128. https://doi.org/10.1016/j.ssci.2018.04.010

Wong, W. (2004). The critical decision method. In D. Diaper & N. Stanton (Eds.), *The handbook of task analysis for human-computer interaction* (pp. 327–346). Mahwah, NJ: Lawrence Erlbaum Associates.

Wood, J. R., & Wood, L. E. (2008). Card Sorting: Current Practices and Beyond History and Assumptions. *Journal of Usability Studies*, *4*(1), 1–6.

Woods, D. D., & Hollnagel, E. (2006). *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*. Boca Raton, FL: CRC Press.

Wright, P., & McCarthy, J. (2005). The value of the novel in designing for experience. In A. Pirhonen, P. Saariluoma, H. Isomäki, & C. Roast (Eds.), *Future Interaction Design* (pp. 9–30). London, UK: Springer-Verlag.

Wright, P., & McCarthy, J. (2010). Experience-Centered Design: Designers, Users, and Communities in Dialogue. *Synthesis Lectures on Human-Centered Informatics 3(1)*.

Wright, P., McCarthy, J., & Meekison, L. (2003). A framework for analysing user experience. In M. Blythe, A. Monk, P. Wright, & K. Overbeeke (Eds.), *Funology: from usability to user enjoyment*. Dordrecht, The Netherlands: Kluwer.

Wright, P., Wallace, J., & McCarthy, J. (2008). Aesthetics and experience-centered design. *ACM Transactions on Computer-Human Interaction*, *15*(4), 1–21. https://doi.org/10.1145/1460355.1460360

Wurhofer, D., Fuchsberger, V., Meneweger, T., Moser, C., & Tscheligi, M. (2015). Insights from User Experience Research in the Factory: What to Consider in Interaction Design. In J. Abdelnour-Nocera, T. Clemmensen, B. Barricelli, A. Lopes, & P. Campos (Eds.), *IFIP Advances in Information and Communication Technology* (pp. 39–56). Springer.

Appendix A: Practitioner guidelines on UX goals in human-centred design of safety-critical systems (see Väätäjä et al., 2012; Varsaluoma et al., 2015a)

UX goals and other requirements specification

- In the identification of an initial set of possible UX goals, earlier gathered user data, appropriate theoretical underpinnings, and trend studies can be utilised
- Early-stage user research should be conducted in as realistic settings as possible to validate the identified UX goals and possibly find out more relevant UX goals
- The gathered broad set of identified UX goals should be narrowed down to the most important goals to be set for the design work based on the user research results
- Once the UX goals are set for design, it should be defined in detail what the chosen goals mean in the specific context of use (i.e., defining their design implications)
- Possible requirements connected with the set UX goals should be considered based on the gained user research results
- UX goals should be specified before the actual product development takes place, but can also be iteratively modified during design

The description and communication of UX goals

- The specified UX goals should be described so that all stakeholders can create a shared understanding of their definition and meaning
- On the one hand, the UX goals should be described precisely enough to make them actionable for designers, but on the other, broadly enough to allow room for creativity
  - The reasoning behind the goals (i.e., *why*) should be described, as designers need to select the proper means (i.e., *how*) for conveying the experience (i.e., *what*)
- It should be planned what means (i.e., design outcomes or artefacts) are used to communicate the UX goals to relevant stakeholders

UX goals' utilisation in concept design

- The UX goals should be presented in the beginning of the possible co-design workshops and work as guiding stars in the generation of design ideas
- The goals can be iterated when more about the context domain and the relevance of the goals is learned throughout the design process

UX goals in evaluation

- It should be planned how to trace the design solutions back to the defined UX goals to be able to evaluate the fulfilment of the goals in different stages of the design work
- The UX goals should be operationalised and appropriate metrics for their reflective evaluation should be selected

ORIGINAL PAPERS

I

HIDDEN ROLES OF THE TRAIN DRIVER: A CHALLENGE FOR METRO AUTOMATION

by

Hannu Karvonen, Iina Aaltonen, Mikael Wahlström, Leena Salo,
Paula Savioja, & Leena Norros

Interacting with Computers vol 23, 289–298

https://doi.org/10.1016/j.intcom.2011.04.008

HIDDEN ROLES OF THE TRAIN DRIVER: A CHALLENGE FOR METRO AUTOMATION

Hannu Karvonen[a]*, Iina Aaltonen[a], Mikael Wahlström[b], Leena Salo[a], Paula Savioja[a], Leena Norros[a]

[a] VTT Technical Research Centre of Finland, P.O. Box 1000, 02044 VTT, Finland; firstname.lastname@vtt.fi

[b] Helsinki Institute for Information Technology HIIT, P.O. Box 19800, 00076 Aalto, Finland; firstname.lastname@hiit.fi

* Corresponding author

Telephone: +358 400 216 396

Fax: +358 20 722 6027

ABSTRACT

In the year 2014, the Helsinki Metro is planned to be fully automated. This automation means that the metro trains will be computer-driven and monitored remotely from a stationary control room. To investigate the challenges related to this scenario, we decided to study the ways in which the current train drivers contribute to the metro system. We conducted three separate but interrelated studies, which were based on the Core-Task Analysis method. Our results suggest that there is much more to driving the metro train than meets the eye. The drivers do not only operate the train on track and its doors at stations, but they also contribute to a variety of other important, albeit more hidden functions in the metro system. For example, the drivers anticipate, observe, interpret, and react to events in the surrounding environment. Furthermore, they are a significant interaction link between different actors of the metro system. Our conclusion is that if the identified critical roles of the drivers are not accounted for, a migration to a fully automated metro system can affect the quality of service and raise safety issues. In addition to automated metros, the results of this research can be applicable to automation implementations also in other domains.

KEYWORDS: Core-Task Analysis, Metro train driver work, Automated metro, Human-centered automation, Complex systems, Safety

1. INTRODUCTION

An increasing number of metro lines are being automated in different cities around the world. The main reasons for metro automation are cost-effectiveness, high traffic frequency, and flexibility. For example, experiences from France [1] show that the implemented automated metro lines have met these expectations.

Automated metros are complex IT-based systems where computing is embedded into people's everyday environment. Research on automated metros has been mostly concentrating on the technological side of the domain and studies from cognitive ergonomics point of view have been scarce. However, human interaction with these kinds of safety-critical systems should be taken under careful investigation, because otherwise people's lives might be endangered.

In the field of conventional metros, there has been a fair amount of human interaction-related studies investigating some specific parts of the domain, for example, the work of traffic controllers [2–4] and crowd management in operation rooms [5]. However, to the best of our knowledge little public research has been published regarding metro train driver's work, which is the interest of this study.

In lack of comprehensive human-oriented driver studies from the metro domain, we looked into studies from traditional rail traffic. For example, Wilson and Norris [6] stated that conventional train driver studies have been conducted widely related to the driver's vigilance, perception, and recognition of and acting upon signs and signals.

More relevantly to our approach, Jansson, Olson, and Kecklund [7] have investigated conventional train driver's work using Cognitive Work Analysis (CWA). According to them, train driver's work is a mixture of observation of the surrounding environment, maintenance of situation awareness, automated cognitive processes, recognition, working memory limitations, and dynamic decision making. Furthermore, a train driver must combine information from many sources, which are, for example, trackside signals, Automatic Train Control (ATC), route book, timetables, rulebook, and trackside environment.

Regarding automated metros, their operation models can differ in the level of automation and the role of metro personnel. Georgescu [8] has listed the following operation models for automated metros:

- Semiautomatic train operation (STO): The driver is responsible for safe departure from stations, but the train will drive automatically to the next station. The driver observes the track and is able to stop the train if a hazardous situation occurs.
- Driverless train operation (DTO): The driver is absent from the train cockpit, but there is operating staff inside the train. Safe departure of the train from the station (including closing the doors) can be the responsibility of the operating staff or may be done automatically.
- Unattended train operation (UTO): Operating staff is not required inside the train. Safe departure of the train from the station is done automatically. Additional systems such as intrusion detection systems and onboard CCTV are usually installed to support the detection and management of hazardous conditions.

The case of our study, the Helsinki Metro, will be automated by the end of year 2014 according to the current plans. From the above presented options, Helsinki's automated metro will be STO during the testing phase of the system. After the automatic system proves to work correctly, UTO model is taken into use. Such fully automated metro lines are already in place, for example, in Copenhagen, Denmark and Nuremberg, Germany. However, public reports about experiences from these kinds of systems are rare. In Helsinki, the fully automated metro will mean that there will no longer be a metro train driver controlling the train inside the cockpit. Instead, an on-board computer drives each train and a central computer controls the overall traffic. Human operators from a stationary control room will monitor the traffic situation. Furthermore, staff will supervise and guide passengers in stations and trains.

The authorities have presented several reasons on why the Helsinki Metro should be automated. To mention a few, the renewal of metro technology in Helsinki is unavoidable in any case because the components in use currently are no longer manufactured and there is a shortage of spare parts. Added to this, system support is not provided any more. The advantages of the future Helsinki automated metro compared to the current conventional one are said to be, for example, higher frequency of traffic, less expenditures in the long term, energy savings, maintenance savings, and increased safety according to a report [9] by HKL (Helsinki City Transport). It is argued that with an automatic system, a more punctual and efficient metro from a technical standpoint can be achieved.

However, previous human-centred automation related-research suggests that automation may also have harmful effects [10]. For example, Billings [11] has examined the effects of automation in the aviation domain where many of the pilots' tasks are automated. Billings emphasised that automation should not be done too quickly before the effects that the automation has on human operators are fully understood. The human operators should have a good situation awareness of what the automation is doing, and why. The final responsibility remains with the human operator even if automation performance leads to dangerous situations. In aviation, the human operators are pilots and air traffic controllers; in a fully automated metro, they are traffic controllers and supervisory personnel in stations and trains. Therefore, one should consider what can be lost when the human element – the driver – is no longer onboard the train.

To study this issue empirically, we conducted interviews, observations, and a mirror data workshop [12] with the Helsinki metro operating personnel. The main goals with these qualitative research methods were to investigate the drivers' work, their different roles in the metro system, and to identify possible challenges related to the planned automation concept of the Helsinki metro system. By metro system, we mean the whole complex socio-technical system which is required to run the trains safely and on time. By automation concept, we mean the concept of operation where the automation is planned to work, the coupling of the human and automated functions, the level of automation, and the roles different actors have in this concept.

2. METHODS

In the preparatory phase of the studies, we had several meetings with the management representatives of the target organisation to understand the work context of the metro train drivers. These meetings included, for example, walking through the basic concepts of the Helsinki Metro. In addition to the official meetings, we also had several free-form discussion sessions with the traffic control room personnel. All of these meetings were essential in order to learn about the domain and its actors, language, and terminology. To acquaint ourselves with the driver's working environment, we visited the trains' cockpits in the depot. This allowed us to gain insight to what the different equipment in the cockpit is used for and how the train is operated.

After this preparatory phase, we conducted three different but interrelated qualitative studies in the following chronological order: an interview study, an observation study, and a mirror data workshop. Each study was followed by an analysis phase, which helped us refine our goals. Furthermore, a simple card sorting method was used after all the studies to categorise the found challenges under common themes.

The approach of these studies as a whole was based on Core-Task Analysis [13] method. The aim of Core-Task Analysis is to identify the core task of a specific work. Core task is the main result-oriented content of the work that can be derived by analysing the objective of work and the demands that the objective lays on workers both in general and in specific situations. Core-Task Analysis has been used earlier in several domains, for instance, in maritime piloting [14].

In addition to the previously mentioned studies, we also did a quantitative analysis related to the statistical data of occurred exceptional situations in the Helsinki Metro. In this way, especially the potentially challenging situations related to the forthcoming automation could be elucidated.

2.1. Interview study

Altogether 12 metro train drivers and traffic controllers of the Helsinki Metro took part in the interviews. Three of them were drivers, five of them were traffic controller-drivers, and four of them were traffic controllers. The traffic controllers' perspective on the metro system and drivers' work was deemed important as they have an overview of the metro system and interact with the drivers regularly. In addition, most traffic controllers are former train drivers.

The interviews were conducted in the main traffic control room of the Helsinki Metro. Each participant was interviewed individually by two researchers, one of whom mainly concentrated on taking notes by typing the main points of each answer with a laptop computer. The interviews were also recorded and afterwards the written notes were supplemented with the audio recordings.

The interviews were conducted as semi-structured theme interviews. The themes were the nature of work and core task in the work, professional skills and co-operation, tasks and user-interfaces, and the changes caused by metro automation. In each of these themes, there were from six to 15 specific questions but the interview was not restricted to only pre-defined questions. Overall, the interview included 42 questions, based on the Core-Task Analysis method. If the interviewee brought up new issues or themes, those were also discussed.

In the analysis phase, the interview answers were categorised under emerging groups of frequently occurring answers. These answer groups were then linked to the core-task demands of the metro domain.

2.2. Observation study

In the field observation study, we observed four Helsinki metro train drivers during their normal daily work shifts. During observation, there were usually two researchers with the driver in the train's cockpit (Fig. 1), one using a video camera and asking possible questions, and the other typing notes.

Insert Fig. 1 here.

At the beginning of each observation session inside the train, we first agreed with the driver that he or she would drive the whole route from beginning to end in silence, so that we would not interfere in any way. This procedure allowed us to observe the driver's work in natural settings without disturbance. After this first round, while the driver was still driving the train, we discussed the driver's work in practice and asked questions regarding the actions and events we saw during driving and the driver's role in general in different exceptional situations, such as technical problems, accidents, or passenger misbehaviour.

During analysis, relevant episodes from the collected video material were sought based on the taken notes. After identifying the driver's tasks and communication with the other actors of the metro system from the video material, we also looked for the challenges related to the automated metro, i.e., to the situation when the driver is no longer present in the cockpit of the train.

2.3. Mirror data workshop

After conducting both the interview and observation studies, we extracted episodes unclear to us from the collected data. These episodes were further discussed in a mirror data [12] workshop, which we organised with the same metro train drivers and traffic controllers who had taken part in the previous studies. We asked the participants about our observations using illustrations such as photos and video. The purpose of the method was to stimulate the participants with mirror data material regarding their own work to get information about a certain situation that had happened [12]. This mirror data allows the participants to reflect on the presented issues on the grounds of previous experience. Possible exceptional situations, such as accidents, were also discussed in more detail during the workshop. The workshop sessions were video recorded so that they could be analysed later in more detail.

The mirror data workshop benefited from the interview and observation studies done earlier. The methods supplemented each other so that issues identified in the interviews could be verified in practice in real settings during the observations and further discussed in the mirror data workshop. The different studies provided diverse perspectives to the work of the metro train drivers and their role in the metro system. In the analysis phase, the workshop video recordings were analysed and the results were combined with the results of the interview and observation studies.

2.4. Card sorting

After identifying the challenges related to metro automation during analysis, we wanted to categorise the challenges under common themes. For the categorization, we used a simple card sorting method by writing the description of each challenge on a Post-it note and grouping them together to logical groups. The themes that emerged are listed as subsections in section 3.5.

2.5. Exceptional situation log analysis

In order to have a clear view on what are the most common exceptional situations in the current metro system, we asked our target organisation to provide us with statistical data about previous exceptional situations. We processed this data into two tables, which can be seen in section 3.2.

3. RESULTS AND DISCUSSION

We have organised our results in the following order. First, we present the core-task characteristics and demands for the driver's work and the metro domain in general, and present means for managing these demands. Second, we give an overview on what are the most common exceptional situations in the Helsinki Metro currently. Third, we define the metro train driver's basic tasks in more detail, and fourth, present a model of the interactions between the main actors of the metro system from the driver's point of view. Finally, we identify and discuss challenges the driver's different roles pose to the forthcoming metro automation and its automation concept.

3.1. Core-Task Analysis

The participants of the studies identified the ultimate goal of their work to be safe metro traffic, which runs uninterruptedly according to the schedule. Metro train driver work's practical goals were seen to be, for example, driving the train, observing the environment, and interacting with passengers.

To achieve these goals, the drivers have to take into account many characteristics of the domain, which are associated with its control demands. The characteristics of the metro domain from the driver's perspective are presented in Table 1. The description method is based on the tools of Core-Task Analysis.

Insert Table 1 here.

According to the Core-Task Analysis, a domain is characterised by control demands related to dynamism, complexity, and uncertainty. Factors related to dynamism are associated especially with work's temporal demands, such as timing, duration, and delays. In the metro system, the trains should be run according to a schedule and certain tasks are repeated regularly. Large-scale interferences to the flow of traffic happen rarely. The lack of exceptions and the need for running according to the schedule make the driver's work monotonous and routine. However, the tight schedule is vulnerable to problems and quick decision-making and action are required in exceptional situations.

Factors related to complexity originate from the large number of different elements in the system, which have more or less direct interactions with each other. According to the results gained from the Core-Task Analysis, there are less cogent reasons for complexity than to dynamism and uncertainty. For example, currently the Helsinki metro track network is not very complex since it consists of only one line. This line is approximately 15 km long and forks into two branches in the eastern end of the line. Furthermore, the basic driving task of the metro train itself is not very difficult. The complexity of the work, however, derives from other aspects.

The challenge for the drivers is that exceptional situations are usually dissimilar and therefore any ready-made procedures or routines for them are in general nonexistent. These situations require the driver to divide his or her attention appropriately between fixing the exception, taking care of passengers, and normalizing the traffic, i.e., getting the train to run back on schedule. Another factor affecting the complexity of work is that even though the metro as a system is more closed than, for example, tram traffic, it still has to be considered open as it is affected by different environmental factors such as weather and passenger behaviour. The impact these factors have on the work must be understood and managed by the drivers.

Factors related to uncertainty are, for example, related to the used technology and surrounding environment. In the metro system, the appearance and duration of exceptional situations is difficult to predict. The time required for solving a problem situation depends on the severity and location of the problem. Exceptional situations can be related, for example, to the technical functioning of the system or passenger misbehaviour and they are presented in more detail in section 3.2.

Fig. 2 presents means for managing the core-task characteristics related to dynamism, complexity, and uncertainty and their identified subsections in Table 1. These factors are managed by collaboration, skill, and knowledge. When each of the core-task characteristics is examined through these means, the general core-task demands of the work emerge. The red dots with listed items next to them in Fig. 2 refer to these core-task demands.

Insert Fig. 2 here.

It can be presumed that dynamism is primarily managed by collaboration and different types of skills. An example of such a skill would be maintaining the train on schedule by controlling the driving speed and stopping time at stations. Managing factors related to complexity requires different types of knowledge, for instance, about the functionality of technology and impact of the surrounding environment. Collaboration between different actors of the metro system and combining their expertise is also very important, especially in problem-solving situations.

Drivers' skills related to managing uncertainty include, for example, observing the surrounding environment and anticipating dangerous situations. The drivers can manage uncertainty also with knowledge by preparing themselves for unexpected situations by fault situation training, repetition, and mental image training (see Fig. 2).

The forthcoming automation of the metro affects all the core-task characteristics and means of managing them. New technology naturally requires learning new skills and knowledge. The staff must also understand and manage the new uncertainty factors that potentially come with the automatic system. Increasing the number of technical components in the system increases also the complexity of the system. If the functionality of the automation is not clear to the involved actors, understanding the system and its functioning logic can become more difficult. The automated metro makes shorter train intervals possible, but at the same time it can also increase the interference vulnerability and complicate the normalization of traffic after a stoppage. In addition, big changes to the collaboration will result when one actor group, the drivers, is withdrawn completely. Therefore, the other actor groups have to learn how to cope with work without someone being present onboard the train. More challenges related to the fully automated metro scenario will be discussed on a detailed level in section 3.5.

3.2. Most common exceptional situations

In Table 2, we present the descriptions of the most common exceptional situations in the Helsinki Metro. This data is based on the log analysis of reports from different exceptional situations between years and 2006–2010.

Insert Table 2 here.

In Table 3, we see the amount of different exceptional situations each year during this five-year period. When looking at these amounts it should be accounted for that these are only the exceptional situations, which are reported forward from the field. It can be assumed that in reality, much more minor exceptional situations happens, for example, related to passengers preventing the train's doors from closing, but they are not reported due to their insignificant nature.

Insert Table 3 here.

As can be seen from the data, emergency handle pullings occur on average over twice a week. Other more common exceptional situations are technical train door problems and unauthorised persons on rails. This data serves as a basis for analysing what are the situations where the driver is especially important. For example, with emergency handle pullings the driver (or the guard) conducts the needed actions quickly. In case of door problems, the driver can sometimes fix the problem on the spot. If an unauthorised person is on the rails, the driver can stop the train and inform traffic control or the guards about the situation.

3.3. Driver's current tasks

Based on the data collected from the studies, we identified four different main tasks of the drivers:

1. Operating the train
2. Taking care of passengers
3. Observing events outside the train
4. Acting in exceptional situations

These main tasks consist of subtasks that can overlap between the different main tasks. An analysis on different subtasks and to which main tasks they contribute to is presented in Table 4.

Insert Table 4 here.

The drivers' routine work tasks include driving the train safely on schedule according to the trackside signals, stopping at each station, and opening and closing the train doors. In addition to these visible tasks, the driver contributes to a variety of other important but less obvious functions such as anticipating, observing, interpreting, and reacting to events outside the train. The driver monitors the surrounding environment for factors affecting the train, including, for example, the weather conditions, passenger behaviour, and possible objects on the track.

In exceptional situations, such as a case of a technical fault, the driver is not only the other metro actor groups' "eyes on the field", but also an active actor in fixing the problem. In these kinds of situations, the driver contacts traffic control, the guards, or the depot's maintenance. The driver can not see the whole traffic situation, and therefore the traffic controllers decide how the situation is handled. The drivers are in general unfamiliar with the work of traffic controllers, but co-operation works better if the driver has knowledge on traffic control. Guards are needed if there is a passenger causing trouble inside the train. Maintenance's task is to guide the driver in fixing technical faults and, in case of more severe faults, possibly go on-site to fix them. If the fault causes the train to break up entirely or compromises safety, the driver has to guide the passengers out of the train to the closest platform.

The drivers are expected to provide passenger service, which was seen by the participants of our studies to include soft driving style, safety, and enjoyment of the trip, friendly and clear announcements, and general respect towards passengers. Drivers are also responsible for answering the train's emergency phone calls from the passengers.

The demanding aspects of driver's work are taking into account weather conditions, acting in exceptional situations, and maintaining attention during long working hours. In addition, the drivers brought up the stress of driving the train according to the tight schedules.

3.4. Driver's interactions in the metro system

The driver's and the metro system's other main actors' current interaction relationships are illustrated on a general level in Fig. 3. These interactions can be, for instance, verbal communication with other actors (rectangles in the figure) or the management of technical objects (ellipses in the figure). The figure demonstrates the driver's central role as an intermediator between the other main actors and metro system's technical objects such as the train and the track. If we look at the interactions in more detail, we see that it is actually only the interaction between the driver and the train, i.e., the basic controlling of the train, which can be automated by computers without deeper consideration. A central question with the automatic system is that how the driver's interactions with the metro system and its actors are taken care of without the drivers. In specific, it has to be considered how the driver's role, interactions, tasks, and viewpoint to the metro system will be implemented by means of automation and other actors? Issues regarding this question will be discussed in more detail in section 3.5.4.

Insert Fig. 3 here.

3.5. Challenges with the automation concept

The benefits of metro automation were described in the Introduction section of this article. In contrast, the fact that the drivers are withdrawn from the trains' cockpits may have disadvantageous effects. In the analysis and card sorting phase of the material collected from the conducted studies, different kinds of challenges for the automated system related to the following themes emerged: driving the train on the track, stopping at a station, passenger care, and interactions of the driver with other actors of the metro system.

These themes include the situations in which the drivers currently have hidden but significant roles. The challenges for the planned automation concept deriving from these situations should be addressed before taking the fully automated metro with the UTO model (see section 1) into use. Some considerations are applicable with the DTO model as well. Furthermore, some of the presented challenges are not limited only to metro automation, but can be seen universal to all automation implementations.

Next, these challenges are presented and discussed according to the identified themes. Direct quotes from the participants are used to illustrate the described challenges.

3.5.1. Driving the train on the track

The relatively fast speed of the metro trains with quick stops creates a challenge of its own for achieving a safe environment. The driver has a unique view to the track environment from the train cockpit. The drivers are concentrated on the driving task and observing the environment. Therefore, in spite of the rapid speed, the driver can react promptly to potential threatening situations in the environment. Any other actor in the metro system would have difficulties in achieving such a fast reaction as the driver, for example, in case of obstacles in front of the train. Therefore, even though there would be some supervisory staff patrolling the train, they might not be able to react early enough to stop the train before an obstacle.

The Helsinki metro track includes outdoor-portions, which are a great source of uncertainty due to several reasons. First, the Finnish four seasons imply varying weather conditions with high rainfall and snowfall. The drivers themselves highlighted the importance of having an intuitive feel of a long train and its behaviour in bad weather conditions. In an automatic system, accounting for all possible weather conditions can be problematic. Helsinki's forthcoming automated metro will be the most northern automated metro in the world and the technology will be put to test in the difficult weather conditions, especially during winter.

"Frost and snow affect a lot of things, you can, for example, never be sure that the railway points turn correctly during winter". –Traffic Controller

In addition to taking into account the weather conditions, driving in the outdoor-portions involves observing the traffic on wayside roads and predicting potential incidents. One participant of our studies reported of a situation where the driver had succeeded in doing a panic stop when a truck drove off an over-head bridge and crashed down onto the track. Similarly, a driver can spot a vandal dropping objects onto the track, make a decision whether it is needed to stop the train, and report the location of the vandal immediately to the security personnel. Furthermore, a driver can also aptly identify and react to intruders climbing over safety fences or otherwise wandering on the track. In these kinds of situations it is beneficial to have the human operator inside the cockpit to interpret whether the situation is dangerous or not and conduct the needed actions. With the automated metro in Helsinki, the current plan is not to have any intrusion detection systems because of the harsh weather conditions, which would not allow the systems to operate correctly. Therefore, in case of an obstacle or person on the track, the train will just keep on going and crash into the object, disregard of its nature and size.

A few of the metro train drivers believed that error-proneness of the automatic system will increase when the drivers are no longer present in the trains. Currently, a driver can solve quickly and easily small technical faults, such as train door problems, which are quite common (for annual amounts, see Table 2), especially during winter. With an automatic system, the small technical problems with one train could cause the entire automatic system to halt while waiting for repair coming from elsewhere. Therefore, in the future the need for technical reliability grows because there are no longer drivers to fix the small problems fast on-site.

”Bigger problems occur almost every week. They originate from either technology or passengers. Most of the time it’s metro technology, door problems, or the train just stops working”. –Traffic Controller-Driver

“I would say that the driver has been the number one in fixing those technical problems already on the spot”. –Traffic Controller

”There are a lot of errors where I only inform the traffic control that I am going to go and fix the problem”. –Driver

Therefore, the driver has a significant role in keeping a practical touch also for the other actors of the metro system regarding the equipment in the field and to the surrounding environment. The driver can also observe potential risks beforehand, such as little technical problems growing bigger, decaying infrastructure, or holes in the fences of the outdoor-portions.

Furthermore, the rail greasing method in Helsinki’s metro is different from the systems used in other countries. Driving on a newly greased track requires train-controlling expertise from the driver. If the automated system cannot take the greasing into account after it has been applied, trains might not be able to brake in time before the correct section at the station.

3.5.2. Stopping at a station

While observing the drivers at their work, we noticed the vast amount of information the drivers receive when the train arrives at a station. The drivers could tell at once if there is something or someone looking abnormal on the station among the passengers. The reason for this is that the drivers have seen the same arrival situation numerous times in their work and their practised glance saw immediately if something was out of ordinary. This phenomenon was confirmed at the mirror data workshop as we showed the participants a photo shot from the train cockpit when arriving at a station. The participants of the workshop could tell several factors, which affect the situation. When approaching a station, the driver can detect a person standing on the danger zone of the platform or otherwise acting in a threatening way. In addition, for example, children, youngsters, and visually impaired or drunken people were said to constitute the passenger groups that need to be monitored with special care. Furthermore, passengers with skateboards, animals, wheelchairs, pushchairs, bicycles, or walking frames were said to stand out from the crowd.

The driver can therefore predict possible problems deriving from the needs or problems related to certain passenger groups while stopping at a station. The issue, which needs to be considered before withdrawing drivers from the trains, is how this can be done when the automatic system is in use.

Currently, a person can get stuck between the closing doors of the train because the doors will not open again after they have reached the proximity of 10 cm from each other. This causes situations where, for example, passengers' wrists, ankles, or personal objects can be left between the doors while the train departs from the station. The driver is the one who can recognise and act in these situations by looking at the back mirror from the train's cockpit and reopening the doors manually. Many participants of our studies were worried about how does the automatic system work if the doors close and the system fails to recognise a passenger stuck accidentally between the metro train doors.

In addition to unintentional problems caused by passengers, an onboard driver is able to respond instantly to intentional small-scale misbehaviour on the platform (e.g., a passenger detaining the train by forcing the doors to stay open) by sounding the train's horn or announcing a warning through the loudspeakers. While watching the drivers' normal driving and listening to their comments about departing a station, it became obvious that some of the metro passengers of Helsinki have a habit of trying to rush between the doors when the alarm indicating door closing is heard. Furthermore, for example, sometimes especially young people try to keep the doors open so that their late-coming friends can still reach the train.

In the planned automation concept, there will be additional doors installed to the platforms next to the track. These platform doors, which are said to be more sensitive than the train doors, and can therefore increase the number of previously mentioned problems. Other problems with the platform doors can be caused by snow, ice, and sand being stuck between them. If the doors are reopened too carefully, the extra stopping time caused by the reopening will directly reflect to the functioning of the entire metro system.

Regarding delays in the current metro system, the metro train driver can minimise the stopping time according to the number of passengers while stopping at a station. This is especially useful if the train is late, since in this way the driver can try to get the train back on schedule. With the automated metro, this activity is not possible, because the minimum stopping time at stations is always fixed.

3.5.3. Passenger care

Fast evacuation of passengers from a metro tunnel can be difficult even with a conventional metro system. If the automated train stops in the tunnel between stations with no driver onboard, problems can arise when evacuation is time-critical. Currently in such situations, for example, in the case of a fire, flood, act of terrorism or a bomb threat, the metro train driver has a critical role in calming down the passengers and escorting them to safety as fast as possible without further danger.

"You have to have patience and make announcements to the passengers so that they don't panic, especially in tunnels". –Driver

In evacuation situations, metro train drivers are familiar with each section of the underground tunnel, and can guide passengers to the nearest exit or emergency evacuation tunnel without hesitation. This knowledge is especially relevant in the Helsinki metro track network since the tunnels lack an evacuation platform next to the track, which would help the passengers with self-evacuation. With an automatic system, evacuation from a difficult location would require sending workers to the train from the traffic control room or the nearest station, which can be time-consuming.

The participants of our studies expressed their concern about leaving passengers under remote surveillance with no onboard staff in the automated metro. In exceptional situations, it is expected that the passengers themselves contact the traffic control about the situation. The participants feared that the quality of information coming from the passengers could be too unclear or imprecise. Passengers of the Helsinki Metro were also considered too negligent to report of small, but relevant incidents that would be reported by the metro train drivers as a part of standard procedure.

”It would feel strange if there were not any staff inside the metro train. It is weird that the passengers would be left all by themselves under remote surveillance”. –Driver

”Passengers want that there is a driver present so that they can focus their frustration at someone”. –Driver

“Is it acceptable to leave it up to the passengers to report of the actual situation in the field?”. –Traffic Controller-Driver

3.5.4. Interactions of the driver with other actors of the metro system

Currently, the metro train driver forms an important link among the different metro actors as can be seen from Fig. 3. For instance, the driver provides the traffic control room operators first-hand information about the status of the train, track, and platforms. This kind of communication from the driver to other actors of the metro system about essential field factors affecting train operation is part of their normal work. Information flows both ways: the drivers can also rely on the supervision of the control room operators in all situations.

In the case of serious accidents (such as derailments or fires), the metro train driver is the primary source of information regarding the situation in the field. In these kinds of situations, the metro traffic controllers will be constantly in contact with the driver to receive expert information about factors affecting the situation on the scene of the accident. This raises the question what will happen when the driver is no longer present there. In the planned automation concept, if an accident happens a specialised team would be dispatched from the traffic control room. Depending on the location of the accident, it could take several minutes until the team arrives. The participants of our studies emphasised that timely expert information from the scene immediately after the accident can be critical. However, a trained team is a better choice in case the driver is severely injured or dies in an accident.

"I would guess that operation in accident situations is considerably more difficult with the automated metro. The driver is our eyes on the field. When there is no longer the driver

present, I guess someone has to go on the scene and see what has happened". –Traffic Controller-Driver

During observation, we discovered that in the more modern train models, the driving system gave many alarms while driving normally on the track. When asked about these alarms, the drivers explained that they are false alarms and the driver does not have to react to them in any other way than merely acknowledge them and possibly inform traffic control. The drivers knew that they were false because they appeared at specific sections in the track. Without drivers, these alarms go directly to metro traffic controllers, who after receiving the alarm might not be aware of or able to infer the exact preceding conditions affecting a particular train. Therefore, making judgements about the falsity of the alarms with the automated metro can be problematic. Even worse, if the automatic system is adjusted to stop all the metro traffic too carefully because of false alarms, it may cause many unnecessary traffic stoppages.

During driver shift changes, we noticed that the drivers exchange information between each other regarding the current traffic situation, possible faults in the system, and other relevant factors related to the metro system. Driver shift changes happen between every two hours with each driver. Although the shift changes are over quite fast, the drivers could still communicate all the relevant information regarding the traffic situation. This kind of unofficial exchange of information regarding the current state of the metro system might be lost with no drivers onboard.

4. CONCLUSIONS

The analysis of core-task demands presented in this article describes the metro domain's and metro train driver work's requirements on a generic level, which is irrespective of the tools used. This kind of formative description works as a framework, which allows the migration from current work to the new one when technological tools are being developed. It is also possible that entirely new demands emerge when tools evolve. It is crucial that this kind of analysis is conducted at the beginning phase of design of complex safety-critical socio-technical systems. If the analysis is excluded, some important safety-supporting functions might be missed and people's lives can be put at risk.

Our results indicate that the metro train driver has a direct view to the track, stations, platforms, and passengers while operating the train. This view gives the driver a unique possibility to anticipate, observe, interpret, and react to events in the surrounding environment. The driver is also an important link to other actors of the metro system, such as traffic controllers and guards. The descriptions of the interactions between the driver and the other main actors of the Helsinki metro system serve as a basis for understanding what kinds of interactions are required to make the metro system function fluently. The illustration of driver's interactions in the metro system make it clear that new arrangements have to be implemented to compensate the lost vantage point of the driver when the automation is introduced. Similar illustrations have also been found beneficial, for example, in case of traditional railways [15].

As can be seen from the results, the metro train driver's working environment is complex and includes more diverse roles than merely the basic driving task. If the current roles of the drivers are not properly understood, it will be very difficult to design, handle, and manage the automation system taking over some of these roles. Therefore, before migrating to a fully automated metro, it needs to be answered how the IT-based automated system can fulfil these roles. Trivial tasks for human can be difficult tasks for the automated system. Furthermore, with safety-critical systems it is important to emphasise that computers and automatic systems excel only at handling monotonous and repetitive basic tasks, which do not require complex problem-solving skills [16]. The fact that the driver is a human, like passengers, means that he or she is therefore a natural interface towards the passengers. The logic and control of the automation system does not apply to passengers, who are independent actors and sometimes behave unpredictably. If the passengers do not understand what the automated system does in different situations, distrust and disuse towards the system can become a problem [17]. It is crucial that in the beginning phases of the use of the future automated metro everything works fluently so that the system can gain passengers' trust.

Methodologically our study entails some limitations. First, the study participants for both the interview and observation studies were chosen in advance by the public transport company and therefore might represent a biased sample of workers. Additionally, only four participants took part in our train driver field observation study, i.e., with more participants additional results could have been acquired.

Moreover, although it has been promised that none of the drivers will be let go, drivers' opinions on the automation might be overtly critical because the change nevertheless renders their profession obsolete and therefore might contradict with their pride in the work. Our results, however, rely only partially on the drivers' interviews, and the possible bias is likely to affect only the interview answers concerning the automation, not the drivers' tasks per se. Additionally, our sample featured drivers with differing attitudes: some expressed very critical stance on the issue while others were moderate.

In this article, we have identified challenges that the full automation of a metro system might bring out. Some of these challenges are related to quality of service (e.g., keeping the train on schedule) and others to safety (e.g., resolving accident situations). It should be noted that safety issues in accident situations are especially challenging to study in the metro and rail domain because accidents happen rarely compared, for example, with road transportation [15]. Consequently, if something happens, the reason for the accident is likely to be very unpredictable. Therefore, it is difficult to decisively determine whether an automated metro is more or less safe than a conventional metro.

One of the findings that can be generalised to other domains is the need for addressing the uncontrollable variables of the domain environment when designing the automated system. In the case of future automated metro of Helsinki these are, for example, harsh weather conditions and unpredictable behaviour of passengers. Safety problems may be expected if automation is viewed as a closed system with disregard to these surprising and uncontrollable variables. In future, this issue might be increasingly relevant if automated systems will be more and more integrated into people's daily environments.

With increasing automation, there is a possibility that the people involved with the process might become too separated from the practical reality in the field. Previous research [10,11] has shown that if the practical skills are not sustained in an environment with high automation level, there is a danger of losing control of the system. Furthermore, it has to be emphasised that increase in automation does not necessarily mean that you need less personnel. Cost-effectiveness measures often presented with automation projects promote that some of the staff can be replaced by automation. However, for example, in the case of an automated metro, it is expected that the amount of supervisory and guidance staff on stations and trains needs to be increased when the drivers are withdrawn from the trains. Yet another issue, which needs to be taken into account, is that in addition to knowledge on the metro domain, the new automated system might require the staff to learn how the automation itself works technically. Understanding this kind of deep technical knowledge requires a lot of special training and sophisticated information systems [10].

Overall, our study suggests that there is more to driving a metro train than meets the eye. We have presented several roles in which the metro train driver serves in addition to the obvious ones. If not accounted for, the existence of these hidden but important roles can cause problems with the operation of the fully automated metro.

ACKNOWLEDGEMENTS

This work was conducted jointly by VTT Technical Research Centre of Finland and Helsinki Institute for Information Technology HIIT in the Amoveo project funded by the Academy of Finland. The authors would like to thank HKL for offering the possibility to conduct the studies.

REFERENCES

[1] J.M. Erbin, C. Soulas, Twenty years of experiences with driverless metros in France, in: Proceedings of the 19th Dresden Conference on Traffic and Transportation Science, Dresden University of Technology, Dresden, 2003, pp. 1–33.

[2] C. Heath, P. Luff, Team work: Collaboration and control in London Underground line control room, in: C. Heath, P. Luff (Eds.), Technology in Action, Cambridge University Press, Cambridge, 1991, pp. 88–124.

[3] P. Smith, A. Blandford, J. Back, Questioning, exploring, narrating and playing in the control room to maintain system safety, Cogn. Technol. Work. 11 (2009), 279–291.

[4] J. Theureau, G. Filippi, Analysing cooperative work in an urban traffic control room for the design of a coordination support system, in: P. Luff, J. Hindmarsh, C. Heath, Workplace Studies – Recovering Work Practice and Informing System Design. Cambridge University Press, Cambridge, 2000, pp. 68–91.

[5] C. Heath, P. Luff, M.S. Svensson, Overseeing organizations: configuring action and its environment, Br. J. Sociol. 53 (2002), 181–201.

[6] J.R. Wilson, B.J. Norris, Rail human factors: past, present and future, Appl. Ergon. 36 (2005), 649–660.

[7] A. Jansson, E. Olsson, L. Kecklund, Acting or reacting? A cognitive work analysis approach to the train driver task, in: J.R. Wilson, B. Norris, T. Clarke, A. Mills (Eds.), Rail Human Factors – Supporting the Integrated Railway, Ashgate, Surrey, 2005, pp. 40–49.

[8] M.P. Georgescu, Driverless CBTC – specific requirements for CBTC systems to overcome operation challenges, in: J. Allan, C.A. Brebbia, A.F. Rumsey, G. Sciutto, S. Sone (Eds.), Computers in Railways X: Computer System Design and Operation in the Railway and Other Transit Systems, WIT Press, Southampton, 2006, pp. 401–413.

[9] S. Vepsäläinen, Automated metro, Series: Thoughts on Developing the Helsinki Public Transportation, C 2/2003, HKL, Helsinki, 2003. (in Finnish).

[10] L. Bainbridge, Ironies of automation, Automatica 19 (1983), 775–779.

[11] C.E. Billings, Aviation automation: The search for a human-centered approach, Lawrence Erlbaum Associates, Mahwah, 1997.

[12] Y. Engeström, Putting Vygotsky to work: The Change Laboratory as an application of double stimulation, in: H. Daniels, M. Cole, J.V. Wertsch (Eds.), The Cambridge Companion to Vygotsky, Cambridge University Press, Cambridge, 2007, pp. 363–304.

[13] L. Norros, Acting Under Uncertainty – The Core-Task Analysis in Ecological Study of Work. VTT Publications, Espoo, 2004.

[14] M. Nuutinen, L. Norros, Core task analysis in accident investigation: analysis of maritime accidents in piloting situations, Cogn. Technol. Work. 11 (2009), 129–150.

[15] A.R. Hale, T. Heijer, Is resilience really necessary? The case of railways, in: E. Hollnagel, D. Woods, N. Leveson (Eds.), Resilience Engineering: Concepts and Precepts, Ashgate, London, 2006, pp. 125–148.

[16] S.W.A. Dekker, C.E. Billings, Humans and the evolution of industrial automation, in: C.E. Billings (Ed.), Aviation Automation: The Search for a Human-Centered Approach, Lawrence Erlbaum Associates, Mahwah, 1997, pp. 51–64.

[17] R. Parasuraman, V.A. Riley, Humans and automation: Use, misuse, disuse, abuse, Hum. Factors. 39 (1997), 230–253.

Figure 1 (color)
Click here to download high resolution image

Figure 2 (color)

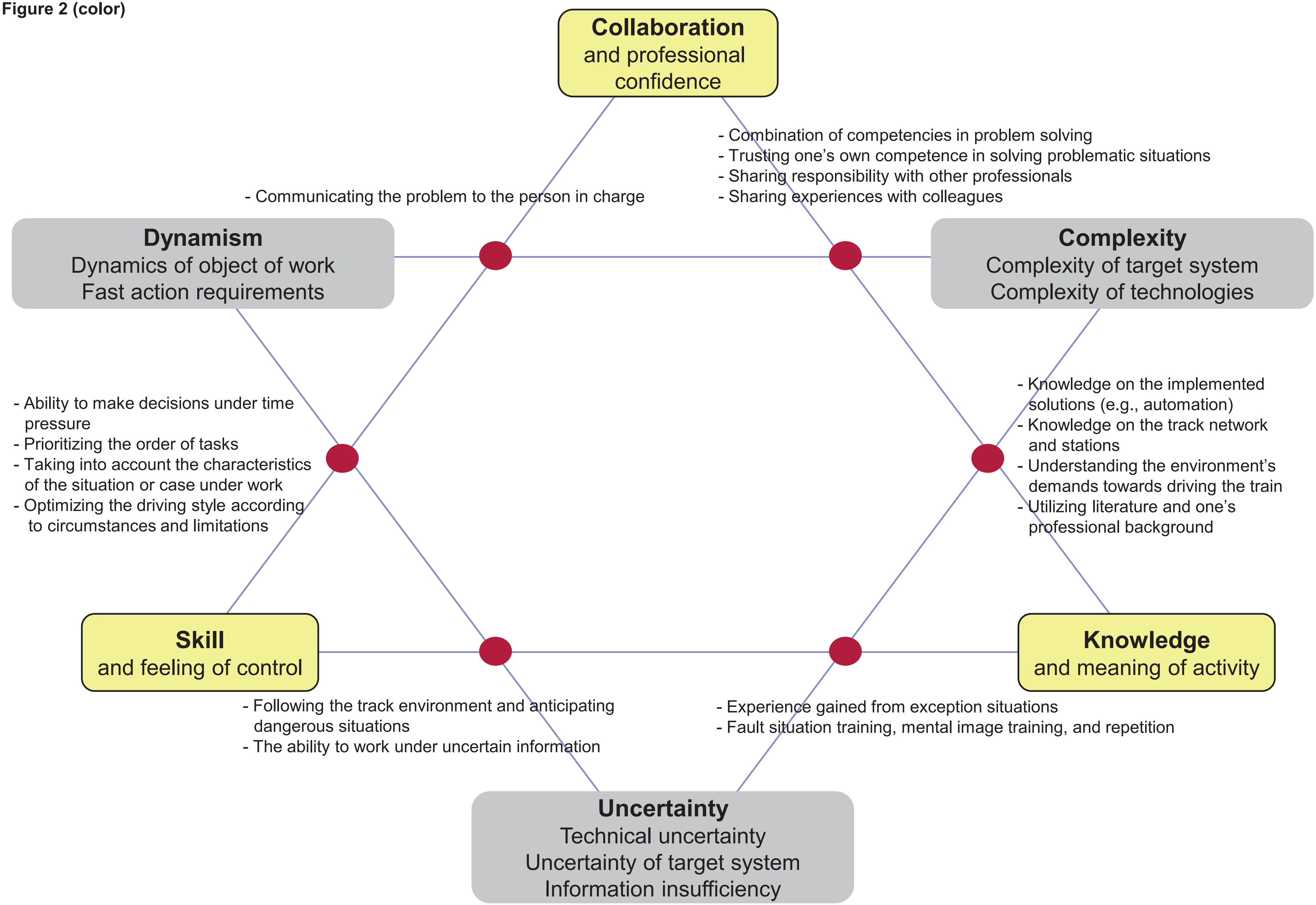


Figure 3 (color)

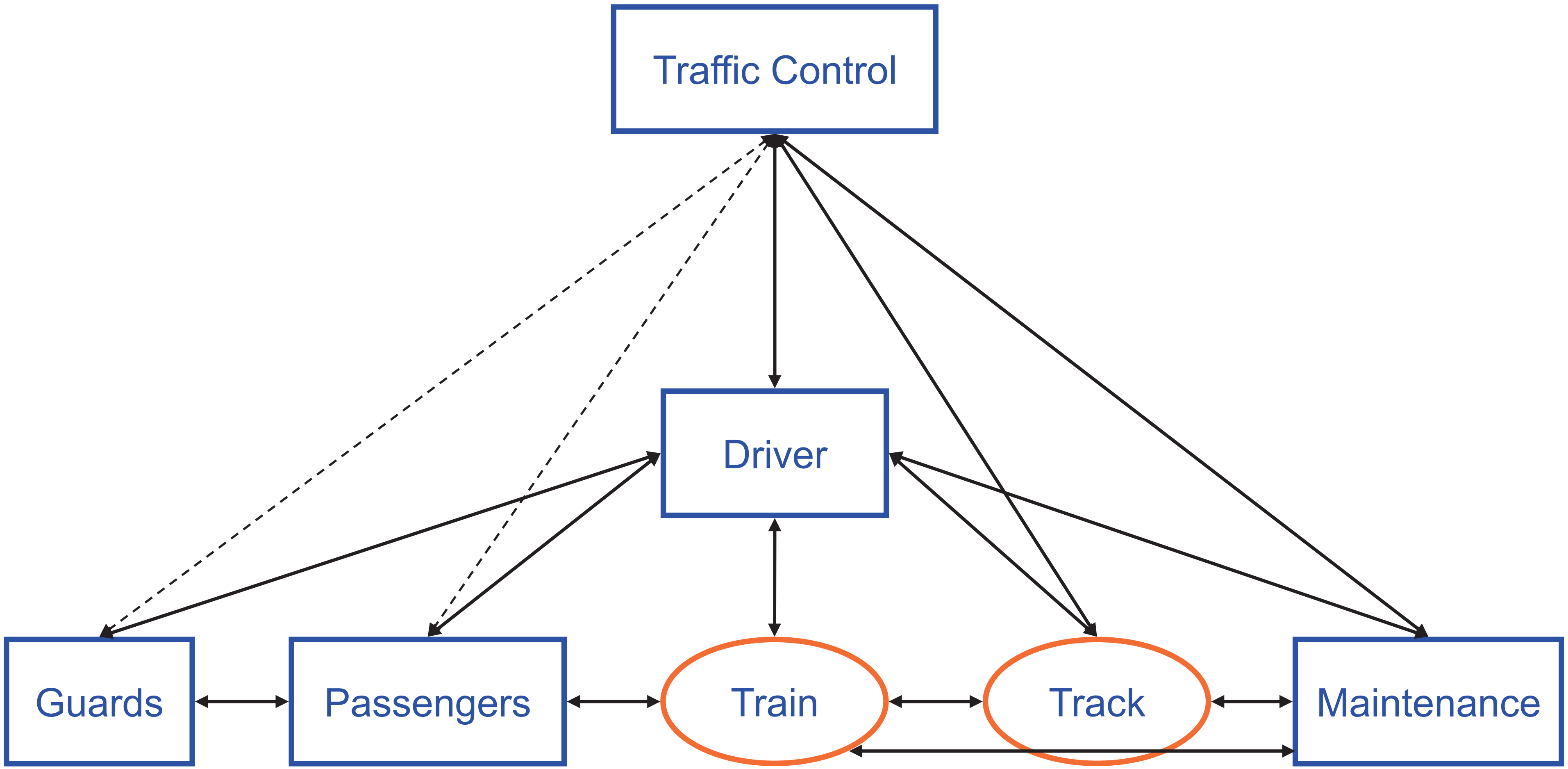


Table 1

Dynamism	Uncertainty	Complexity
Dynamics of object of work - Traffic has to run according to the schedule → certain tasks are repeated regularly - The trains run at predefined speed, the “driving profile” specifies the pace - Exceptions happen rarely → regularity and scarcity of events cause monotonicity - There is more traffic and passengers during rush hours → effects of this to workload - Technology faults and the surrounding environment cause delays	***Technical uncertainty*** - Small objects on the track cause false alarms - Faults in the interlocking system - The functioning of the current automation - Expiring technology, e.g., train flaws - Faults in devices, dysfunctional communication connections - Power failures - Usability problems	***Complexity of target system*** - Track network is not very complex - Driving the train is not difficult - Exceptional situations are rarely similar → there are not necessarily ready-made procedures or routines - Exceptional situations require taking care of both the traffic and passengers - Driver has to know which partner has the responsibility over what issues - Driver has to know environmental factors affecting driving the train - Over time, there has been an increase in technologies used, number of stations, track length, and the amount of passengers
Fast action requirements - Exceptional situations require fast reaction to prevent the traffic from jamming → the traffic has to be normalized quickly - In accident situations, passengers must be escorted to safety quickly - Getting help to the scene of the situation takes time depending on the location - The driver has to be constantly alert inside the train’s cockpit because of the continuously changing environment	***Uncertainty of target system*** - Passenger behaviour - Mischief, obstacles on track - Varying weather conditions - Fires	***Complexity of technologies*** - There are differences in the operation of different train models - Driver is able to fix small technical problems fast on-site in the train
	Information insufficiency - Fault duration is hard to predict - It is difficult to develop a routine for certain exceptional situations because they are rarely similar	

Table 2

Exceptional situation	**Cause**	**Disruption Duration**	**Driver action**	**Additional comments**
Pulling of an emergency handle	Passenger	3–10 minutes	Checks the track and acknowledges the emergency handle	In 90% of cases, the pulling of an emergency handle is either mischief or the ignorance of children
Passenger casualty	Passenger	30–75 minutes, causes also cancellation of departures	Contacts traffic control	Rescue department and police always come to the scene
Passenger preventing the train doors from closing	Passenger	Time it takes to reopen and close the train doors	Makes an announcement to the passenger	Preventing the doors from closing can also cause a technical fault in the door. Then, the disturbance will last about 4–7 minutes.
Passenger seizure	Passenger	A few minutes	Calls for rescue help	-
An unauthorised person on rails	Passenger	3–35 minutes	Stops the train if necessary; contacts traffic control	-
Technical train door problem	Train faults	From a few minutes to a cancelled departure	Contacts maintenance; fixes the door problem if possible	Most common train fault
Train has no drive	Train faults	From a few minutes to a cancelled departure	Contacts maintenance	Second common train fault
Train's brakes don't come loose	Train faults	From a few minutes to a cancelled departure	Contacts maintenance	Third common train fault
Remote control fault	Security device faults	0–20 minutes	No driver action: a traffic control problem	-
Rail circuit fault	Security device faults	3–15 minutes	No driver action: a traffic control problem	-
Turning point fault	Track maintenance faults	5–30 minutes	No driver action: a traffic control problem	Heavy snowfall in the beginning and end of year 2010 caused a lot of trouble for the traffic: turning point faults, rail circuit freezes and technical door problems

Table 3

	2006	2007	2008	2009	2010
Pulling of an emergency handle	140	107	114	134	126
Passenger casualty	5	6	6	4	3
Passenger preventing the train doors from closing	12	16	23	12	24
Passenger seizure	2	2	3	12	15
An unauthorised person on rails	23	47	51	38	46
Technical train door problem	63	64	40	43	59
Train has no drive	18	13	13	8	26
Train's brakes don't come loose	10	8	6	11	12
Remote control fault	6	13	2	8	11
Rail circuit fault	30	23	13	14	10
Turning point fault	4	6	8	4	10

Table 4

#	Subtask	Contributes to main task(s)	Situation/Time	Precondition(s)	Loops
1.1	Driving preparations	1	Every time at a terminal station		-
1.2	Driving the train on the track	1 & 2	Whenever the train is moving	After subtask 1.1	-
1.3	Stopping at a station	1 & 2	At every station	After subtask 1.2	-
1.4	Opening and closing the train doors	1 & 2	Whenever the train stops at a station	After subtask 1.3	After subtask 1.4, goes back to subtask 1.2 and loops until the terminal station
1.5	Changing the direction of the train	1	Every time at a terminal station	Subtask 1.1 and looping subtasks 1.2.–1.4. at every station	Performed at each terminal station
1.6	Decoupling the additional train cars	1	After afternoon rush hour	Arrival to a terminal station	Daily with each train
2.1	Keeping the train on schedule	1 & 2	All the time	-	-
2.2	Making announcements to the passengers	2 & 4	When necessary	Something worth announcing appears	-
2.3	Answering the train's emergency calls	2 & 4	When necessary	Passenger uses the emergency phone in the train	-
2.4	Guiding passengers out	2 & 4	When necessary	Passengers are in danger or the train breaks down	-
3.1	Interpreting events in the environment	3	All the time	Events happening in the environment	-
3.2	Reacting to events in the surrounding environment	3	When necessary	Something worthwhile reacting to in the environment	-
3.3	Co-operating with guards	3	At stations when necessary	Guards are on the platform or boarding the train	-
3.4	Co-operating with ticket inspectors	3	At stations when necessary	Ticket inspectors are on the platform or boarding the train	-
4.1	Reacting to exceptional situations	4	When necessary	An exceptional situation occurs	-
4.2	Acknowledgement of faults	4	When necessary	A fault in the train occurs	-
4.3	Contacting traffic control	4	When necessary	Any exceptional situation worthy of reporting occurs	-
4.4	Contacting maintenance	4	When necessary	Some part of the train breaks down	-
4.5	Fixing small faults in exceptions	4	When necessary	A small fault such a train door problem occurs	-

II

ENHANCING THE USER EXPERIENCE OF THE CRANE OPERATOR: COMPARING WORK DEMANDS IN TWO OPERATIONAL SETTINGS

by

Hannu Karvonen, Hanna Koskinen, & Jaakko Haggrén

In Proceedings of the 30th European Conference
on Cognitive Ergonomics (pp. 37–44)

Enhancing the User Experience of the Crane Operator: Comparing Work Demands in Two Operational Settings

Hannu Karvonen[1]
VTT Technical Research Centre of Finland
P.O. Box 1000
02044 VTT, Finland
hannu.karvonen@vtt.fi

Hanna Koskinen[1]
VTT Technical Research Centre of Finland
P.O. Box 1000
02044 VTT, Finland
hanna.koskinen@vtt.fi

Jaakko Haggrén
Konecranes Plc
P.O. Box 661,
Koneenkatu 8
05801, Hyvinkää, Finland
jaakko.haggrén@konecranes.com

ABSTRACT

Motivation – To understand the different crane operation experiences by analysing the work demands in both conventional and remote operation settings. On the basis of this analysis, we aim to find out ways to enhance the operator's experience of sense of control and feeling of presence when operating remotely.

Research approach – We conducted qualitative field studies in two different operating environments. The studies were based on the Core-Task Analysis method and included altogether 12 operator interviews and observations.

Findings – The results suggest that remote operation brings more uncertainty factors to the operator's work, whereas in the conventional cabin operation setting there is a stronger emphasis on dynamism. Based on the field studies, two user experience targets – sense of control and feeling of presence – were chosen and the design implications of these user experience targets for the development of a new remote operation station were elaborated. We suggest that in the design of remote crane operation solutions special attention should be paid for example to the creation of a comprehensive and coherent operating view as well as the development of a rich and realistic feel of operation.

Take away message – Remote crane operation system should provide the operator with an enriched hands-on experience to the crane on the field.

Keywords

Container crane operation, conventional cabin operation, remote operation, Core-Task Analysis, Systems Usability, user experience target

1 INTRODUCTION

Standard containers are an essential part of international sea freight transportation (Steenken, Voss & Stahlbock, 2004).

ECCE '12, August 28 - 31 2012, Edinburgh, United Kingdom

Containers are transhipped onwards between different modes of transportation (e.g., from cargo ships to road trucks and vice versa) in container terminals (CTs) located at ports. This container handling requires novel solutions from the terminal operator in order to answer to nowadays increasing efficiency, safety, and employee well-being demands.

As a one solution to these needs, more and more CTs are taking new remote operation systems with automated functionalities into use. One of the advantages of remote crane operation is that it can increase productivity, because one human operator can handle multiple cranes remotely. Another important benefit from the crane operators' point of view is that the level of work safety and ergonomics increases as the operators do not have to be physically inside the crane's cabin and also move around in the port yard. Instead, they can control the crane from a distant office environment, which is more safe and comfortable compared to the conventional cabin operation.

However, the introduction of remote crane operation solutions can include several dangers as well, especially from the cognitive ergonomics and safety perspective. For example, the remote operator has to rely on limited video feeds as their main operating view to the container loading zone, which can affect the detection of dangerous events considerably. In addition, the remote operator cannot realistically feel and hear what happens within and around the crane one is operating. This kind of a comprehensive multimodal on-spot experience is only possible in conventional crane operation, where the operator is inside the cabin that is located in the crane structure.

In the research presented in this paper, we have studied container gantry crane operation both in conventional and remote contexts. We conducted interview and observation studies based on the Core-Task Analysis (CTA) method (Norros, 2004; Norros & Norros, 2009) in two different international CTs: one with conventional rubber tired gantry (RTG) stacking cranes and another with remotely operated automated stacking cranes (ASCs) in use. The objective of these studies was to understand the factors affecting the crane operators' user experience (UX) by analysing the work demands in both environments. This analysis was needed in order to understand how the UX of future remote operation systems could be improved. Based on the studies' results, also UX targets (see e.g., Hartson & Pyla, 2012) for guiding the design could be defined.

2 BACKGROUND

The development of modern remote operation (i.e., teleoperation) systems started in the mid-1940s, when Goertz developed the first mechanically controlled master-slave teleoperator (Sheridan, 1989). Since then, a plethora of different teleoperation applications have been developed from remote mining (e.g., Hainsworth, 2001) to space operation systems (e.g., Sheridan, 1993).

In general, these kinds of remote manipulation systems are useful in safety-critical environments where it is beneficial for the human operator to be interacting with the object environment from a safer and/or economically more reasonable location than on the spot. In recent years, teleoperation has been taken into use in some CTs' container handling as well (see e.g., Speer, John & Fischer, 2011).

Containers are handled in CTs with different kind of equipment. CTs with manned equipment can include for example regular quay cranes (see the three large cranes next to the ship in the top part of Figure 1), rubber tired gantry cranes (see e.g., the smaller crane lowering a container down on a truck chassis in the foreground of Figure 1), straddle carriers (see the equipment in the far left part of Figure 1), and lift trucks (see the top-lift spreader equipment operating a container in the bottom part of Figure 1). Once a container arrives to a CT, it is unloaded and usually moved to wait for some time inside the port for shipment onwards. On the port yard, the containers are stored in tightly ordered stacks (see the visualisation of lines of white boxes in Figure 1) where the gantry cranes can move. From the stacks, they are loaded for example to road trucks, ships, and trains for further transportation. Both the stack operation and unloading and loading of containers require accuracy from the used gantry and the crane operator.

Conventionally, container handling equipment is operated from a cabin that is located in the equipment. The actual operation is carried out with a combination of joysticks and steering wheels, depending on the used equipment and task that is worked on. In this study, we concentrated on RTGs, which are operated with two joysticks and dedicated buttons and switches in the control consoles of the cabin. The joysticks are for controlling the main movements (hoisting, trolleying, and gantrying) of the crane and the buttons are used for secondary controls (such as operating the driving lights). Biggest challenges for crane operators working from the cabin are the non-optimal working ergonomics (because of the downright working posture), physical stress from the noisy and bouncing cabin, limited visibility from the cabin, and weather-related problems (sun, heat, moist, cold, etc.). All these challenges are exacerbated when the crane operator needs to be inside the cabin for long periods of time.

Nowadays, in some CTs also unmanned machinery is in use (for more information see e.g., Speer, John & Fischer, 2011). This means that there is no human operator inside the machine and it can be operated either manually from a remote location or be fully automatic. For example, automated stacking cranes and automated guided vehicles (AGVs) are already commonly used machinery in CTs where parts of the container handling process are automated (e.g., Vis & De Koster, 2003). In this paper, we focus on ASCs and in particular their remote operation from the perspective of the remote operator's work.

ASCs are semi-automated rail-mounted gantry cranes, which can operate only inside one specific container stack. The cranes are automated to a certain extent, for example to handle the water-side loading zone and stack operation automatically without human intervention in normal operating situations. In this kind of a setup, the cranes are operated manually only during loading and unloading of external road trucks and other type of chassis in the land-side loading zone (see the fenced area in the middle foreground of Figure 1).

Instead of the crane operator being physically in the crane's cabin, the manual operation is in today's ASC solutions carried out from a distant control room in an office environment. In these control rooms, several remote operator stations (ROS) are located. During one working shift, each operator operates their own ROS. One operator can connect through her ROS to any of the cranes in operation. The ROS includes a control console and related software user interfaces. Furthermore, the ROS has both displays for receiving information from the field and different physical user interfaces (UIs) for controlling the crane remotely.

The concept of operation (see e.g., Karvonen et al., 2011) in road truck loading and unloading between different CTs can be somewhat different. In the conventionally operated CT we studied, the truck loading and unloading was carried out on a separate truck lane along the side of the stack (i.e., side-loading). The truck driver positioned the truck to the correct place in (non-verbal) collaboration with the crane operator before loading or unloading (with horn signals). In the semi-automated remote operation CT of our study, the loading zone was at the land-side end of the stack where the trucks backed up and parked to the lane without contact with the remote operator (i.e., end-loading, see the end of the stacks in the foreground of Figure 1).

Figure 1. Visualisation of a container terminal layout and used machinery

3 METHOD

The purpose of this research was to find out the core-task demands of conventional and remote container crane operation by studying the crane operator work in both of these contexts. From these results, we strive to identify implications for the design of remote crane operation workstations, which can to a certain extent be applied to any remote operation application. On the basis of this analysis, we aim to find out ways to enhance the crane operator's experience of sense of control and feeling of presence when operating remotely.

3.1 Theoretical background

Our theoretical background includes influences from both cognitive engineering (see e.g., Rasmussen, 1986) and activity-oriented approaches (see e.g., Engeström, Miettinen & Punamäki, 1999). We see constraint-oriented analysis of work domain as an important part of design of working processes, in the same vein with Vicente (1999). Leaning on these approaches, and also to the American pragmatist tradition (Peirce, 1998), we have developed a method called Core-Task Analysis. CTA aims to identify the core task of a specific work. Core task refers to the main result-oriented content of work that can be derived by analysing the objective of work and the demands that the objective lays on workers both in general and in specific situations (Norros, 2004). Core-Task Analysis has been used earlier as an analysis tool in several domains, for instance, in nuclear power plant operation (Norros, 2004), maritime piloting (Nuutinen & Norros, 2009), and metro train driving (Karvonen et al., 2011).

According to the CTA method, a safety-critical domain is characterised by control demands related to dynamism, complexity, and uncertainty. Factors related to dynamism are associated especially to work's temporal demands, such as timing, duration, and delays. Complexity related factors originate for example from the large number of different elements in the system that have more or less interaction with each other. Factors related to uncertainty arise, for example, from the used technology and the surrounding environment (Norros, 2004; Karvonen et al., 2011).

When the above-mentioned control demand factors have been identified, it is possible to recognize the core-task demands of the work. Core-task demands refer to the domain's critical functions, which the workers must fulfil. In addition, they work as functional requirements that manifest themselves in specific forms in particular situations and set constraints and possibilities for tool-using interactions (Norros, 2004). We explain the usage of this framework in detail later in the Results section of this paper, where we use the method to present the results of our crane operator studies.

4 DATA COLLECTION

4.1 Interview studies and field observation

We conducted interview and observation studies in two different international container terminals during a one week period. Overall capacity (throughput) of the first CT we visited was approximately 2.5 million TEUs (twenty-foot equivalent unit, i.e., one 20 ft. container) per year. In this CT, the container unloading, loading, and stack operation on the port yard were carried out on-spot with conventional cabin operation. In this CT, we interviewed and observed six crane operators (ages between 30 and 59 years and working experience in container handling between five and 35 years) from which three were working in day shift and three in night shift. The operators worked in eight hour shifts around the clock. We interviewed and observed each crane operator individually. The interviews were organized in a separate office building and afterwards it was possible to go to the crane's cabin with the interviewed operator to observe her during her regular working. Three crane operators could be studied simultaneously as each of the researchers had an assigned operator from day and night shifts.

The overall capacity of the second studied CT was approximately one million TEUs per year. We conducted the studies in this CT after two days from the first CT studies. In the second CT, the land-side loading and unloading of containers was carried out by the crane operators through a remote operation system with no direct eye-sight to the crane. In normal situations, the stack and water-side operation was carried out automatically by the cranes. However, in case of an exceptional situation (e.g., a sensor failure), the automated system handed over also the stack operation task to be carried out manually by the remote operator. In this CT, altogether five day shift remote operators and one maintenance operator (ages between 32 and 49 years and working experience in container handling between five and 22 years) were interviewed and the operators' daily working in a centralized control room was observed.

All the interviews were conducted as semi-structured theme interviews. The first part of the interview focused on themes such as understanding the general characteristics of the work and the working environment as well as the core task of the operator. In the second part, the interview continued on a more detailed level into challenging operating situations and the information needs and co-operation and communication demands in those situations. This part of the interview was facilitated with printed colour photos, which showed the different tasks and phases of a normal operation task. In this connection, also the theme of user experience of crane operator's working environment and tools were elaborated in more detail. Under each theme, several more detailed questions were addressed, but the discussion was not restricted to only predefined questions.

While most of the questions in the interviews were based on the CTA method, the Systems Usability Framework (Savioja & Norros, 2012) offered a wider background for the user experience related questions. Overall, the interview consisted of 34 predefined questions and each interview was estimated to take approximately 1½ hours. The interviews were audio recorded and also notes were written up during the interview for later analysis purposes.

In the conventional setting, a 1½ hour field observation session during operator's normal working was conducted after each interview. In the remote operation setting, the observations happened more sporadically over 2 days. During the observation, the crane operators were encouraged to think aloud while operating and for example go through and explain the controls they were using in more detail. While the crane operator was doing her work, it was also possible to ask questions which were still left open from the interviews. The observations were freehand video recorded for later analysis purposes.

5 ANALYSIS

After the studies in both container terminals, the interviews were transcribed and notes were written up. In addition, relevant data from the video material was sought for. In line with the CTA method, we extracted from the transcriptions the crane operators' control demands related to dynamism, complexity and uncertainty in both conventional and remote operation settings. Then, we identified the core-task demands of both settings and compared them with each other. Finally, we drew design implications for remote crane operation systems

6 RESULTS

The main results of the study presented in this paper are organized below in the following order. First, we present the control demand categories of both environments. Second, we explicate separately on a general level the characteristics of both conventional cabin operation and remote operation based on the control demands. Third, we describe the core-task demands of crane operator work in both domains and present the differences when moving from conventional to remote operation. Fourth, we suggest design implications as a basis for developing remote operation solutions in CT context. These design implications can be applied to a certain extent to remote applications in general as well.

6.1 Control demands

Based on the analysis, the following control demand categories in crane operation related to dynamism, complexity, and uncertainty were identified:

- Dynamism
  - Dynamics of object of work
  - Setting the pace for one's own work
  - Fast action requirements
- Complexity
  - Complexity of technologies
  - Complexity of object of work
- Uncertainty
  - Technical uncertainty
  - Environment related uncertainty
  - Uncertainty related to other actors
  - Information insufficiency

Under each of these categories, we recognized several control demands in both settings, which are not presented here in detail due to space limitations. Instead, we discuss next the findings regarding these demands on a general level in both settings.

6.2 Characteristics of conventional cabin operation

Based on the CTA, it was discovered that the conventional cabin operation was characterized strongly by dynamism. Even though the complexity and uncertainty related factors are also present, the control demands related to dynamism seem to be more dominant.

There are many reasons why dynamism is emphasised in the results. First of all, the variety of operations and tasks that are in the line of duty for conventional cabin crane operators is more diverse compared to remote operation. In addition to the loading and unloading operations in the truck lane, the crane operators are responsible also for operations in the stack side, i.e., locating, picking up, and sorting of containers. This means that the operators need not only to master more extensive variety of operational phases and control techniques, but also move from one operational phase to another more often. Furthermore, control demands for the operations in the truck lane differ greatly from the stack operations.

Another factor that creates dynamism in the conventional cabin operation is the amount of other actors and traffic in the port area that the crane operator needs to pay attention to. During the day time there can be a lot of external traffic in the port and they might not always be aware of the port-specific rules and ways of working. There are also ground personnel working around the cranes as well as other cranes and equipment moving around the port yard, which need to be looked out for.

Lastly, the crane operators seemed to be very well aware that the port's objective is to be as productive as possible and they wanted to fulfil this objective in their work. To achieve efficient numbers, they for example tried to optimize the movements of the cranes by multitasking in different situations.

6.3 Characteristics of remote operation

In the analysis of remote operation's control demands, it was noticed that in the remote operation, there seems to be more emphasis on the uncertainty factors compared to dynamism and complexity related factors. There are several reasons for this, from which we mention here only a few.

First of all, the scope of operations carried out through the remote operation is diminished and not very complex. The main work activity contains only remote loading and unloading of containers in the land-side loading zone, where the crane is located. However, sometimes exceptional situations can get quite complicated and are difficult to address remotely. Dynamism and fast action requirements were not emphasized by the remote operators to be characteristic for their work. On the contrary, they seemed to always emphasize the importance of calm and safe operation, no matter how much traffic pressure there is on the port. One reason for the strong emphasis on uncertainty factors in the interview answers is that the remote operators have more qualms when operating the crane remotely through computer displays with only live video feeds from the loading zone as their main operating view. Thus, during operation they are not physically on the spot where they could have direct eye contact and see everything what happens in the loading zone.

6.4 Core-task demands

Figure 2 presents the means (light yellow boxes) for managing the control demands related to dynamism, complexity, and uncertainty (their categories in crane operation are presented in the dark grey boxes). These demands are managed mainly by skill, knowledge, and collaboration. When the associated control demands are examined through these means, the general core-task demands of the work emerge. The numbered dark red dots in Figure 2 refer to these core-task demands. The titles of these core-task demands are based on the CTA method and are presented in Table 1 with the number references. In addition, Table 1 uses summarized lists to present what the core-task demands mean in detail in both conventional and remote crane operation settings.

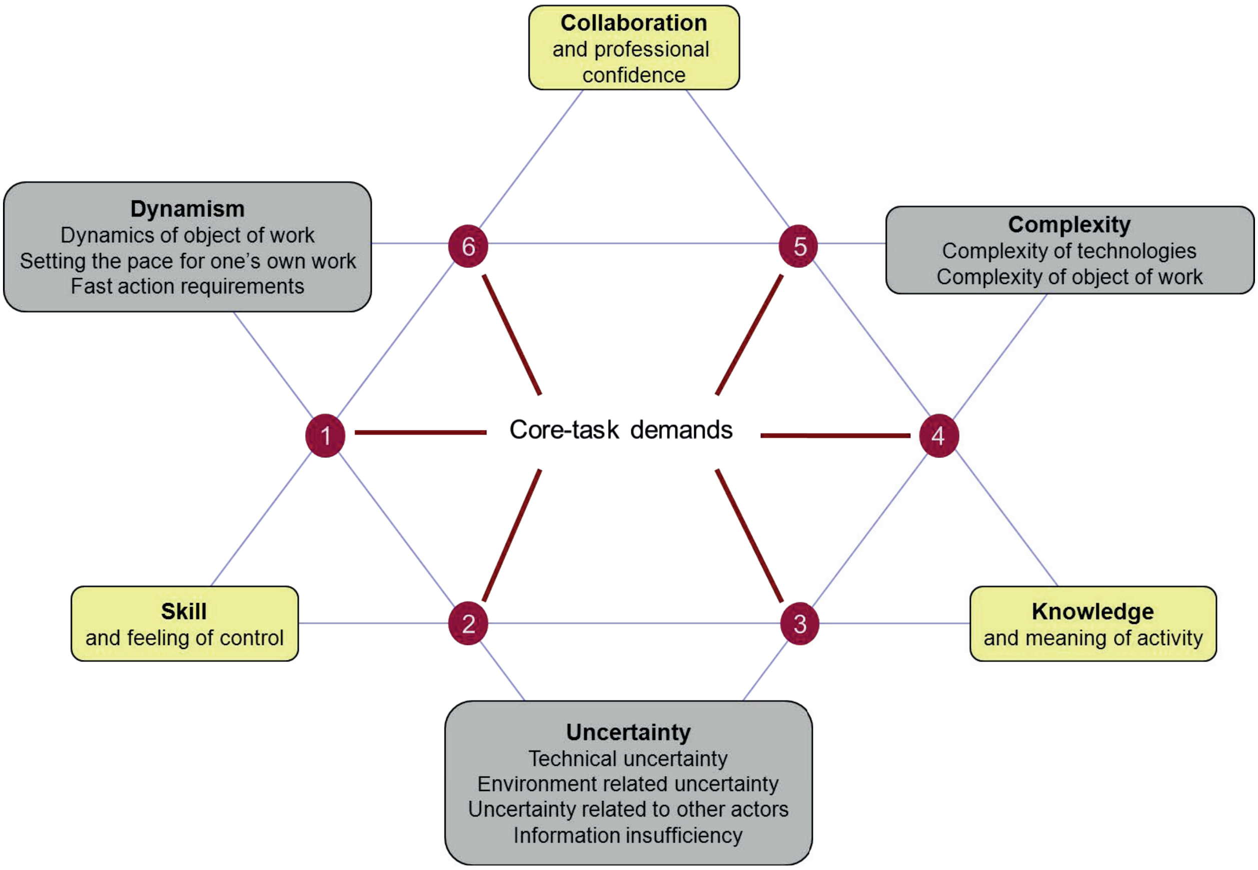


Figure 2. A core-task model of the generic environmental constraints on action. The model includes control demands (the dark grey boxes), means of managing them (the light yellow boxes), and the associated core-task demands (the dark red dots, see Table 1 for descriptions).

Core-task demand	Conventional cabin operation	Remote operation
1. Readiness to act	• Ability to handle a broad variety of situations • Good depth perception to understand the distances while operating • Anticipatory activity to optimize the work flow (e.g., "digging" before a truck arrives)	• Sensitivity for characteristics of different operational situations viewed from the ROS • Good hand-eye coordination and spatial perception for adjusting operation and reaction to operational events • Confidence and calmness in ROS operation
2. Flexibility of acting and reorienting	• Mastering certain physical techniques (e.g., "doglegging" when not being able to see sufficiently) to operate safely and smoothly • Being alert and aware of the crane's surroundings on the ground all the time • Use of reference points (in the port yard) in operation for positioning	• Mastering special controlling techniques with the ROS joysticks and routines of remote operation in a variety of operational situations (e.g., lift-and-shift) • Anticipation and detection of relevant (e.g., safety and production related) information for operation • Ability to re-orient and focus quickly on the specific characteristics of the received new task
3. Interpretativeness of acting	• Sufficient training and operating experience on crane's behaviour and possible exceptional situations • Knowledge on when to commit anticipatory actions (e.g., slowing down) in order to ensure safety • Cross-checking and confirmations are needed once the job is received	• Sufficient training and knowledge on the technical implementation's restrictions on container handling (e.g., regarding override functionalities) • Knowledge on how should a normal operational situation look like (in order to detect deviations) • Utilizing prior ASC problem-solving experience and other crane operating experience in exceptional situations
4. Conceptual mastery of crane operation	• Knowledge on the crane's technical solutions (e.g., assisting functionalities) • Knowing how to apply the operating instructions correctly • Understanding the environment's demands and related physical forces in operation	• Knowledge on the implemented technical solutions in remote operation (e.g., possible delays) • Knowing how to apply the operating instructions correctly • Understanding the environment's demands and related physical forces in operation
5. Creating shared awareness	• Radio communication with other operators in order to maintain situation awareness • Clear and understandable radio communication • Informing maintenance about different technical problems	• Sharing experiences with the colleagues in the control room on complex operating situations • Clear and understandable communication with different stakeholders in exceptional situations in order to confirm safe operation • Combination of competencies in problem solving
6. Optimal sharing of efforts	• Collaboration and communication with the truck driver is essential for fluent work flow • Asking for help from nearby crane operators in busy and demanding situations • Trusting one's own competence in solving problematic operating situations	• Communicating an occurring problem to the correct person in the organisation • Sharing responsibility with other operators in the control room in order to optimize the work flow • Trusting one's own competence in solving problematic operating situations

Table 1. Comparison of summarized core-task demands of conventional and remote crane operation

Change from cabin to remote operation

In the following, we aim to explicate how the stacking crane operator work differs when carrying out container crane operation remotely compared to conventional cabin operation and highlight the main differences between these two. We do this by providing practical examples regarding each core-task demand from the studied environments.

1. Readiness to act: One of the most important things in container handling throughout the whole CT is to know the exact location of each container. This is also something what the crane operators are trying to do when carrying out their part of the operation; the container is all the time viewed in relation to its surroundings. To have and maintain an understanding of this relation is a prerequisite for making any control actions. In the conventional cabin operation the crane operators are constantly utilizing different references points on the container and the surrounding stack environment in order to make estimations about the distances and spatial dimensions. This is also something that the remote operator is trying to do through her ROS. However, because of the limitations of the ROS's video feeds (e.g., reduced spatial information: the depth perception is limited by the quality of the video feeds), the crane operators are not always able to utilize the same reference points as in the conventional operation. Thus, in remote operation, the operators need to learn new ways to look, perceive, and orient towards the operating environment.

2. Flexibility of acting and reorienting: As previously mentioned, the remote operator carries out mainly unloading and loading operations. Making a one move takes on average only 30 seconds. If it is a busy situation in the port, the operator might receive the tasks in a constant flow of loading and unloading operations. In the conventional cabin operation, carrying out a one move contains more operational phases than in remote operation and the crane operator handles the container all the way from the stack to the truck and vice versa. In this kind of operation, he is active in the whole process and has for example more time and insight to anticipate the upcoming situations. Therefore, in the remote setting it is essential that the operator has an ability to reorient and focus on the specific characteristics of the received task. The task can be every time different and comes without any information about how the operation has proceeded so far.

3. Interpretativeness of acting: In order to interpret the information provided through the ROS, the remote operators should have sufficient training and knowledge on the remote operation's technical implementation and its restrictions in container handling. In conventional cabin operation the control operations have a natural link to the crane's movements and behaviour. However, in remote operation the link is artificial and mastering it demands the operator to understand how it is implemented.

4. Conceptual mastery of crane operation: When the operation is carried out remotely through the ROS, part of the feel of operation (e.g., weight and movement speeds) might disappear. Therefore, it is essential for the remote operators to understand the environment's demands and related physical forces in operation, so that in all times they can have safe control over container handling.

5. Creating shared awareness: In the conventional setting, the crane operators are in close co-operation with the truck drivers during loading and unloading of containers. In these situations, the crane operator also keeps contact with the other crane operators on the yard about the overall situation of the stack operation. In the remote setting, communicating with the other remote operators is easy since all the remote operators are seated in the same operation room. In both operating environments, sharing the experiences with colleagues on complex operating situations seemed to be an important part of the work and this kind of communication should be supported even when moved to remote operation.

6. Optimal sharing of efforts: In the conventional loading and unloading operations in the truck lane, the crane operators actively communicate with the truck drivers with the horns of their cranes. In fact, the truck driver is responsible for moving the truck in the direction of the truck lane to adjust and fine tune the location of the truck in order to help setting the container correctly on the truck chassis. In remote operation, the operators mostly communicate with the truck drivers in problematic situations (e.g., such as if the truck driver had forgotten to unlock all the pins of the truck chassis) and the truck driver is expected to stay in her designated booth during the operation.

DESIGN IMPLICATIONS

A general fear is that in remote operation some essential part of the operating experience is lost. One of our aims in this study was to understand what does the change from the conventional cabin operation to the operation carried out remotely mean from the crane operators' point of view. The Systems Usability Framework (Savioja & Norros, 2012) recognizes user experience as one important element in the development of new tools for work. Therefore, in the design of new remote control solutions for crane operation also the aimed user experience targets should be defined. These UX targets set out the experiential qualities to which the development of a new crane remote operation system should aim at.

In the following, we are going to discuss two UX targets that were identified fundamental to remote crane operation and the design implications that they have for the development of remote crane operation systems. The two chosen UX targets are 1) sense of control and 2) feeling of presence. Based on the analysis of crane operator work and the core-task demands, the factors affecting to the chosen UX targets were identified and the following design implications for remote operation workstation design are suggested.

Sense of control: 1) Creating a coherent and unrestricted operating view over the remotely operated crane is maybe the most important design objective. The ability to visually validate the state of the operating environment enhances the crane operator's sense of control and is especially important when the operator is not physically present in the object environment. 2) Moreover, estimating the different aspects of operation (e.g., distances, turn of operation speed and loads) comes more difficult when operating remotely through video feeds. The design should take this into account and aim to support the operator to judge on these operationally important factors. 3) Finally, the crane operator's sense of control could be supported by allowing the operator to flexibly adjust and do decisions on one's own operation and its rhythm. In this way, the design should enhance the good flow of operation.

Feeling of presence: 1) The quality of interaction i.e., how realistically the crane operator can interact with the system, has a significant influence on the feeling of presence. In the design of a remote operation environment there are many quality factors that may be altered. Such factors are for example the feel of operation, clearness of the operating view, and the fluency of communication. 2) One thing that might affect negatively to the feeling of presence is the system delays, which are often present in remote operating systems. Therefore, although it is usually beneficial to have a rich data (e.g., good quality video feeds) from the remotely operated object environment, it should not compromise the system's responsiveness. 3) Furthermore, against our initial conception before the studies, the social environment where the operation is carried out seems to be important in both conventional and remote operation. Currently, the ROS designs seem to be usually focusing on only one operator performing her operations in the ROS. However, according to our results, the operators are in constant collaboration and communication to each other and other parties. The ROS design should support this kind of social presence, whether inside the control room or between the remote operator and external partners (e.g., truck drivers).

In addition, our results indicate that the design of remote crane operation UIs should take into account the so called game effect. As a consequence of this effect, the operators might lose their connection to the real situation in the field and to the safety-critical nature of the remotely operated task. This kind of experience can be induced for example if the operating environment is not realistically represented (e.g., with real-time video feeds) in the ROS end, the operating user interfaces have playful features, or the physical UI (e.g., joysticks) design does not correspond to the essence of the operated machine. Therefore, if the remotely operated machine is massive in size and has safety-critical tasks, it should also show and be appropriately expressed in the ROS's user interfaces.

7 CONCLUSIONS

We have presented the results of conventional and remote crane operator studies, which were based on the Core-Task Analysis method. These studies were the first time the CTA method was used for this kind of comparative research to facilitate design. According to the results, the method seemed to serve as a useful tool for comparing different operational settings such as conventional and remote operation. Furthermore, it suited well for understanding user experience in both environments. In this way, the method also worked in guiding the design of future remote crane operation solutions to enhance the remote operators' UX.

After the analysis of control and core-task demands of both work environments, it became evident that conventional cabin operation has more dynamism related factors than remote operation. Remote operation on the other hand had an emphasis on uncertainty related factors deriving e.g., from the fact that the operators are not physically on the spot of the operation.

We view, that user requirements for remote crane operation can be defined on the basis of the results of the CTA and the chosen UX targets for guiding the design. Furthermore, in the future we need to evaluate the applied design process from a methodological perspective in order to analyse further how it suits for this kind of user experience-driven design.

8 ACKNOWLEDGMENTS

We would like to acknowledge the studied ports and especially the operators for sharing their valuable experiences on crane operation. This work was conducted in the UXUS (User Experience and Usability in Complex Systems) research program under Tekes-funded FIMECC competence cluster, which concentrates on metals and engineering industries

.

9 REFERENCES

[1.] Engeström, Y., Miettinen, R., and Punamäki, R.-L. (Eds.). (1999). *Perspectives on Activity Theory*. Cambridge, UK: Cambridge University Press.

[2.] Hainsworth, D. (2001). Teleoperation user interfaces for mining robotics. *Autonomous Robots 11*(1), 19-28.

[3.] Hartson, R., and Pyla, P.S. (2012). *The UX Book – Process and Guidelines for Ensuring a Quality User Experience.* Waltham, MA: Elsevier.

[4.] Karvonen, H., Aaltonen, I., Wahlström, M., Salo, L., Savioja, P., and Norros, L. (2011). Hidden roles of the train driver: a challenge for metro automation. *Interacting with Computers, 23*(4), 289-298.

[5.] Norros, L. (2004). *Acting under uncertainty – The core-task analysis in ecological study of work*. Espoo: VTT Publications.

[6.] Norros, I. & Norros, L. (2009, September). Human factors in the dependability of IP networks. *ENISA Quarterly Review, 5*(3), 20-21.

[7.] Nuutinen, M., and Norros, L. (2009). Core task analysis in accident investigation: analysis of maritime accidents in piloting situations, *Cognition, Technology & Work, 11*(2), 129-150.

[8.] Peirce, C. S. (1998). The Harvard Lectures on Pragmatism. In Peirce Edition Project (Ed.), *The Essential Peirce. Selected Philosophical Writings, Volume 2* (pp. 133-241). Bloomington and Indianapolis, IN: Indiana University Press.

[9.] Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction*. Amsterdam: North-Holland.

[10.] Savioja, P., and Norros L. (2012). Systems usability framework for evaluating tools in safety–critical work. *Cognition, Technology & Work*, In print.

[11.] Sheridan, T. B. (1989). Telerobotics. *Automatica, 25*(4), 487–507.

[12.] Sheridan, T.B. (1993). Space teleoperation through time delay: review and prognosis, *IEEE Transactions on Robotics and Automation, 9*(5), 592-606.

[13.] Speer, U., John, G., and Fischer, K. (2011). Scheduling Yard Cranes Considering Crane Interference. *Lecture Notes in Computer Science: Computational Logistics, 6971*, 321-340.

[14.] Steenken, D., Voss, S., and Stahlbock, R. (2004). Container terminal operation and operations research – a classification and literature review. *OR Spectrum, 26*(1), 3-49.

[15.] Vicente, K. J. (1999). *Cognitive Work Analysis. Toward a Safe, Productive, and Healthy Computer-Based Work.* Mahwah, NJ: Lawrence Erlbaum Publishers.

[16.] Vis, I.F.A., and De Koster, R. (2003). Transshipment of containers at a container terminal: an overview, *European Journal of Operational Research, 147*, 1-16.

III

USER EXPERIENCE TARGETS AS DESIGN DRIVERS: QA CASE STUDY ON THE DEVELOPMENT OF A REMOTE CRANE OPERATOR STATION

by

Hanna Koskinen, Hannu Karvonen, & Helena Tokkonen

In Proceedings of the 31st European Conference
on Cognitive Ergonomics (Article No. 25)

User Experience Targets as Design Drivers: A Case study on the Development of a Remote Crane Operator Station

Hanna Koskinen
VTT Technical Research Centre of Finland
P.O.Box 1000
02044 VTT, Finland
firstname.lastname@vtt.fi

Hannu Karvonen
VTT Technical Research Centre of Finland
P.O.Box 1000
02044 VTT, Finland
firstname.lastname@vtt.fi

Helena Tokkonen
University of Jyväskylä P.O.Box 35
FI- 40014, Finland
firstname.lastname@jyu.fi

ABSTRACT
In recent years, the notion of user experience, or UX, as an essential aspect to be addressed in the design and development of technologies has been increasingly discussed. In this paper, we present a case study in which we have used UX targets as the main design drivers and focus areas in developing a new remote operator station user interface for container cranes. UX targets describe the experiential qualities to which the product design should aim at. However, taking UX targets into consideration during product design is challenging, because only little is known about how they would be best operationalized to serve the different phases of the design process. Through our case study, we describe how we identified relevant UX targets, how their content was defined, and how they were implemented into a new remote operator station concept that was then tested and evaluated by expert users.

Authors Keywords
User experience; user experience target; concept design; container crane operation; remote operator station.

ACM Classification Keywords
H.5.m. [**Information interfaces and presentation**]: *Miscellaneous*.

General Terms
Human Factors; Design.

INTRODUCTION
In recent years, the notion of user experience (UX) as an essential aspect to be addressed in the design and development of technologies has been increasingly discussed. As a consequence of this on-going trend, the focus of human computer interaction (HCI) design has shifted from traditional usability to user experience [8]. One reason for this has been that the traditional usability approaches have not been experienced as a sufficient quality attribute to measure all the aspects of HCI.

ECCE '13, August 26 - 28 2013, Toulouse, France

http://dx.doi.org/10.1145/2501907.2501956

While usability focuses on the cognitive aspects [8] and tries to make technology accessible for all, UX is seen to include also attributes related to aesthetics, joy, emotions, and affective aspects of technology use [5, 1]. To ensure a good UX, both pragmatic (i.e., fits to behavioural goals) and hedonic (i.e. supports be-goals) aspects of products must be fulfilled [7]. Furthermore, the experiential aspects of technology usage and the concept of UX have become more important in HCI as computers have become an integral part of our everyday life.

In previous literature, for example, Blythe et al. [2] have argued that usability is only one of the many values that user-cantered design should contribute to. Following this line of thought, Donald Norman shifted HCI researchers' and practitioners' attention towards more experience-oriented design approaches in his book Emotional Design [12]. Designing for an experience calls for a holistic and richer approach, which can address everything that implies in the relationship between human and technology as a whole [11]. Furthermore, it is a way to utilize understanding of users to create a design solution that has a potential to support the full range of different human experiences. Already Winograd (1996) recognized that a more holistic UX orientation in design calls for new kinds of approaches and methods, and expressed a demand for next generation design approaches [18]. Some have also proposed scenario-based storytelling and introduction of personas in design as a way for designers, as well as users, to express their needs, wishes, and desires concerning the designed objects [4, 14].

It is of course not possible to design experiences as such, but with a careful investigation of users and their worlds, it can be possible to design for a better experience. Therefore, the designer should also focus on the desired experience and analyze the relations between its parts to understand how technology can support in the creation of a satisfying experience [19].

UX in work environments
It has often been claimed that in work systems, the experience aspect is generally in the background. At least, it has not been a central area of concern in the domain of complex work systems design. It can be even questioned if the concept of UX is meaningful in the context of work- related studies. However, in our view, the functionality of the work system is tightly intertwined with the experience- related aspects of work activity. Separating or confronting these two against each other does not benefit the design of a work system. It may be that in the work context, a different set of user experiences becomes important than in the most often discussed context of the UX studies, that is products for consumer markets. For example, carrying out the work tasks successfully with the

assistance of professional tools will most likely lead to experiences such as joy and professional pride, whereas passing a level in a computer game can cause for example feelings of triumph and excitement. At their best, well-designed work tools can enhance and facilitate the engagement in work, learning, motivation, and emergence of flow [3] at work. Therefore, the concept of UX is as important in work systems design as it is in the design of commercial products.

From the UX point of view, the focus of design should be on the potential of technology to enhance and transform people's lives [20]. Professional users, who may have many years of experience of working in the domain, may often also have certain intuitive abilities when assessing the appropriateness of new tools proposed to support them in their work. For this reason, the UXs of individual users concerning a developing tool are a factor that could be taken into account already in the early phase of the design process. This idea is supported by Systems Usability framework [15], which bases its understanding of the concept of UX on activity theory. In activity theory, subjective experiences are understood as profoundly embedded in the activity itself. Accordingly, the Systems Usability framework emphasizes the user's experience of the development potential of an evolving tool or a solution within the work activity. Feedback on the design features and solutions, which the professional users experience as promising, provides important indications to inform the further development of the tool. Furthermore, as UX is a subjective phenomenon that concerns the overall status of the activity and its objectives, it also can provide general information regarding the final degree in which the new design or tool can contribute to the development of an appropriate UX. Involving the future users of the work system in the design process can better help to grasp the potentiality of the evolving tool to address both the functional and emotional aspects of the work.

UX targets as design drivers

One way to address the notion of UX in designing new technologies is proposed to be the use of UX targets. [6] This means that not only the instrumental tool properties, such as efficiency, reliability, and ease of use, are considered and aimed at in the design, but also qualities that support certain user experiences. From our viewpoint, the function of the UX targets is to describe the experiential qualities to which the product design should aim at. More concretely, we see that UX targets describe what kind of positive experiences the designed product should evoke in the user [10]. In UX target-informed design, the identified relevant targets and their design implications for the design challenge at hand should guide the product development in its different phases (e.g., domain and user data gathering, concept design, user interface [UI] design, implementation, and iterative evaluation throughout the whole development process). However, accounting the experiential aspects in the design of technology can be seen as a new knowledge area and, thus a methodological challenge that still lacks examples and clear guidelines.

In this paper, we describe the design process of an automated stacking crane's (ASC) remote operator station (ROS) from the user experience point of view and demonstrate how the UX targets were utilized during this process. The remainder of the paper is organized as follows. First, we describe the problem space for the design task of developing a ROS concept and the general goal of considering user experience aspects during the process of design. Second, we introduce how we went about to the practical process of designing the new ROS concept and how the aim of using the UX targets in a guiding role in this design work was achieved. Finally, we discuss the lessons learned from the specific case of UX target-driven design approach of a new ROS concept and its more general implications to apply the concept of UX in the design and product development practices.

DESIGN TASK OF REMOTE CRANE OPERATOR STATION

In remote operation (i.e., teleoperation) a machine is operated from a distant location from which there is no direct human sensory contact to the operated machine [16]. Thus, the human operator depends for example on video camera feeds, sensors, and other technical means to receive information about the remotely operated machine and its immediate environment.

An increasing number of container terminals are taking remote operation systems of container cranes into use (see e.g., [17]). Remote crane operation has been claimed to allow for example an increased productivity, a safer work environment, and better physical ergonomics for the human operators compared to the conventional on-the-spot operation from the crane's cabin. However, this on-going remote operation trend presents a huge challenge from the human operator point of view, most remarkably because the remote operation introduces an entirely new concept of operation to which the operators need to adjust themselves and their operation practices. This is due to the fact that a new information technology-based remote operation system replaces the conventional cabin systems and presents the operators with a fully different viewpoint on crane operation (e.g. from the physical work environment, the work practice, and professional experience point of views).

Furthermore, from the user experience point of view, designing a novel UI concept for the remote operation of cranes, in this case ASCs (see Figure 1), is a challenging task. Existing remote crane operation solutions have already shown that these kinds of systems can be technically built and operated in a real-life context, but the main design challenge in our case lies in how the user experience of the crane operators could be enhanced by introducing a new ROS design.

Figure 1. A visualization of remotely operated container cranes in the truck loading zone. The smaller cranes in the top part of the figure are ASCs.

FROM CONTEXTUAL INQUIRY OF CRANE OPERATION WORK TO UX TARGETS

The design task at hand often determines also to some extent the kinds of experiential targets that should guide the development of that specific product. Even though preliminary UX targets for a future product can be defined based on for example common knowledge about the domain or identified business goals, we see that especially in the case of work systems design, the final UX targets should be the result of a systemic analysis of the a specific work domain characteristics and the users' work activity and tasks. That is why early in and throughout the design process, it is important to carry out field investigations and contextual inquiries in the environment and the specific context in which the technical system or artefact will be situated in the future.

Discovering the relevant UX targets

In the first phase of our study, the main aim was in finding and defining the UX targets that are relevant to facilitate and inform the given design task of developing a new ROS UI. Both a study of the technical possibilities and trends in container handling and a field study on the crane operators' work were utilized as sources for relevant UX targets.

Available technical solutions and new possibilities

In order to understand what is technically feasible and what kind of remote operation solutions already exist and are implemented in different ports, a benchmarking study was carried out. For this purpose, freely available online material regarding a variety of container handling crane suppliers and ports was utilized. Material from altogether four international crane equipment suppliers and seven existing semi-automated ports were investigated in the benchmarking study.

The crane suppliers were benchmarked on their role, existing customer terminals, offering of different type of machines, the automation level both from the system point of view and on the degree that the human operators participate in the operation. Also, the UI technologies used in the suppliers' control systems and the variety of nowadays real-life implementations was investigated. In addition, it was reviewed if the suppliers were mentioning any distinct business drivers regarding the increasing use of remote operation and the benefits that it might provide for the operation (e.g., business promises in form of slogans and other marketing material). Particularly, we were interested if the suppliers were describing their technological abilities or infrastructures from the users' benefit point of view or with clearly experiential wordings (e.g., safer operation, because in remote operation there is a more limited possibility for accidents).

The ports were compared in terms of their responsible terminal operator, equipment and facilities, ports' equipment suppliers, and in their overall capacity. The special interest was in the ports' descriptions of their container handling concepts (e.g., level of automation) and the used manpower and shift patterns of how their operation was organized. For example, the ECT Delta terminal in Maasvlakte, The Netherlands, applies an operational concept in which the truck drivers play an active role in the operation and the cranes are remotely controlled from the central control room only in the case of exceptional situations (e.g., in a sensor failure). Otherwise, the control room operators just monitor the progress of the operations. This benchmarking analysis supported the identification of relevant domain-specific UX targets.

Characteristics of remote crane operation and core-task demands

Gaining an understanding only of the technical possibilities and trends in the field is often not enough to inform the design of a new product. In addition to the technical investigations, studying also the work activity in its real context is essential. As the main focus of our case study was on enhancing the UX of the remote crane operators, we studied the crane operators' work in real port environments.

In order to facilitate the identification of relevant UX targets in the design of the new ROS, we conducted qualitative field studies in two different international container terminals during a one week period. The studies were based on core-task analysis method [13] and included altogether 12 interviews and field observations with professional crane operators in their real work environment.

In the first visited container terminal, the overall capacity (container throughput) was approximately 2.5 million TEUs (twenty-foot equivalent unit, i.e., one 20 ft. container) per year. The container handling, unloading, loading, and stack operation on the port yard were carried out on-spot with conventional cabin operation. In this terminal, we interviewed and observed six crane operators, out of which three were working in day shifts and three in night shifts. We used a semi-structured theme interview frame and each operator was interviewed and observed individually. The overall capacity of the second studied container terminal was approximately one million TEUs per year. In this terminal, the landside loading and unloading operation of containers was carried out by the remote crane operators through a remote operation system providing video feeds to the truck loading area. In addition to these camera views, the remote operators did not have direct eye-sight to the operated crane and its environment. During normal operation, the stack and waterside operation of the terminal were carried out automatically by the cranes. In this terminal, we interviewed altogether five day shift remote operators and one maintenance operator by using a similar semi- structured theme interview frame. In addition, operators' daily working in a centralized control room was observed during the two-day study period.

In the core-task analysis of remote ASC operation's control demands (see [9] for details), we noticed that factors related to uncertainty were particularly emphasized in the results compared to dynamism or complexity factors. This is obviously due to the fact that the remote crane operators are not physically present on the spot where the operation is carried out and cannot directly feel and hear everything what is happening with and around the crane. This, in turn, brings a certain level of cautiousness to the operation.

In the analysis, we also identified the core-task demands related to readiness to act, flexibility of acting and reorienting, interpretativeness of acting, conceptual mastery of crane operation, creating shared awareness, and optimal sharing of efforts. A more detailed description of the main results regarding the individual core-task demands can be found from [9]. From this analysis, we gained valuable insights regarding which UX targets would be the most relevant for the design of the new ROS UI.

Table 1. A list of recognized UX targets and their sources. In bold and italics the final chosen UX targets.

UX target	Source of the target
Feeling of presence	Benchmarking
Sense of control	Systems Usability framework
Experience of fluent co-operation	Field studies
Better ergonomics	Benchmarking
Feeling of safe operation	Field studies
Experience of appropriate functioning of the tool	Systems Usability framework
Experience of fit for own use	Systems Usability framework
Experience of own competence	Systems Usability framework
Experience of interesting and rewarding work tasks	Systems Usability framework
Feeling of having a professional tool to operate	Systems Usability framework and Benchmarking
Professional pride and motivation	Systems Usability framework
Feeling of being an essential part of the work community	Systems Usability framework
Appropriate trust in technology	Systems Usability framework

Defining the guiding UX targets

As a result of the contextual investigations, a preliminary and relatively extensive list of relevant UX targets were identified (see Table 1). However, as extensive as the list is, it was not focused enough to guide the actual concept design phase and set the suitable experiential targets for it. Therefore, it was essential to narrow the list further down into limited number of distinct UX targets to be used in guiding the ROS UI design task.

The final chosen UX targets were: 1) feeling of safe operation, 2) sense of control, 3) feeling of presence, and 4) experience of fluent co- operation. The importance of the feeling of safe operation and the experience of fluent co-operation targets became evident in the discussions with the crane operators, i.e. in the field studies, and they were the most recurring themes in the core-task analysis results. The Systems Usability framework provided us with a theoretical background of human activity in complex work environments as well as an assumption about the UXs that are salient in safety-critical work context. Based on this knowledge, sense of control was selected to be one of the main targets when developing the new ROS UI. Finally the feeling of presence was recognized in the benchmarking study, which indicated that in the crane operating field there is an on-going trend towards remote operation due to the fact that more efficient data communication and automation technologies are nowadays available. These new ROS UI solutions should enable the crane operators to feel and empathize with the operational situation in the remotely operated object environment.

Even though the selection of the UX targets should always be based on a comprehensive understanding of the design task at hand, the final decisions will always remain design team specific. Meaning that there is not only a one defined set of particular UX targets that could be utilized in the design of remote operation solutions.

UX TARGETS GUIDING THE CONCEPT DESIGN AND THE DESIGN OF INNOVATIVE FEATURES

In this section, we demonstrate and draw a link between the chosen UX targets and, the specific design features and the solutions built into the new ROS UI concept for the remote operation of ASCs. We both describe how we operationalized the selected UX targets for the purposes of the concept design phase and give a design example of an operating view solution development from the user experience point of view (i.e., intended support in the form of a particular design feature for the defined UX targets).

Articulating the design implications of the UX targets

Being able to set the four specific UX targets (feeling of safe operation, sense of control, feeling of presence, and experience of fluent co-operation) to be aimed at in the design of the new UI for the ROS was a good starting point. However, the meaning and content of the defined UX targets need to be understood in the object context before the targets can really be utilized as drivers in the design. To do this, a careful investigation and analysis of the conceptual meaning of the chosen UX targets and the issues raised around them was carried out. The aim of this work was to describe the UX targets as clearly and understandably as possible and in this way to operationalize them for the purposes of concept design.

In the following, the four selected UX targets are discussed in a detailed manner and specific content is given for each of them in the remote crane operation context.

Feeling of safe operation

In our analysis, the UX target feeling of safe operation was found to be one of the most important targets. One reason for this is that during crane operation there are huge forces in question, which in the worst case can cause even a loss of human life in a fraction of a second. Therefore, a possibility to visually validate the state of the operating environment, for example, with good quality live video feeds from the object environment, is essential. This makes it possible that the location of the workers in the loading area is seen clearly at all times.

Delivery of sufficient and relevant data from the object environment is another important aspect of remote operation that should be supported in order to fulfil the subjective feeling of safety target. For example, meaningful operational values need to be provided for the remote operator.

Experiencing feeling of safe operation was also connected to the accurate perception of operation kinetics. For example, it is critical to know how fast the crane is moving and into what direction. Therefore, the design of the ROS should include UI solutions that support the understanding of operation kinetics.

Sense of control

From the very beginning of the project, we had already recognized sense of control as one of the key experiential targets to be aimed at. The question from the design point of view was that what this target precisely means in this particular design case. We found that a coherent and unrestricted operating view over the remotely operated environment contributes significantly to the experience of control. For example, the provided camera angles should be at such a position that they are not in contradiction with each other and that they constitute a uniform operating view.

A possibility to decide one's operating rhythm is surely an aspect of work that affects how much control one is experiencing in operation. Therefore, the remote operator needs to have the final authority to decide when to start and stop operating. In this way, the ROS design could enhance the good flow of operation.

Also, the ROS design should give support for the correct estimation and understanding of different relevant aspects of operation. For example, the actual weight of the load needs to be presented illustratively to the remote operator in the UI.

Feeling of presence

Feeling of presence is closely connected to the experience of concentration and flow [3] in one's work. This experience relies much on the quality of interaction, i.e., on how naturally the remote operator can interact with the crane. For example, the provided camera views need to be sharp enough in order for the appropriate feeling of operation to be supported.

As the current remote operation solutions reduce the operation view from three-dimensional live view to a two-dimensional video feed-based presentation, a support for the comprehension of the physical dimensions in the object environment becomes essential. Therefore, the ROS UI design should assist in understanding the dimensions of the loading area and the crane, such as the distance between the spreader (the lifting device of the crane) and the container to be picked up.

Furthermore, the feeling of presence was found to be affected by the availability of rich data from the object environment without disturbing delays. Therefore, although it is usually beneficial to have a rich data (e.g., good quality video feeds) from the remotely operated object environment, it should not compromise the system's responsiveness.

Experience of fluent co-operation

The real importance of the UX experience of fluent co-operation became evident in the field of observations. The work of the crane operator, regardless of it being conventional or remote operation, is a highly collaborative effort. Therefore, the facilitation of activity between the different professionals and stakeholders is an important quality attribute that the new ROS design should have. For example, the design should support social presence, whether inside the control room or between the remote operator and external partners (e.g. truck drivers).

The experience of a fluent co-operation also demands making the ease of communication certain. For example, improvements on the communication tools and facilities can promote this teamwork-oriented UX target.

In the most general sense, the experience of fluent co-operation requires support for the knowledge about the used domain specific terminology, rules and responsibilities. For example, the ROS design should make the implications of the utilization of the override functionality (i.e. deviating from the set safety limits) evident.

Defining user requirements for the ROS

The UX targets' conceptual investigation revealed the meaning of each individual target in this particular design context and enabled them to function in an informative manner in the further design. Based on the chosen UX targets and the core-task analysis, we also defined a set of 30 concept-level user requirements for the new ROS UI design under themes such as automation, operating view, operational data, communication and collaboration, and the physical ROS. While the UX targets worked as general level descriptions of what things are important and eligible in the crane operation, the user requirements were more specific in nature by providing hand-on guidance for the UI development. The data regarding each requirement included for example a detailed description of the content of the requirement, its priority, its background (from which data it was drawn), and the related UX target(s) to which the requirement is expected to contribute to. The user requirements elicitation and definition was crucial in this phase also for the design team's attempt to proceed from the investigation of UX targets to their implications for the new ROS UI design.

Design and implementation of solutions supporting the UX targets

The aim of this concept design phase was to transform the UX targets into corresponding design implementations, i.e., design solutions that would incorporate support for the defined UX targets. This concept design activity was carried out in several design workshops in which all the project partners had a possibility to present their ideas and intentions. In order to provide the design team with more focus, the insights of the fieldwork, the operationalization of the UX targets, and the user requirements and their design implications were presented at the beginning of the design workshops. After these presentations, the team made a collaborative effort to materialize some of the most promising ideas in a form of a variety of low fidelity mock-ups and prototypes. During this process, many alternative design solutions and technological possibilities were considered and iterated. The main topics around which the design solutions were constituted were the operating view, the operational data, communication and collaboration facilities, and the physical setup of the ROS. The iterative approach assessing the design outcomes enabled the design team to move relatively quickly towards a one basic ROS. UI concept within which the different partial features were altered to fine-tune the concept. To support this activity, a functioning virtual simulator was built.

Operating view

To demonstrate the above-described UX target-driven concept design phase, we provide a design example regarding the development of the operating view solution to be included as part of the new ROS UI concept (see Figure 2). As mentioned earlier, the properties of the operating view were expected to have a major effect on the UX targets of feeling of safe operation, sense of control, and feeling of presence. Our aim was to provide the crane operators with an operating view that would complement all their main information needs when taking care of the container handling tasks.

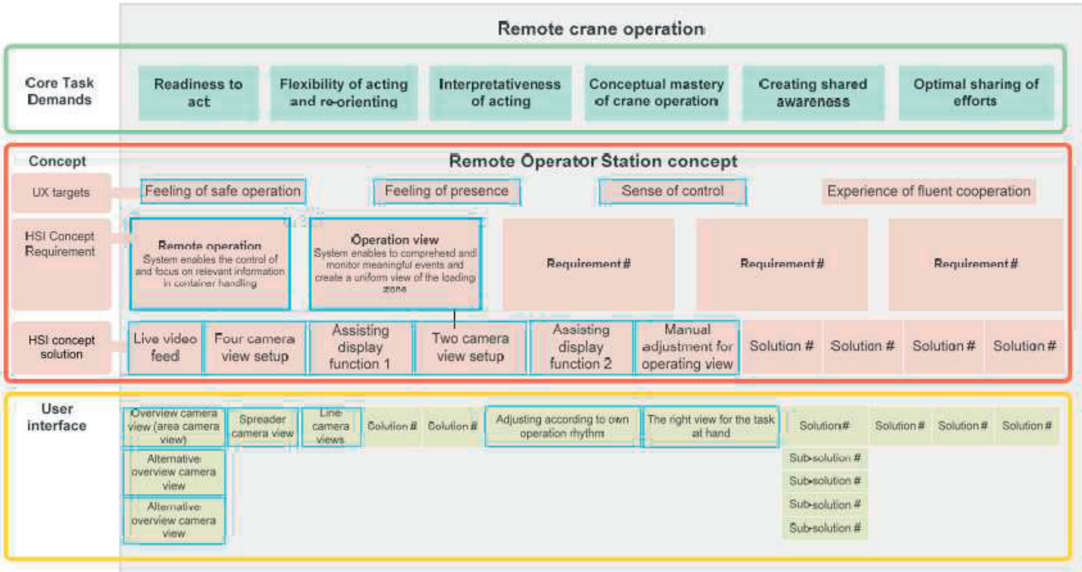


Figure 2: A description of the operating view solution and its innovative features (framed with a turquoise colour)

Maybe the most fundamental part of the operating view solution was the decision to use real-time video feeds from the cameras placed on the truck loading area as the primary source of information. Therefore, the placement of the cameras and the views that they provided were the defining characteristic of the new operating view solution. The use of video feeds also corresponded to the operators' need for the possibility to visually validate the state of the operating environment, which supported the feeling of safe operation.

Another part of the solution was to decide how these camera views would be presented and compiled together in the ROS display so that they would best serve the task at hand. For this purpose, several different optional layouts for arranging the camera views were generated and considered and finally two main layout solutions were defined for further development and investigation; a setup with four camera views and a setup with two camera views. In the four-view setup, three of the camera views were arranged in a line on the top-most part of the display and one of the views in the middle of the lower part of the display. In the two-view setup, the two main camera views were arranged next to each other in the top part of the display. An unavoidable consequence of the four-view setup was that the sizes of the camera views became smaller, but on the other hand more information (video feeds) could be displayed continuously. In turn, the sizes of the views with the two-view setup were substantially larger, but all the needed camera views could not be presented at once in the UI (see Figure 3). Another design challenge that rose from the chosen display layout was to manage the positioning of the different camera views and the relation between them: for example, how to design the camera angles in such a way that they are not in contradiction with each other and that they constitute a uniform and coherent operating view that facilitates operators' focusing and engagement on the task at hand. The quality of this operating view solution related also closely to the sense of control and the feeling of presence UX targets.

The operating view solution encompassed also ways to control and manipulate the operating view. Both alternative view setups included assisting display functions that were meant to satisfy the general aim of the new design to support operators' concentration in their primary task, container handling, while relieving them as much as possible from the secondary task to control and manage the operating view. In the two-view setup, the problem of not being able to display all the necessary camera views simultaneously was attempted to be solved by introducing a special assisting functionality that would help the operators in gaining always the right view for their task at hand. In addition to the assisting display functionalities, we made an effort to improve the manual control of the operating view in a situation where additional adjustments were required.

Figure 3. ROS concept illustration with the two camera-view setup.

BRINGING THE NEW REMOTE OPERATOR STATION INTO USE: CONCEPT EVALUATION

In UX target-driven design, it is important to assess to what extent the proposed new concepts have succeeded in embodying support for the targeted UXs. Therefore, the main aim of this phase in our case was to evaluate the alternative concept designs in light of the aimed UX targets. In the UX target evaluation phase, the design process becomes exposed to questions like "does the new design enhance those experiential qualities that the design team was originally aiming at?" and "do these experiences become actualized in the use of the new system?"

Concept and UX target evaluation

The main goals of the concept evaluation were to compare the UXs of the two different ROS UI concepts in order to verify how the defined user requirements were fulfilled with the current prototype system and how well the UX targets feeling of safe operation, sense of control, and feeling of presence were realized with the current UIs. The fourth UX target, experience of fluent co-operation, was excluded from the evaluations, because the concept features that aimed to support and facilitate communication and collaboration were not implemented into the present prototype system. This was due to the fact that it was not possible to simulate a collaborative operation situation in the evaluation studies.

The concept evaluations were conducted with a virtual simulator. The virtual 3D simulator replicated the dynamics of the container handling operations of an ASC. Thus, the simulator enhanced realism of the concept evaluations in a situation when testing in a real-world setting was not obtainable. The test setup of the ROS included two joysticks, a touch screen-based tablet, and a 32-inch television screen as the main operation display. The user interface of the main operation display consisted of camera views, i.e., virtual reality graphics that were mimicking the actual camera views, and operational data provided by the system.

Method

Six crane experts participated in the concept evaluations. In order to evaluate how the aimed UX targets are realized and how the defined user requirements are fulfilled with the concept prototypes, we used a combination of different evaluation methods and approaches, which included the following ones: Simulator runs with varying level of task difficulty, a semi- structured interview, and two separate questionnaires after the operational tasks. The questionnaires were scaled with a 5-point Likert scale (1=strongly disagree, 5=strongly agree).

Operational task

The operational tasks that were executed with the virtual simulator covered the most common container handling tasks; container lifting and landing operations. In the beginning of the test sessions, the participants executed a very basic operational task to learn how to use the system and the different control functions. The other operational tasks were more challenging and introduced disruptive factors such as for example strong wind, nearly similarly coloured container chassis to reduce visibility of container sides, other containers in the surrounding lanes, a truck driver walking in the loading area, and a locked chassis pin.

Evaluation procedure

In the beginning of the evaluation, the test participants were provided with a brief description of the aims of the evaluation, the functionalities of the virtual simulator system, and the features of the two proposed ROS UI concepts. The evaluation sessions were video-recorded to aid data analysis later on. The test leader asked the test participants to think-aloud if possible while executing the operational tasks in order to understand how the participants experienced the concepts.

Before starting with the operational test tasks, some general questions regarding the professional background of the participants and their previous experiences of gantry crane operations were asked. The participants could conduct the given operational tasks at their own pace. After each task, the test leader asked additional questions related to the task, the operation situation, the camera views, and other features of the ROS UI. When all the operational tasks with the first ROS UI concept had been executed, the test participants were asked to fill in a questionnaire consisting of twelve user experience statements. After completing the questionnaire, the other ROS UI concept was evaluated by conducting a set of operational tasks and by answering to the additional questions that were asked after each operational task. When the operational tasks with the second UI concept had been conducted, the test participants filled in the same kind of user experience questionnaire as with the first ROS UI concept. The order in which the two alternative ROS UI concepts were presented to the users was varied between the test participants so that every other participant started with the two-view setup and every other with the four-view setup.

Some general questions related to the proposed ROS UI concepts were asked before the participants were requested to select the ROS UI concept that he preferred and that in his opinion provided most appropriate user experience. Finally, a Systems Usability questionnaire [15] consisting of 36 systems usability statements for the selected ROS UI concept was filled in.

Evaluation results

Both of the ROS UI concepts received positive feedback and were accordingly selected for further development by the participants. However, neither of the concepts was clearly preferred over the other by the participants and in both concepts they also identified clear areas of improvements that should be addressed before the appropriate operation experience could be achieved.

The outcomes of the evaluations concerning the alternative operating view solutions showed that the UX target sense of control was supported in an appropriate degree in both concepts. However, the evaluations also revealed some areas of improvements. For example, the participants were not able to form a clear picture of the situation in the operating area when handling the container in the outmost truck lanes. Therefore, they needed to adjust the camera views in order to gain a better view to the position of the truck and corners of the container. In addition, the overview camera was not sharp enough when zoomed in for the participants to be able to validate for example the registration number of a truck. The assisting display functionalities that were intended to address the above-mentioned problems were experienced positively and promising, although there were still needs for fine-tuning them further.

The responses concerning the feeling of safe operation UX target indicated that it was only partly fulfilled. The frequently used views of both concepts provided relevant information for executing the basic tasks of container handling and, for example to observe people in the loading zone. However, the use of the virtual simulator made it hard to assess the safety aspect of the concepts realistically. For that reason, in the next development phase, it would be beneficial to test the concepts with real-life video feeds to address appropriately this significantly important aspect of remote operation.

The feeling of presence UX target was also supported to some degree, since the participants felt that both ROS UI concepts were quite realistic and natural to use, since the camera views were clear and large enough. However, the feeling of presence was reduced due to the fact that some of the camera views were difficult for the participants to understand and orientate themselves into. Therefore, in the future it would be essential to pay attention to the support that can be given to the operators via providing them with a coherent ROS UI and enabling them to engage and immerse to the situation in the loading zone.

CONCLUSIONS

We have presented a UX target-driven design process through a case study example regarding the design of a new remote operator station concept and UI. In this design process, we first identified a large set of relevant UX targets. During this phase, we noticed that the early stages of defining UX targets are an interplay between suitable background theories and domain knowledge and user data, gathered for example from user interviews and field observations. After a careful consideration, we chose four particularly relevant UX targets that should guide our design efforts in creating a new remote crane operator station. Specific UX targets in focus early in and throughout the design process can help all the involved stakeholders to find common ground and concentrate on developing support for those user experiences recognized to be the most important ones in the design context at hand. We have also described how we see UX targets relate to the different phases of concept design in general. During these phases, the UX targets worked as the guiding stars for the design process and they were utilized for example in user requirements elicitation, concept level solution implementation, and concept evaluation and testing.

Furthermore, we have elaborated in detail some of the design implications of these UX targets in the context of our case environment related especially to the operation view solution of the new ROS concept. For example, the importance of real-time video feeds for the visual validation of the operation environment was identified as an essential factor and therefore implemented to the new design to support the feeling of safe operation. Finally, we have evaluated how the selected UX targets and user requirements were met with the current concept prototypes of the new ROS UI.

Also, when developing the final system in later phases of the product development, the UX targets and the associated user requirements need to be iteratively evaluated in order to find out how they are met at each stage and what they mean in detail regarding each implemented solution. In the future development of the ROS UI , special attention should be paid to the experience of fluent co-operation UX target and different aspects related to it (e.g., the interaction between the co-workers or the truck drivers) as in the present study it was not possible to address this target appropriately.

ACKNOWLEDGMENTS

We would like to acknowledge the studied ports and especially the interviewed crane operators for sharing their valuable experiences on crane operation. We would also like to thank the domain experts that participated in the concept evaluation and helped us to set guidelines for the further development of the ROS concept. This research was carried out as part of the Finnish Metals and Engineering Competence Cluster (FIMECC)'s UXUS program.

REFERENCES

[1] Beauregard, R. and Corriveau.P. User Experience Quality: A Conceptual Framework for Goal Setting and Measurement. In *Proceedings of the 1st international conference on Digital human modelling*, V. G. Duffy (Ed.). Berlin, Heidelberg , Springer-Verlag, 2007, 325-332. DOI= 10.1007/978-3-540-73321-8_38

[2] Blythe, M. A., Overbeeke, K., Monk, A. F., and Wright P. C. (Eds.). *Funology: From usability to enjoyment.* Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003

[3] Csikszentmihalyi, M. and Csikszentmihalyi, I.S. *Optimal experience: Psychological studies of flow in consciousness*. Cambridge: Cambridge University press. 1988

[4] Forlizzi, J. and Ford, S. The building blocks of experience: an early framework for interaction designers. In *Proceedings of 3rd conference on Designing interactive systems.* DIS'00, D. Boyarski, W.A. Kellogs (eds.), ACM Press, 2000, 419-423, DOI= 10.1145/347642.347800

[5] Forlizzi, J. and Battarbee, K. 2004. Understanding Experience in Interactive Systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques,* DIS'04, J. Forlizzi,

K. Battarbee (Eds.), ACM, 2004, 261-268. DOI= 10.1145/1013115.1013152
[6] Hartson, R.and Pyla,P.S. 2012. The UX Book: Process and guidelines for ensuring a quality user experience. Morgan Kaufmann.
[7] Hassenzahl, M. User Experience (UX): Towards an experiential perspective on product quality. In *Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine*, E. Brangier, G. Michel, J.M.C. Bastien, N. Carbonell (Eds.), ACM, 2008, 11-15. DOI=10.1145/1512714.1512717
[8] Hassenzahl, M. & Tractinsky, N. User experience - a research agenda, *Behaviour & Information Technology*,, 2006, 25:2, 91-97.
[9] Karvonen, H., Koskinen, H. and Haggrén, J.Enhancing the user experience of the crane operator: comparing work demands in two operational settings. In P. Turner and S. Turner (Eds.) *Proceedings of the European Conference on Cognitive Ergonomics,* ECCE2012, 2012, ACM Press, 37-44. DOI= 10.1145/2448136.2448144
[10] Karvonen, H., Koskinen, H., & Haggren, J. Defining User Experience Goals for Future Concepts. A Case Study. *In:, NordiCHI2012 UX Goals 2012 Workshop proceeding*s, Väätäjä, H., Olsson, T., Roto, V. & Savioja, P. (Eds.), TUT Publication series, Tampere University of Technology, Department of Software Systems, (2012), 14-19. http://URN.fi/URN:ISBN:978-952-15-2955-9
[11] Kuutti, K. Hunting for the lost user: From sources of errors to active actors – and beyond. *Paper present at Cultural Usability Seminar*, Media Lab, University of Art and Design, Helsinki, 24.4.2002. http//mlab.uiah.fi/culturalusability/papers/kuutti_paper.html.
[12] Norman, D. Emotional Design: Why we love (or Hate) Everyday Things. New York, Basic Books, 2004
[13] Norros, L. Acting under uncertainty – The core-task analysis in ecological study of work. Espoo: VTT Publications, 2004.
[14] Pruitt, J. and Grudin, J. Personas: practice and theory. In *Proceedings of Designing for User Experiences, DUX'03,* number 6 in Informing DUX, New York. ACM Press, 1-15. DOI=10.1145/997078.997089
[15] Savioja, P. and Norros L. Systems usability framework for evaluating tools in safety–critical work. *Cognition, Technology & Work*, In print. (available in OnlineFirst). DOI=10.1007/s10111-012-0224-9.
[16] Sheridan, T. 1992. Telerobotics, automation and human supervisory control. Cambridge, MA: MIT Press.
[17] Speer, U., John, G., and Fischer, K. Scheduling Yard Cranes Considering Crane Interference. *Lecture Notes in Computer Science*: *Computational Logistics*, 6971, 2009, 321-340.
[18] Winograd, T. *Bringing Design to Software*. New York, ACM press, 1992.
[19] Wright, P., McCarthy, J. and Meekison, L. 2003. Making Sense of Experience. In *Funology: From Usability to Enjoymen*t, Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003, 43-54
[20] Wright, P. and McCarthy, J. Experience-Centered Design: Designers, Users, and Communities in Dialogue. Morgan & Claypool Publishers, 2010

IV

DEFINING USER EXPERIENCE GOALS TO GUIDE THE DESIGN OF INDUSTRIAL SYSTEMS

by

Eija Kaasinen, Virpi Roto, Jaakko Hakulinen, Tomi Heimonen, Jussi P. P. Jokinen, Hannu Karvonen, Tuuli Keskinen, Hanna Koskinen, Yichen Lu, Pertti Saariluoma, Helena Tokkonen, & Markku Turunen

Behaviour & Information Technology vol 34, 976–991

https://doi.org/10.1080/0144929X.2015.1035335

Defining User Experience Goals to Guide the Design of Industrial Systems

Eija Kaasinen[1], Virpi Roto[2], Jaakko Hakulinen[3], Tomi Heimonen[3], Jussi Jokinen[4], Hannu Karvonen[1], Hanna Koskinen[1], Tuuli Keskinen[3], Yichen Lu[2], Pertti Saariluoma[4], Helena Tokkonen[4] and Markku Turunen[3]

[1]VTT Technical Research Centre of Finland, P.O.Box 1300, FI-33101 Tampere, Finland. Email {eija.kaasinen; hannu.karvonen; hanna.koskinen}@vtt.fi. Phone +358 20 722 111

[2]School of Arts, Design, and Architecture, Aalto University, P.O.Box 31000, FI-00076 AALTO, Helsinki, Finland. Email {virpi.roto; yichen.lu}@aalto.fi. *Phone +358 09 47001*

[3]School of Information Sciences University of Tampere, Kanslerinrinne 1, FI-33014 Tampereen yliopisto, Tampere, Finland. Email {jaakko.hakulinen; tomi.heimonen; tuuli.keskinen; markku.turunen}@uta.fi. Phone +358 3 355 111.

[4]University of Jyväskylä, P.O.Box 35, FI-40014 Jyväskylän yliopisto, Jyväskylä, Finland. Email {jussi.p.p.jokinen; helena.tokkonen; pertti.saariluoma}@jyu.fi. Phone +358 14 260 1211.

Corresponding Author:

Eija Kaasinen, VTT Technical Research Centre of Finland, P.O. Box 1300, FI-33101 Tampere, Finland. Phone: +358 40 557 1005. Email: eija.kaasinen@vtt.fi

Defining User Experience Goals to Guide the Design of Industrial Systems

Abstract

The key prerequisite for experience-driven design is to define what experience to design for. User experience (UX) goals concretise the intended experience. Based on our own case studies from industrial environments and a literature study, we propose five different approaches to acquiring insight and inspiration for UX goal-setting: Brand, Theory, Empathy, Technology, and Vision. Each approach brings in a different viewpoint, thus supporting the multidisciplinary character of UX. The *Brand* approach ensures that the user experience goals are in line with the company's brand promise. The *Theory* approach utilises the available scientific knowledge of human behaviour. The *Empathy* approach focuses on knowing the actual users and stepping into their shoes. The *Technology* approach considers the new technologies that are being introduced and their positive or negative influence on UX. Finally, the *Vision* approach focuses on renewal, introducing new kinds of user experiences. In the design of industrial systems, several stakeholders are involved and they should share common design goals. Using the different UX goal-setting approaches together brings in the viewpoints of different stakeholders, thus committing them to UX goal-setting and emphasising user experience as a strategic design decision.

Keywords: user experience, user experience goal, experience-driven design, industrial systems

1. Introduction

Good user experience (UX) is nowadays the goal of most products and services intended for the consumer market. UX is also receiving increasing attention in the development of industrial products and services. Hassenzahl and Tractinsky (2006) claim that the notion of UX has been so well adopted because the previous narrow focus on interactive products as tools did not capture the variety and emerging aspects of technology use. According to Hassenzahl (2003), user experience consists of both the pragmatic and hedonic aspects of product use. Similarly, Mahlke (2005) sees user experience as stemming from the instrumental and non-instrumental qualities of product use. The pragmatic or instrumental refers to the utilitarian aspects, such as usefulness and ease of use, and hedonic or non-instrumental to the emotional and experiential aspects of product use.

Experience-driven design focuses on the non-instrumental, meaning its function is not so much utilitarian as experiential (Hekkert, Mostert and Stompff 2003). Experiential issues have been included in earlier approaches, but rarely as the main objective of the design process. For example, trust has been an important factor in many e-commerce user studies (e.g. Järvenpää and Tractinsky 1999; Karvonen 2000; Gefen 2000). Usability and user acceptance studies include some experiential elements, while in experience-driven design, emotional and experiential elements are the main focus.

Our work focuses on industrial environments, and especially on the use of tools in workplaces. We base our user experience definition on the UX White paper by Roto et al. (2010), in which UX refers to the experience(s) derived from encountering systems. We define user experience at work as: "The way a person feels about using a product, service, or system in a work context, and how this shapes the image of oneself as a professional".

The field of Human-Computer Interaction (HCI) has defined a process for ensuring product usability, where the key is to define usability requirements in the early phases of product development. When designing for good usability, the general usability criteria from the ISO 9241-210 standard (effectiveness, efficiency, satisfaction) can be taken as the starting points, and precise user requirements for functionalities can be defined accordingly. No similar lists of universally applicable qualities are available for good user experience, as different products may target entirely different experiences. The ideology behind experience-driven design is first to define the intended experience and only then to think about the possible designs that might evoke such an experience: "One of the basic claims of Experience-driven design is to consider the experience before products" (Hassenzahl 2010, 63). Thus, the key prerequisite for successful experience-driven design is to define what experience(s) to design for.

An early example of experience-driven design is Kansei engineering, used proficiently in the Japanese car industry from the 1970s onwards (Nagamachi 2002). However, the research on experience-driven product design started to boom only in the late 1990's, probably due to the establishment of the Design and Emotion Society[1] in 1999. Since then, the importance of designing for emotions and experiences has been acknowledged by several design experts (Sanders and Dandavate 1999; Jordan 2000; Shedroff 2001; Hekkert, Mostert and Stompff 2003; Norman 2004 among the early ones). Experience-driven design "takes an intended user experience as the primary objective of the design process" (Hekkert, Mostert and Stompff 2003). It is naturally impossible to force people to have a specific experience, but designers can aim at facilitating a certain type of experience, i.e. they design *for* an experience rather than design an experience (Sanders and Dandavate 1999; Wright, McCarthy and Meekison 2005). Terms Experience Design (Hassenzahl 2010) and Experience-Centered Design (Wright, Wallace and McCarthy 2008) also refer to designing for user experience.

From the literature, we have found several different approaches to experience-driven design, each with a different process for defining the intended user experience.

---

[1] *www.**designandemotion**.org*

For example, Sanders and Dandavate promote co-designing in order to gain access not only to what people say and do, but also to their experiences and dreams (Sanders and Dandavate 1999). Hekkert, Mostert and Stompff, in contrast, leave the experience to be defined by the designer (Hekkert, Mostert and Stompff 2003; Hekkert, van Dijk and Lloyd 2011). Hassenzahl (2010) utilizes a list of basic psychological needs when defining experiential goals for design, while Wright and McCarthy (2008, 2010) emphasize a dialogue and co-production to build empathy. We have not found publications that would analyse the differences of these experience-driven design approaches, although they seem to introduce striking disparities in their starting points.

In this article, we focus on those approaches in which the design is driven by the intended experience, which we call a "UX goal". The first academic workshop to collect cases of UX goal utilisation was organised in 2012 (Väätäjä et al. 2012). Even if there were several approaches to experience design reported in academic publications at the time, there were few workshop submissions in which researchers would have concretised the targeted experience as experiential goals. The lack of UX goals in academic experience-driven design cases may be due to the small scale of academic experience design cases, in which the whole team consists of experts in experience design, the mind-set is relatively similar, and the outcomes are concepts rather than actual products. Concrete UX goals may be most useful in experience-driven design in an industry context, where various stakeholder groups need to agree on what to design. UX goals can help keeping user experience in focus through the multidisciplinary product development and marketing process.

This article is based on our experiences in four different case studies focused on work environments: moving within office buildings, working in metal workshops, and operating cranes in factories and ports. The cases shared the aim of experience-driven design with concrete UX goals. Otherwise, each case used its own methods and approaches. The cases started at the same time and lasted from 9 to 22 months. The variation in the length is due to the industrial environments where, for example, organising user studies requires a suitable time window. As the cases had each defined their own UX goals, we gathered together to integrate the results and to learn from each other. We found that, even if the cases were using different design approaches, they used similar sources for insight or inspiration in order to define user experience goals. From the literature, we did not find studies that had studied the process of defining UX goals. We decided to extend our focus more widely to related research: the kinds of approaches to experience goal-setting that we can find in the literature. We chose two research questions that focus on the first phases of experience-driven design:

Research question 1: What kinds of approaches are there for defining UX goals?

Research question 2: What kind of contribution do these approaches make in defining UX goals: What kinds of UX goals do they produce? What are the benefits and challenges of the approaches?

In this article, we first discuss user experience goals in Section 2, namely, what these goals are and how they are used. Then, in Section 3, we describe the four case studies that we have carried out. We describe the UX goals utilised in each case, and how these UX goals were defined. In Section 4, we widen the perspective to related research, and we identify experience goal-setting approaches from earlier research. Based on our own work and the literature, we present a framework that includes five approaches to defining UX goals. In Section 4, we also aim to find answers to the second research question: What kinds of UX goals do each approach produce? What are the benefits and challenges of the approaches? Finally, in Section 5, we analyse and discuss our findings and propose directions for future work.

2. User Experience Goals

An experience goal describes the intended momentary emotion or the emotional relationship/bond that a person has towards the designed product or service (Lu and Roto, 2014). We prefer to use the term goal instead of the term requirement for the experiences to design for, because a designer can only facilitate, not guarantee, a certain user experience. Experiences with interactive products and services are context-dependent, dynamic, and subjective (Law et al. 2009, Roto et al. 2010). What a designer can do is design *for* an experience (Sanders and Dandavate 1999). As Desmet and Schifferstein state, it is challenging to find the right experience to design for (Desmet and Schifferstein 2011). In this article, we focus on this challenge: how to get insight and inspiration to define UX goals that concretise the intended experience.

There are similarities between UX goals and other concepts used as the starting point for design. Lu and Roto (2014) analyse how experience goals differ from the earlier concepts: from user requirements (ISO 2010) by focusing on the emotional aspects; from value propositions (Rintamäki, Kuusela and Mitronen 2007) by leaving cost-benefit thinking behind; from a design brief by stating the wanted experiences in a compact form; and from a design driver (Wikberg and Keinonen 2002) by focusing on experiences. In industry, UX goals are often defined on a very abstract level, such as “superior UX” or “wow”. “Good user experience” as such does not guide design; to design for user experience, more specific, concrete user experience goals should be defined. In the following, we will review the literature in which the design goals have focused on the experiential aspects.

Hassenzahl (2003) introduces hedonic be-goals that differ from pragmatic do-goals, and calls for the definition of the be-goals before the functional do-goals. Rogers, Sharp, and Preece (2011) list several user experience goals that describe different emotions and felt experiences. In both the above views, UX goals are concerned with how users experience interactive products from their personal perspective. This is different to usability goals that define how useful or productive a system is from its own perspective. Usability goals do not address the overall quality of the user experience (Rogers, Sharp, and Preece 2011), nor the higher-level concerns that have become widely recognised as part of UX literature (Beauregard and Corriveau 2007). As a consequence, an increasing amount of interest has been focused on UX goals (see e.g. Hartson and Pyla 2012).

In the UX Goals workshop by Väätäjä et al. (2012), a good UX goal was seen to guide design towards a positive experience, to help in communicating objectives, and to be measurable. However, it is hard to define a UX goal that would both give guidance for design and, at the same time, be measurable. This can be seen from the UX goals presented in the workshop cases, such as sense of control, feeling of presence, stimulation, competence, self-efficacy, freedom from pain and distress, freedom to express natural behaviour, comfort, and various playful experiences (captivation, submission, fellowship, humour, good mood, amusement, relaxation). The sources for defining these goals were user studies, theory, standards or guidelines, or common sense, and the cases presented in the workshop combined several of these sources.

UX goals guide the substance of design, but within business contexts, the UX goals can also be used as a means of communication between decision-makers and UX professionals. As shown by Olsson et al. (2013), general-level UX goals can serve well as design inspiration and guidance; for example, they can form fruitful starting points for brainstorming, as well as constant reminders of the rationale of design. As the design process proceeds to a more specific level, the UX goals should be defined at a more specific level that can be interpreted in terms of design implications. During the later design phases, each design solution implementation should be traceable back to the originally defined UX goals (Karvonen, Koskinen, and Haggrén 2012a).

3. The Case Studies

In this section, we will describe four case studies where we have applied experience-driven design in designing industrial systems. For each case, we will describe the general setup of the case and how the experience-driven design process proceeded. Then, we describe the UX goals and how they evolved in the design process. We focus on the early phases of the design where UX goals were set before the actual implementation activities.

3.1 Mobile Interaction with Elevators

In complex environments such as office building blocks, moving between buildings and floors can be challenging and time-consuming due to several issues. For example, each block can consist of several buildings, which in turn may contain multiple elevators. The elevators are further divided into segments, carrying people to different floors/parts of the building. Thus, people often need to use multiple elevators to reach their destinations during a day. Typically, elevators are also in constant use. Finally, each building usually contains several access control points. In this study, we aimed to address some of these challenges by introducing a mobile application for elevator control.

Design Process

We first analysed the problems faced by elevator users based on earlier studies. There are several challenges with current systems that could potentially be improved with mobile elevator control:

- People may not know how elevators work in a building (which elevator goes where)
- People do not know in advance if there is space available in the elevator
- People may not know the optimal way to their destination
- Normal elevator door closing times do not support special (slow) movement patterns (e.g. heavy load, wheelchair users, etc.)

User experience goals for our mobile application were identified based on these challenges. An agile development process was utilised in this case. We iteratively designed and developed a prototype application that enables users to place elevator calls remotely to real elevators inside an office building. The mobile application communicates wirelessly with the elevator scheduling system in the building. The design and development process was continuously informed by feedback from elevator industry professionals, who also provided us with the opportunity to evaluate the prototype application in a real context of use. More details of the application can be found in (Turunen et al. 2013).

For the first application prototype we organized an initial user experience evaluation and subsequent long-term evaluation with four participants. The second prototype was evaluated long-term with 29 participants, 12 of whom were interviewed in detail about their experiences.

UX Goals

The UX goals address the identified challenges in current elevator systems through the lens of supporting 'people flow', which is the brand promise of the corporation:

(1) Expediting movement in large buildings

(2) Feeling of control of elevator action

(3) Reduced feeling of waiting

(4) Possibility for remote operation of elevator

Expediting movement facilitates a positive experience of the overall indoor journey from entry to the destination. The moments of waiting tend to cut the movement flow, thus a specific goal is to affect positively the *feeling of waiting*. *Feeling of control* is important to facilitate that the user has influence on elevator actions. *Remote operation* further extends the feeling of control and promotes a more personalized feeling. Findings from the user studies indicated the value of personalized scheduling options that take into account daily movement patterns.

3.2 Gesture–based Interaction in Metal Workshops

In a factory automation system, the loading station environment is dedicated to loading and unloading machining pallets. The load can be transferred, lowered, rotated and/or tilted to give the operator the best possible access to the work pieces. Traditionally, the operator controls the movements of the loading station by push buttons or switches placed away from the pallet for reasons of operator safety. In crowded workshop conditions, the controls can be hard to reach, and their operation requires constant movement from the pallet to the controls and back.

This case study aimed to address these challenges with a radically new gesture-based interaction concept. The focus of the design was to provide a natural interaction concept for controlling the loading stations and to investigate how different design requirements (naturalness of gestures vs. robustness of gesture detection) can be accounted for in the design of the gesture set.

Design Process

An agile development process was also utilised in this case. The design process consisted of an examination of the metal workshop domain, including the context of use, current interaction methods and the work process, followed by an iterative development cycle. Domain experts from the participating company were used as informants in order to form an understanding of user requirements. A set of preliminary gestures was analysed in laboratory studies to show that performing the gestures was associated with emotional user experience. This understanding was utilized later in the field studies.

A design workshop was conducted to form the basis for the robust gesture set used in the prototype. This gesture set, and the accompanying visualization, was then refined iteratively until the final prototype stage was reached (Figure 1). During this process, researchers demonstrated features of the gesture recognition technology through interactive prototypes, and domain experts proposed changes and provided feedback. The user acceptance and user experience of the concept were evaluated in real contexts of use in metal workshops. More details about the findings are presented in (Heimonen et al. 2013).

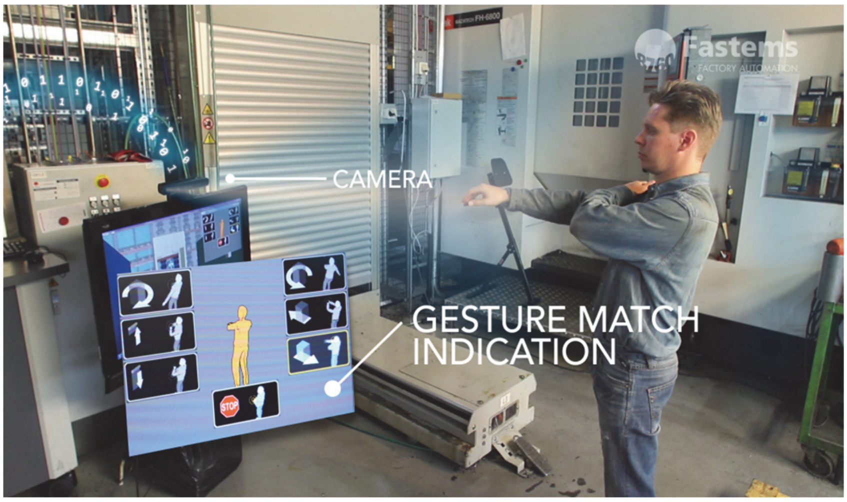


Figure 1. Loading station environment with gesture-based interaction.

UX Goals

The user experience goals defined for the gesture-based concept were:

(1) Using the system feels like magic

(2) Sense of control over the system

The *feels like magic* goal indicates a need to provide something radically new that would surprise the user. Entertaining and intuitive interaction should not require excessive effort. However, the user should still have *sense of control*. This goal indicates the need for gestures that are easy to learn, simple to perform and whose detection is robust. Both user experience goals contribute towards desirable customer values of increased productivity, attractiveness of the workplace, and a cutting edge image of the company.

3.3 Smart Interaction with a Crane

The goal of the Smart interaction with crane (SmartGUI) case was to understand how automated smart features of an electronic overhead travelling (EOT) gantry crane affect user experience, and how this should be taken into account when designing new user interfaces for the crane. The EOT crane is a crane with a hoist travelling along a girder between parallel runways (Figure 2, crane controller in Figure 3). EOT cranes are typically used for material handling in industrial processes.

Figure 2. Operating an EOT crane in a factory.

Figure 3. The crane controller

Design Process

The design framework in this case was based on user psychology (Saariluoma and Oulasvirta 2010). The core of this design approach is that every design solution should be based on psychologically valid and coherent concepts and theories of the problem domain. The case started with 11 semi-structured interviews with crane operators. The main finding from the interviews with crane operators was a set of subjective experiential goals and problems relating to crane operation, especially in the context of increasing automation.

Positive and negative experiences were analysed separately, and were given an emotional theme. This resulted in defining two user experience goals: supporting competence and avoiding anxiety. To understand the goals in more detail, a laboratory study was conducted with 20 users that were not familiar with operating cranes.

After the UX goals were defined, their experiential aspects were assessed, and a set of heuristics was proposed. In a design workshop the participants were presented with the UX goals, their experiential aspects, their relation to crane automation, and a set of heuristics to be utilised in the conceptualising process. The workshop resulted in multiple concepts, which were evaluated against the user experience goals and heuristics. The most suitable concepts were implemented, and the prototype was evaluated in a field experiment with four crane operators and one designer. The field experiment revealed that the interface supported the set user experience goals, but also suggested a set of improvements for the next iteration of the interface.

UX Goals

In this case we had two high-level user experience goals:

1. Supporting competence

2. Avoiding anxiety

Competence refers to the user's ability to conduct tasks efficiently and skilfully and the feeling that results from an understanding of how one's own skills led to efficient task completion (Saariluoma and Jokinen 2014). Anxiety, on the other hand, is the result of not being in control of the automation and being obstructed from an efficient task accomplishment (Saariluoma and Jokinen 2014). *Supporting competence* UX goal indicates that all design decisions should support a positive understanding of one's own abilities. This combines such experiential goals as determination, motivation, and freedom of choice. *Avoiding anxiety* goal indicates that possible user experience problems, such as being alarmed or nervous during crane operation should be foreseen and avoided with the design decisions.

3.4 Remote Operation of a Container Crane

In this case study, we developed a new remote operator station (ROS) user interface (UI) concept for the remote operation of semi-automated container cranes. The cranes are operated manually from a remote office environment through dedicated ROSs during loading and unloading of external road trucks and other types of chassis in the landside loading zone (see the fenced area in the mid right-hand side of Figure 4).

Figure 4. A visualisation illustrating cranes in a port environment.

Design Process

The aim of this case was to differentiate the new ROS UI design from the existing solutions by focusing particularly on the crane operator's UX in the design. The main vision for the new ROS was defined to be hands-on experience in remote operation, as we wanted the remote operation with the UI to feel as vivid and safe as it would be carried out on-site where the crane is located. The design activities of the case were conducted in a similar way to most concept development processes (e.g., Keinonen and Takala 2006), but with a particular focus on UX-related matters, as depicted for example in the Understand-Envision-Create process by Desmet and Schifferstein (2011).

In defining the UX goals, we first used the systems usability framework (Savioja and Norros 2013) as the starting point. In particular, the framework's 'User experience: The development potential of use' (Savioja, Liinasuo and Koskinen 2013) perspective on activity was utilized. These considerations resulted in a first set of UX goals, which included, for example, the goals of feeling of a well-functioning tool, appropriate trust in technology, and sense of control (see Koskinen, Karvonen and Tokkonen 2013 for a complete list of UX goals in this phase)

Next, the concept specification phase was embarked on. In this phase, we first familiarized ourselves with the domain environment and the crane operation work by conducting literature-based investigations. The literature review included, for example, a benchmarking study of other similar remote operation solutions. After this phase, we created an initial and broad set of possible UX goals, which included, for example, feeling of presence in addition to the previously mentioned goals. In order to validate and refine the generated broad set of UX goals, we conducted pilot interviews with two domain experts. Based on the results of these interviews, sense of control and feeling of presence were chosen as the main goals to be investigated in the field studies.

The actual field studies (see Karvonen, Koskinen and Haggrén 2012b for a detailed description) were conducted in two international container terminals with altogether 12 crane operators. The studies focused on the analysis of the chosen UX goals (i.e., what they actually mean in the operators' everyday work) and on the analysis of the domain and crane operation work activity. Methodologically, the studies included interviews and observations, which were based on core-task analysis (Norros 2004) and critical decision method (e.g. Wong 2006). The field studies resulted in adding the feeling of safe operation and experience of fluent co-operation to the list of potential UX goals, since the study results highlighted the importance of these goals.

After the field studies, we analysed the gathered data according to the core-task analysis framework and, based on this analysis, chose the final UX goals to guide the concept development work. On the basis of the UX goals and user requirements, a virtual reality-based prototype system of the ROS (see Figure 5 for a concept illustration) was built in the project.

Figure 5. Concept illustration of the ROS system.

UX Goals

The final set of UX goals to guide the concept design work in this case were:

(1) Feeling of safe operation,

(2) Sense of control,

(3) Feeling of presence, and

(4) Experience of fluent co-operation

Feeling of safe operation is especially important in this context as the cranes are lifting heavy loads, and human lives can be in danger if something goes wrong. *Sense of control* is crucial as the remote operator does not have direct touch with the crane. Similarly, *Feeling of presence* is important as the remote operator is not physically present at the site and (s)he still has to perceive the prevailing conditions in the object environment vividly and at a sufficient level of realism. Finally, experience of *fluent co-operation* was also chosen, because the crane operation work is – against our initial conceptions – a very social activity with a great deal of communication between different professionals.

3.5 Analysis of the Case Studies

The case studies were each using several approaches to UX goal setting. All four cases focused on developing radically new interaction concepts by introducing new technologies to the usage context. It is no wonder that the possibilities and challenges of new technology can in all the cases be identified as a source of UX goals. The anticipated possibilities offered by new technology can be identified in UX goals such as *possibility for remote operation* (Mobile interaction with elevators), and *feels like magic* (Gesture-based interaction). Technology also influences UX goals so that the goals aim to prevent or minimize threats raised by the technology, e.g. *feeling of competence* that automation and smart features may reduce (Smart GUI). Another example of preventing the threats of technology is *feeling of presence* and *sense of control* that remote operation may tend to reduce (Remote operation of a crane).

A common denominator for the cases was also a strong emphasis on user needs, values and preferences. Thorough user understanding was a source for user experience goals in all the cases. The cases aimed at stepping into the users' shoes and understanding the users' world with empathy. The empathy was gained from user observations and interviews, as well as interviews with domain experts. Empathy-based UX goals can be identified e.g. in the Smart GUI case, where emotional aspects are clearly present in the high level UX goals *avoiding anxiety* and *competence support*. In the case of Remote operation of a crane, empathy was crucial in understanding the importance of the UX goals *feeling of safe operation* and *fluent communication*.

In addition to understanding the users with empathy, a theory-based approach to user understanding can also be identified in the cases. Emotional UX (Saariluoma and Jokinen 2014) was used as the theoretical background in the cases of gesture-based interaction and Smart GUI. Theoretical background on human activities in work environments based on systems usability and core-task analysis (Savioja and Norros 2013) helped in identifying an initial broad set of UX goals in the case of Remote operation of a crane. Theory-based approaches helped to set up a framework for UX goals, whereas empathic understanding of the particular users helped in identifying the most crucial UX goals in the individual cases.

In two cases, the company brand can be identified as a source of UX goals. In the case of mobile interaction with an elevator the 'People flow' brand of the company is, as such, user experience-oriented: it describes how the company wants the users to feel about using their elevators. In the case Gesture-based interaction in metal workshops, company brand can be seen as a source for UX goals in another way: the company wanted to emphasize their image as an innovative forerunner company with a radically new interaction concept, which is reflected in the *feels like magic* UX goal.

Another source for the *feels like magic* UX goal is a vision of renewal. Renewal was also a common theme in the cases: we wanted to show that UX can be a source of radical renewals. The aim of radically renewing current interaction or operational practices can also be identified in the cases Mobile interaction with elevators and Remote operation of a crane. The time span of renewal can vary, from products that could be realised quickly on the market, to more futuristic concepts. For example, the gesture-based interaction case was an exploration of future interaction possibilities without an immediate plan to shift into product development. The other three cases were targeting actual products in the near future, and in one case (Mobile interaction with elevators), the concept development launched product development within the partner company, which resulted in a commercial product.

The UX goal-setting proceeded in the cases in different ways. In the Gesture-based interaction case and in the Mobile interaction with elevators case, different research activities and participating individuals produced knowledge that was analysed, and UX goals were defined accordingly. The cases Smart interaction with a crane and Remote operation of a crane refined UX goals gradually, based on successive research activities. Multiple viewpoints were used in the UX goal-setting, to integrate the views of users: what kinds of experiences they value; of designers: what kinds of experiences can be facilitated; and of the company: what kinds of experiences the company wants to provide for customers.

The case studies revealed that there are several different approaches to defining UX goals, and that using different approaches together brings in the viewpoints of different stakeholders. In the following, in Section 4, we will analyse our case study findings further, and we will integrate them with related research to identify, classify, and analyse different UX goal-setting approaches.

4. Approaches to Defining User Experience Goals

Partly parallel to the case studies, we carried out a literature study to ascertain what kinds of approaches have been used to define goals for user experience. Three researchers independently searched publications that focused on design rather than mere evaluation of the user experience. We studied publications from 1995–2013, and used “user experience” and “design” as the main search criteria. Most of the papers did not use the term “UX goal”, but in the papers, we could still identify design targets related to user experience. The findings were shared and then each researcher studied selected papers in more detail. We finally chose 46 papers that dealt with UX goal-setting. In three consecutive workshops, the three researchers discussed the similarities and differences between the approaches, and iterated a framework in which the papers could be positioned. We started with a framework that included four viewpoints:

(1) UX inspiration from a designer's empathic understanding of the users' world
(2) UX inspiration from user experience in a different field
(3) UX targets identified starting from basic needs and user values
(4) UX targets identified based on possibilities and challenges of a new technology

During the iteration, we complemented the framework with an approach based on company brand. There were not many papers focused on this approach but we clearly identified it as an approach on its own. We also refined the definitions of the other approaches. We considered whether co-design, meaning user involvement in the goal-setting, should be defined as an approach on its own, but we decided to include this viewpoint as part of the empathic understanding of the users' world, according to the original idea of co-design, to better understand users' dreams and experiences (Sanders and Dandavate 1999).

We ended up with a framework of five different approaches to getting insight or inspiration for UX goal-setting:

(1) Company or brand image (Brand)
(2) Scientific understanding of human beings (Theory)
(3) Empathic understanding of the users' world (Empathy)
(4) Possibilities and challenges of a new technology (Technology)
(5) Reasons for product existence and envisioning renewal (Vision)

In our own cases, we could identify these approaches as illustrated in Table 1. The approaches were used in parallel or sequentially. When used in parallel, each approach contributed information to setting UX goals. When used sequentially, each approach refined the UX goals defined in the earlier phases.

Table 1. An overview of the approaches used in our four cases to define UX goals.

Case **UX approach**	**Mobile interaction with elevator**	**Gesture-based interaction in metal workshops**	**Smart interaction with a crane**	**Remote operation of a crane**
Brand	People flow brand	Company image as an innovator	-	-
Theory	-	Emotional UX (Saariluoma and Jokinen 2014)	Emotional UX (Saariluoma and Jokinen 2014)	Systems usability (Savioja and Norros 2013) Core-task analysis (Norros 2004)
Empathy	Existing understanding of users' challenges in complex environments	Existing understanding of user tasks and context of use	User interviews	Field observations and user interviews based on the core-task analysis method
Technology	Mobile interaction	Gesture-based interaction	Automated, smart features	Remote operation technologies
Vision	Remote elevator operation	Freeing the user from physical control devices	-	Hands-on experience in remote operation

An overview of our literature study findings is presented in Table 2, showing the papers that we connected to each approach. Based on our analysis, it seems that different experience design 'schools' lean on one chosen approach rather than combine the approaches above. The reason may be that scientific papers typically only report the most effective, influential, and context-dependent starting points of experience design (e.g., Hassenzahl, Diefenbach and Göritz 2010; Korhonen, Montola and Arrasvuori 2009). Thus, it is difficult to trace the origins of the study to reveal how many sources of experience goals have actually been explored and trialled in a specific design case. Therefore, we can only conclude that the most reported starting point in the scope of our literature study is empathic understanding of the users' world and scientific understanding of human beings. On one hand, the way we categorise an individual study in one category is helpful in identifying which source is dominantly reported in current experience design research. On the other hand, we assume that experience goal-setting leans on more than one approach in practice, but the current lack of knowledge about these approaches makes it hard to analyse and report this process. Our present work addresses this lack.

Table 2. Approaches to UX goal setting identified from the literature study

Approach	Case studies	Other papers
Brand	Roto and Rautava 2008; Stompff 2003	Schifferstein, Kleinsmann and Jepma 2012
Theory	Lucero and Arrasvuori 2010; Olsson et al. 2012; Savioja and Norros 2013	Abeele and Zaman 2009; Desmet and Hekkert 2007; Hassenzahl 2003; Hassenzahl, Diefenbach and Göritz 2010; Korhonen, Montola and Arrasvuori 2009; Saariluoma and Jokinen 2014
Empathy	Blythe et al 2006; Gruen et al. 2002; Kujala 2008; Mattelmäki and Battarbee 2002; Nielsen 2002; Özçelik Buskermolen, Terken and Eggen 2012; Sanders and Stappers 2012; Väätäjä et al. 2012; Gaver, Dunne and Pacenti 1999; Leonard and Rayport 1997	Edvardsson et al. 2011; Kaasinen, Koskela-Huotari et al. 2012; Kouprie and Sleeswijk Visser 2009; Sanders and Dandavate 1999; Sanders and Stappers 2008; Sleesswijk Visser, 2009; Wright and McCarthy 2008; Wright, Wallace and McCarthy 2008
Technology	Bowman and MacMahan 2007; Häikiö et al. 2007; Jumisko-Pyykkö, Weitzel and Strohmeier 2008; Ljungblad 2008; Mäntyjärvi et al. 2004; Olsson 2012	Kaasinen, Kymäläinen et al. 2012; Väänänen-Vainio-Mattila, Väätäjä and Vainio 2009
Vision	Desmet and Schifferstein 2011; Hekkert, Mostert and Stompff 2003	Hekkert, van Dijk and Lloyd 2011
Other	Buxton 2007; Shedroff 2001; Sweet 1999	Forlizzi and Battarbee 2004; Roto et al. 2010

In what follows, we will define and describe the five approaches to UX goal-setting in more detail. We will give examples of how these approaches showed in our own case studies and in related research.

4.1 UX Goals Derived from Company and Brand Image (Brand)

Perhaps the most obvious source of UX goals is company and brand identity. In the mobile interaction with elevator case we utilised the brand promise that was focused on user experience (People Flow) as a source for UX goals. In the gesture-based interaction case, we identified a more general need to highlight the company as an innovative forerunner, and this was also shown in UX goal setting.

The brand-based approach is based on the idea that user experience of products should be in line with brand experience, the image that a company wants to convey to its customers. The web sites of many companies are good examples of how brand identity is visible in design. Stompff (2003) addresses the problem that brand values are typically not visible in physical products. Stompff sees that a long-term relationship between the company and the designers is needed until the brand values become visible in products. Roto and Rautava (2008) describe how Nokia's brand promise can be taken into account when defining UX goals for all of the company's products. They include both instrumental and non-instrumental aspects in four high-level UX goals.

Schifferstein, Kleinsmann and Jepma (2012) talk about experience-driven innovation rather than plain product design. They claim that it is not enough to change the UX-driven product design process, but user experience has implications at three levels of an organization: at the level of the company, at the level of the brands within the company, and at the level of the individual product or service offerings. Experience-driven innovation aims at a consistent company, brand, and product experience.

In the research, there seems to be a gap as regards the brand-driven approach to experience-driven product design. We believe this is because academic UX research has been relatively distant from brand experience research. In industrial cases brand should be a self-evident source for UX goals.

4.2 Deriving UX Goals from Scientific Understanding of Human Beings (Theory)

Psychological theories can be used to explain why some experiences are satisfying and engaging for a user. In our studies we were using emotional UX (Saariluoma and Jokinen 2014) as well as systems usability and core task analysis (Savioja and Norros 2013) as theoretical sources for UX goals. From the literature, we identified many other theoretical frameworks that have been used to define the goals for user experience. In the following we discuss some of those.

In 2003, Hassenzahl presented an influential hedonic-pragmatic model of user experience that highlights the importance of pleasurable experiences, such as stimulation, identification, and evocation, in addition to the traditional pragmatic, i.e. instrumental, aspects, such as efficiency and effectiveness (Hassenzahl 2003). According to him, the hedonic aspects address a person's be-goals, such as being competent, being related to others, or being special. In their recent work, Hassenzahl, Diefenbach and Göritz (2010) found that be-goals, or rather the universal psychological needs, are related to positive affect. Seven of these needs in particular: competence, relatedness, popularity, stimulation, meaning, security, and autonomy, are sources of positive experience with interactive technologies (Hassenzahl, Diefenbach and Göritz 2010).

Desmet and Hekkert (2007) introduced a general framework for product experience applying to the affective responses that can be experienced in human-product interaction. They discuss three distinct components or levels of product experiences: aesthetic experience, experience of meaning, and emotional experience.

As an example of a practical tool for setting UX goals based on scientific understanding human beings, we take the framework of playful experiences (PLEX) (Korhonen, Montola and Arrasvuori 2009). Based on the 22 different categories of playful experiences in this framework, Lucero and Arrasvuori (2010) introduced PLEX cards to help the different stakeholders in the design process. An example of utilizing PLEX cards as the starting point in designing for playful experiences is reported by Olsson et al. (2012).

The existing scientific UX frameworks include several user experience factors that can be employed as the basis for setting UX goals for design. Since the factors in the frameworks are at different abstraction levels, they may need either to be generalized or specified to serve as UX goals.

4.3 UX Inspiration from Designer's Empathic Understanding of Users' World (Empathy)

By understanding users with empathy, the designers can obtain inspiration for products and services that provide good user experience. In our cases we used observation and interviews to gain empathic understanding of the users. User studies are, indeed, a frequent way to determine UX goals (Väätäjä et al. 2012). Empathic design was introduced as a concept by Leonard and Rayport as early as in 1997, even if at that time they did not use the term user experience. They saw empathic design as a complementary approach to marketing research, contributing to the flow of ideas that still need further testing. When a company representative explores their customers' worlds with the eyes of a fresh observer, the company can redirect existing organizational capabilities to new markets. Wright and McCarthy (2008) see empathic approach to design as a part of the broader pragmatist approach to design. They see that “knowing the user in their lived and felt life” involves understanding what it feels like to be that person, and what their situation is like from their own perspective; i.e. empathy.

Wright, Wallace and McCarthy (2008) remind that good experience-centred design requires designers to engage with the users and their culture in rich ways in order to understand how the users make sense of technology in their lives. Empathy is at the heart of this approach. Kouprie and Sleeswijk Visser (2009) propose a framework for empathy in design “Stepping into and out of the user’s life”. Based on psychological literature, they distinguish two components of empathy: affective and cognitive. The affective component includes emotional response, feeling and identifying with the user: becoming the user. The cognitive component includes understanding, perspective taking and imaging the other: staying beside the user. Mattelmäki and Battarbee (2002) propose empathy probes to induce design empathy. With empathy probes the users can document their physical and social context, life style, attitudes, and experiences. The probes can be used to create an empathic and respectful dialogue between users and designers, and the probes support designers’ empathic understanding of users. Sleeswijk Visser (2009) emphasizes that knowing the users’ world is important for designer motivation, and stories are good tools to contribute to this understanding. Successfully communicated user information provides empathy for users and inspiration for product ideas.

Our user studies often revealed negative feelings such as anxiety, uncertainty or feeling alienated. These negative feelings can be interpreted to positive user experience goals such as avoiding anxiety in the smart crane operations, or sense of control and feeling of safe operation in the remote crane operation case. In the remote operation case, the user interviews emphasized the importance of fluent communication. Many work tasks include cooperation with teammates and fluent communication with them is a source of good user experience. Thus, especially when considering work environments, the viewpoint should cover also the work team in addition to the individual.

Co-design can be seen as one form of empathic design. In co-design the user's role changes from that of a passive research subject to that of an active design partner. Sanders and Dandavate (1999) were among the first to discuss "design for experiencing", and their work has inspired the co-design movement. They introduced Make Tools to access people's feelings, dreams and imaginations in order to gain inspiration for experience-driven design. Kujala (2008) showed that user involvement not only provides useful information about users' needs but also increases the understanding of users' values. Kaasinen, Koskela-Huotari et al. (2012) propose that co-design can be supported with inspiring physical or virtual spaces in which users, designers and other actors can meet informally and participate in design activities as equals.

All the above empathic approaches can provide information and inspiration for UX goal setting. Empathic understanding of the user's world makes it possible to step into the user's shoes and make decisions on the design details throughout the design phase. Furthermore, co-design enables making design decisions with the users.

4.4 UX Goals Identified Based on Possibilities and Challenges of a New Technology (Technology)

Technology push was one driver for change in all our case studies, as we were seeking for renewals through novel interaction concepts. We saw that, with user experience goals, we can ensure a smooth introduction of new technologies to the usage context. User experience goals help in drawing one's attention to the positive experiences that the technology can facilitate and, on the other hand, user experience goals can focus on minimising the anticipated negative experiences such as a lost sense of control or a lost feeling of presence. Technology-driven design, or "blue sky" technology research as described by Rogers and Bellotti (1997), is focused on developing novel technological solutions that often look beyond immediate commercialization. Friction exists between these technology-oriented design approaches and the need to ground the designs in the practical needs and wants of users. For example, Ljungblad (2008) summarizes previous criticism on the design of ubiquitous computing systems, noting that research often investigates the development of novel technological solutions, and that the actual scenarios are not properly justified or based on existing practice.

Our case studies give evidence that technology-based approaches can support UX goal setting if at an early stage one studies the positive and negative user experiences that the technology can cause. UX goals can then be set to strengthen the positive experiences (such as *feels like magic* in one of our case studies) and to overcome the negative experiences (such as *feeling of presence* or *sense of control* as UX goals to minimize the negative experiences or remote control). There are various studies in which user experience research has been carried out in relation to the development of new technologies, such as the studies by Kaasinen, Kymäläinen et al. (2012) of intelligent environments, Olsson's (2012) studies of mobile augmented reality, the studies by Väänänen-Vainio-Mattila, Väätäjä and Vainio (2009) of service user experience of Web 2.0 and the studies by Bowman and MacMahan (2007) of immersion in virtual environments. The above-mentioned studies aim to identify UX issues related to a certain technology by concluding findings from several studies. Even if the studies are based on evaluation results, they do introduce challenges and possibilities that can be utilised in UX goal setting in forthcoming design activities.

As the above findings show, there are quite a lot of research results of the possibilities and threats of different interaction technologies. These results provide a good basis for defining user experience goals in order to utilise the possibilities and minimise the threats. However, focusing on those possibilities and threats alone may be too narrow a view of overall user experience.

4.5 UX Inspiration from Investigating the Deep Reasons for Product Existence and Envisioning Renewal (Vision)

Sometimes UX inspiration comes from investigating the deep reasons for product existence and envisioning renewal: vision from desirable possibilities, often taking inspiration from other fields. In our case studies, the case of Remote operation of a crane looked for inspiration from other fields such as space operations, telesurgery and mining. Mobile interaction in other fields was an inspiration for the case of Mobile interaction with elevator.

Hekkert, Mostert and Stompff (2003) use the Vision in product design (ViP) approach for experience-driven design. They propose that innovative product design can be achieved by first abandoning presuppositions about the product and then developing the product by formulating three visions: a context vision at an appropriate level of abstraction; this is then advanced to an interaction vision, which states how the user interacts with the product, and finally to a product vision. ViP forces designers to free themselves from apparent restrictions or requirements and, instead, look for desirable possibilities. The designer empathizes with the future user, but the user is not involved in the design process. Hekkert, Mostert and Stompff state that, in this way, undesirable constraints resulting from the user fixations on familiar solution directions are avoided.

Desmet and Schifferstein (2011) do not provide a specific process for designing for experience, but rather a set of activities that one can utilize as needed. They divide these activities into three categories: Understand, Envision, and Create. Activities in the Understand category aim at understanding the user and usage situation. Envision activities help define the UX goal, while Create activities help conceptualize, materialize, and test new concepts. Envision activities include envisioning the UX goal and user-product interaction, as well as formulating the target product appraisal and the target product character.

The vision-based approach has good potential in creating something totally new, but as the connection to the user's world is quite loose, user acceptance of the visionary solution may not be guaranteed.

4.6 Analysis of the UX goal-setting approaches

In the following, we will further analyse the five UX goal-setting approaches. We aim to answer Research question 2: "What kind of contribution do these approaches make in defining UX goals: What kinds of UX goals do they produce? What are their benefits and challenges?".

The Theory and Empathy approaches aim to understand future users and their world, and to find a vision for desirable UX goals from there. The Brand, Technology, and Vision approaches are not as directly focused on users, but these approaches aim to find inspiration for the UX vision from the brand identity, technology, or the reason for existence of the product. Theory and Technology aim to find insights into setting UX goals and measurable targets, whereas Brand and Vision are more focused on finding inspiration. The Empathy approach is focused on both insight and inspiration. The five approaches bring in different viewpoints to UX goal setting and the resulted UX goals differ in the following way:

- *Brand*-based approaches can produce focused and easy-to-share UX goals such as “Connecting People” or “People flow”. The high-level user experience goal may be directly available as the brand promise. The brand promise may need interpretation, as it may not be self-evident how the brand promise should show in an individual product.
- *Theory*-based approaches often provide a set of user experience goals that cover thoroughly different aspects. The most important UX goals have to be chosen from the alternatives, for example based on the results of user studies or brand identity. UX goals in work environments may be different from goals typically used for consumer systems. General user experience goals may mean different things in different domains and to different users, so they need interpretation for each specific design case.
- *Empathy*-based approaches have potential in giving access to the deep emotional aspects of the users’ world. When designing for work environments, empathy alone is not enough. Thorough domain and work analysis is needed, based on extensive studies of work activities and domains.
- *Technology*-based approaches may provide focused UX goals, but focusing on the possibilities and threats of a certain new technology does not necessarily cover all aspects of the overall usage situation.
- *Vision*-based approaches can help to define positive user experience goals that can renew the product. UX goals may be different when exploring future possibilities versus developing something to be put into actual use in the near future. Challenges can arise when the vision is far from the users’ current practices, as user acceptance is difficult to foresee.

In Table 3, we present a summary of the contributions that the five approaches make to UX goal-setting. In Table 4, we describe the benefits and challenges of the five approaches, based on our findings.

Table 3. The contributions of the five approaches to defining UX goals.

Approach	**Contribution to UX goal setting**
Brand	A high-level UX vision to unite products under the same brand
Theory	A collection of possible UX goals to choose from
Empathy	A mind-set focusing on the users' world
Technology	UX possibilities and UX challenges raised by a technical enabler
Vision	Getting rid of fixations on familiar solutions, inspiration from other domains

Table 4. Benefits and challenges of the five approaches to defining UX goals.

Approach	**Benefits**	**Challenges**
Brand	Pre-defined, focused, easy-to-share UX vision	Interpretation of the vision of UX goals for different products under the brand
Theory	Science-based evidence for the UX Goals	Choosing the ones to focus on from a wide set of possible UX goals
Empathy	Mind-set focus supports decision-making beyond the goal-setting phase	Gaining insight into the deep emotional aspects of differing users' worlds
Technology	UX goals support successful adoption of new technologies	Focusing on a certain technology may not cover all aspects of use
Vision	Support for renewal with UX goals	User acceptance of the visionary goals

The interplay between the different UX goal-setting approaches supports the multidisciplinary nature of UX and gives different stakeholders possibilities to contribute to the goal-setting. Using multiple approaches can produce multiple, even conflicting UX goal candidates. On the other hand, the different approaches may reveal similar goals, which gives evidence of the importance of those goals. In our cases, we finally chose a set of 2–4 UX goals, which kept the design work focused. Focusing on a few user experience goals helps in sharing the selected user experience goals, committing the design team to those goals, and keeping those goals in everyone's mind throughout the design process.

5. Discussion

Experience design is gaining ground as an approach to designing interactive systems that address the emotional, not only utilitarian, aspects of product use. The starting point and the core of experience design is the definition of UX goals. Still, there is no prior research on different approaches to defining these goals. According to our literature review, the current approaches to UX goal-setting seem to lean on one of several approaches, which keeps the different schools of experience design separate. We see that a more thorough understanding of the different approaches would strengthen the core activity of UX goal-setting, make reporting of experience design cases more systematic, and bring clarity to this growing but fragmented research field.

Based on the above goal, our first research question was to identify different approaches for defining UX goals. By reviewing the literature and by studying four cases of our own, we identified five approaches: Brand, Theory, Empathy, Technology, and Vision. The second research question focused on the contribution that these approaches make in defining UX goals. The different approaches each bring a different viewpoint. The *Brand* approach ensures that the user experience goals are in line with the company's brand promise. The *Theory* approach utilises the available scientific knowledge of human behaviour. The *Empathy* approach focuses on knowing the actual users and stepping into their shoes when defining UX goals. The *Technology* approach considers the new technologies that are being introduced and their positive or negative influence on UX. Finally, the *Vision* approach focuses on renewal – whether a new kind of product experience can be introduced. Due to the multidisciplinary nature of UX, it is beneficial to use as many of the approaches as possible. This, however, may conflict with the need to have a limited set of UX goals for practical design purposes. Thus, UX goal-setting requires consolidation of the contributions from the different approaches, so that the selected UX goals represent not just one, but several viewpoints. The way in which these viewpoints are emphasised in UX goal selection depends on the case at hand, and may be driven by the perspectives of the stakeholders.

Concrete UX goals are especially useful in industrial contexts, where various stakeholders need to agree on what to design. Without clear UX goals, user experience is easily left as a good intention without any concrete influence. Shared UX goals ensure that all who contribute to the design process have a clear conception of the targeted experience, and can make design decisions accordingly. In the design of industrial systems, the concrete UX goals help in keeping user experience in focus throughout the complex, multidisciplinary product development and marketing processes.

User experience is a multifaceted concept, and it can be questioned whether it is acceptable to try to narrow it down to specific UX goals. Is there a danger that we lose the idea of thinking widely about how people feel in different usage situations, if we focus too closely on a set of predefined UX goals? As in any design activity, focusing and concretising is necessary in order to clarify and communicate the design goals. However, we suggest that, at the same time, the wide UX viewpoint should be embraced when carrying out user and customer studies, so that it is possible to refine the initial UX goal setting when needed.

The targeted user experience should show at the different touch points with the user, such as marketing, maintenance, and customer service. This emphasises that setting UX goals is a strategic decision that will require representation from the whole organisation; the designers cannot make the decision alone. Based on earlier research, it might not be easy to convince all stakeholders about the importance of experiential aspects as design goals, especially when dealing with work-related products (Abramov and Roto 2012). Setting user experience goals can benefit from multidisciplinary cooperation with the key stakeholders, such as experts from design, marketing, and maintenance. Involving the different units fosters understanding of the different perspectives on product design, builds commitment for the UX goals throughout the organisation, and helps in planning marketing that is in line with the UX goals. In this paper we have focused on UX goal-setting. How the set UX goals serve the actual design and marketing processes will require further studies.

Our four case studies each followed their own UX goal setting process. However, the cases were part of a common research programme, where the researchers and practitioners shared a common vision of the necessity of concrete, focused UX goals in the design process. Each case involved a multidisciplinary design team, in which the participants brought into the UX goal-setting the viewpoints that they, according to their experience, felt necessary. In future design cases, utilising the proposed framework of the five approaches will ensure that different viewpoints and their contribution to the UX goal-setting can be considered even in less multidisciplinary design teams.

Our results indicate that the five identified approaches can be used for UX goal-setting. There may also be other approaches that future research can reveal. An interesting path for future research is to study the different time spans of user experience, namely anticipated, momentary, episodic, and cumulative UX (Roto et al. 2010). Moreover, additional case studies that systematically consider which approaches to use in UX goal-setting can provide more evidence of the benefits and challenges of each approach.

Once the user experience goals have been defined, the next challenge is to communicate them to the design team and other stakeholders, make the team commit to the goals, and utilise them in the design process. User experience goals are not the only goals guiding the design, and there may be other goals from other parts of the organisation regarding maintenance, price, compatibility, and so on. In practice, UX goals need to be integrated with these other goals, in order to make sure that there are no conflicts. Further work is needed in studying the practical consequences of the UX goals in the design solutions. Our future plan is to investigate how user experience goals are utilised in the later phases of experience-driven design, how the goals serve the design and marketing processes, and what kinds of challenges are encountered.

Acknowledgements

This study was carried out in User Experience and Usability of Complex Systems (FIMECC UXUS) programme by Finnish Metals and Engineering Competence Cluster (FIMECC). The programme has been partly funded by Tekes – the Finnish Funding Agency for Technology and Innovation. We wish to thank FIMECC and Tekes as well as the participating companies for their good cooperation.

References

Abeele, V. V., and B. Zaman. 2009. "Laddering the user experience!" http://www.kuleuven.be/facdep/social/com/mediac/cuo/admin/upload/Laddering%20the%20User%20Experience.pdf

Abramov, V. and V. Roto, 2012. "Accounting for Intermediate Parties in Experience-driven Product Design for Business-to-Business Environment." In *Proceedings of Human Work Interaction Design (HWID) working conference*. Copenhagen.

Beauregard, R., and P. Corriveau. 2007. "User experience quality: a conceptual framework for goal setting and measurement." In *Digital Human Modeling,* edited by V. G. Duffy, 325-332. Berlin Heidelberg: Springer.

Blythe, M. A., and P. C. Wright. 2006. "Pastiche scenarios: Fiction as a resource for user centred design." *Interacting with Computers* 18(5): 1139-1164.

Bowman, D. A., and R. P. McMahan. 2007. "Virtual reality: How much immersion is enough?" *Computer* 40(7): 36-43.

Buxton, W. 2007. *Sketching user experiences: getting the design right and the right design.* Morgan Kaufmann.
Desmet, P. M. A., and H. N. J. Schifferstein (eds). 2011. *From Floating Wheelchairs to Mobile Car Parks: a Collection of 35 Experience-Driven Design Projects*. Den Haag, NL: Eleven Publishers.
Desmet, P. M., and P. Hekkert, P. 2007. "Framework of product experience." *International Journal of Design* 1(1): 57-66.
Edvardsson, B., P. Kristensson, P. Magnusson, and E. Sundström, E. 2012. "Customer integration within service development—A review of methods and an analysis of insitu and exsitu contributions." *Technovation* 32(7): 419-429.
Forlizzi, J., and K. Battarbee. 2004. "Understanding experience in interactive systems." In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, 261-268. ACM.
Gaver, B., T. Dunne, and E. Pacenti. 1999. "Design: Cultural probes." *Interactions* 6(1): 21-29.
Gefen, D. 2000. "E-commerce: the role of familiarity and trust." *Omega 28.6* (2000): 725-737.
Gruen, D., T. Rauch, S. Redpath, and S. Ruettinger. 2002. "The use of stories in user experience design." *International Journal of Human-Computer Interaction* 14(3-4):503-534.
Häikiö, J., A. Wallin, M. Isomursu, H. Ailisto, T. _Matinmikko, and T. Huomo. 2007. "Touch-based user interface for elderly users." In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services*, 289-296. ACM.
Hartson, R., and P. S. Pyla. 2012. *The UX Book: Process and Guidelines for Ensuring a Quality User Experience*. Waltham, MA: Elsevier.
Hassenzahl, M. 2003. "The Thing and I: Understanding the Relationship Between User and Product." In *Funology: From usability to user enjoyment*, edited by Blythe, M.A., K. Overbeeke, A. F. Monk, and P. C. Wright. Netherlands: Kluwer Academic Publishers.
Hassenzahl, M. 2010. *Experience Design – Technology for All the Right Reasons*. Morgan & Claypool.
Hassenzahl, M. and N. Tractinsky. 2006. "User experience - a research agenda." *Behaviour & Information Technology* 25(2): 91-97.
Hassenzahl, M., S. Diefenbach, and A. Göritz. 2010. "Needs, affect, and interactive products–Facets of user experience." *Interacting with Computers* 22(5): 353-362.
Heimonen, T., J. Hakulinen, M. Turunen, J. Jokinen, T. Keskinen, and R. Raisamo. 2013. "Designing Gesture-based Control for Factory Automation." In *Proceedings of INTERACT 2013, Part II, LNCS 8118*, 202-209. Springer.
Hekkert, P., M. Mostert, and G. Stompff. 2003. "Dancing with a machine: a case of experience-driven design." In *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces (DPPI '03)*, 114-119. New Your: ACM.
Hekkert, P., M. van Dijk, and P. Lloyd. 2011. *Vision in Product Design: handbook for innovators*. Amsterdam: BIS Publishers.
ISO. 2010. "ISO 9241-210. Human-centred design for interactive systems. International Standard." International Organization for Standardization.

Järvenpää, S.L., and N. Tractinsky. 1999. “Consumer Trust in an Internet Store: A Cross-Cultural Validation.” *Journal of Computer-Mediated Communication* 5(2).

Jordan, P.W. 2000. *Designing Pleasurable Products*. London: Taylor & Francis.

Jumisko-Pyykkö, S., M. Weitzel, and D. Strohmeier. 2008. "Designing for user experience: What to expect from mobile 3D TV and video?" In *Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video*, 183-192. ACM.

Kaasinen, E., K. Koskela-Huotari, V. Ikonen, M. Niemelä, and P. Näkki. 2012. “Three approaches to co-creating services with users.” In *Proceedings of 1st International Conference on Human Side of Service Engineering*, San Francisco: USA Publishing. DVD publication.

Kaasinen, E., T. Kymäläinen, M. Niemelä, T. Olsson, M. Kanerva, and V. Ikonen. 2012. “A user-centric view of intelligent environments: User expectations, user experience and user role in building intelligent environments.” *Computers* 2(1): 1-33.

Karvonen, H., H. Koskinen, and J. Haggrén. 2012a. ”Defining User Experience Goals for Future Concepts. A Case Study.” In: *NordiCHI2012 UX Goals 2012 Workshop proceedings*, edited by Väätäjä, H., T. Olsson, V. Roto, and P. Savioja, 14-19. Tampere: TUT Publication series. http://URN.fi/URN:ISBN:978-952-15-2955-9.

Karvonen, H., H. Koskinen, and J. Haggrén. 2012b. “Enhancing the User Experience of the Crane Operator: Comparing Work Demands in Two Operational Settings.” In *Proceedings of the 30th European Conference on Cognitive Ergonomics*, 37-44. Edinburgh.

Karvonen, K. 2000. "The beauty of simplicity." In *Proceedings on the 2000 conference on Universal Usability*. ACM.

Keinonen, T. and R. Takala. 2006. *Product concept design: A review of conceptual design of products in industry*. Springer.

Korhonen, H., M. Montola, and J. Arrasvuori. 2009. “Understanding Playful Experiences Through Digital Games.” In *Proceedings of DPPI’09*. ACM Press.

Koskinen, H., H. Karvonen, and H. Tokkonen, H. 2013. ”User experience targets as design drivers: a case study on the development of a remote crane operator station.” In *Proceedings of the 31st European Conference on Cognitive Ergonomics*, 25. ACM.

Kouprie, M., and F. S. Sleeswijk Visser. 2009. “A framework for empathy in design: Stepping into and out of the user's life.” *Journal of Engineering Design* 20(5): 437-448.

Kujala, S. 2008. “Effective user involvement in product development by improving the analysis of user needs.” *Behaviour & Information Technology* 27(6): 457-473.

Law, E. L. C., V. Roto, M. Hassenzahl, A. P. Vermeeren, and J. Kort. 2009. “Understanding, scoping and defining user experience: a survey approach.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 719-728. ACM.

Leonard, D., and J. F. Rayport. 1997. “Spark innovation through empathic design.” *Harvard Business Review* 75:102-115.

Levy, P. 2013. “Beyond kansei engineering: The emancipation of kansei design.” *International Journal of Design* 7(2): 83-94.

Ljungblad, S. 2008. “Beyond users: grounding technology in experience.” PhD diss. Stockholm University.

Lu, Y., and V. Roto. 2014. “Towards Meaning Change: Experience Goals Driving Design Space Expansion.” In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction*, 717-726. ACM.

Lucero, A., and J. Arrasvuori. 2010. “PLEX Cards: a source of inspiration when designing for playfulness.” In *Proceedings of the 3rd International Conference on Fun and Games*, 28-37. ACM.

Mahlke, S. 2005. “Understanding users' experience of interaction.” In *Proceedings of the 2005 annual conference on European association of cognitive ergonomics*, 251-254.

Mäntyjärvi, J., J. Kela, P. Korpipää, and S. Kallio. 2004. "Enabling fast and effortless customisation in accelerometer based gesture interaction." In *Proceedings of the 3rd International Conference on Mobile and Ubiquitous Multimedia*, 25-31. ACM.

Mattelmäki, T., and K. Battarbee. 2002. “Empathy probes”. Paper presented at the PDC, 266-271.

Nagamachi, M. 2002. “Kansei engineering as a powerful consumer-oriented technology for product development.” *Applied Ergonomics* 33: 289–294.

Nielsen, L. 2002. "From user to character: an investigation into user-descriptions in scenarios." In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*, 99-104. ACM.

Norman, D. A. 2004. *Emotional design: Why we love (or hate) everyday things*. Basic books.

Norros, L. 2004. *Acting under uncertainty – The core-task analysis in ecological study of work*. Espoo: VTT Publications.

Olsson, T. 2012. “User expectations and experiences of mobile augmented reality services.” PhD diss. Tampere University of Technology.

Olsson, T., E. Lagerstam, T. Kärkkäinen, and K. Väänänen-Vainio-Mattila. 2013. ”Expected user experience of mobile augmented reality services: A user study in the context of shopping centres.” *Personal and Ubiquitous Computing* 17(2): 287-304.

Olsson, T., K. Väänänen-Vainio-Mattila, T. Saari, J. Arrasvuori, and A. Lucero. 2012. ”Playful Experiences as Targets in Concept Design.” In *Proceedings of the Workshop on How to Utilize User Experience Goals in Design*, 27-31.

Özçelik Buskermolen, D., J. Terken, and B. Eggen. 2012. "Informing user experience design about users: Insights from practice." In *Proceedings of the 2012 ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems*, 1757-1762. ACM.

Rintamäki, T., H. Kuusela, and L. Mitronen. 2007. ”Identifying competitive customer value propositions in retailing.” *Managing service quality* 17(6): 621-634.

Rogers, Y. and V. Bellotti. 1997. “Grounding blue-sky research: how can ethnography help?” *Interactions* 4(3): 58-63.

Rogers, Y., H. Sharp, and J. Preece. 2011. *Interaction Design: Beyond Human-Computer Interaction*. Chichester, UK: John Wiley & Sons.

Roto, V., and M. Rautava. 2008. “User experience elements and brand promise.” In *Proceedings of International Engagability & Design Conference, in conjunction with NordiCHI’08*.

Roto, V., E. L-C. Law, A. P. O. S. Vermeeren, and J. Hoonhout (eds.). 2010. “Demarcating User eXperience”. In *Abstracts Collection of Dagstuhl Seminar 10373*, Dagstuhl, Germany.

http://drops.dagstuhl.de/opus/volltexte/2011/2949/pdf/10373_AbstractsCollection. 2949.pdf.
Saariluoma, P., A. Oulasvirta. 2010. “User psychology: Re-assessing the boundaries of a discipline.” *Psychology*, 1(317).
Saariluoma, P., and J. P. P. Jokinen. 2014. “Emotional dimensions of user experience: a user psychological analysis.” *International Journal of Human–Computer Interaction* 30(4): 303–320.
Sanders, E. B.-N. and P. J. Stappers. 2012. *Convivial Design Toolbox: Generative Research for the Front End of Design*. Amsterdam: BIS Publishers.
Sanders, E. B.-N., and U. Dandavate. 1999. “Design for experiencing: New tools.” In *Proceedings of the first international conference on Design & Emotion*, edited by Overbeeke, C. J., and P. Hekkert, 87—92.
Sanders, E.B.-N. and P. J. Stappers. 2008. "Co-creation and the new landscapes of design." *CoDesign: International Journal of CoCreation in Design and the Arts* 4(1): 5-18.
Savioja, P., and L. Norros. 2013. “Systems usability framework for evaluating tools in safety–critical work.” *Cognition, Technology & Work* 15(3): 255-275.
Savioja, P., M. Liinasuo and H. Koskinen. 2013. “User experience: does it matter in complex systems?” *Cognition, Technology & Work*. doi: 10.1007/s10111-013-0271-x.
Schifferstein, H. N., M. S. Kleinsmann, and E. J. Jepma. 2012. “Towards a conceptual framework for Experience-Driven Innovation.” In *Proceedings of 8th International Design and Emotion Conference*, London.
Shedroff, N. 2001. *Experience design*. Indianapolis: New Riders.
Sleeswijk Visser, F. 2009. “Bringing the everyday life of people into design.” PhD diss. TU Delft.
Stompff, G. 2003. “The forgotten bond: Brand identity and product design.” *Design Management Journal (Former Series)*14(1): 26-32.
Sweet, F. 1999. *Frog: Form Follows Emotion*. New York: Watson- Guptill Publications.
Turunen, M., H. Kuoppala, S. Kangas, J. Hella, T. Miettinen, T. Heimonen, T. Keskinen, J. Hakulinen, and R. Raisamo. 2013. ”Mobile Interaction with Elevators – Improving People Flow in Complex Buildings.” In *Proceedings of Academic Mindtrek ’13*, 43-50. ACM.
Väänänen-Vainio-Mattila, K., H. Väätäjä, and T. Vainio. 2009. ”Opportunities and challenges of designing the service user eXperience (SUX) in web 2.0.” In *Future interaction design II*, 117-139. Springer.
Väätäjä, H., T. Olsson, V. Roto, and P. Savioja (eds.). 2012. *How to Utilize User Experience Goals in Design? UX Goals 2012 Workshop proceedings*. Tampere: Tampere University of Technology.
Wikberg, H., and T. Keinonen. 2002. Ergonomics of Wearability as a Design Driver. A Case Study of User-Centered Design Process of Designing Mobile Phones and Accessories for Active Use. *HF*.
Wong, B. L. W. 2006. “The critical decision method.” In *International encyclopedia of ergonomics and human factors*, edited by Karwowski,W., 3067-3073. CRC/Taylor & Francis.
Wright, P., and J. McCarthy. 2008. “Empathy and experience in HCI.” In *Proceedings of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems*, 637-646.

Wright, P., and J. McCarthy. 2010. Experience-centered design: designers, users, and communities in dialogue. In *Synthesis Lectures on Human-Centered Informatics* 3(1): 1-123. Morgan & Claypool Publishers.

Wright, P., J. McCarthy, and L. Meekison. 2005. "Making sense of experience." In *Funology: From usability to enjoyment*, edited by Blythe, M., Overbeeke, C., Monk, A. F., and Wright, P. C., 43-53. Netherlands: Springer.

Wright, P., J. Wallace, and J. McCarthy. 2008. "Aesthetics and experience-centered design." *ACM Transaction on Computer-Human Interaction* 15(4): Article 18.

V

RADICAL INNOVATION BY THEORETICAL ABSTRACTION – A CHALLENGE FOR THE USER-CENTRED DESIGNER

by

Mikael Wahlström, Hannu Karvonen, Leena Norros, Jussi Jokinen, & Hanna Koskinen

The Design Journal vol 19, 857–877

https://doi.org/10.1080/14606925.2016.1216210

*Manuscript title page (with author names)
Click here to view linked References

Radical Innovation by Theoretical Abstraction – a Challenge for Design Anthropologists

Mikael Wahlström[a], Hannu Karvonen[a], Leena Norros[a], Jussi Jokinen[b], Hanna Koskinen[a]

[a] VTT Technical Research Centre of Finland

[b] University of Jyväskylä, Department of Computer Science

Author Note

Correspondence pertaining to this article should be addressed to Mikael Wahlström; P.O. Box VTT Technical Research Centre of Finland; Vuorimiehentie 3, Espoo (P.O. Box 1000); FI-02044 VTT, Finland; email mikael.wahlstrom@vtt.fi; tel. +358 50 360 7272; fax +358 20 722 7001.

This study is part of the research and development programme User Experience and Usability in Complex Systems (UXUS), which is one of the research programmes of the Finnish Metals and Engineering Competence Cluster, FIMECC.

The authors would like to thank all those who have provided input to the design of future ship bridges and to this manuscript, these being Iiro Lindborg (Rolls-Royce), Eija Kaasinen (VTT), Jussi Lehtiniemi (Troll VFX), Zhang Daoxiang (Aalto University), Petri Mannonen (Aalto University), Göran Granholm (VTT), Frøy Birte Bjørneseth (Rolls-Royce), Karno Tenovuo (Rolls-Royce), and Lassi Liikkanen (Helsinki Institute for Information Technology HIIT) among others.

Abstract

It is generally accepted that scientific disciplines such as psychology, sociology, and anthropology contribute beneficially to design by providing understanding of users' needs, experiences, and desires. Arguably, however, these disciplines have more to contribute, because they include theories and models that can be applied as design frames and principles. More specifically, goal-setting, visualisation, thematisation, and conceptual reconfiguration are general mechanisms through which theories translate into design contributions. Actualising radical design solutions via these mechanisms is discussed: theories provide appropriate means of abstraction, which allows "distance" from user data; departure from the existing design and user paradigms toward "what has not yet been imagined" is thereby possible. These suggestions draw from and are exemplified by a ship bridge design case.

Keywords: framing, design theory, design knowledge, case study, social theory

Radical Innovation by Theoretical Abstraction – a Challenge for Design Anthropologists

Design is generally regarded as purposeful and creative activity: it involves creating something new, such as a new physical object or service, for a certain human need (Buchanan, 2001; Cross, 2011). The creativity needed for these designs involves reasoning (Dorst, 2011) while creative activity, as does any activity, also involves the social and physical environment in which the creativity takes place (Glaveanu, 2013). The purpose of this study is to explore various ways in which sciences that explore human activity and experience, such as sociology, psychology, and anthropology, may provide input for the creative process of design. The aim is to pinpoint general "translation mechanisms" through which theoretical thinking and models in these research disciplines are transformed into design ideas. By mechanisms we here mean general practices that offer beneficial social and mental connections for design: we demonstrate how theories and theoretical models provide new design aims, design themes, visual models, and conceptualisations incorporated into the design process, by contrasting literature on "design thinking" (Cross, 2011; Sterwart, 2011) against a case study of a user-experience-oriented ship bridge design project.

The typical way of thinking about the potential input of sociology, psychology, or anthropology for design is that they provide means for making sense of the "people factor". Customer needs are identified, design ideas are collected from users, usability is tested, and so on. Indeed, in Kelley and Littman's (2005) description of good design teams, a professional social, or cognitive scientist takes the role of "anthropologist", the main purpose being to observe people. It is assumed that the design team's anthropologist is able to reframe the design problem through intuition and empathy drawing from the observations. Additionally, "user-centred design" (Vredenburg, Mao, Smith, & Carey, 2002) has been an influential trend. It suggests that user needs, desires, and capabilities should be central to the design process. For example, the importance of "user experience" has been emphasised, especially in design of computerised systems (Hassenzahl & Tractinsky, 2006): "experience design" suggests that the experience of users, rather than the products themselves, should be among the foci of design (Hassenzahl, 2010). The importance of understanding human experiences and capabilities implies that means of studying people may yield benefits for design.

It is noteworthy, however, that user-centred design has been criticised for not providing design solutions that surprise people by offering them new possibilities. It predominantly gives users what they knew they wanted, and the design solutions correspond to people's current activities, simply responding to the issues identified when the users were studied (Keinonen, 2009; Norman & Verganti, 2012). In other words, it seems that examining users' activities and wants does not provide "game-changing", "radical" innovations. In this, one may identify a problem to be solved: understanding users seems to provide beneficial design input, yet whether this input might stimulate radical design solutions remains unclear.

Several authors (Frascara, 2002; Karwowski, Soares, & Stanton, 2011; Saariluoma, 2005) have discussed applying psychological and social scientific research methods and theories in design. Some authors (e.g., Kuutti, 2005) discuss the design implications of certain specific theories. To the best of our knowledge, however, few have pinpointed general processes involved when radical design ideas are created in synthesis of user study findings with theory from social and cognitive sciences. Our case study features two theoretical approaches, core-task analysis (CTA) (Norros, 2004) and joint cognitive systems (Woods & Hollnagel, 2006), but these are used to exemplify the general use of theoretical thinking in design rather than discussed specifically. Our main research question is this: what are the processes through which theories of cognitive and social sciences provide design input? We do not propose a comprehensive answer but provide specific suggestions, which are new to design science and are especially relevant for professionals striving to generate radical design ideas rooted in exploring of users or customers.

Design literature emphasises that good designers view the element under design from a certain perspective, or frame, that gives structure to the creative process. Additionally, good designers employ systemic thinking as they draw connections between different and dissimilar phenomena, such as human behaviour, technology trends, culture, and physical realities (Cross, 2011). Further, it has been suggested that design creativity benefits from deliberate abstraction; that is, appropriate “reduction of possibilities” via guidance-providing models is essential: meaningfully generated simplified sketches, for example, have served as a basis for the discovery of architectural design solutions (Dogan & Nersessian, 2010). Drawing from these findings and in trying to find a solution to the problem outlined above, we propose, firstly, that social and cognitive sciences might provide useful frames for design via theories that afford systemic thinking. These research disciplines could provide theory-based design frames related to the phenomena that the theories are explaining and describing. Secondly, studying users may allow radical design solutions when the data are viewed through theoretical concepts and models: these provide abstracted interpretations of the data; that is, the findings are, in a sense, reduced from specifics through reformulation in accordance with the concepts and models of the theories. In tandem with the process of abstraction, the process of concretisation is visible when the theories are applied in design – the theoretical content is concretised through goal-setting, visualisation, and thematisation, processes with the capacity to guide the design process. With these reformulations, a “reasoned leap” beyond existing comprehension of users’ behaviour is possible.

These conclusions became clear during a recent design task with mixed targets: the design solutions were to be user-oriented yet innovative. The first three authors of this paper were part of a design team designing new interaction practices and ways of operating in the maritime domain. These design suggestions reflect psychological theory on work and interaction with work tools. Thus, the design case demonstrates how design thinking (Cross, 2011) can draw from cognitive and social theory.

1 The Case Study: Future Ship Bridge Concept Design

The design case had its origins in an extensive research and development programme aiming at promoting user experience and usability oriented thinking in the metal and engineering industry by the means of collaboration between corporations and research institutes. The aim in the design case taken as a case study for this paper was to create innovative ship bridge concepts. Positive user experience (Hassenzahl & Tractinsky, 2006) was the general aim in the concept design. Three expectations for the bridge concepts were indicated by the corporate participant, Rolls-Royce: the concepts would have to evoke a “wow feeling” in potential users and other stakeholders; the concepts would need to reflect radical design (i.e., reflect future technologies and trends); and they would have to be “user-centred” (i.e., make sense for the current users and aid in their work tasks).

The research team that designed the future bridge concepts together with the maritime experts of Rolls-Royce had background in product design, human–computer interaction, human factors, psychology and social sciences. Below, we describe how the theories were applied; the full design process, however, showed greater variety and featured many phases not described here. The main message illustrated here is that understanding of users may be translated into design ideas in a two-way manner: via abstraction as empirical data are viewed through theoretical lenses and by concretisation with themes, visualisations, and design aims.

Firstly, for domain and work analysis, the researchers used a specific analysis method. Several mariners and other maritime experts were interviewed, and the bridge work was analysed through the core-task analysis approach (Norros, 2004; Norros, 2013), which provides a theoretical model of human–environment interconnection. We do not describe the data and the analysis process in detail here but do briefly demonstrate how the model can be applied in design. The CTA model assumes that challenging and safety-critical work activity always entails generic control demands related to 1) dynamism (i.e., temporal demands, such as a need to make quick decisions), 2) uncertainty (i.e., unexpectedness of events, which implies that decisions must be made with insufficient information), and 3) complexity (i.e., multiple, reciprocally connected influencing elements, such as weather, technology, and human behaviour). In addition, it takes three basic features of work activity to be the resources with which these control demands are addressed: 1) skill, 2) collaboration, and 3) knowledge. Work activity can be analysed through exploration of how these control demands and resources connect one with another, where the connections found are called core-task demands of the particular work domain. The findings represent both enacted (i.e., as expressed in the interviews or observed by the researchers) and potential (i.e., as inferred or suggested by the researchers or interviewees) ways in which the control demands are addressed. Thereby, an “analytical grid” of these interrelations is formed. The model can be visualised in both pictorial (Figure 1) and tabular (Table 1) format.

Insert Figure 1 about here

Insert Table 1 about here

While the CTA model as depicted in Table 1 involves nine analytical connection points, greater detail is provided here only for the “uncertainty” control demand (cells 7, 8, and 9 in Table 1), for brevity. In the table, the underscored, numbered elements of work activity are assumed by the model, and some basic findings are shown, with alphabetic labels. In all, four design implications are discussed here, one to do with the data as a whole and three related to the uncertainty aspect of maritime activity specifically.

Firstly, one general finding was that “togetherness” and unity are important for the mariners at embodied, cognitive, and social levels: they operate the vessel with the intuitive feel in their bodies of how the boat “interacts” with the environment as the boat rocks, they have a profound understanding of the features of both the environment and the vessel, and they feel a strong social unity among the crew. It was recognised that the idea of togetherness corresponds to the theoretical approach of joint cognitive systems (Woods & Hollnagel, 2006; Norros & Salo, 2009). The central idea of the joint cognitive systems theory is that human(s) and technology form a functional unit.

This implies that the human–technology system, rather than certain tasks or individual cognitions, is the unit of analysis. In other words, the joint human–technology activity is viewed “holistically” rather than “atomistically”; it is inferred from diverse interrelated elements, such as physical settings, learned meanings and notions, communication practices, and usability issues with the tools used. Overall, on the basis of the findings and the joint cognitive systems approach, togetherness and holism were considered to be general design approaches and aims. To concretise these ideas, an inspirational theme for the concept design process was chosen: “being one with the ship and the sea”. The concept design solutions, of which three are presented below, reflect this theme, holism, and the general design target of togetherness: instead of focusing on specifics in design, such as, forms or colours, we considered the basic activities through which mariners relate with the environment. While the project covered diverse vessel types, these concepts were designed for tugboats.

Secondly, the CTA model was applied in arrangement and visualisation of the more specific findings and the related design goals. The CTA figure and the tabular format were applied in the design process, because they had been concretely employed during the concept-making. Three design solutions can be pinpointed that addressed the uncertainty-related control demand (cells 7, 8, and 9 in Table 1).

One of the basic findings was that radio communication is essential, since the tug operator needs to communicate efficiently with the deckhands and the machine operator in extremely dynamic tugging operations (cell 7 in Table 1). This finding was translated into a design-related question: could there be a solution that might supersede radio communication? In other words, “enhanced communication” was considered as a design target. In examination of the various goals and findings, a design solution called Telepresent Crew was envisioned (see Figure 2). The idea is to enhance dialogue between the captain and the crew. It involves the crew wearing augmented-reality goggles equipped with cameras and displaying of the video feed originating from the goggles on a touchscreen used by the captain. Hence, the screen allows the captain to see what a crewmember is gazing at. Furthermore, when the captain touches the touchscreen, information is presented for a selected crew member.

Insert Figure 2 about here

Another basic finding was that escorting larger boats with a tugboat requires skill and readiness to anticipate the movements of the bigger boat constantly. Massive cargo boats turn especially slowly; therefore, the tugs have to assist in an anticipatory manner – that is, long before the cargo boat reaches an undesired location and course. This requires communication between the tug and the cargo boat; typically, the cargo ship's crew provide steering-relevant information, such as speed and turning rate, for the tug by radio communication. All of this points to a design goal of "enhanced anticipation of the escorted boat" (cell 8 in Table 1). This, in turn, was addressed with the concept idea of Intelligent Towing (see Figure 3). The concept involves a direct data link between the two boats; that is, the relevant information is provided automatically for the tugboat. Discussions are replaced with clear indications. Furthermore, the concept involves enhanced visualisation of tugging-relevant activity, particularly the strain on the tugboat rope.

Insert Figure 3 about here

Finally, the third design solution, Sea-Ice Analyser (see Figure 4), was created in a similar manner. A basic finding here was that in maritime activity, features of the environment require interpretation based on learned skills, hence the design aim of "enhanced interpretation of the environment" (cell 9 in Table 1). In icy sea conditions, it is sometimes difficult to know whether a ship is able to break the ice in front of it, especially amid darkness. The design solution's intent is to provide assistance in this estimation: the thickness and strength of the ice around the vessel are calculated, and the computer estimates whether it is possible for the ship to proceed, then presents this information on the command bridge.

Insert Figure 4 about here

The design solutions exemplify ways in which general theoretical ideas and theory-based modelling can be applied for the purpose of design. Abstract models and theories are re-formed as goals, visualisations, and themes. Findings from field studies may have a dual role here. Firstly, they may indicate a feasible theoretical option (as when our findings corresponded with the joint cognitive systems approach), yet they may also point to more concrete design goals (as visible in Table 1). Below, we explicate how this theory-based abstraction and reformulation of data can be the key for production of good and perhaps radical design solutions.

2 Implications: Social and Psychological Theories in Design Thinking

We now compare the design activities described in the case study with the concepts applied in design literature. Proceeding from interview-based studies, Cross (2011) has identified "three key aspects of design thinking" that can be applied in a model via which our case study can be considered and an argument formulated. First, Cross found that expert designers draw from "first principles" providing general guidance and structure to the design work. For example, in a case study he examines (pp. 31–51), reflection on laws of physics was the first principle for a sports car designer: ultimately, the physical realities would determine how fast the car could go. For another designer interviewed (pp. 53–66), who designed home appliances, functionality or ease of handling was the first principle structuring the design work. The designer consistently thought about what is practical from the user's perspective – i.e., considered the practical principles in domestic work activity.

The second key aspect of design thinking, according to Cross (2011), is establishing a certain frame or perspective that prestructures a particular design task. For example, the car designer specifically wanted the ride to be especially low, because this allows higher speeds, through the physical mechanism of the "ground effect". Regulations, however, specified a certain (6 cm) gap and banned devices that would reduce this gap "on the fly", during a race. The ground effect achieved through a low ride height was the frame that guided the design work. Eventually, the designer found a way to “circumvent” the regulations. An especially slow suspension system was invented, letting the car “sink” during the race and return to full height so slowly that, in fact, the race car adhered to regulations mainly when at a standstill – this was acceptable, however, since all cars in this formula bounce up and down (to a gap of under 6 cm) to some extent during race.

The example above exemplifies also the third key aspect of design thinking: taking a broad systems approach to the problem instead of accepting narrow criteria for it. Cross (2011) explains that expert innovators have a whole system of interacting elements in their mind. In the case of the formula car design, these included the physical forces affecting the car and the way in which regulations are interpreted. The regulations and physical realities were not two distinct spheres; only by considering both simultaneously could one achieve an innovative and radically faster car for the formula circuits.

The concepts of framing and first principles are not precise and both delimit the design intention. Frames, however, are more specific than first principles. In the example above, the general intention was to harness the laws of physics (first principle), while the more specific goal was to apply the ground effect (frame). Secondly, a first principle seems to refer to a beneficial and relevant approach that a designer applies across various design cases whereas a frame is typically specific to a particular design case. Overall, the Cross (2011) model suggests that expert design thinking entails structure but also flexibility to perceive the design problem in various alternative terms. Thirdly, consistently with the idea of a systems approach to design, Cross suggests that good designers are able to identify “surprising connections” by drawing from distant domains.

In the case of designing ship bridge concepts, taking user experience into account could be considered an underlying first principle giving structure to the design thinking: the users would have to perceive the concepts positively, and use of the final product would have to be pleasant. This implies that, similarly to how “physical principles” must always be borne in mind if one hopes to design fast cars, “psychological principles” should be kept in mind in designing for user experience – that is, creating something that is comfortable, pleasant, or evocative of positive emotions of any sort.

In itself, however, the idea of user experience is especially broad and can be given any number of interpretations. It is a buzzword of recent times that emphasises the importance of enjoyability in interacting with computers and other devices (Hassenzahl & Tractinsky, 2006). Positive experience is a complex phenomenon, which can be viewed through the lenses of various psychological and social theories, and on its own, it provides only a non-specific aim for the design work. In other words, specifying frames for human satisfaction are necessary for bringing more precise guidance to design. For the ship bridge design, relevant human scientific theories and models were applied to create frames. These ideas, core-task analysis and joint cognitive systems, can be considered applicable for understanding preconditions for positive user experience in certain work domains.

Our case study implies that theories translate into design ideas through certain distinct pathways or translation mechanisms. First, it is apparent that a theory-based frame, as any design frame, may provide a new design aim. These aims can be manifested at different levels; that is, they can be more general or specific. The frame derived from the joint cognitive systems approach implies the broad design aim of “increasing unity” among human, machine, and environment, while core-task analysis implies more specific aims in accordance with the model it provides, one example being the aim of reducing uncertainty by promoting collaboration (see Table 1). This mechanism can be called goal-setting.

Secondly, theories can provide visual models (see Figure 1). As the description of the design case explains, the concept ideas of Telepresent Crew, Intelligent Towing, and Sea-Ice Analyser correspond with the analytical grid provided by the core-task analysis model. This second mechanism can be called visualisation. Assumedly, visualisation aids in conceiving of the model as a whole; also, visually oriented comprehension has been found to be characteristic of design thinking (Findeli, 2001; Schön, 1992).

It seems that psychological or social scientific theories can be turned into inspirational themes before being transformation into design ideas; e.g., the “being one with the ship and the sea” theme reflected the idea of joint cognitive systems. Thereby, thematisation is the third mechanism identified through which social and cognitive sciences may take part in design. Thematisation communicates the ideas and value propositions drawn from the theories in an emotion-oriented manner; the themes can be applied in design and also in communication with potential users.

Fourthly, psychological and social scientific theories may provide conceptual reconfiguration – that is, change how the various concepts that can be applied in creative work are related to one another. The core-task analysis model implies connections between challenges and opportunities in the work. More generally, theories may change the scope in which the phenomenon subject to design is viewed, and they may entail a systemic approach, which is arguably beneficial in design. A systems perspective is integral to the joint cognitive systems approach: it implies a “whole” composed of people and technology together (Woods & Hollnagel, 2006), a holistic approach that can be seen in the Telepresent Crew, Intelligent Towing, and Sea-ice Analyser concept ideas.

The above-mentioned four distinct pathways through which cognitive and social scientific theories may provide design input – goal-setting, visualisation, thematisation, and conceptual reconfiguration – involve both abstraction and concretisation. Abstraction can be seen as design-relevant information is translated to the terms of the concepts and “conceptual structures” of the theories, as when field study findings were assembled as content for the core-task analysis model and its visual presentation. Concretisation, in turn, refers here to processes that make the theoretical approaches more tangible for design. Goal-setting, visualisation, and thematisation can be seen as involving both abstraction and concretisation: they both distance the designer from the design-relevant data and present the data in an accessible and simplified manner.

Table 2 shows the general ideas that stem from comparing the case study with Cross's (2011) concepts of essential elements in design. One can see from the table that the processes of goal-setting and conceptual reconfiguration are involved with the key elements of first principles, framing, and solution criteria. To include the processes of visualisation and thematisation, however, Table 2 also brings in a fourth element relevant to design, "design artefact creation". This refers to generation of "boundary objects" and "modalities" channelling the design activity. The concept of boundary object refers to objects that assist in sharing of understanding among people with different backgrounds (Gasson, 2006; Star & Griesemer, 1989). "Modality", in turn, refers here to a category of representation or other "channel category" through which design activity takes place. For instance, invention of the telephone required both visual representation of sound waves and auditory representation of sound. The inventor, observing an experiment translating speech into variation in a gas flame, realised that speech can be visualised and that this idea of speech having a shape might be applicable for construction of a device that relays sound by translating it into variation in electrical signals (Bruce, 1973). Visualisations and themes can be considered boundary objects and modalities applicable in design: they help to communicate the theories in such a manner that the theoretical ideas are more easily accessible in real-world design work; they synthesise multiple vantage points and render them applicable for dialogue.

Insert Table 2 about here

3 Discussion

Kelley and Littman (2005) appear to express a common way of thinking about social and cognitive sciences in design when suggesting that in good design teams, a cognitive or social scientist is a metaphorical anthropologist, with the main goal of making sense of the potential users and customers. The takeaway message here, however, is that these design anthropologists have opportunities to provide more to design than the typical vantage point for understanding users and customers. This is enabled through application of the theoretical knowledge inherent to the scientific education of these professionals. The design problem can be conceptualised and theoretically considered in a manner that facilitates creation of insightful connections relevant in design activity. There are, however, some specifications and expansions to our argument.

First, one may consider what kinds of theories from cognitive and social sciences might be beneficial for design. Our understanding is that design frames and first principles (Cross, 2011) should provide meaningful and useful design-related vantage points for design thinking. While there is no way of precisely determining what is meaningful or useful for any given design case, it can be assumed that theoretical approaches drawing on careful academic elaboration may provide meaningful knowledge content also for design. This is because these approaches explain and describe phenomena meaningfully and plausibly, if one assumes that academic discussion and research validates the theories' utility and plausibility in the long run. We do not believe, however, that all cognitive or social scientific approaches found valuable in academic discussion and research would provide design input as explained in this paper. A great deal of psychology and social sciences relies on statistically testable theories and hypotheses. In their generalisations, those studies seem of greater value for general design guidelines (e.g., when findings indicate that visual presentation should be preferred over textual in user interfaces [Saariluoma & Sajaniemi, 1989]) than as perspectives for generation of work-domain-specific design goals and themes. Furthermore, holistic and systemic theories can be considered especially beneficial in providing these domain-related design indications if one accepts Cross's (2011) argument according to which a systemic approach to design criteria is beneficial. In addition to our case study's application of holistic and systemic approaches, examples of their exploration of actual work settings include "distributed cognition" (Hutchins, 1995) and "cultural-historical theory of activity" (Engeström, 1987; Nardi, 1996; Tan & Melles, 2010).

Secondly, one may elaborate on what kinds of design cases might benefit from the theoretical content of social and cognitive sciences. In the case study, positive user experience and user-centredness were the general aims alongside innovativeness of solutions. It is worth noting, however, that there are other relevant first principles and design goals, in addition to the user-centred approach, that could benefit from psychological or social scientific knowledge. For example, Wright and McCarthy (2010) have discussed the "humanist tradition of human–computer interaction", which emphasises values such as, agency, democracy, and equality in design activity and design solutions. The humanist tradition hence implies broad aims for design activity: design solutions should support democracy and similar values as important. However, just as there is no simple answer to what constitutes human (or user) experience, there is no simple definition of democracy or equality here. These issues are discussed and explored by social scientists and philosophers. Accordingly, if "humanism" is considered the first principle for design, social scientists and philosophers might provide positive input for the design process by offering more specific design frames and aims. Applying their theoretical knowledge to the subject matter, they might elaborate on what kind of democracy or what interpretation of democracy the design solution should support. In other words, non-technical theoretical thinking has potential to expand design thinking related to the matters that the theories address. We do not, however, believe that every design case can benefit from social or psychological theories: they are applicable in those cases related to the phenomena that the theories actually address.

Thirdly, we could discuss further how theories of social and cognitive sciences may provide new kinds of innovations. Indeed, Norman and Verganti have argued that design activity based on understanding and exploration of existing user and customer needs may be insufficient to provide "business-changing" radical innovations; truly innovative design solutions induce change and surprise people by allowing them to "do what they did not do before" rather than giving them what they knew they wanted, the latter implying merely "doing better what is already done" (Norman &

Verganti, 2012, p. 5). The present paper already points to some means by which user studies could generate radical innovations. Firstly, the notion that psychology and social sciences may provide systemic approaches and conceptual reconfigurations implies the possibility of more radical design solutions. This argument stems from Cross's (2011) finding according to which a systemic approach characterises good design thinking: expert designers are able to make connections that draw from quite different domains, among them legislation, culture, technological advances, and physical realities. Holistic and systemic theories about humans provide lenses that allow perceiving connections between human activity and various other phenomena. Applying the core-task analysis model (Norros 2004), for example, allows perception of human–environment connections from various relevant vantage points (see Table 1). Secondly, theory-based modelling and conceptualising arguably allows designers who apply user studies to "distance themselves" from the user data in a manner beneficial for achieving new innovations. As our case study demonstrates, a higher level of abstraction can go hand-in-hand with reformulation via various means of concretisation. An "innovative leap" of sorts is possible when a transition between abstraction levels, modalities, and/or vantage points takes place: the data are translated into design ideas not directly but after these reformulations. The phenomenon of designers being "trapped within the current paradigms" (Norman & Verganti, 2012, p. 11) may be avoided through this distance to data. These suggestions echo studies by Nersessian (2002) according to which scientific discoveries are achieved by means of generalisation and visual modelling of ideas derived from existing theories of analogous source domains. For example, abstract models derived from fluid and machine mechanics were combined in generation of the theory of electromagnetic fields. Similarly, we propose synthesis of visual and schematic abstractions derived from social and cognitive theories for the purpose of innovation.

Finally, one may consider <u>what kinds of radical innovations</u> are possible with psychological and social theory. Norman (2010) has argued that influential life-changing innovations, such as aeroplanes, computers, and mobile phones, have nothing to do with studying people or society and just reflect technological experimentation. It is unlikely that social or psychological theories would lead to breakthrough innovations of this type. However, as Verganti (2009) argues, an additional way of inducing life-changing radical innovation is by changing the meaning associated with existing technologies. He cites the Nintendo Wii, which, instead of focusing on expert players, featured simplified controllers and provided activities typically not associated with video games – physical exercise and entertainment for the whole family. According to Verganti, designers aiming at radical innovations of this type must consider people's potential activities – that is, consider what kind of activities would be meaningful in a possible future wherein the meanings associated with a certain device (or service) have changed. It seems plausible that introducing social or psychological theories to the design process could yield innovations of this type. Providing greater abstraction and "forcing" the data into the categories and vantage points of the theories could be ways to identify potential activities of the future; theories can be translated into design goals in the manner seen in the case study. The Sea-Ice Analyser and Intelligent Towing concepts may represent incremental innovation in the sense that they augment existing activities. The Telepresent Crew concept, in contrast, is more radical, introducing a new kind of activity for the sailors.

The overall contribution of this paper is naming and illustration of some general mechanisms through which various kinds of psychological and social scientific theories may provide design input, specifically goal-setting, visualisation, thematisation, and conceptual reconfiguration. These, in turn, feature the even more general mechanisms of abstraction and concretisation. Admittedly, the notion that these mechanisms might allow radical design is hypothetical rather than proven: future studies are necessary for greater certainty. On the other hand, we believe that the general processes described here are widespread in actual design work as various design anthropologists intuitively apply their theoretical background knowledge. This implies that the process of design artefact creation (see Table 2) may be manifested in more varied ways than exemplified in our case study. Stories (Turner & Turner, 2003) and user “personas” (Cooper, 1999; Miaskiewicz & Kozar, 2011), for instance, have been applied to concretise studies’ findings and, presumably, also theoretical thinking. By naming these mechanisms, we hope to clarify and expand the understanding of how social and cognitive sciences may provide design input.

References

Bruce, R. V. (1973). *Bell: Alexander Graham Bell and the conquest of solitude.* Ithaca, New York: Cornell University Press.

Buchanan, R. (2001). Design research and the new learning. *Design Issues, 17*(4), 3–23.

Cooper, A. (1999). *The inmates are running the asylum.* Indianapolis: Morgan Kaufmann.

Cross, N. (2011). *Design thinking: Understanding how designers think and work.* Oxford: Berg.

Dogan, F., & Nersessian, N. J. (2010). Generic abstraction in design creativity: The case of the Staatsgalerie by James Stirling. *Design Studies, 31*(3), 207–236.

Dorst, C. H. (2011). The core of "design thinking" and its application. *Design Studies, 32*(6), 521–532.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research.* Helsinki: Orienta-Kosultit.

Findeli, A. (2001). Rethinking design education for the 21st century: Theoretical, methodological and ethical discussion. *Design Issues, 17*(1), 5–17.

Frascara, J. (2002). *Design and the social sciences: Making connections.* London: Taylor & Francis Group.

Gasson, S. (2006). A genealogical study of boundary-spanning IS design. *European Journal of Information Systems, 15*(1), 26–41.

Glaveanu, V. P. (2013). Rewriting the language of creativity: The five A's framework. *Review of General Psychology, 17*(1), 69–81.

Hassenzahl, M. (2010). Experience design: Technology for all the right reasons. *Synthesis Lectures on Human-Centered Informatics, 3*(1), 1–95.

Hassenzahl, M., & Tractinsky, N. (2006). User experience – a research agenda. *Behavior & Information Technology, 25*(2), 91–97.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, Massachusetts: MIT Press.

Karwowski, W., Soares, M., & Stanton, N. (2011). *Human factors and ergonomics in consumer product design: Methods and techniques.* Boca Raton, Florida: Taylor & Francis Group.

Keinonen, T. (2009). Immediate and remote design of complex environments. *Design Issues, 25*(2), 62–74.

Kelley. T., & Littman, J. (2005). *The ten faces of innovation.* New York: Doubleday.

Kuutti, K. (2005). Defining an object of design by the means of cultural-historical activity theory. In P. Chow & M. Jonas (Eds.), *Proceedings of the EAD06 conference* (CD-ROM). Bremen: University of Art Bremen.

Miaskiewicz, T., & Kozar, K. A. (2011). Personas and user-centered design: How can personas benefit product design processes? *Design Studies, 32*(5), 417–430.

Nardi, B. (1996). *Context and consciousness: Activity theory and human–computer interaction.* Cambridge, Massachusetts: MIT Press.

Nersessian, N. J. (2002). Abstraction via generic modeling in concept formation in science. *Mind & Society, 5*(3), 129–154.

Norman, D. A.. & Verganti, R. (2012). Incremental and radical innovation: Design research versus technology and meaning change. Paper submitted to *Design Issues*. Retrieved from http://www.jnd.org/dn.mss/incremental_and_radi.html

Norros, L. (2004). *Acting under uncertainty: The core-task analysis in ecological study of work.* Espoo, Finland: VTT Publications.

Norros, L. (2013). Developing human factors/ergonomics in a design discipline. *Applied Ergonomics, 45*(1), 61–71.

Norros, L., & Salo, L. (2009). Design of joint systems: A theoretical challenge for cognitive systems engineering. *Cognition, Technology & Work, 11*(1), 43–56.

Saariluoma, P. (2005). Explanatory frameworks for interaction design. In A., Pirhonen, H. Isomäki, C. Roast, & P. Saariluoma (Eds.), *Future interaction design* (pp. 67–83). London: Springer.

Saariluoma, P., & Sajaniemi, J. (1989). Visual information chunking in spreadsheet calculation. *International Journal of Man–Machine Studies, 30*(5), 475–488.

Schön, D. (1992). Designing as reflective conversation with the materials of a design situation. *Research in Engineering Design, 3*(3), 3–14.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, “translations” and boundary objects: Amateurs and professionals in Berkeley’s Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science, 19*(3), 387–420.

Sterwart, S. C. (2011). Interpreting Design Thinking. *Design Studies, 32*(6), 515–520.

Tan, S., & Melles, G. (2010). An activity theory focused case study of graphic designers’ tool-mediated activities during the conceptual design phase. *Design Studies, 31*(5), 461–478.

Turner S., & Turner, P. (2003). Telling tales: Understanding the role of narrative in the design of taxonomic software. *Design Studies, 24*(6), 537–554.

Verganti, R. (2009). *Design-driven innovation.* Boston: Harvard Business Press.

Vredenburg, K., Mao, J., Smith, P., & Carey, T. (2002). A survey of user-centered design in practice. In *Proceedings of the ACM CHI 2002 Conference on Human Factors in Computing Systems conference* (pp. 471–478). New York: ACM Press.

Woods, D., & Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering.* Boca Raton, Florida: Taylor & Francis Group.

Wright, P., & McCarthy, J. (2010). *Experience-centered design: Designers, users, and communities in dialogue.* London: Morgan & Claypool.

Table 1

The Elements of the Core-Task Analysis Model, with Basic Results, Design Aims, and Design Solutions in Selected Cells (7, 8, and 9)

	Resources		
Control Demands	Collaboration	Skill	Knowledge
Dynamism	1. Optimal sharing of efforts	2. Readiness to act	3. Anticipation and detection of weak signals
Complexity	4. Shared awareness and problem-solving	5. Focus on what is essential	6. Concepts' mastery
Uncertainty	7. Dialogue-based communication a. Experience of unity and good interaction with others increases work motivation b. Radio communication allows knowing the location and activities of fellow crewmembers Design aim: better communication Design solution: Telepresent Crew	8. Flexibility in action and reorientation a. Escorting ships with a tugboat requires anticipation of the movements of the escorted boat b. This anticipation requires skill and radio communication between the tug and the escorted boat Design aim: enhanced anticipation on the escorted boat Design solution: Intelligent Towing	9. Interpretative nature of activity a. Training and work experience allow operation-readiness b. Ice conditions are difficult to interpret, especially in the dark Design aim: better interpretation of the environment Design solution: Sea-Ice Analyser

Table 2

Potential Benefits of Social and Cognitive Sciences in Design Thinking as Derived from the Case Study

Elements in Design Thinking	Examples from the Case Study	Takeaway Message on the Benefits of Cognitive and Social Sciences
Relevant first principles – general underlying guidelines applied in various design tasks by the relevant designer(s)	The idea of "user experience" (that use of a product should provide positive emotional impact) was the underlying principle guiding the design processes	Psychological or social scientific theories can provide new design principles somewhat similarly to natural sciences: they provide broad design aims and concepts for conceiving of the design issue and the domain wherein the design would be applied; that is, they exert influence through the mechanisms of goal-setting and conceptual reconfiguration
Framing – a certain approach or relatively specific goal related to the design problem	More specific goals were derived from the theoretical models applied (i.e., core-task analysis and joint cognitive systems)	Psychological or social scientific theories can provide new frames through which the design issue is viewed: they provide specific design aims and concepts for conceiving of that issue and the domain wherein the design would be applied; that is, they exert influence through the mechanisms of goal-setting and conceptual reconfiguration
Approach to solution criteria – how constraints (such as physical realities, regulations, and customer desires) are considered; a systemic approach is considered beneficial	The criteria of user-centredness and innovativeness were addressed in view of the essential features of work contexts with the theoretical models applied; the models emphasise a systemic approach	The systemic approach benefiting design can be derived from psychological and social scientific theories; that is, they exert influence through the mechanism of conceptual reconfiguration
Design artefact creation – how various boundary objects and modalities are present in design thinking	The "joint cognitive system" idea was concretised with the theme "being one with the ship and the sea". The core-task analysis model was translated into design ideas through a pictorial presentation	Scientific approaches to humans can be translated into design ideas via the mechanisms of visualisation and thematisation

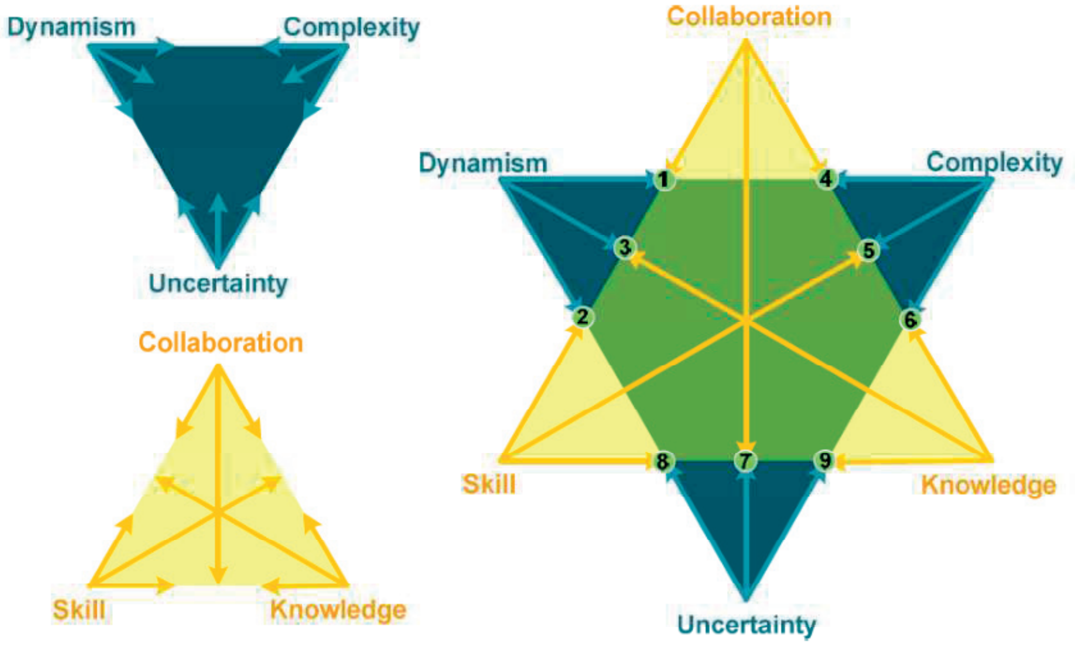


Figure 1. Pictorial presentation of the core-task analysis model – the control demands of dynamism, complexity, and uncertainty are met with the resources of collaboration, skill, and knowledge, as indicated by the arrows. The numbers correspond to Table 1.

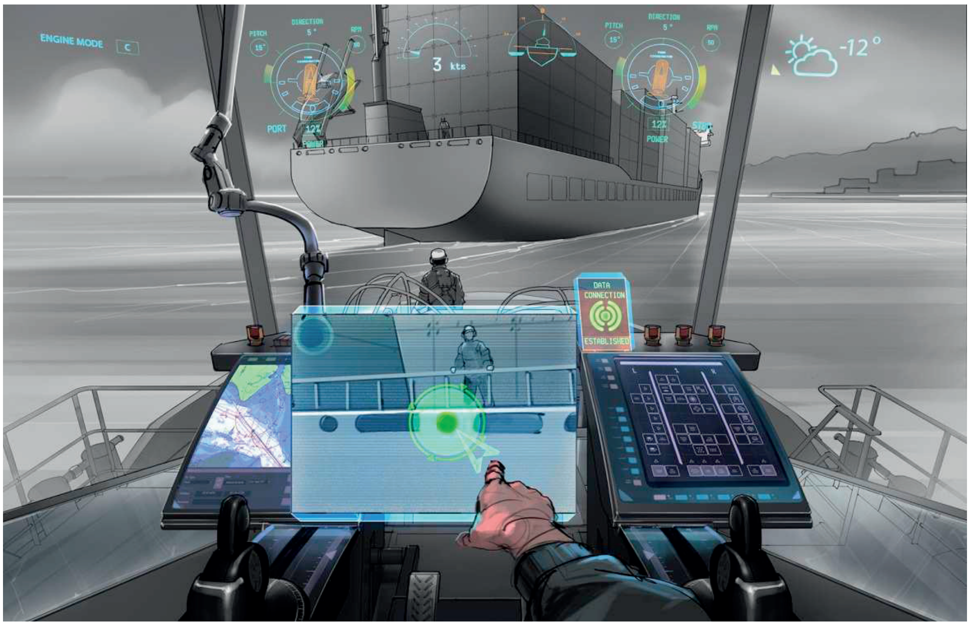


Figure 2. The Telepresent Crew concept presented from the tug operator's perspective – the touchscreen shows what the deckhand (standing behind the touchscreen) is looking at and allows the tug operator to draw visual indications for the deckhand. By drawing on the screen, the tug operator points out a hawsehole of the container ship to indicate to the deckhand on the deck of the tugboat where a connection rope should be attached. (© 2013 Rolls-Royce plc)

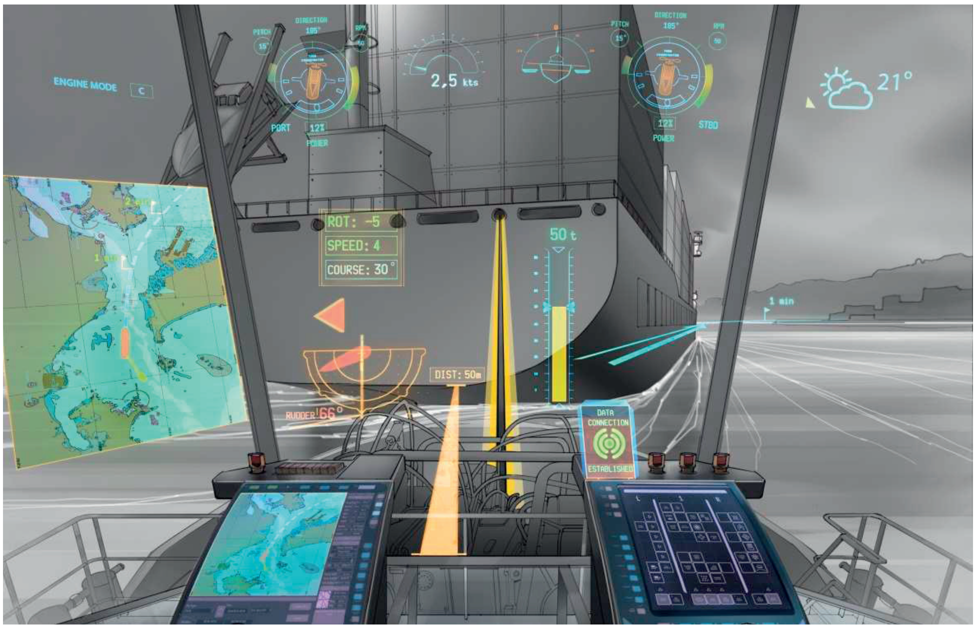
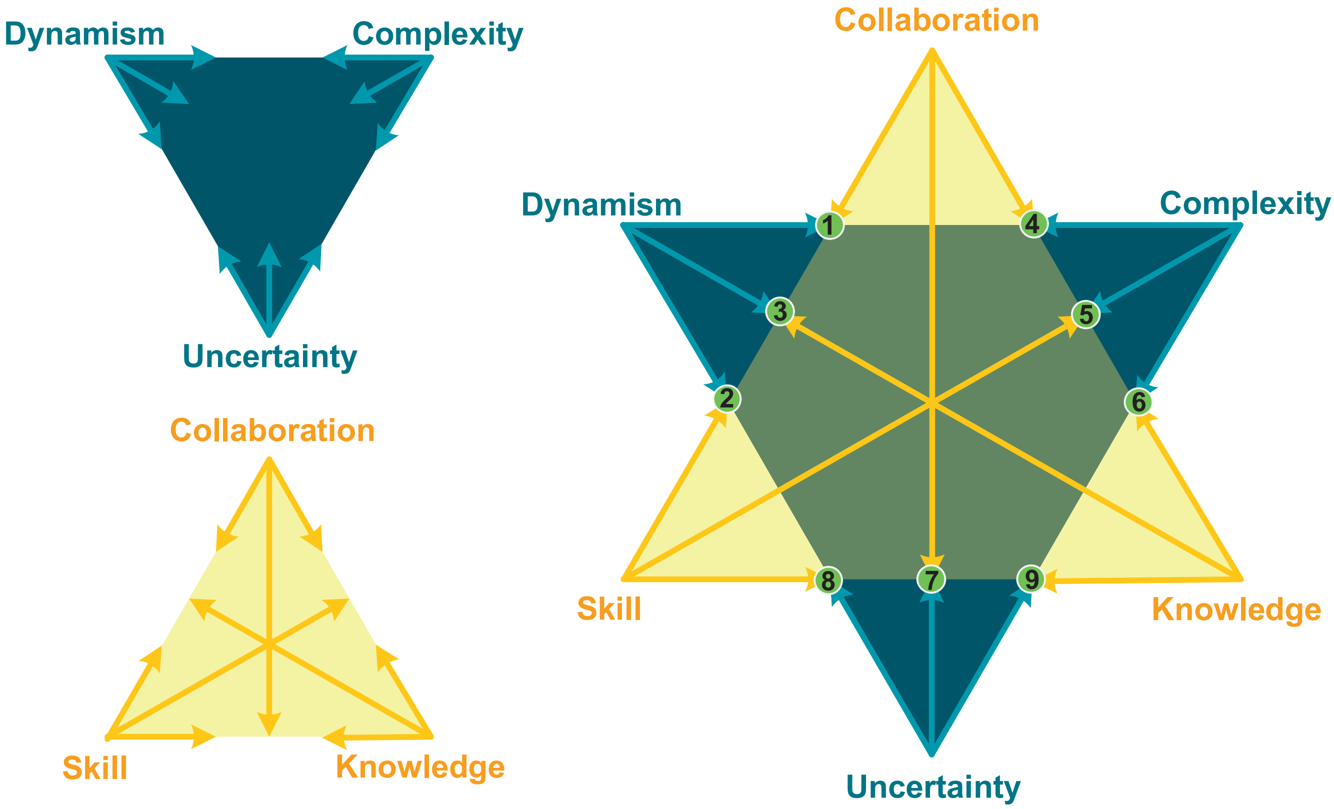


Figure 3. The Intelligent Towing concept – towing-relevant information is presented on the windows of the tugboat; this includes the speed, turning rate, and course of the assisted cargo boat, along with the distance between the boats and the strain directed to the tugging line. (© 2013 Rolls-Royce plc)

Figure 4. The Sea-Ice Analyser concept – the ship bridge presents information on ice thickness and strength and indicates whether or not it is safe to proceed. (© 2013 Rolls-Royce plc)

Figure 1



Dynamism
Complexity
Uncertainty
Collaboration
Skill
Knowledge
Collaboration
Dynamism
Complexity
1
2
3
4
5
6
7
8
9
Skill
Knowledge
Uncertainty

Figure 2

Figure 3

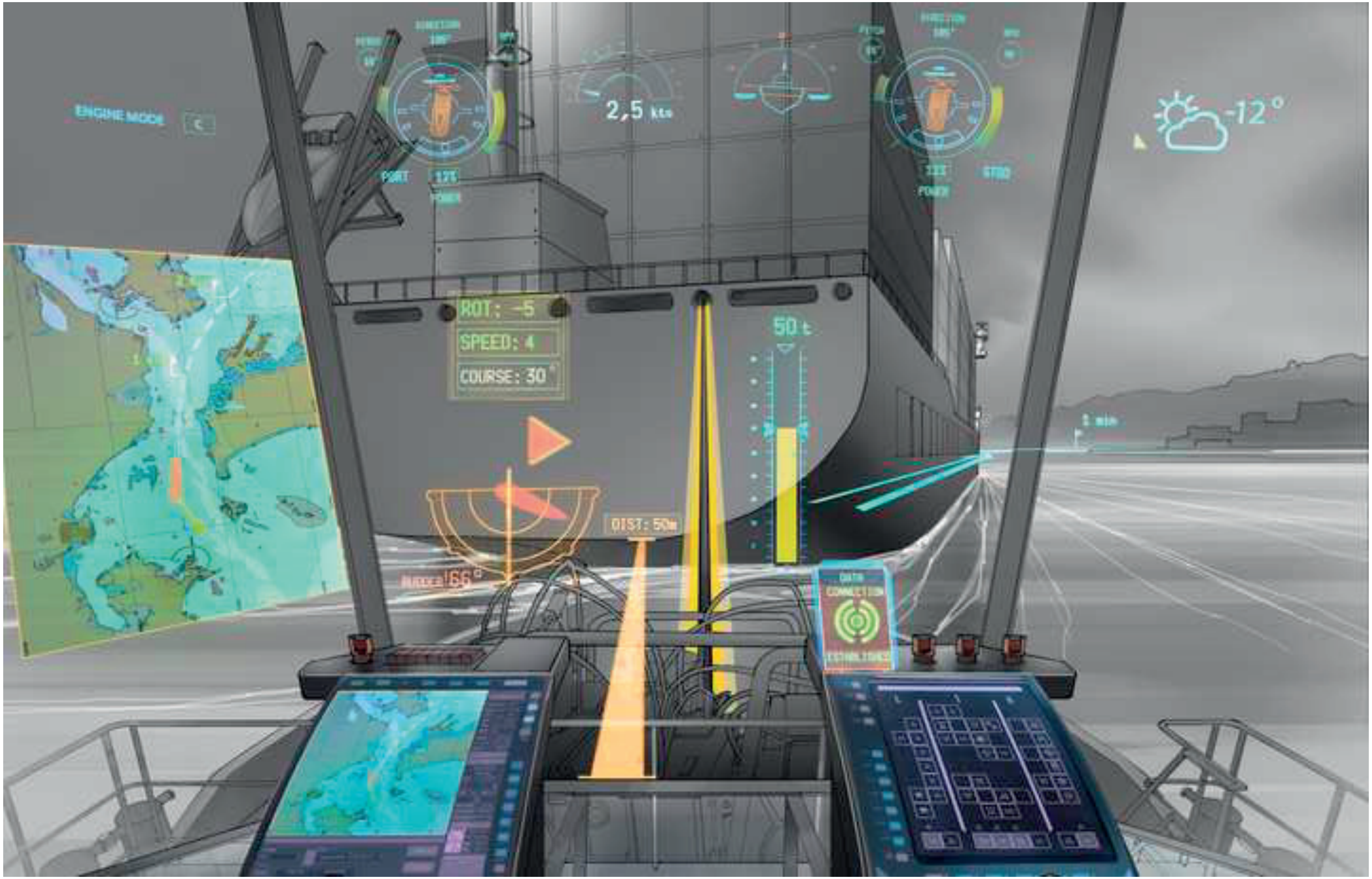

ENGINE MODE
2,5 kts
-12°
PORT
STBD
POWER
ROT: -5
SPEED: 4
COURSE: 30°
50 t
DIST: 50m
RUDDER 66°
DATA
CONNECTION
ESTABLISHED

Figure 4

Table 1

Table 1

The Elements of the Core-Task Analysis Model, with Basic Results, Design Aims, and Design Solutions in Selected Cells (7, 8, and 9)

	Resources		
Control Demands	Collaboration	Skill	Knowledge
Dynamism	1. Optimal sharing of efforts	2. Readiness to act	3. Anticipation and detection of weak signals
Complexity	4. Shared awareness and problem-solving	5. Focus on what is essential	6. Concepts' mastery
Uncertainty	7. Dialogue-based communication a. Experience of unity and good interaction with others increases work motivation b. Radio communication allows knowing the location and activities of fellow crewmembers Design aim: better communication Design solution: Telepresent Crew	8. Flexibility in action and reorientation a. Escorting ships with a tugboat requires anticipation of the movements of the escorted boat b. This anticipation requires skill and radio communication between the tug and the escorted boat Design aim: enhanced anticipation on the escorted boat Design solution: Intelligent Towing	9. Interpretative nature of activity a. Training and work experience allow operation-readiness b. Ice conditions are difficult to interpret, especially in the dark Design aim: better interpretation of the environment Design solution: Sea-Ice Analyser

Table 2

Table 2

Potential Benefits of Social and Cognitive Sciences in Design Thinking as Derived from the Case Study

Elements in Design Thinking	Examples from the Case Study	Takeaway Message on the Benefits of Cognitive and Social Sciences
Relevant first principles – general underlying guidelines applied in various design tasks by the relevant designer(s)	The idea of "user experience" (that use of a product should provide positive emotional impact) was the underlying principle guiding the design processes	Psychological or social scientific theories can provide new design principles somewhat similarly to natural sciences: they provide broad design aims and concepts for conceiving of the design issue and the domain wherein the design would be applied; that is, they exert influence through the mechanisms of goal-setting and conceptual reconfiguration
Framing – a certain approach or relatively specific goal related to the design problem	More specific goals were derived from the theoretical models applied (i.e., core-task analysis and joint cognitive systems)	Psychological or social scientific theories can provide new frames through which the design issue is viewed: they provide specific design aims and concepts for conceiving of that issue and the domain wherein the design would be applied; that is, they exert influence through the mechanisms of goal-setting and conceptual reconfiguration
Approach to solution criteria – how constraints (such as physical realities, regulations, and customer desires) are considered; a systemic approach is considered beneficial	The criteria of user-centredness and innovativeness were addressed in view of the essential features of work contexts with the theoretical models applied; the models emphasise a systemic approach	The systemic approach benefiting design can be derived from psychological and social scientific theories; that is, they exert influence through the mechanism of conceptual reconfiguration
Design artefact creation – how various boundary objects and modalities are present in design thinking	The "joint cognitive system" idea was concretised with the theme "being one with the ship and the sea". The core-task analysis model was translated into design ideas through a pictorial presentation	Scientific approaches to humans can be translated into design ideas via the mechanisms of visualisation and thematisation

VI

DESIGNING USER-ORIENTED FUTURE SHIP BRIDGES – AN APPROACH FOR RADICAL CONCEPT DESIGN

by

Mikael Wahlström, Hannu Karvonen, Eija Kaasinen, & Petri Mannonen

In Ergonomics in design: Methods and techniques (pp. 217–231)

Designing User-Oriented Future Ship Bridges – An Approach for Radical Concept Design

Mikael Wahlström [a], Hannu Karvonen [a], Eija Kaasinen [a], Petri Mannonen [b]

[a] VTT Technical Research Centre of Finland Ltd
P.O. Box 1000; FI-02044 VTT; Espoo, Finland

[b] Aalto University
P.O. Box 11000; FI-00076 AALTO; Espoo, Finland

ABSTRACT

The paper proposes a design approach intended for generation of radical design concepts for professional activity. This approach encompasses analysis of domain-specific work activity, user experience goal-setting, and work-domain and technology-trend foresight. The intention with the approach is that the concept solutions reflect the existing means and professional resources applied in work activity, the potential benefits of future technology, and the existing and future challenges specific to a certain work domain. This approach is explored through the lens of a ship command-bridge design case in which future bridge concepts were generated for three ship types: tugboats, platform supply vessels, and cargo ships. Several interviews with subject-matter experts and observation rounds were conducted for accumulation of the understanding required in generation of the concept ideas. Drawing from existing user-centered design approaches, our approach combines elements of experience design, co-design, and contextual design, but certain features are added for the purpose of creating radical instead of incremental design solutions, the latter typically being the result of user-centered design. These features include futurology; generation of broad systems usability design goals via the Core-Task Analysis method; reformulation of user experience goals into themes, stories, or personas; and co-design in the final part of the concept design process.

Keywords: design methods, user-centered design, radical design, ship bridges

INTRODUCTION

In design for industrial work, it is typically vital that the design ideas generated support the existing aims of the industrial workers. This is because the aim of the activity in designing for work activities, as opposed to consumers, is usually fixed. Whereas consumers may engage in activities that are wholly new to them, such as playing new games or embarking on a new dietary regimen, industry workers respond to certain basic needs and goals, among them habitation, energy, security, health care, production, and logistics. When considering the shipping industry, for example, one may assume that operations will continue to serve their current main mission – the circulation of goods via the oceans of the world. All this implies the usefulness of studying the existing work activities when one is designing for industry professionals: the design solutions should correspond with the existing aims of the activity. In other words, user-centered design (Gould & Lewis, 1985), in which the needs, desires, and capabilities of the users are taken into account, appears clearly justified.

It is noteworthy, however, that user-centeredness in design has been criticized for not providing new types of design solutions with the potential to surprise the people involved and offer them new possibilities (Norman & Verganti, 2014). Indeed, it typically offers users what they already knew they wanted, since the design solutions respond to the issues identified by the users (Keinonen, 2009). That is, because the solutions correspond too closely with the existing models of activity or "user paradigms," studying users' activities and needs does not seem to provide new kinds of solutions or radical, revolutionary innovations. Assumedly, when work-related innovations are under development, the users are not aware of all the forthcoming technical opportunities or trends. In addition, they may be too wedded to current practices (perhaps even heavily invested in them) to ideate radical changes. It is thought, therefore, that user studies predominantly provide incremental or evolutionary design solutions – that is, mere modifications to existing designs. In contrast, new kinds of design solutions should offer greater potential to provide business advantage by significantly enhancing or modifying the existing work activities. Accordingly, one may identify a topical problem to be solved: how to generate design solutions that both 1) support the existing activities of professional workers and 2) surprise the users with innovativeness by offering new possibilities. One can conclude that applying a design approach that solves this problem creates potential for positive renewal of industries.

The design approach presented here is aimed at addressing that problem. It is based on the following procedures:

1) Reformulation of user-study findings in such a manner that sufficient "distance" is achieved for the study's findings, by not "directly dictating" the design solutions, thereby allowing for radical design

2) Foresight of technology trends and future developments, enabling future-oriented design solutions

3) Co-design and co-evaluation with actual users and experts in the relevant field of application after creation of the initial design ideas, thereby allowing higher quality and better specification of design ideas

In the following discussion, we illustrate each of the procedures employed through examination of a ship-bridge concept design case. The bridge is the place from which the ship is commanded, navigated, and maneuvered. The ship's surroundings are observed from the bridge, with the watch-keeper looking out the bridge windows to note any potential dangers, such as rocks or other vessels. In addition, the vessel's steering and communication devices are found on the bridge. Modern ships' bridge equipment includes, for example, electronic chart displays (in ECDIS format, used for navigation), radar displays, and dynamic positioning systems.

In the design assignment reported upon here, the aim stated by the industry partner with whom the concepts were generated, Rolls-Royce, was for the solutions to stimulate the field of maritime operations in general by providing future-oriented ship-bridge alternatives. An additional aim was for the concepts to be user-oriented in the sense that, while providing highly novel types of solutions, they would still have to be accepted and appreciated by mariners who possess practical knowledge of hands-on maritime activity. It was, however, agreed that the project need not take maritime legislation into consideration: the objective was to supply alternatives – options for possible futures – instead of strictly accommodating existing realities. The goal was that the design concepts would represent the ship bridges of the year 2025. Three ship types were considered, these being tugboats, platform supply vessels (PSVs), and cargo ships. The results of the design project are described online (Rolls-Royce, 2014; 2015).

Overall, the design approach resembles those termed "contextual design" (Beyer & Holtzblatt, 1998), "experience design" (Hassenzahl, 2010), and "co-design" (Sanders & Stappers, 2008): it is based on 1) studying of the work context, 2) user-experience-related goal-setting, and 3) collaborative design involving the users, respectively. The deviation from these approaches is found in the purposeful aim of future-orientation and radical design. In addition, we employed a specific analysis method called Core-Task Analysis (Norros, 2004; 2013) to make sense of the relevant professional activity. We contrast our design approach against others' in the "Discussion" section, below.

OUR APPROACH FOR RADICAL YET USER-ORIENTED DESIGN SOLUTIONS

Our design approach combines thorough understanding of the users and their situations, visions of the future from several fields, and co-creation to produce radical design solutions. The process included a group interview, 12 one-on-one interviews and four field observations. Interviewed were maritime experts from diverse fields, with varied professions, among them designers, researchers, shipping company directors, trainers, officers, seamen, and sea captains. The interviews were conducted both before and after expression of the initial concept ideas. The field studies, which took place in Finland, Norway and Estonia, included observation of actual tugboat operation and observation of PSV operations in a simulator context. Additionally, a future-studies workshop was held, wherein researchers from various disciplines came together to discuss emergent trends related to maritime transportation, maritime technology, interaction technologies, and general societal developments.

Reformulation of user-study findings

As has been discussed elsewhere (Wahlström et al., under review), creating a "reasoned departure" from user-study findings can serve as a useful means of avoiding the phenomenon of designers being "trapped within the current paradigms" (Norman & Verganti, 2014). In other words, we propose that in the preliminary phases of design, the design indications drawn from the users' explications should be meaningful but also purposefully broad. This broadness allows room for the users' ideas not to dictate the creative process of design too directly and specifically. This is important for its facilitation of creation of ideas that are new to the users. In practice, the idea is for user activity to be modeled and understood rather than for users' ideas to be directly applied as design indications.

Another mean of creating radical design ideas in work proceeding from study of users is to focus on the user experience rather than on product features. This approach draws from experience design (Hassenzahl, 2010), whose output serves the purpose of users' ideas not directly translating into product ideas. The key prerequisite for experience-driven design is determination of what experience to design for. User experience (UX) goals concretize the intended experience (Kaasinen et al., in press). An experience goal describes the momentary emotion that is intended to be experienced during use of the product or service or, alternatively, a person's emotional relationship to the designed product or service (Lu & Roto, 2014). In the design of industrial systems, there are several stakeholders involved, and they should hold shared design goals. The UX goal-setting serves the objective of bringing together the divergent viewpoints of the stakeholders and gaining their commitment to said goal-setting as a strategic design decision (Kaasinen et al., in press). User experience goals reflect intended user feelings; in so doing, they do not inhibit radical design by directly dictating the design idea creation.

Product use can be conceptually divided into the inherent instrumental and non-instrumental qualities (Mahlke, 2005). The former involve utilitarian elements such as usefulness and ease of use, and non-instrumental elements have to do with the emotional aspects of product use. Similarly, in design, the aim could be a certain feeling, such as a sense of comfort, or it might be a certain practical task, such as efficient communication between individuals. We assume Core-Task Analysis, or CTA (Norros, 2004; 2013), to be a useful method for pinpointing instrumental task- and activity-related design goals for certain work domain and at systemic level. This assumption appears justified because in studies of risk-intensive work it has been used to identify interconnected elements influencing the way in which the aims of a specific work activity have been reached, where these identified issues include elements related to the work activities, the tools used, and the work environment in general. In other words, the CTA method is useful for identifying pertinent systems usability issues (Savioja, 2014). As has been discussed by Norros, Savioja and Koskinen (in preparation) analyzing systems usability is beneficial in varied phases of design, including ideation and concept design. Under the CTA model, challenging and safety-critical work activity entails generic control demands related to 1) dynamism (i.e., temporal demands, such as the need to make quick decisions), 2) complexity (i.e., multiple, reciprocally connected influencing elements, such as weather, technology, and human behavior), and 3) uncertainty (i.e., unexpectedness of events, which implies that decisions must be made in the absence of sufficient information). In addition, the CTA model assumes three basic features of work activity to be the means (i.e., resources) through which these control demands are managed: 1) skill, 2) knowledge, and 3) collaboration. Work activity can be analyzed through examination of how these control demands and resources connect one with another; the connections found are referred to as core-task demands of the relevant work domain. The core-task demand findings represent both enacted ways in which the control demands are addressed (i.e., as expressed in the interviews or observed by the researchers) and potential ways (i.e., as inferred or suggested by the researchers or interviewees). In this way, an "analytical grid" is formed of these interrelations (see Figure 1); the interrelations can be used as indications of the instrumental user experience and systems usability goals. Indeed, this was done in the ship-bridge design case. The model can be visualized in both graphical (see Figure 1) and tabular (see Table 1) form.

Figure 1 and Table 1 present the core-task demands identified when we studied PSV operations for the purpose of concept design ideation.

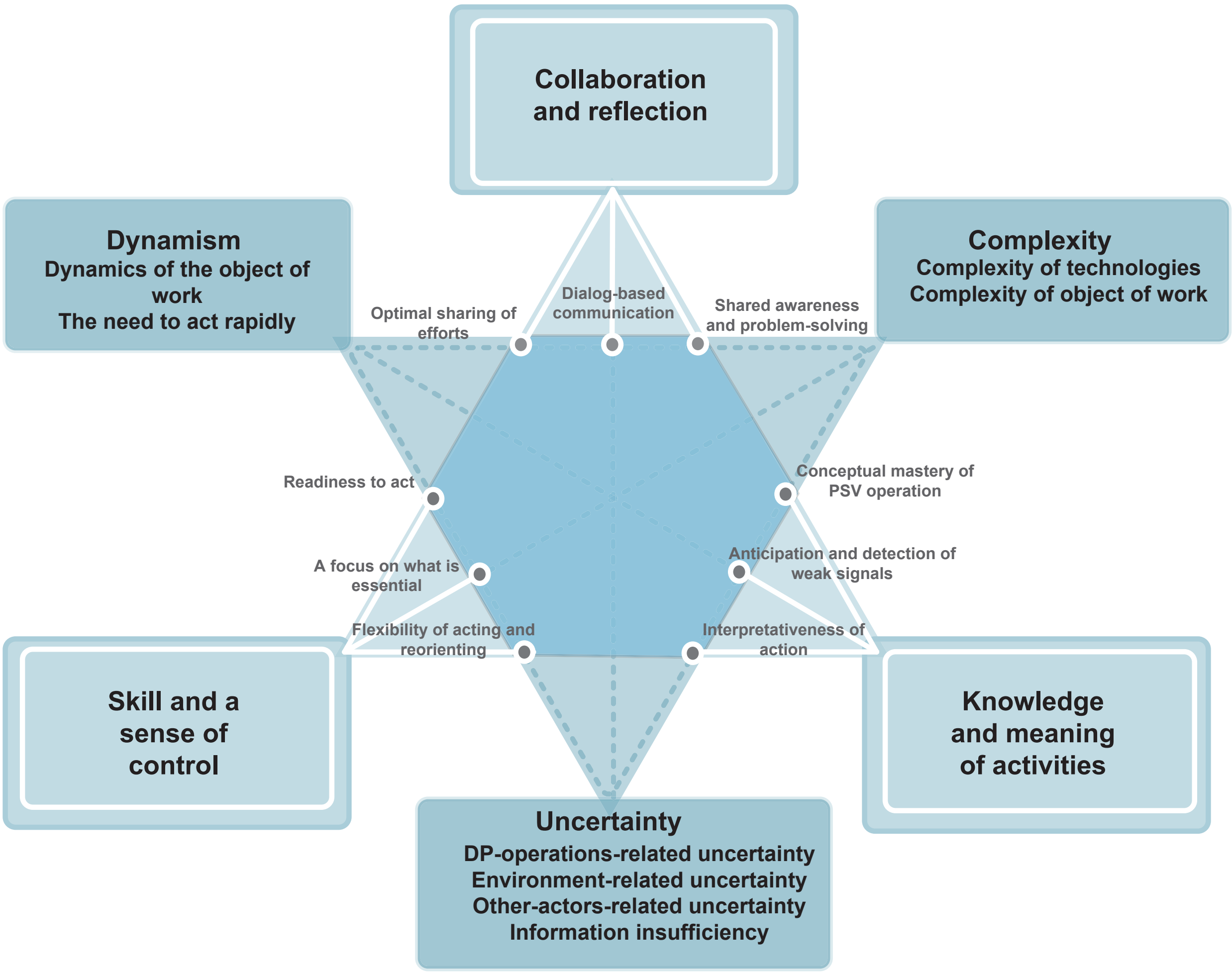


Figure 1. The core-task demands of PSV operation. The model includes control demands, means of managing them, and the associated core-task demands; see Table 1 for more detailed descriptions of the core-task demands in PSV operation.

Table 1: Description of the core-task demands of PSV operation, where control demands run down the leftmost column (in **boldface**) and means of managing the control demands are along the top (also **bolded**), with the core-task demands underscored in the description cells.

Resources ⇨ Demands ⇩	**Collaboration**	**Skill**	**Knowledge**
Dynamism	Optimal sharing of efforts - Sharing responsibility with others - Trusting in one's own competence in resolving problematic operation situations - Switch-offs in who is in change of DP operations, to counter boredom/fatigue - Good relationships among the people on-board and trust in their expertise	Readiness to act - Development of new skills and situational sensitivity through practical experience from both simulator and real-world conditions - Taking situational attributes into account - Confidence and calmness in PSV operation at all times - Good depth perception and spatial awareness	Anticipation and detection of weak signals - Understanding of chains of events (e.g., those related to GPS satellites lining up non-optimally) - Checklists that prepare both the crew and the vessel for the upcoming DP operations
Complexity	Shared awareness and problem-solving - Radio communication (e.g., with the deck crew and the rig crane operator) in order to maintain situation awareness - Clear communication to confirm safety of operation - Combination of competencies in demanding situations - Appreciation for others' work contribution - Sharing of experiences with colleagues	A focus on what is essential - Rapid comprehension of the essential information provided by the displays - Ability to "shut down" surrounding distractions on the bridge, to enable focusing solely on the task at hand - Utilization of tacit knowledge - Utilization of professional experience from diverse situations	Conceptual mastery of PSV operations - Knowledge of the technical solutions implemented and of their limitations - Understanding of the environmental demands and the physical forces relevant in the operations - Knowledge of the surrounding operation environment
Uncertainty	Dialogue-based communication - Professional appreciation for others and oneself - Experience of unity and good interaction as a motivational factor - Radio communication that enables following the input of all parties involved in DP operations	Flexibility of action and reorientation - Training that creates preparedness to act in various situations - Maintaining a standard bridge environment, to enable effective application of routines and skills - Confirmations and checks to ensure safe operation - Ability to adapt to the situation and perseverance for completion of the tasks at hand	Interpretiveness of action - Readiness to draw knowledge from training - Sufficient training and operation experience in DP operations and exceptional situations - Management of routines (e.g., attending to safety checklists)

We propose that the systems usability issues, which influence user experience on a functional level and can be distinguished via the CTA method, can be directly applied as useful concept design indications. In addition, we believe that feeling-related (non-instrumental) user experience goals may benefit from another kind of reformulation of data. Reflecting on design that applies user experience personas – that is, fictional characters representing certain target demographics (Cooper, 1999) – or stories (Carroll, 2000), one may assume that the goal of designing for a certain emotion is not, in itself, sufficiently inspirational for designers. The stories and user experience personas aid the designers in grasping the abstract emotion-related ideas by giving these a human face (Pruitt & Adlin, 2006).

Somewhat similarly, feeling-related findings were reformulated into an inspirational theme in our ship-bridge design case. A general finding was that feelings of togetherness and unity are important for mariners at embodied, cognitive, and social levels. At the embodied level, the mariners operate the vessel with an intuitive feel in their bodies of how it interacts with the environment as the vessel rocks on the waves and of how it reacts when it is being maneuvered in various conditions. At the cognitive level, the mariners have profound understanding of the features of the environment and the vessel. Finally, at the social level, the crew members feel strong social unity; they are together on these vessels 24 hours a day. Because "a feeling of togetherness" is a rather abstract idea to apply in design, an inspirational theme reflecting this non-instrumental user experience goal was generated in our case study. The theme, referred to as "being one with the ship and the sea," served as a reference to how the mariners' should feel with the aid of the design solutions. In practice, the theme reflects both the instrumental CTA-derived findings and the intuitive interpretations of the non-instrumental emotion related elements in the nautical work; in the practical design work, the instrumental and non-instrumental elements are inferred simultaneously and cannot be distinguished one from another.

The Augmented Crane Operations Concept (see Figure 2) exemplifies this design theme and use of the CTA model in design: it reflects an enhanced sense of unity between the ship operator and the environment, and it addresses some of the systems usability issues identified. The aim with the concept was to generate a solution that could support the creation of collaborative understanding during container-lifting operations involving rigs. The rigs have a container crane that is used when a PSV is positioned in the correct place. Currently, the collaboration between the crane and PSV operator is conducted mostly via radio communication. The concept idea was for the PSV operator to see exactly the same view (the bottom-left display in Figure 2) that the rig crane operator sees from the crane cabin and vice versa. Furthermore, the PSV operator should see where the container is supposed to land on the aft deck – this presentation would take place via augmented reality (AR) lines on a heads-up display (HUD) (see the middle portion of Figure 2). In the concept, the rig crane operator has the same view available in the cabin. Assumedly, through all these supportive systems, the feeling of continuity between self and environment could be increased: the operators would be more aware of one another's view of the situation. This concept design solution draws from the "shared awareness and problem-solving" and "dialogue-based communication" core-task demands (see Table 1), identified in the core-task analysis of PSV operators' work.

Figure 2. The Augmented Crane Operations Concept (© 2013 Rolls-Royce plc).

Future studies

In parallel with user studies, we carried out studies of technological, maritime, and general societal trends. Visions of new user interaction tools were of special interest, since it was apparent that these might change work processes by providing new possibilities. Trends in the maritime industry shed light on what kinds of vessels the future might hold, where they may operate, and what purposes they could serve. All these issues affect the work on ship bridges. Our aim was to create a shared understanding within the project group of relevant trends. Project participants studied relevant future studies from their own perspective. The results were used in combination with the user-study findings to inspire ideation in a co-creation workshop.

Commissioned from a consulting company, a study of emerging interaction technologies and techniques was performed. The user interaction technologies explored included personal projection, large display areas, deformable devices, wearable devices, gestures, tangible user interfaces, hover sensing, tactile feedback, touch input, brain–computer interaction, augmented reality, HUDs, motion simulators, and gaze tracking, along with speech, ambient sound, biosignals, and implanted user interfaces. For each technology, pros and cons were assessed, as was technology readiness.

Maritime transportation and technology trends were analyzed by a researcher focusing specifically on this area and by the company participant (Rolls-Royce). The most important trends identified were these: 1) globalization of markets (greater competition), 2) environmental concerns (entailing a need to reduce emissions and risks), 3) exhaustion of natural resources (bringing about a quest for new sources and a need to reduce consumption), and 4) navigation in Arctic conditions (since new shipping lanes in the Arctic region are gradually opening).

The general societal trends found had to do with future users in particular. Trends were identified on the basis of forecasts by Frost & Sullivan (2010), Gartner Research (2009), JWT Intelligence (2012), and Frog (2012). These general trends encompassed, for example, silence and minimalism as counter-forces to information, media, and technology overflow; ubiquitous and embedded computing; and the values of today's young adults.

After all the trend analysis work had been carried out and the results were shared within the group, we organized a workshop in which both researchers and company representatives participated. The participants were assigned a task for completion before the workshop: they were asked to choose their "favorite" trends (1–2 user trends, 1–2 collaboration trends, 1–2 interaction trends, and 1–2 other trends) and prepare to indicate why these are important. At the workshop, the technology and general societal trends were presented each on one sheet of paper on the wall. The workshop participants were asked to mark their favorite trends, after which each participant could distribute five votes among the trends so marked. Even if the ranking of the trends was of interest, a more important result was characterization of the process that engaged the participants in considering the trends and discussing them.

After the voting, maritime trends, trends in user interaction tools, and the findings from the field studies were presented and discussed. The workshop continued with the participants divided into two groups, each of which was asked to identify themes (challenges/possibilities/solutions/concepts) to process further. The groups were also to choose related trends and discuss how these could affect the theme. Though their task was the same, the two groups chose slightly different approaches. The first group proposed radically new, challenging "what-if" concepts and solutions rooted in domain trends and technical enablers. In contrast, the second group identified new ways of carrying out today's tasks, resolving the core-task demands identified, and defining user experience targets. By combining the results produced by the two groups, we identified the following ideas for further work:

- What if there were no joysticks? Could the operations be carried out with other kinds of tangible objects?
- What if there were no separate screens? Instead of screens, the necessary information could be presented on a head-up display integrated into the bridge's windows or on AR glasses worn by the operator.
- Visibility could be enhanced via placement of navigation/steering workstations in a "glass bubble" at the front of the bridge. The vertical view downward could be enhanced with a glass floor.
- Data overflow is an issue on the bridge. How could this be reduced?
- The bridge environment could be a less "machine-like" and "working environment" type of place, that is, it could be more cozy, homelike, and personalized.

Several design concepts reflect the work in the trend workshop. One example is the Sea-ice Analyzer Concept (see Figure 3). The capability of operating in Arctic conditions was one of the main maritime domain trends identified; therefore, we aimed to enhance the associated operations through novel technologies. Indeed, in icy conditions, it can be difficult to know whether a ship is able to break the ice in front of it, especially in darkness or fog. The intent with the Sea-ice Analyzer Concept is to assist in this estimation: the thickness and strength of the ice around the vessel are calculated, and a computer estimates whether it is possible for the ship to proceed and shows the best route. This information is presented on a large HUD in the front window. The augmented reality HUD data are organized such that the display shows the data needed for decision-making overlaid on the actual view of the outside.

Figure 3. The Sea-ice Analyzer Concept (© 2013 Rolls-Royce plc).

Co-design and co-evaluation

By intuition, one might think that design in collaboration with users is not beneficial from the standpoint of generating radical design solutions. This is because, by definition, radical design ideas are those that provide users with wholly new kinds of activities – one cannot assume that these new kinds of possibilities can be easily imagined by users submerged in the existing modes of work in day-to-day operations (Norman & Verganti, 2014). During the creation of future ship-bridge design solutions, however, it was found that applying co-design and expert users' evaluations in the final part of the concept design process aids in promoting radical design. This is because, firstly, the awareness that the preliminary design solutions is going to be evaluated by expert end users frees the designers to imagine even potentially "bad" design solutions; it does not matter if some of the design solutions do not yield potential, since the experts will cull these non-functional or uninteresting solutions from the pool of design ideas. By affording diminished self-censorship, this knowledge promotes generation of more ideas more rapidly – and, eventually, since many ideas will arise, there will also be solutions appreciated by the actual users. The future ship-bridge design case produced several design ideas that were rejected by the end users and therefore not refined further.

The second way in which expert users provide beneficial design input is via the refinement of design ideas. A good example of this is the Intelligent Towing Concept. The initial idea was that a head-up display would present the unit of tug plus towed ship (as visible in the map box at the left in Figure 4). We had discovered that when a tug pushes or tows another vessel, where that other vessel starts to rotate and head is not always self-evident. To assist in estimating this, the bridge could indicate the forces influencing the pushed or towed ship. Upon discussion with the real-world tugboat operators, however, the concept idea was developed further. It was explicated that actually the tugboats and the escorted cargo ships often share towing-relevant information via radiophone. The tugboat operator requests relevant information from the cargo ship’s crew, such as rate of turn (ROT), speed, and course. An immediate design implication of these accounts was that the Intelligent Towing Concept should also incorporate direct presentation of these verbal exchanges. In other words, a direct data link between the vessels would provide the tugboat with indications of the ROT, speed, rudder direction, and course of the cargo ship (see the yellow box and the yellow semicircle at the center of Figure 4).

In the future ship-bridge design case, the initial concept solutions were presented to the expert users by means of pictures and user scenarios. The creation of scenarios was itself an iterative and collaborative process. We first imagined certain kinds of scenarios and then discussed these with certain users. If the scenarios seemed plausible, they were applied when the concepts were discussed with a larger sample of users. For instance, the Intelligent Towing Concept was explained to the users via a scenario in which the rudder of the towed vessel becomes jammed, thereby steering it in an undesired direction; in such a situation, communication between the two vessels is essential.

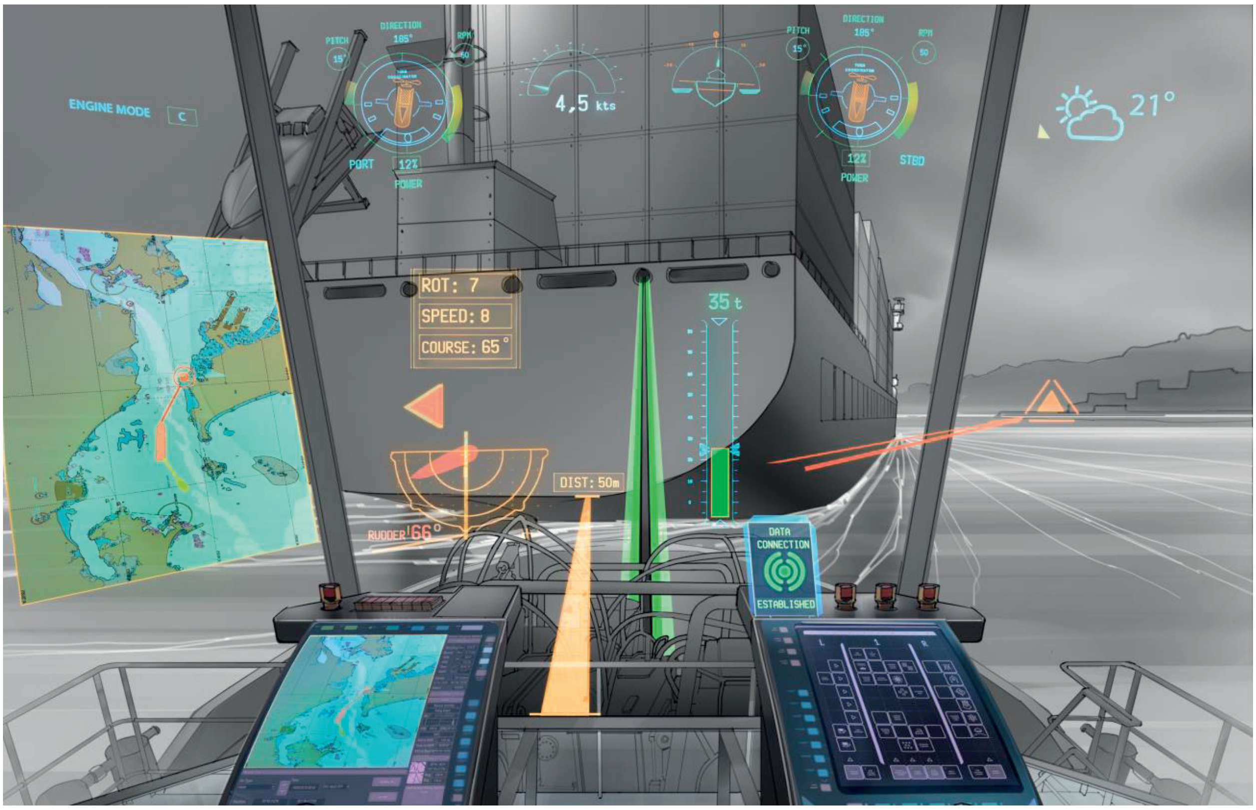


Figure 4. The Intelligent Towing Concept (© 2013 Rolls-Royce plc).

DISCUSSION

Figure 5 presents the overall workflow of the design approach suggested in this chapter. The approach is based on the following premises, which are presented in the figure. Firstly, we assume that it is useful to distinguish the challenges and strengths involved in the activity that the design solution is to serve; this can be accomplished by study of users (Step 1 in Figure 5) and with Core-Task Analysis (Step 2). We believe that it is beneficial for the design solutions to draw on and support these existing strengths that reflect professionalism and human capabilities. Assumedly, current industry professionals are likely to take a positive view of a design solution if that solution allows the user to apply his or her existing potential and/or provide support for resolving real-world challenges in the work – that is, if the so-called systems usability design goals are addressed (Step 3a). These goals can then be arranged visually – for instance, in tabular form as shown in Table 1 (Step 4a). Furthermore, we find it beneficial if the design solutions reflect the most important emotion-specific elements related to the work domain (that is, addressing issues such as work-related identities and emotion-laden ideals). We do not suggest a particular method for this. Instead, we encourage intuitive consideration of of interview- and observation-based findings; empathy is needed for derivation of emotional user experience goals from the various accounts given by professionals (Step 3b). Reformulation of data into broad function-oriented goals and into emotion-related stories, themes, and/or user personas (Step 4b) allows that measure of "distance" from user data that is necessary for radical design. Furthermore, the design solutions should not merely address the existing domain-specific challenges; future challenges and needs too are important (Step a). This further assures that the design solutions are future-oriented, which, again, leaves room for more radical design alternatives. The future studies should include technology foresight (Step b1), industry-domain-related foresight (Step b2), and examination of general trends and likely future values within the society (Step b3); the future studies may take place in parallel with user studies. Applied together, these procedures lay a foundation for initial design ideas (Step 5), which can be visualized or prototyped for purposes of user evaluations. It might be beneficial if the visualizations are embedded in user scenarios (that is, stories) so that users may more easily imagine themselves applying the concept solutions (Step 6). By these means, co-creation and co-evaluation take place; considering the solutions alongside the users enables the better ideas to be distinguished far more readily from the worse. Those deemed worthy can be further developed (Step 7). This ultimately leads to the final concept design solutions (Step 8).

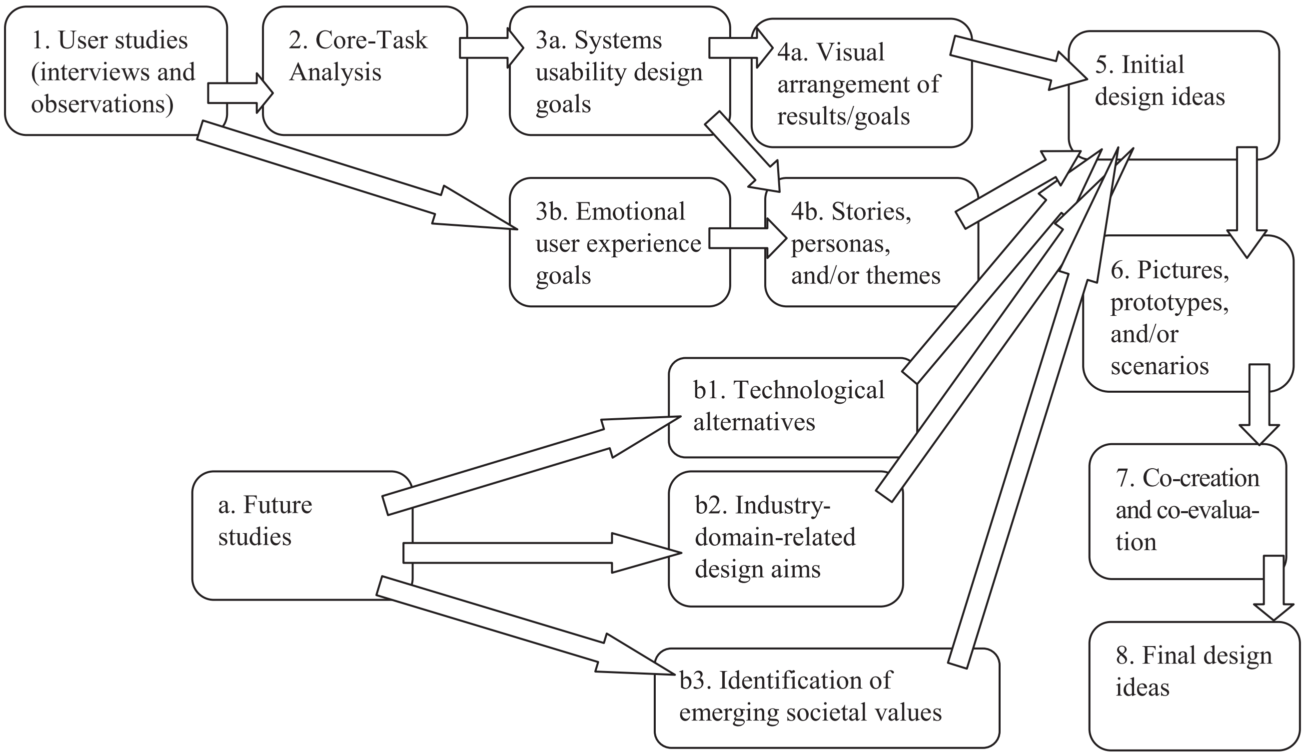


Figure 5. Overall workflow for user-oriented radical concept design.

The approach presented in Figure 5 reflects some existing design approaches, contextual design (Beyer & Holtzblatt, 1998) chief among them. Contextual design assumes that good design solutions are created through profound understanding of the use (or work) context. The idea is that "contextual inquiry" precedes the actual design process. This includes studying the work or use context that the new design is to serve. Among the methods are interviews and observations: thereby, shared understanding of the content of the work is developed with the users. Notess (2005) summarizes contextual design in general by suggesting that it applies four core principles: First is the assumption that data on work activity are largely contextual and, therefore, it is the actual work context that is to be studied. The second principle is that designers should work in partnership with users who act as experts. Thirdly, contextual design applies visualizations; that is, the findings from contextual inquiry are presented via diagrams for the purpose of aiding the design process. The fourth principle, that of iteration, implies that the design process is not entirely linear. Instead, it involves paper prototypes that may lead to further refinement of the product. The approach presented in Figure 5 is largely in line with these principles of contextual design. Some noteworthy differences from the contextual design approach do, however, exist, these serving the purpose of generating radical design ideas.

The first key difference is that the design approach we present applies the CTA method as a means of analyzing the contextual inquiry data. Typically, CTA has been applied for the purpose of studying various work contexts, such as nuclear power plant (Norros, 2004) and metro train (Karvonen et al., 2011) operation, without the explicit purpose of providing the relevant domains with new concept design solutions. Nonetheless, there are elements of the method that render it useful for provision of indications for radical design. This is because the method can be used to generate broad design goals. The core-task demands pinpointed – that is, the ways in which the control demands and resources uncovered connect one with another – can be seen as design goals. These general design goals can be considered to be systems usability goals, ideas as to how the overall work system, including its users, the environment, and the technologies employed, could and should function together. Furthermore, the method provides visualizations of the findings, thereby responding to the common conclusion in design studies (Findeli, 2001; Schön, 1992) that visual representations are especially beneficial in design activity. In our experience, the visual models aid in conceiving of and communicating the model and the findings in the concrete design work.

Secondly, the approach we suggest here emphasizes the importance of considering user experience in design. The latter reflects the experience design approach. It suggests that, rather than products, it is the user experience that should be at the focus in design work (Hassenzahl, 2010). "User experience goal driven design" as described by Karvonen and colleagues (2012) and by Koskinen et al. (2013), in turn, is more specific in suggesting that "user experience goals" should be defined at the very beginning of the design process. In other words, the designers should, in principle, begin by determining what kind of activity or emotion should be supported by the design; only after this should product-related design ideas be generated. This, indeed, might allow "thinking outside the box," since the design process is not bound to the existing product when one considers the future design. Both CTA-derived systems usability goals and the feeling-related non-instrumental user experience goals draw the design focus from the pre-existing product to workers' potential future activity. As has been suggested elsewhere (Wahlström et al., under review), user studies may provide a basis for radical innovations if the resultant findings are reformulated as models, goals, or themes in such a manner that sufficient distance from the findings themselves is achieved. This implies the findings not directly dictating the design solutions so much as informing them by providing broad but justified design indications. These assumptions are in line with the thinking of Hekkert et al. (2003), who propose that innovative product design can be achieved when one first abandons presuppositions about the product and then develops the product by formulating three "visions": an initial vision of the user and the context of use is elaborated into an interaction vision, which describes how the user interacts with the product, then a product vision. The latter approach arguably forces designers to free themselves of apparent restrictions or requirements and, instead, encourages seeking desirable possibilities. The approach of Hekkert and colleagues also entails the designer empathizing with the future user; however, the user is not directly involved in the design process. They state that in this way undesirable constraints resulting from users' fixation on familiar solution directions are avoided.

Indeed, design through empathetic understanding of users has been a common theme in the design literature. Leonard and Rayport (1997) introduced empathic design as an approach complementary to marketing research, one that contributes to the flow of ideas that still need further testing. When a company representative explores the customers' worlds through the eyes of a fresh observer, the company can redirect existing organizational capabilities to new markets. Wright et al. (2008) remind that good experience-centered design requires designers to engage with the users and their culture in rich ways in order to understand how the users make sense of technology in their lives. Kouprie and Sleeswijk Visser (2009) propose a framework for empathy in design, formulated as "stepping into and out of the user's life." Proceeding from psychological literature, they distinguish two components of empathy: the affective and the cognitive. The affective component includes the emotional response, feeling and identifying with the user: “becoming the user.” The cognitive component, on the other hand, includes understanding, adopting a perspective, and imagining the other: “staying beside the user.” Sleeswijk Visser (2009) emphasizes that knowing the users’ world is important for designer motivation, and stories are tools that contribute well to this understanding. Successfully communicated user information provides empathy and inspiration for product ideas.

A third way in which our design approach differs from typical contextual design is in our emphasis on the inclusion of future studies in derivation of design indications. Foresight is addressed in innovation management literature and as a separate field, but it seldom features explicitly in user-centered design literature or in the design literature generally. Design literature usually relies on brainstorming and reframing methods and approaches for creating novel ideas and breaking free from current constraints (e.g., Krippendorff, 2006). However, an understanding of technology and societal trends can be crucial for a product or company’s success. For example, Christensen (1997) has identified multiple cases of existing market leaders failing to understand the effect of emerging technologies. Overall, our approach reflects the technology research of Nieminen and Mannonen (2005) and more general trend analysis carried out by Salovaara and Mannonen (2005). Nieminen and Mannonen suggest that a separate element of technology research should be part of user-centered concept development projects, for tying the design ideas to a meaningful level of technology. Salovaara and Mannonen, in turn, attempt to achieve balance between the future-orientation and user-centeredness requirements of concept development by dividing the design-supporting information into information on upcoming changes (i.e., trends in society and working life) and stable features of the context. Our foresight involved a wider view of the future than pure technology review in that it also covered general trend analysis. However, as has been discussed elsewhere (Wahlström et. al, 2014), very central to our approach is the combination of user-study findings with future studies (of interaction technologies in particular); it can be assumed that novel interaction technologies are those with the most potential to enhance work activities.

Fourthly, in a similarity to contextual design, our approach accepts the utility of working alongside the users in design. The terms “participatory design,” “co-creation,” and “co-design” have been used in the design literature. The associated concepts are employed with varying meanings, but we follow the interpretation by Sanders and Stappers (2008), in which co-creation refers to any creative act involving more than one individual while co-design refers to co-creation in a design process specifically. Participatory design, in turn, can be seen as a Northern-Europe-based design movement that views the user as a partner in the design process. We embrace the underlying message implied by these concepts but suggest that users should be involved only after the generation of the initial design ideas. However, the design ideas do have to be based on thorough and empathic understanding of the users’ work and work environment. This is vital for the purpose of creating radical design ideas – i.e., it is important because the aim is to create solutions not yet imagined by the users. Here, the users are involved in two phases of the concept design: in the beginning as informants on the work domain and in the final stage as co-designers evaluating and enhancing the initial design concepts. In our experience, the visually illustrated design ideas have served well as “boundary objects” that deliver understanding of a joint object across boundaries between disciplines (Star & Griesemer, 1989). Star and Griesemer emphasize the communicative nature of boundary objects in enabling collaborating parties to represent, transform, and share knowledge. In our co-design activities, wherein the stakeholders had very different backgrounds, the visualized future concepts facilitated and encouraged communication. The future concepts served well in the three roles of boundary objects proposed by Star (2010): offering personalized value to each party, facilitating understanding of the task at hand in the same manner, and offering an understanding of all relevant options related to the task.

Our contribution with the design approach presented in this chapter has potential to yield concept designs that are radical yet remain grounded in the actual activity and work-domain-specific needs at hand. As is arguably visible in the ship-bridge design solutions presented here, the approach has worked for us. Future studies, however, would be needed to confirm the approach's generalizability, its broader utility.

ACKNOWLEDGMENTS

This study is part of the User Experience and Usability in Complex Systems (UXUS) research and development program, which is one of the research programs of the Finnish Metals and Engineering Competence Cluster, FIMECC.

This study was preliminarily reported in a conference paper (Wahlström, Karvonen, Kaasinen & Mannonen, 2014). The present paper further elaborates the argument and includes new information on the ship-bridge design project.

The authors would like to thank all those who have provided input to the design of future ship bridges and to this article, among these being Leena Norros (VTT), Iiro Lindborg (Rolls-Royce), Jussi Lehtiniemi (Troll VFX), Zhang Daoxiang (Aalto University), Hanna Askola (VTT), Göran Granholm (VTT), Frøy Birte Bjørneseth (Rolls-Royce), and Karno Tenovuo (Rolls-Royce).

REFERENCES

Beyer, H., Holtzblatt, K. (1998), *Contextual Design: Defining Customer-centered Systems*. San Francisco: Morgan Kaufmann.

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

Christensen, C. M. (1997), *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston, MA: Harvard Business School Press.

Cooper, A. (1999), *The Inmates Are Running the Asylum*. Indianapolis: Morgan Kaufmann.

Findeli, A. (2001), "Rethinking design education for the 21st century: Theoretical, methodological and ethical discussion," *Design Issues*, Vol. 17, No. 1.

Frog (2012), "12 tech trends that will define 2012, selected by Frog's design minds." Fast Company Web site: http://www.fastcodesign.com/1668912/12-tech-trends-that-will-define-2012-selected-by-frogs-design-minds.

Frost & Sullivan (2010), "World's top global mega trends to 2020 and implications to business, society and cultures*."* Frost & Sullivan

Gartner Research (2009), "A day of your life, 2028." Gartner Web site: https://www.gartner.com/doc/933312/day-life-.

Gould, J., Lewis, C. (1985), "Designing for usability: Key principles and what designers think," *Communications of the ACM*, Vol. 28, No. 3.

Hassenzahl, M. (2010), "Experience design: Technology for all the right reasons," *Synthesis Lectures on Human-centered Informatics*, Vol. 3, No. 1.

Hekkert, P., Mostert, M., Stompff. G. (2003), "Dancing with a machine: A case of experience-driven design," in: *Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces (DPPI '03)*. New York: ACM.

JWT Intelligence (2012), "100 things to watch in 2012." JWT Intelligence Web site: http://www.jwtintelligence.com/2011/12/100-things-to-watch-in-2012/

Kaasinen, E., Roto, V., Hakulinen, J., Heimonen, T., Jokinen, J., Karvonen, H., Koskinen, H., Keskinen, T., Lu, Y., Saariluoma, P., Tokkonen, H., Turunen, M. (in press, 2015), "Defining user experience goals to guide the design of industrial systems," *Behaviour & Information Technology*.

Karvonen, H., Koskinen, H., Haggrén, J. (2012), "Defining user experience goals for future concepts. A case study," in: *Proceedings of 7th Nordic Conference on Human–Computer Interaction (NordiCHI2012), UX Goals 2012 Workshop – How to Utilize User Experience Goals in Design?*. Tampere, Finland: TUT Publication series.

Karvonen, H., Aaltonen, I., Wahlström, M., Salo, L., Savioja, P., Norros, L. (2011), "Hidden roles of the train driver: A challenge for metro automation," *Interacting with Computers*, Vol. 23, No. 4.

Keinonen, T. (2009), "Immediate and remote design of complex environments," *Design Issues*, Vol. 25, No. 2.

Koskinen, H., Karvonen, H., Tokkonen, H. (2013), "User experience targets as design drivers: A case study on the development of a remote crane operation station," in: *Proceedings of the 31st European Conference on Cognitive Ergonomics (ECCE2013)*. Toulouse, France: ECCE.

Kouprie, M., Sleeswijk Visser, F. S. (2009), "A framework for empathy in design: Stepping into and out of the user's life," *Journal of Engineering Design*, Vol. 20, No. 5.

Krippendorff, K. (2006), *The Semantic Turn: A New Foundation for Design*. New York: Taylor & Francis.

Leonard, D., Rayport, J. F. (1997), "Spark innovation through empathic design," *Harvard Business Review*, Vol. 75.

Lu, Y., Roto, V. (2014), "Towards meaning change: Experience goals driving design space expansion," in: *Proceedings of the 8th Nordic Conference on Human–Computer Interaction*, 717–726. ACM.

Mahlke, S. (2005), "Understanding users' experience of interaction," in: *Proceedings of the 2005 Annual Conference of the European Association of Cognitive Ergonomics (EACE '05)*. Athens, Greece: University of Athens.

Norman, D. A., Verganti, R. (2014), "Incremental and radical innovation: Design research versus technology and meaning change," *Design Studies*, Vol. 30, No. 1.

Norros, L. (2004), *Acting under Uncertainty: The Core-Task Analysis in Ecological Study of Work*. Espoo, Finland: VTT Publications.

Norros, L. (2013), "Developing human factors/ergonomics as a design discipline," *Applied Ergonomics*, Vol. 45, No. 1.

Norros, L., Savioja, P., Koskinen, H. (in preparation). *Core-Task Design – a practice-theory approach to human factors.* Manuscript in preparation.

Nieminen, M. P., Mannonen, P. (2005), "User-centered product concept development," in: *International Encyclopedia of Ergonomics and Human Factors*, 2nd edition, Karwowski, W. (ed.), 1728–1732.

Notess, M. (2005), "Using contextual design for digital library field studies," in: *Proceedings of ACM/IEEE–CS Joint Conference on Digital Libraries Workshop Studying Digital Library Users in the Wild: Theories, Methods, and Analytical Approaches*. Denver, CO: ACM.

Pruitt, J., Adlin, T. (2006), *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. San Francisco, CA: Elsevier.

Rolls-Royce (2014), "Rolls-Royce presents the future of tug bridge controls." Available at: https://www.youtube.com/watch?v=27uCL90s20o (accessed on January 8, 2015).

Rolls-Royce (2015), "Ship intelligence for cargo vessels." Available at: https://www.youtube.com/watch?v=_nApv-C7qSg (accessed on January 8, 2015).

Salovaara, A., Mannonen, P. (2005), “Use of future-oriented information in user-centered product concept ideation,” in: *IFIP TC13 International Conference on Human–Computer Interaction (Interact 2005)* (Lecture Notes in Computer Science, Vol. 3585). Springer-Verlag.

Sanders, E., Stappers, P. (2008), “Co-creation and the new landscapes of design,” *CoDesign: International Journal of CoCreation in Design and the Arts*, Vol. 4, No. 1.

Savioja, P. (2014), *Evaluating systems usability in complex work*, doctoral dissertation. Available via VTT Web site: http://www2.vtt.fi/inf/pdf/science/2014/S57.pdf

Schön, D. (1992), “Designing as reflective conversation with the materials of a design situation,” *Research in Engineering Design*, Vol. 3, No. 3.

Sleeswijk Visser, F. (2009), *Bringing the Everyday Life of People into Design*, doctoral dissertation. Available via TU Delft Web site: http://www.narcis.nl/publication/RecordID/oai:tudelft.nl:uuid:3360bfaa-dc94-496b-b6f0-6c87b333246c.

Star, S. L. (2010), “This is not a boundary object: Reflections on the origin of a concept,” *Science, Technology, & Human Values*, Vol. 35, No. 5.

Star, S. L., Griesemer, J. R. (1989), “Institutional ecology, “translations” and boundary objects: Amateurs and professionals in Berkeley’s Museum of Vertebrate Zoology, 1907–39,” *Social Studies of Science*, Vol. 19, No. 4.

Wahlström, M., Karvonen, H., Kaasinen, E. (2014), “InnoLeap – creating radical concept designs for industrial work activity.” Presented at NordiCHI 2014 Workshop WS4: The Fuzzy Front End of Experience Design, October 26, 2014, in Helsinki.

Wahlström, M., Karvonen, H., Kaasinen, E., Mannonen, P. (2014). ”Designing for Future Professional Activity – Examples from Ship-Bridge Concept Design,” in: *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*, Kraków, Poland 19-23 July 201

Wahlström, M., Karvonen, H., Norros, L., Jokinen, J., Koskinen, H. (under review), *Radical Innovation by Theoretical Abstraction – a Challenge for Design Anthropologists*, under peer review.

Wright, P., Wallace, J., McCarthy, J. (2008), “Aesthetics and experience-centered design,” *ACM Transactions on Computer–Human Interaction*, Vol. 15, No. 4.

VII

UTILIZING EXPERIENCE GOALS IN DESIGN OF INDUSTRIAL SYSTEMS

by

Virpi Roto, Eija Kaasinen, Tomi Heimonen, Hannu Karvonen,
Jussi P.P. Jokinen, Petri Mannonen, Hannu Nousu, Jaakko Hakulinen,
Yichen Lu, Pertti O. Saariluoma, Tiina Kymäläinen, Tuuli Keskinen,
Markku Turunen, Hanna Maria Kaarina Koskinen

In Proceedings of the 2017 CHI Conference on Human Factors
in Computing Systems (pp. 6993–7004)

https://doi.org/10.1145/3025453.3025620

Utilizing Experience Goals in Design of Industrial Systems

Virpi Roto[1], Eija Kaasinen[2], Tomi Heimonen[3], Hannu Karvonen[2], Jussi P.P. Jokinen[1], Petri Mannonen[1], Hannu Nousu[4], Jaakko Hakulinen[5], Yichen Lu[1], Pertti O. Saariluoma[6], Tiina Kymäläinen[2], Tuuli Keskinen[5], Markku Turunen[5], Hanna Maria Kaarina Koskinen[2]

[1]Aalto University, Helsinki, Finland, [first.last]@aalto.fi
[2]VTT Technical Research Centre of Finland Ltd, Tampere, Helsinki, Finland, [first.last]@vtt.fi
[3]University of Wisconsin-Stevens Point, Stevens Point, Wisconsin, United States, [first.last]@uwsp.edu
[4]KONE Corporation, Helsinki, Finland, [first.last]@kone.com
[5]University of Tampere, Tampere, Finland, [first.last]@uta.fi
[6]University of Jyväskylä, Jyväskylä, Finland, [first.last]@jyu.fi

ABSTRACT
The core idea of experience-driven design is to define the intended experience before functionality and technology. This is a radical idea for companies that have built their competences around specific technologies. Although many technology companies are willing to shift their focus towards experience-driven design, reports on real-life cases about the utilization of this design approach are rare. As part of an industry-led research program, we introduced experience-driven design to metal industry companies with experience goals as the key technique. Four design cases in three companies showed that the goals are useful in keeping the focus on user experience, but several challenges are still left for future research to tackle. This exploratory research lays ground for future research by providing initial criteria for assessing experience design tools. The results shed light on utilizing experience goals in industrial design projects and help practitioners in planning and managing the product design process with user experience in mind.

Author Keywords
Experience goal; Experience-driven design; Experience design tools; User experience; Industrial systems.

ACM Classification Keywords
Human-centered computing ~ Empirical studies in interaction design.

INTRODUCTION
In early 2000s, researchers introduced the experience-driven design approach, in which an intended user experience (UX) is the primary objective of a design process [10,30], in contrast to problem- or technology-driven design. Designers first define the experience they aim to enable in the users of the design, and only then decide what kind of product, service, or system would best enable the intended experience [4,5,6]. The design space is determined by the intended experience rather than by the technology at hand [18]. However, it seems difficult for technology-driven industry to adopt this kind of an experience-driven design approach, since technological skills in a company often dictate the design space. For example, an experience goal of social connectedness can be implemented in various ways, but a company developing mobile apps typically limits the design opportunities to social apps and does not start to market, say, cruises for singles. The core proposition in experience design, 'experience before product' [5, p.63], is thus rarely realized in industry.

While human-centred design (HCD) [11] and goal-directed design [3] have been widely studied in industry contexts, few works investigate the integration of experience-driven design into product development processes. Most studies about integrating UX design activities to industrial product development examine current practices, such as how the UX work in general has been integrated to the agile development process [28], or which tools the industry uses for experience design [25], rather than the integration of a new tool into a product development process. Rozendaal [25] shows that experience design processes in industry include largely similar properties as HCD, such as iterative development, a somewhat unpredictable process, user insights, prototyping, and empathy tools such as personas. Narratives were used as a tool to explain the intended experience, but how they were used in the different phases of the design process was not studied.

The aim of this research was to understand how experience design could be introduced to product development in companies. Although purely experience-*driven* design was not always possible, we wanted to advance the design process from traditional human-centered design towards an experience-centric approach, and from addressing utilitarian user needs to psychological needs [7]. We used experience goal (Xgoal) [12,32] as the central conceptual tool in transferring research knowledge to practice. An Xgoal is a

CHI 2017, May 06 - 11, 2017, Denver, CO, USA.

ACM 978-1-4503-4655-9/17/05…$15.00.
DOI: http://dx.doi.org/10.1145/3025453.3025620

design goal that states "the intended momentary emotion or the emotional relationship/bond that a person has towards the designed product or service" [18]. Thus, an Xgoal is more specific than a generic 'good' or 'pleasant' experience [4, p.11]. Table 1 provides examples of Xgoals defined in the design cases described in this paper. Earlier work has studied how to define what experience to design for, i.e., Xgoal setting [12], so this paper focuses more on the other challenge of experience-driven design: how to design something that could evoke that experience [4, p.11].

This research suggests that for experience design methods, the same applies as for usability testing methods: we should focus on ingredients (design techniques) rather than recipes (design methods) [33]. By focusing on Xgoals, we address the need stated by Woolrych et al. [33]: "HCI needs to focus more on what gets cooked, and how it gets cooked, and not just on how recipes suggest that it could be cooked." Thus, Xgoals focus on what gets cooked, and integrating them to industrial product development focuses on how it gets cooked. Companies have different design practices and the design projects are different. Rather than providing a full recipe by following the experience-driven design process by the book, we aimed to understand if Xgoals can be integrated to existing design processes to help product development teams focus on experiential aspects.

The objective of this research was to integrate Xgoals into design process and study their benefits and drawbacks. This paper specifically examines how Xgoals are used by multi-disciplinary development teams in as realistic product development contexts as possible. By examining four such case studies, this work sheds light on utilizing Xgoals in industrial design projects and proposes criteria for assessing experience design tools.

METHODOLOGY

Since studies on introducing Xgoals to industrial product development appear to be rare, and only one study has investigated the experience design process in industry [25], we only had a tentative idea on how Xgoals could work in industry; no quality criteria existed for experience design tools such as Xgoals. Therefore, exploratory research was chosen as the research approach.

In a research program with metals and engineering industry in Finland, we worked together with several companies to introduce experience-driven design into new product development. The company partners of the research program identified suitable topics for experience design, and a group of researchers worked with each company, following design approaches that suited the case and matched the team's expertise (Table 1). This setup resembled real life design projects in the sense that the project staff was not highly experienced in experience-driven design, and the projects were restricted by several practical limitations.

Our cases were active in different phases of the product development process, although all of them covered phases before the productization only. The cases followed different types of design processes, varying from scenario-based design to iterative agile development. This is why the cases were analyzed based on the different types of activities during the process, rather than distinct development phases. Following the structure of the traditional HCD process [11], we name the main activities as *Investigation* (ISO: Understand and specify of context of use), *Design* (Produce design solutions), and *Evaluation* (Evaluate).

The first activity category is Investigation. All activities aiming at improving team's understanding of the task at hand belong to this category, such as background research (interviewing users of current systems, literature review), defining and analyzing Xgoals or other types of design requirements. The second category, Design, includes the generative activities of ideation and prototyping the actual product concept. Evaluation, the third category, includes assessment of the concepts to identify whether they evoke the intended experiences, and evaluation of the feasibility of candidate Xgoals.

In earlier phases of the research program, we had developed methods for setting Xgoals, but the teams had little experience in using them in design projects. The viewpoints of different stakeholders were included in the Xgoal-setting, where insight and inspiration was sought from company brands, theoretical knowledge about users, stepping into the users' shoes with empathy, possibilities and challenges of new technologies, and vision of renewal [12].

The design cases were executed independently and in different points in time, but some researchers were involved in several cases. Although lessons learned from an earlier case informed and influenced subsequent cases, we are not analyzing the learning process or the design outcomes, but rather aim to understand how Xgoals were utilized in the different activities of design process. The Xgoal utilization analysis of each case was done retrospectively by the researchers who participated the case.

After the project was completed, the groups of researchers reported their Xgoal utilization in the main project activities. They were asked to report general information about the design case, i.e., the columns in Table 1, and how Xgoals were utilized in the three HCD activities. The utilization and impact of Xgoals was analyzed in several meetings with representatives from all project teams, and additional details were added as needed to make the descriptions easier to compare.

DESIGN CASES

This paper analyses four design cases from three companies: *Future Factory, SmartGUI, Remote Operation Station (ROS), and Remote Elevator Control* (Table 1). In the *Future Factory* case, the design project aimed at far reaching future concepts where envisioning the context of

Case	Design brief	Xgoals	Outcome	Approach	Team
Future Factory	An extensive vision of a factory process control room and work practices targeting 10-15 years ahead.	E.g., Trust in automation, Sense of freedom, Ownership of the process, Relatedness to the work community	A science fiction prototype in the form of scenario videos and interaction demonstrators	Future trends analysis and field visits as the basis to set Xgoals. Experience-driven science fiction prototype developed in a series of multidisciplinary co-design workshops [16]. Evaluation with user interviews and a web survey.	2 R&D staff members from the company; 5 researchers (systems usability, experience design); a video production professional
SmartGUI for crane operation	Analyze how the existing crane operation user interface supports Xgoals and what features are still missing, and investigate how the missing features could be implemented.	Supporting competence; Avoiding anxiety	A touch screen graphical user interface for crane operation (functional prototype)	Crane operator interviews and the framework of emotional UX [26,27] as the basis to set Xgoals. Scenarios guided design. 1-day concept design workshop with company staff. Qualitative evaluation of concepts.	1 engineer from the company; 2 researchers (psychology); 6 engineering staff members in concept design workshop
Remote Operator Station (ROS) for cranes	A novel remote operator station concept for container cranes in ports with 'hands-on experience for remote operation'.	Feeling of safe operation; Sense of control; Feeling of presence; Experience of fluent co-operation	Virtual reality based container crane remote operator station simulator (a functional prototype).	Core-Task Analysis [22,23] for work and domain analysis and the Systems Usability Framework [29] to set Xgoals and user requirements. From these, design implications and then design solutions were derived. Xgoals' fulfillment evaluated via user requirements in user tests.	1 usability expert and 1 designer from the company; 5 researchers (systems usability, psychology, interactive technologies)
Remote Elevator Control	A remote control solution for elevators in complex buildings as a mobile application.	E.g., Feeling of control of elevator action; Reduced feeling of waiting	A mobile application for controlling elevators (functional prototype)	The company and researchers jointly defined the Xgoals. Agile development of the mobile application until sufficient for user testing.	2 R&D staff members from the company; 4-6 researchers (computer science, interaction technology)

Table 1. An overview of the design cases.

future factory work was the key. The project started with future trend analysis and by collecting understanding of users and context of use from experts. Based on these, the initial Xgoals were defined in expert workshops and evaluated with potential users. In addition to Xgoals, an overall experience vision “Peace of mind” was set. The team could not have defined Xgoals without an idea of the future context of factory work, and therefore there was a need for a holistic future vision of the concept, i.e., a science fiction prototype that illustrated the Xgoals, which was used as a conversional tool in the evaluations [16]. In summary, a large part of the *Future Factory* case consisted of investigation activities, and a set of Xgoals was one of the main outcomes of the case.

The *Elevator Control* case was quite different from the Future Factory case. Since the company already had a clear vision of the system to be designed, the Xgoals were agreed in the very beginning of the project. Xgoals were based on team members’ knowledge of user needs gained from years of domain expertise. Thus, the investigation activities focused on defining how the Xgoals can be met in the given context. In this case, more specific design implications were defined as the outcome of the investigation activities.

In the *Remote Operator Station (ROS)* and *SmartGUI* cases, field studies with actual users were done before setting the Xgoals, and more focused studies followed to indicate the design implications of each Xgoal. In the *ROS* case, also an overall experience vision “Hands-on experience of remote operation” was set.

UTILIZING XGOALS IN PRODUCT DESIGN PROCESS

The following subchapters will expand the case descriptions to the different activities, i.e., Investigation, Design, and Evaluation. We will describe these activities in each case and identify similarities and differences in the utilization of Xgoals in each design process.

Xgoal utilization in Investigation activities

Through Investigation activities, the development team aims to gain an understanding of users, other stakeholders, context of use, future trends, competitors, and anything that helps to design a good interaction concept for the given purpose. The outcome of these activities can include not only Xgoals but also other user requirements, as well as understanding of the context. These outcomes can further be interpreted to concrete design implications. Compared to the ordinary requirements that define *what* to design, Xgoals define *how* the design outcome should feel like.

In our four cases, we could identify three different roles of Investigation activities as illustrated in Figure 1. When the investigation is done in order to understand what could be the best possible experiences, Xgoals can be the final outcome of the analysis (Figure 1, type 1). Xgoals can also be defined mid-way of the analysis, and they are further specified during additional investigation (Figure 1, type 2). If Xgoals are defined in the initial design brief, all analysis activities can focus on understanding how the Xgoals could be realized in the specified context (Figure 1, type 3). The general interplay between analysis and Xgoal identification could follow any of these paths – and Xgoals can be continuously defined and refined during design and evaluation activities.

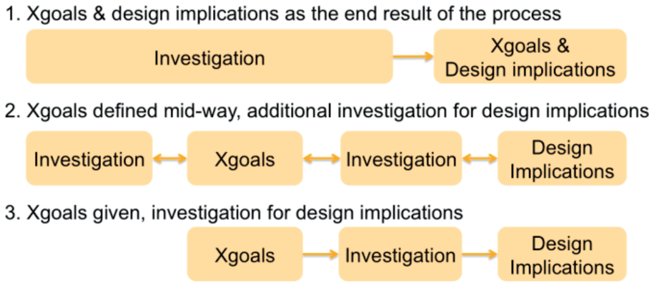


Figure 1. Interplay between investigation activities and Xgoals

Although Figure 1 is simplification of the actual, more iterative processes, it may clarify the overall differences between the cases in the way of Xgoal definition. The *Future Factory* case was of the first type, *ROS* and *SmartGUI* cases the second, and *Elevator Control* case the third type.

In the *Future Factory* case, a preliminary user study included studies of workers’ current experiences of process control work and the positive experiences they expect to have in the future. Consequently, Xgoals were defined in a series of multidisciplinary co-design workshops that involved researchers and company partners [13].

In the Remote operator station (*ROS)* case and the SmartGUI cases first phase investigation activities produced an initial set of UX Goals, which was refined based on second phase investigation activities. In the ROS case, the defined Xgoals were used as the basis for some of the questions asked in the interviews of the crane operators in field studies (for details, see [14]). For example, regarding the sense of control Xgoal, the researchers asked the operators’ opinion on which factors are important in crane operation to achieve a good sense of control. In this way, the team could also collect feedback on whether the proposed Xgoals were the correct ones. Based on the results of the interviews and field studies, the team defined detailed design implications for the chosen Xgoals (for details, see [15]). The design implications described the specific way to enable each Xgoal in the new product. The *ROS* case went through the following process: Investigation → Xgoals and user requirements → Design implications → Design solutions. Each of the defined user requirements was also connected to the appropriate Xgoal(s). In this way, all the created design solutions based on these requirements were traceable back to the originally defined Xgoals.

The *SmartGUI* team developed an initial set of Xgoals and design heuristics based on 31 crane operator interviews. The Xgoals were “competence support” and “anxiety avoidance". The Design heuristics were defined to interpret the Xgoals for the design, similar to the Design implications in ROS case. For example, Xgoal Competence was connected to design heuristic "design for middle-level users, but offer shortcuts for experts". The connection between Xgoals and heuristics was established after a user study where, e.g., competence was found to be related to understanding the goals of the task, and not being constrained by the system for doing these tasks. The heuristics were general rules for any design solution while the design implications were more requirements of solutions to be included. For instance, "feeling of presence" Xgoal was interpreted to several design implications regarding operation view and auditory feedback.

In *Elevator Control* case, the teams had gained extensive user understanding from earlier user research, hence there were no specific background research studies. However, formulating the existing knowledge into Xgoals helped in crystallizing the goals of the project. Discussing those goals influenced the design team's general mindset and understanding of the requirements for the design.

Xgoal utilization in Design activities

The original purpose of setting Xgoals is to guide design. In practice, Xgoals are not used in vacuum, but various other guiding forces influence design solutions as well. In *ROS* and *SmartGUI* cases, investigation activities interpreted Xgoals to concrete design implications for design activities. In the *Elevator Control* case, Xgoals were used to create and maintain experience mindset in the design team. In *Future Factory* case, Xgoals were design outcomes similar to future scenarios. These different roles of Xgoals in making design decisions are described in the following.

In the *ROS* case, with the aid of the defined design implications, the team managed to produce design solutions that they considered to be in line with the defined Xgoals. The concept design activity included several co-design workshops in which all the project partners had a possibility to present their ideas. In order for the design team to focus on correct issues, insights from the fieldwork, the Xgoals, the user requirements, and the design implications were all gone through at the beginning of each design workshop. After these workshops, some of the most promising ideas were collaboratively materialized into a variety of low fidelity mock-ups and prototypes. During this process, many alternative design solutions and technological possibilities were considered and iterated. The iterative approach, which assessed the design outcomes, enabled the design team to quickly move towards one basic ROS concept within which different features were altered to fine-tune the concept. To support this activity, the team built a virtual reality based prototype, which was evaluated in two different stages. The user interface of this prototype was iteratively developed according to the defined Xgoals and requirements. However, Xgoals were not emphasized in the UI design phase as strongly as in the concept design phase.

The *SmartGUI* team organized a design workshop, with crane designers as participants, to conceptualize the new interface for operating automated Electric Overhead Traveling (EOT) crane features. In the beginning of the workshop, the Xgoals, their detailed descriptions, and the set of heuristics for evaluating how well the possible concepts supported the targets, were presented to the workshop participants. The participants were assigned into teams for specific design problems involved with the controller concept. The teams were asked to justify each solution using the Xgoals.

The result of the workshop was a number of solutions for the given design problems. However, at this point, the justification of the solutions based on the Xgoals was not visible in the concepts. In the next step, the individual solutions were brought together and made into a prototype controller concept for automated EOT crane features. At this point, the total solution was evaluated with the set of heuristics, determined in the analysis phase to be important regarding the Xgoals.

Although the workshop teams were asked to justify their solutions by reference to the Xgoals, the workshop participants were unsure how much the Xgoals affected their ideas when they were asked about this at the end of the workshop. It is probable that the introduction to the Xgoals at the beginning of the workshop gave some ideas and a common framework for the conceptualization, but that the Xgoals were not explicitly present. The reason for this may be that the Xgoals Competence and Avoiding Anxiety were too abstract to be understood as design guidance, whereas the heuristics perhaps were too concrete and detailed to give room for new ideas to serve the operationalization of concrete design solutions to the given problems. It was perhaps not possible to examine how the concepts would exactly affect or otherwise be connected to the Xgoals. Another possibility is that the Xgoals were not contextually rich enough, that is, their explicit application would have required more contextual narrative around them. However, when producing the final concept by combining the workshop results, more careful evaluation of how the concepts related to the Xgoals was conducted, and each concept was explicitly justified in connection to the goals.

In the *Elevator Control* case, the Xgoals were set quickly in the beginning of the project, after which they guided the initial ideation phase and provided a common context and understanding of the project goals. In this case, the overall design solution was decided before the project started (a mobile application). This is why Xgoals did not play as big role in design as technical feasibility or basic usability.

In the *Elevator Control* case, the use of the actual Xgoals in the design phase was limited to providing a common context and understanding of the project goals (a UX mindset of sorts). Although informed by the Xgoals and user feedback during evaluations, actual day-to-day design decisions within the iterative software development process tended to be based more on the practical considerations of the case, such as available technological platforms and their features and defects. For example, the timed elevator call feature, which resonates with most of the stated Xgoals, had to be implemented to call the elevator after a specified time, instead of users being able to set the arrival time of the elevator in the lobby, because the desired functionality could not be technologically supported. While the Xgoals did not often directly influence the design activities, they were linked to the changes that were made as a result of design activities informed by the findings of user evaluations. For example, the addition of a rule-based predictive floor selection, improved feedback on the elevator call status and physical touch interface can be seen as attempts to reduce the feeling of waiting, fostering better the feeling of control over the elevator system, and providing better guidance. One of the challenges in applying the Xgoals was that it was not possible to accomplish them to the extent that would have been desirable. Many ideas motivated by the Xgoals were rejected for their complexity, unrealistic technological requirements or uncertain robustness.

In the *Future Factory* case the Xgoals were expected to guide the design towards positive experiences and help in communicating important objectives [5]. The Xgoals gained quite an important role in the project as they formed the backbone for the future concepts and the science fiction prototype. In the design phase both future scenarios and Xgoals for them were developed. The drivers for the Future Factory concept were increasingly intelligent automation, new technical possibilities for remote control, remote presence of workers, and new collaboration practices. The design outcome was an extensive set of scenarios that formed a coherent future vision, a science fiction prototype, and complementary interaction demos. The Xgoals were prioritized based on user feedback, and product development company's perspectives about future technology, societal and business trends. After the prioritization, a brainstorming session was organized to group the most interesting Xgoals, user needs and future trends and to develop the initial concept candidates. The Xgoals were then used to further develop the concepts and especially to describe the usage scenarios. It should be noted that the initial Xgoals were refined in parallel to co-designing the future scenarios. In practice, it was difficult to use Xgoals to narrow down design options, as Xgoals seemed to always produce new design possibilities and opportunities. As the variety of scenarios represented multiple work situations, also the Xgoals varied in different scenarios. Still, in the end of the process, eight main Xgoals were agreed on, and they were used as the UX evaluation framework. The final Xgoals reflect well how user experiences in process control work focus on work activities instead of mere tools or user interfaces (as predicted in [21]).

Based on the experiences of the four cases, we can conclude that keeping the focus on experience is challenging even if the Xgoals would have been interpreted to concrete design implications as in the ROS and SmartGUI cases. Xgoals may be too abstract for designers to take as an input, whereas too concrete interpretations such as heuristics may not leave room for ideation. However, even if there is no concrete evidence on how Xgoals guide design, they can be beneficial in maintaining experience mindset in the design team. This happened in the Elevator control case and also in the Smart GUI case. In an ideal case, Xgoals form the backbone for whole design as happened in the Future Factory case. Then designing Xgoals gets as important as designing the actual concepts, and Xgoals do not just reflect user interaction but work experience in general.

Xgoal utilization in Evaluation activities

The Evaluation activities cover testing of the generated idea, sketches, prototypes, or the actual product. As a part of the human-centred design process, Xgoals in evaluation can be used to study whether the design evokes the desired experiences and whether the targeted experiences were what the users wanted. All our cases studied the former but the latter was present only in the Future Factory case. where it was studied whether the Xgoals were desired and valuable. A challenge in each case was that experience was not the only aspect to be studied, so experience evaluation methods had to be combined with other methods. As identified in the ROS case, experience evaluation would require a realistic setup of an actual work environment. Operating a prototype of the user interface does not raise work related experiences such as competence or feeling of presence.

Elevator Control

The first prototype of the *Elevator Control* was subjected to an initial user experience evaluation and a subsequent long-term evaluation with four participants. Based on the results of these studies, improvements and new features were implemented to the prototype. After this, a larger scale long-term evaluation was organized.

Questionnaires probing expectations and user experiences were used both in the initial user evaluation and in the larger-scale evaluation. Both times, the SUXES method [31] was utilized. Some aspects of the Xgoals where thus covered already by the existing items in the method. For example, the feeling of being in control over the elevator system will supposedly increase if the user experiences using the application to be fast, pleasant, clear and so on. Hence, the Xgoals were indirectly assessed with these statements, although not explicitly used as measures.

However, in the latter evaluation items to explicitly correspond to the Xgoals were added. The background information questionnaire asked how the participants feel about waiting for the elevator, and do they feel they have to wait for too long. For the expectations and experiences questionnaires, three additional statements were constructed to directly assess the design against the set Xgoals:

- I am able to control the elevator better when using the application.
- Using the application shortens the time I need to spend waiting for the elevator.
- Using the application expedites my daily movement.

As a result of the user studies, additional design requirements were identified that contributed to the realization of the Xgoals: offer real value to users, keep the user informed of system status, and ensure reliability of control. The latter are fairly practical and it is reasonably easy to provide examples of how to operationalize them in effective designs, whereas the first one is analogous with the Xgoal provided in the design brief.

SmartGUI

The *SmartGUI* team evaluated the initial design concepts using the heuristics for the Xgoals, and the final concept was evaluated with a prototype in a field experiment. The data from the field experiments consisted of think aloud protocols collected during the testing of the prototype controller, and of interviews conducted after the experiment. The participants of the field experiment were given tasks, which they had to complete with the prototype controller. The tasks were designed so that each aspect of the prototype would be tested. While the tasks were not designed to directly evaluate either competence support or anxiety avoidance, it was expected that the operators would be able to reflect on their emotional states during the tasks.

Both Xgoals were utilized in creating the interview questions. Themes such as being determined, having clear task goals, understanding each step in the interface use, not having to perform seemingly unnecessary actions, being able to operate the crane freely, being free from doubt and confusion, and not feeling anxious about the automated features, were connected to the competence support and anxiety avoidance Xgoals.

The think aloud protocols were analyzed using protocol analysis, in which focus was placed on thinking errors made by the participants. These errors were primarily connected to the usability of the prototype user interface. The reason for not utilizing the Xgoals in the protocol analysis was that the protocols did not relate to the goals. The protocols were detailed and task-oriented, and the goals were more descriptive of the general emotional states of the users throughout their working days. The interviews following the actual test were more suited for evaluating how the concepts related to the Xgoals. The evaluation tasks were short, of course, and contextually reduced, so much of the feedback concerning the Xgoals was hypothetical; the real results would have required more realistic setting and a longer time period.

As the result of the evaluation, five concrete design changes and five discussion points were presented. While most of the points related to the details of the concept, all could be indirectly connected with the Xgoals. However, no explicit reference to either Xgoals was made in these concluding points of the evaluation.

Remote Operator Station

The first-phase evaluation with 20 university students investigated how, and whether, the user experiences about the ROS simulator interface could be enhanced with either force feedback or visual augmentations. Here, the team emphasized the Xgoals especially in the design of the questionnaires. Since the global number of target users is very small and even domain experts are available only in few select locations, the first phase evaluation was done with 20 university students. The target at this point was to evaluate how, and whether, the user experiences of the ROS simulator could be enhanced. Based on the results received from the first-phase evaluation, for example, the amount of events triggering the force feedback was decreased. Overall, the results affected the improvement of user interface solutions supporting the Xgoals and the selections between the implemented options.

After further development efforts, another evaluation study of the prototype system was conducted with six work-domain experts. The objectives of this study were both to compare the UX of two different user interface concepts and to receive feedback on how well the Xgoals of experience of safe operation, sense of control, and feeling of presence are fulfilled with the developed ROS prototype.

This second evaluation was conducted with a simulator version of the ROS system, which was operated with two industrial joysticks and a tablet computer (see Fig. 2). A 32-inch display placed on the operator's desk provided the main operating view, which included virtual reality camera views and simulated, but realistic operational data.

To evaluate how the originally defined Xgoals and user requirements are fulfilled with the evaluated prototype, the team used a combination of different methods: interviews, questionnaires, thinking-aloud, and task performance

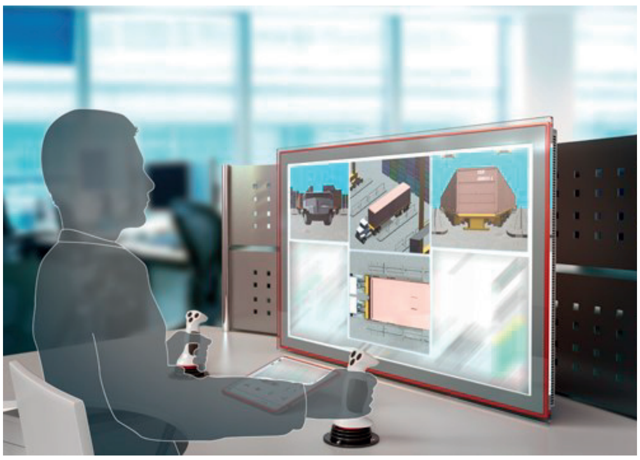


Figure 2. Concept illustration of the ROS system with the four-view setup in the main display

indicators (for details, see [13]). To assess whether the chosen Xgoals were fulfilled with the system, a Usability Case (UC) method [17] was utilized. In line with the UC method, the data gathered from the user studies was carefully analyzed regarding each defined user requirement (i.e., a subclaim in UC) on whether positive or negative cumulative evidence was found about the fulfillment of each requirement. This fulfillment was based on the arguments derived from the evidence. On the basis of the fulfillment of different user requirements, it was possible to determine whether a certain Xgoal (i.e., a claim in the UC) was fulfilled or not. If most of the user requirements connected to a certain goal were met, then also the Xgoal could be said to have been fulfilled. In addition to this kind of evidence-based reasoning, the UC method also provided data on the usability and UX of the concepts under evaluation. These results supported the design work by providing feedback for future development.

The evaluation results indicate that the evaluated concepts had both positive and negative aspects. The design of the final concept solution should be based on the positive aspects taken from both of the evaluated concepts. According to the results, the experience of safe operation and feeling of presence were not supported with the current version of the system. However, it was difficult to assess the fulfillment of these goals with the developed prototype as the operations were conducted in a virtual world where no human lives were at danger and the presented camera views were not real ones. Despite this fact, there was, however, clear support for the fulfillment of the sense of control Xgoal in the results, for example, because the used joysticks were felt to be robust enough and to control the crane with an appropriate feel of operation. In addition, the possibility to freely decide when to start and stop operating and to easily adjust the speed of operation with the joysticks were felt to be positive features supporting sense of control.

In general, it can also be said that the originally defined main experience vision of 'hands-on remote operation experience' was not yet fulfilled with the current prototype system. In the future development, the requirements that were not met should be taken under careful investigation and answered with sufficient solutions. In this way, also the defined Xgoals could be met better with the final ROS system.

Future Factory

For the future factory case two complementary user research setups with expert control room operators and process control workers were established. In the first evaluation setup, the participants were introduced to the Science Fiction Prototype (SFP) via YouTube videos embedded in a Web questionnaire. The questionnaire included a discussion space that was active for a two-month period. In all, 58 experts participated in the Web survey; 16 of whom were active commentators. The participants were selected from among the customer companies of the project's participating company and they had work experience of process control work up to 41 years. The second evaluation setup included interviews conducted in situ in a municipal power. In addition to seeing the SFP via YouTube videos, the participants were also able to try out speech and gesture control demos. The evaluations included six operators (all male) aged 27–34.

The Web survey consisted of both closed and open-ended questions; the interview setup consisted of a video interview with user analysis and a semi-structured interview. In both groups, the participants assessed six video scenes, one at a time; the main difference between the evaluation setups was that in the Web survey the participants could choose which of the six scenes they wanted to see and comment first. The eight Xgoals were used as the evaluation framework in the quantitative part of the user evaluation. The users were requested to assess if they could identify with the Xgoals, after seeing each video scene, by answering a UX significance questionnaire with a 5-point Likert scale. The users in both research setups answered the same open-ended questions related to the SFP; in addition, they were requested to analyze the new interaction methods and deliver new ideas. As a final part, the participants were allowed to give overall feedback on the presented future control room environment.

The UX significance questionnaire worked very well in assessing the strengths and weaknesses of the future scenarios, as by using it there was a possibility to gain an understanding of how the nominated Xgoals were received by the expert process control workers. The interviews and the free comments in the Web survey complemented the results by reasoning the expected experiences. In addition, the interviews gave feedback to whether the presented concepts were feasible, needed, and valuable; whether the future work environment was conceivable and desired; and whether the Xgoals were desired and valuable.

DISCUSSION

We reported four cases where Xgoals were utilized during the three HCD activities: Investigation, Design, and Evaluation. While user experience was important in each case, only the *Future Factory* case was able to define the Xgoals before the product. In other cases, the product was defined in the design brief (Table 1), although the exact functionality was still open for discussion.

Our projects invested very different amount of resources into the Investigation activities: when the team was familiar with the context and users, they could agree on the Xgoals quickly, as in the *Elevator Control* case. The other extreme was the *Future Factory* case that studied the possible futures and developed Xgoals similarly to future scenarios: both Xgoals and scenarios were evaluated, refined or redefined along the process. The process from Xgoals to the final science fiction prototype was not straightforward, since the Xgoals were changed as the future vision was changed. While Xgoals are meant to help focusing on the key experiences, in *Future Factory* case they actually broadened the focus. It seems that the further ahead the project targets are and the more open the end result is, the more time it seems to take to define the Xgoals. Most cases had either a single high-level experience vision (*ROS* and *Future Factory*) or two main Xgoals (*SmartGUI*). These may be easier to share and to keep in the design team's mind than a long list of Xgoals. If there are several Xgoals, defining a unifying experience vision may help in sharing and memorizing the overall goal.

Once Xgoals are defined, they need to be operationalized in generative design activities. In design, we identified three distinct challenges: 1) finding the appropriate abstraction level of Xgoals, 2) translating Xgoals into appropriate guidance for design, and 3) directing and keeping focus on experience. First, our experiences show that Xgoals work well in creating and maintaining an experience mindset within the design team but it is challenging to find the right abstraction level for Xgoals. High-level Xgoals are not tangible enough to guide design but too specific experience based heuristics, as in the *SmartGUI* case, may hinder ideation. Second, the *Elevator Control* case highlights the difficulty in translating Xgoals, and more generally insights generated from user requirements, into actionable design solutions. In the small development team, the original Xgoals were not formally processed into a specific set of design implications although the means to turn Xgoals into design solutions were discussed within the team during development. Hence, the Xgoals provided only generic guidance for design (e.g., providing remote control opportunities and reducing the feeling of waiting) while leaving a lot of freedom for the designers to realize the concept as practical design solutions. In the *ROS* case, on the other hand, an elaborated process was used to get from Xgoals to the design solutions. The benefit of such a process is that it is possible to trace the design solutions back to the Xgoals for validation purposes. However, the potential downside is that the use of a new process that the designers and developers may be unfamiliar with could add overhead to the design process. More research is needed to study in which conditions Xgoals can be translated to design implications, and how complex the relations can be. Third, the *Future Factory* case showed how Xgoals can form the experiential backbone for the developed product concepts. While the use of Xgoals managed to turn the design focus to how the work will feel instead of how the interaction will feel, it will be interesting to see whether this focus can be maintained through the subsequent product development process.

Regarding evaluation activities, Xgoals were successfully used in planning the evaluation. For example, the interview questions in *SmartGUI* case and additional survey questions in *Elevator Control* case were based on the Xgoals, and the UX significance questionnaire used in the *Future Factory* case helped to evaluate whether the intended experience was realized and how the experience was valued. However, evaluating design outcomes against Xgoals was not as straightforward. Xgoals are only a part of a longer list of design requirements; they are not always in the focus of evaluation. For example, some teams were used to using certain questionnaires that provided only indirect feedback of the Xgoals. Also, testing preliminary prototypes against the Xgoals proved to be difficult. Functional, contextual and aesthetic shortcomings directed the attention to pragmatic rather than emotional aspects of experience.

Originally, we expected Xgoals to act as evaluation criteria for meeting experiential goals, but evaluations of the design concepts sometimes revealed the need to reconsider the Xgoals themselves. This validation of the Xgoals turned out to be a more important and challenging activity than expected when planning this research, and ultimately it proved to be useful to ensure the Xgoals are focused on appropriate aspects of experience.

Exploratory research such as the one reported in this paper can help develop frameworks for future studies. Based on our findings, we have developed initial criteria for assessing the usefulness of experience design tools (such as Xgoals) in real life design (Table 2).

Activity	**Assessment criteria: Does the tool help in…**
Investi-gation	1. Stepping into the users' shoes with empathy 2. Sharing experiential design goals to different stakeholders
Design	3. Creating meaningful experiential concepts 4. Tracing design choices back to the intended experiences
Evalua-tion	5. Defining criteria for evaluation of experiential aspects of the design 6. Evaluating whether a design is moving towards the intended experience(s)
Overall	7. Integrating experience aspects with other design aspects (technology, safety, etc.) 8. Making experience design more systematic and enabling continuous improvement 9. Improving user experience of the design outcome

Table 2. Criteria for assessing the usefulness of experience design tools

Limitations

The studied cases had many external limitations affecting the utilization of Xgoals, and thus the cases are not meant to be model examples. However, we deliberately wanted to study Xgoals in realistic contexts and understand how the context influences Xgoal utilization. As such, the real limitations for this study do not lie in the circumstances but rather in the research setup that did not employ a full separation of concerns: rather than assigning the design tasks entirely to the company staff, researchers were doing more than half of the work. In some cases, this was close to subcontracting, which is an increasingly common context of industrial product development. A highly intriguing future research line would be to study Xgoal utilization in companies where experience-driven design is routine.

CONCLUSIONS

The variety of the design projects and real-life constraints in the industry make it challenging for companies to adopt any given experience-driven design method as a rigorous process. Our research suggests that when introducing experience design methods to industrial product development, the same applies as for usability evaluation methods: we should focus on ingredients and meals rather than recipes [33]. We took Xgoals as an ingredient to be added to the design projects and followed how this affected the process. Unfortunately, we cannot analyze if Xgoals made the outcome better, since it is practically impossible to run a controlled experiment comparing industrial design processes with and without Xgoals. Instead, we reported utilization of the Xgoals during the design process by analyzing four design cases against three activities: Investigation, Design, and Evaluation. We address knowledge transfer from design discipline to design practice and therefore our findings have both scientific and practical implications. Below, we summarize the main benefits and challenges in each typical activity and the related research topics for the future.

Although Xgoals were originally meant to aid focusing the design activities on experience, this research found that Xgoals can also serve Investigation activities by providing a framework for user studies or by crystallizing emphatic understanding about users. The main challenge in the investigation activities was related to Xgoal definition. Recent work, which was not available at the time of this research, has identified possible sources for setting Xgoals [12] and defined the Xgoal Elicitation Process [32]. Examples of possible Xgoals may help designers get inspiration, and indeed, there exist many experience/emotion card sets designed for this purpose [1,19,34, see also[1]]. Future research is still needed for investigating the different formats, level of abstraction, and hierarchy of representing experience visions, Xgoals, and design implications.

Regarding the Design activities, high-level Xgoals helped to create and maintain an experience mindset within the design team. The main challenge was the jump from high-level Xgoals to practical design solutions. For seasoned interaction designers or design researchers this may not be a problem, but for developers who are used to solving specific technical challenges, high-level Xgoals need to be processed into more specific design guidelines. In the *ROS* case, a specific process to get from Xgoals to design solutions was successfully trialed [15], and *SmartGUI* concretized Xgoals as scenarios. Other known solutions include experience patterns [8] and design strategies for commonly used Xgoals [19]. Future research on analyzing the role of Xgoals in design should pay attention to fostering an empathic mindset of the development team, the means of deriving design requirements from Xgoals, and tracing the Xgoals throughout the design process.

Finally, Xgoals were successfully utilized in planning Evaluation activities and there is potential in utilizing Xgoals evaluation criteria. However, our evaluation activities faced several of challenges (see Discussion). The main challenge to be addressed in future research is to find means to evaluate the designs against the Xgoals, and do it as early as possible during the design process.

This research addressed the lack of studies on integrating results of methodology research with real-life product development projects. We created project teams consisting of both researchers and industry professionals, and let the project work follow the structure typical for the team. The main intervention was the introduction of Xgoals to the design process. Documenting this research process will hopefully help others conduct similar studies in the future. Based on this exploratory study, we derived criteria for assessing tools aiming to improve experience design processes in the industry. These criteria will help experience design tool developers and provide a basis for researchers to conduct similar studies in the future.

Our design cases were the initial attempt at introducing Xgoals in companies and thus the results are far from optimal. Nevertheless, the research contributed to two products that have been successfully launched on the market[2,3]. The company stakeholders saw Xgoals as a highly promising technique that helped the teams to focus on user experience, and they are motivated to utilize Xgoals in other projects as well. One reason is that focusing on the experience offers remarkable possibilities for organizations to renew and differentiate [9]. However, if future research is planned on introducing the ‘experience before product’ idea [5, p.63] to the industry, studying product development projects might not be the right focus area. Companies plan their product portfolios, technology roadmaps, and market strategies well before the product development starts, which means experience goals should be introduced to the strategic operations. Company-wide experience goals [24] might be one way to get there.

ACKNOWLEDGEMENTS

We are grateful for the partner companies that made this research possible. This research was carried out as part of FIMECC UXUS program and the first author was also supported by Brand New project, both funded by Tekes.

[1] https://hassenzahl.wordpress.com/experience-design-tools

[2] http://www.konecranes.com/remote-operating-station-ros

[3] https://play.google.com/store/apps/details?id=com.kone.mop

REFERENCES

1. Calvo, R. A., & Peters, D. (2016, May). Designing Technology to Foster Psychological Wellbeing. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 988-991). ACM.
2. Carroll, J.M. (2000). Making Use: Scenario-Based Design of Human-Computer Interactions. The MIT Press.
3. Cooper, A. (2004). *The Inmates are Running the Asylum: Why High Tech Products Drive Us Crazy and how to Restore the Sanity (2nd Edition)*. Pearson Higher Education.
4. Desmet, P. M. A., and H. N. J. Schifferstein (eds). (2011). From Floating Wheelchairs to Mobile Car Parks: A Collection of 35 Experience-driven Design Projects. Den Haag, NL: Eleven Publishers.
5. Hassenzahl, M. (2010). Experience design: Technology for all the right reasons. *Synthesis Lectures on Human-Centered Informatics*, *3*(1), 1-95.
6. Hassenzahl, M. (2013). Experiences before things: A primer for the (yet) unconvinced. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 2059-2068). ACM.
7. Hassenzahl, M., Diefenbach, S., & Göritz, A. (2010). Needs, affect, and interactive products–Facets of user experience. *Interacting with computers*, *22*(5), 353-362.
8. Hassenzahl, M., Eckoldt, K., Diefenbach, S., Laschke, M. & Lenz, E. (2013). Designing moments of meaning and pleasure – Experience Design. International Journal of Design, 7 (3), 21-31.
9. Heikkinen, M., & Määttä, H. (2013, December). Design driven product innovation in enhancing user experience oriented organisational culture in B-to-B organisations. In *Design Management Symposium (TIDMS), 2013 IEEE Tsinghua International* (pp. 127-135). IEEE
10. Hekkert, P., Mostert, M. and Stompff G. (2003). "Dancing with a machine: a case of experience-driven design." *In Proceedings of the 2003 international conference on Designing pleasurable products and interfaces (DPPI '03)*, 114-119. New Your: ACM.
11. ISO 9241-210. Ergonomics of human system interaction - part 210: Human-centred design for interactive systems. International Organization for Standardization, Switzerland, 2010.
12. Kaasinen, E., Roto, V., Hakulinen, J., Heimonen, T., Jokinen, J. P., Karvonen, H., ... & Tokkonen, H. (2015). Defining user experience goals to guide the design of industrial systems. *Behaviour & Information Technology*, *34*(10), 976-991.
13. Karvonen, H., Koskinen, H., Tokkonen, H., Hakulinen, J. (2014). Evaluation of User Experience Goal Fulfillment: Case Remote Operator Station. In Proceedings of the 6th International Conference on Virtual, Augmented and Mixed Reality (accepted), LNCS 8526.
14. Karvonen, H., Koskinen, H., & Haggrén, J. (2012, August). Enhancing the user experience of the crane operator: comparing work demands in two operational settings. In *Proceedings of the 30th European Conference on Cognitive Ergonomics* (pp. 37-44). ACM.
15. Koskinen, H., Karvonen, H., & Tokkonen, H. (2013, August). User experience targets as design drivers: a case study on the development of a remote crane operator station. In *Proceedings of the 31st European Conference on Cognitive Ergonomics* (Article No 25). ACM.
16. Kymäläinen, T., Kaasinen, E., Hakulinen, J., Heimonen, T., Manninen, P. Aikala, M., Paunonen, H., Ruotsalainen J. and Lehtikunnas L. (2017). A creative prototype illustrating the ambient user experience of an intelligent future factory, Journal of Ambient Intelligence and Smart Environments, 1(1), 1-17 (in press).
17. Liinasuo, M., & Norros, L. (2007, November). Usability Case–integrating usability evaluations in design. In *COST294-MAUSE Workshop Downstream Utility: The Good* (p. 11-13).
18. Lu, Y., & Roto, V. (2014, October). Towards meaning change: experience goals driving design space expansion. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (pp. 717-726). ACM.
19. Lu, Y., & Roto, V. (2016). Design for Pride in the Workplace. *Psychology of Well-Being*, *6*(1), 1-18.
20. Lucero, A., & Arrasvuori, J. (2010, September). PLEX Cards: a source of inspiration when designing for playfulness. In *Proceedings of the 3rd International Conference on Fun and Games* (pp. 28-37). ACM.
21. McCarthy, J. & Wright, P. (2004). Technology as Experience. The MIT Press.
22. Norros, L. (2004). Acting under uncertainty – The core-task analysis in ecological study of work. Espoo: VTT Publications.
23. Norros, I. & Norros, L. (2009, September). Human factors in the dependability of IP networks. ENISA Quarterly Review, 5(3), 20-21.
24. Roto, V., Lu, Y., Nieminen, H., & Tutal, E. (2015, April). Designing for User and Brand Experience via Company-wide Experience Goals. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts*

on Human Factors in Computing Systems (pp. 2277-2282). ACM.

25. Rozendaal, M. C. (2010). Investigating Experience Design in Industry. In *Proceedings of the 7th International Design & Emotion Conference* (pp. 4-7).
26. Saariluoma P. (2003). Apperception, content based psychology and design. In U. Lindemann (Ed.), Human behaviour in design (pp. 72-78). Berlin: Springer.
27. Saariluoma, P., & Jokinen, J. P. P. (2014). Emotional dimensions of user experience - a user psychological analysis. International Journal of Human-Computer Interaction, 30(4), 303-320.
28. Salah, D., Paige, R. F., & Cairns, P. (2014). A systematic literature review for agile development processes and user centred design integration. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering* (p. 5). ACM.
29. Savioja, P., & Norros, L. (2013). Systems usability framework for evaluating tools in safety–critical work. *Cognition, technology & work*, *15*(3), 255-275. DOI:10.1007/s10111-012-0224-9.
30. Sanders, E. B. N., & Dandavate, U. (1999, November). Design for experiencing: New tools. In Proceedings of the First International Conference on Design and Emotion, Overbeeke, CJ, Hekkert, P.(Eds.), Delft University of Technology, Delft, The Netherlands (pp. 87-91).
31. Turunen M., Hakulinen J., Melto A., Heimonen T., Laivo T., and Hella J. (2009). SUXES – User Experience Evaluation Method for Spoken and Multimodal Interaction. In *Proc.10th Annual Conference of the International Speech Communication Association*. Interspeech '09. ISCA, 2567–2570.
32. Varsaluoma, J., Väätäjä, H., Kaasinen, E., Karvonen, H., & Lu, Y. (2015, December). The Fuzzy Front End of Experience Design: Eliciting and Communicating Experience Goals. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction* (pp. 324-332). ACM.
33. Woolrych, A., Hornbæk, K., Frøkjær, E., & Cockton, G. (2011). Ingredients and meals rather than recipes: A proposal for research that does not treat usability evaluation methods as indivisible wholes. *International Journal of Human-Computer Interaction*, *27*(10), 940-970.
34. Yoon, J., Desmet, P. M., & Pohlmeyer, A. E. (2013). Embodied typology of positive emotions: the development of a tool to facilitate emotional granularity in design. In *IASDR 2013: Proc. 5th International Congress of International Association of Societies of Design Research, Tokyo, Japan*.

VIII

EVALUATION OF USER EXPERIENCE GOAL FULFILLMENT: CASE REMOTE OPERATOR STATION

by

Hannu Karvonen, Hanna Koskinen, Helena Tokkonen, & Jaakko Hakulinen

In Proceedings of the 2014 International Conference on Virtual, Augmented and Mixed Reality (pp. 366–377)

Evaluation of User Experience Goal Fulfillment: Case Remote Operator Station

Hannu Karvonen[1], Hanna Koskinen[1], Helena Tokkonen[2], and Jaakko Hakulinen[3]

[1] VTT Technical Research Centre of Finland, P.O.Box 1000, FI-02044 VTT, Finland
{hannu.karvonen,hanna.koskinen}@vtt.fi
[2] University of Jyväskylä, P.O. Box 35, FI-40014 University of Jyväskylä, Finland
helena.tokkonen@gmail.com
[3] University of Tampere, Kanslerinrinne 1, FI-33014 University of Tampere, Finland
jaakko.hakulinen@sis.uta.fi

Abstract. In this paper, the results of a user experience (UX) goal evaluation study are reported. The study was carried out as a part of a research and development project of a novel remote operator station (ROS) for container gantry crane operation in port yards. The objectives of the study were both to compare the UXs of two different user interface concepts and to give feedback on how well the UX goals experience of safe operation, sense of control, and feeling of presence are fulfilled with the developed ROS prototype. According to the results, the experience of safe operation and feeling of presence were not supported with the current version of the system. However, there was much better support for the fulfilment of the sense of control UX goal in the results. Methodologically, further work is needed in adapting the utilized Usability Case method to suit UX goal evaluation better.

Keywords: remote operation, user experience, user experience goal, evaluation.

1 Introduction

Setting user experience (UX) goals, which are sometimes also referred to as UX targets, is a recently developed approach for designing products and services for certain kinds of experiences. While traditional usability goals focus on assessing how useful or productive a system is from product perspective, UX goals are concerned with how users experience a product from their own viewpoint [1]. Therefore, UX goals describe what kind of positive experiences the product should evoke in the user [2].

In product development, UX goals define the experiential qualities to which the design process should aim at [2,3]. In our view, the goals should guide experience-driven product development [4] in its different phases. The goals should be defined in the early stages of design and the aim should be that in later product development phases the goals are considered when designing and implementing the solutions of the product. In addition, when evaluating the designed product with users, it should be assessed whether the originally defined UX goals are achieved with it.

R. Shumaker and S. Lackey (Eds.): VAMR 2014, Part II, LNCS 8526, pp. 366–377, 2014.

In the evaluation of UX goals in the case study reported in this paper, we have utilized a case-based reasoning method called Usability Case (UC). For details about the UC method, see for example [5]. In order to test empirically how the method suits the evaluation of UX goals, we used it to conduct an evaluation of UX goals of a remote operator station (ROS) user interface (UI) for container crane operation. Next, the details of the evaluation study case and the utilized UC method are described.

2 The Evaluation Study Case

Our case study was carried out as a part of a research and development project of a novel ROS for container gantry crane operation in port yards. These kinds of remote operation systems exist already in some ports of the world and are used for example for the landside road truck loading zone operation of semi-automated stacking cranes.

Both safety and UX aspects motivated the case study. Firstly, taking safety aspects into account is naturally important in traditional on-the-spot port crane operation as people's lives can be in danger. However, it becomes even more important when operating the crane remotely, because the operator is not physically present in the operation area and for example, visual, auditory, and haptic information from the object environment is mediated through a technical system. Secondly, although UX has traditionally not been in the focus of complex work systems development, it has recently been discussed as a factor to be taken into account in this domain also (e.g., [6]).

Hence, the aim of our project was to explore ways to enhance the UX of the remote crane operators by developing a novel ROS operation concept, which also takes into account the required safety aspects. To achieve this aim, we defined UX goals and user requirements based on an earlier field study by us. The field study (for details, see [7]) was conducted in two international ports and included operator interviews and field observations of their work. The UX goals were created in the beginning of the project and then utilized in guiding the design work throughout the development of the new ROS. In addition, altogether 72 user requirements (when counting both main and sub requirements) were defined and connected to the created UX goals.

The overall UX theme for the new ROS was defined to be 'hands-on remote operation experience'. The four UX goals to realize this theme were chosen after a deliberate process to be 'experience of safe operation', 'sense of control', 'feeling of presence', and 'experience of fluent co-operation'. Details about how these goals were chosen and what they mean in practice regarding the developed system can be found in [2] and [3]. In the evaluation study of the ROS reported in this paper, the experience of fluent co-operation goal could not be included as the functionalities supporting co-operation between different actors in operations were not yet implemented to the ROS prototype and the participants conducted the operations individually.

The main objectives of the conducted evaluations were twofold. Firstly, we wanted to compare the user experience of two optional ROS user interface concepts, which were developed during the project. Secondly, we strived to receive data from the evaluations on how well the UX goals experience of safe operation, sense of control, and feeling of presence are fulfilled with the current ROS prototype system.

2.1 The Study Setting

The evaluations were conducted with a simulator version of the ROS system, which was operated with two industrial joysticks and a tablet computer (see Fig. 1 for a concept illustration). A 32-inch display placed on the operator's desk provided the main operating view, which included virtual reality (VR) camera views and simulated, but realistic operational data (e.g., parameters related to the weight of a container).

Fig. 1. Concept illustration of the ROS system with the four-view setup in the main display

The main display's user interface consisted of camera views and operational data provided by the system. In this display, two different user interface setups were implemented to the virtual prototype: a four-view (see Fig. 1 for a simplified concept illustration version) and a two-view setup. Wireframe versions of the layouts of these two user interface setups for the main operating display can be seen in Fig. 2.

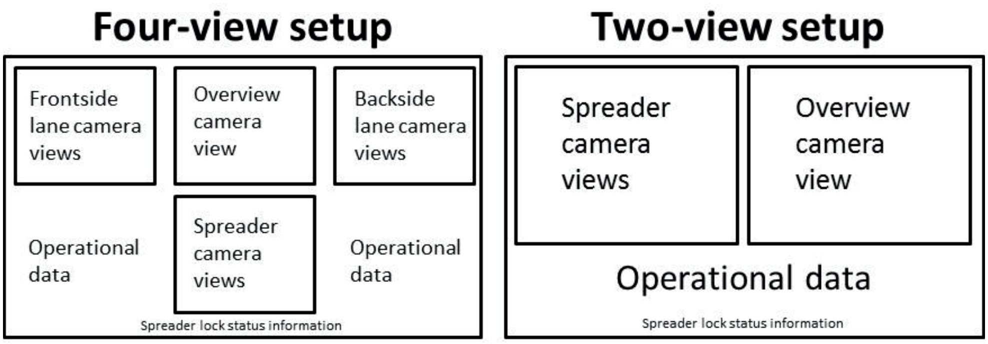


Fig. 2. Wireframe versions of the two alternative main display setups of the concepts

Operation Tasks in Remote Container Crane Operation. Semi-automated gantry cranes in ports are operated manually for example when lifting or lowering containers from and to road trucks, which are visiting the port. These operations happen physically in a specific area called the loading zone. The cranes are operated manually from an ROS after the spreader (device in the cranes used for lifting and lowering the containers) reaches a certain height in the loading zone during the otherwise automated operation. The remote operator utilizes real-time data and loading zone cameras to ensure that the operation goes safely and smoothly.

User Interface of the Four-view Setup. The user interface of the four-view setup (Fig. 1) included four distinct camera views: 1) overview camera view (top-middle), 2) spreader camera view (bottom-middle) that combined pictures of the four cameras attached to the corners of the spreader, 3) frontside lane camera views (top-left), and 4) backside lane camera views (top-right). Both of the lane camera views combined two video feeds from the corners of the truck into one unified view. Three separate camera views could be changed to the overview camera view: an area view (seen in the top-middle view of Fig. 1), a trolley view (a camera shooting downwards from the trolley), and a booth view (a camera showing the truck driver's booth in the loading zone). On the left and right side of the spreader camera view, different types of operational data were displayed.

User Interface of the Two-View Setup. The user interface of the two-view setup (see Fig. 2) consisted of only two, but larger camera views than in the four-view setup: the spreader camera view on the top-left side and the overview camera view on the top-right side. Both of these views could be easily changed to show the relevant camera view at each phase of the task. To the left-side view, also the lane camera views could be chosen. To the right-side view, the aforementioned area, trolley and booth views could be chosen. Under the camera views, there were several crane parameters and different status information displayed in a slightly different order than in the four-view setup.

Control Devices of the Concepts. The joystick functions of the two- and the four-view concepts varied. In the joystick functions of the four-view concept, the left joystick's functions were related to the overview camera (e.g., zoom, pan, and tilt) and for moving the trolley or the gantry. The right joystick was used for special spreader functions such as trim, skew, opening/closing the twist locks (that keep the container attached from its top corners to the spreader), and moving the spreader up- and downwards.

In the two-view concept, the joystick functions were optimized for the operation of the two camera views: the left joystick had controls related to the spreader view (e.g., skew and moving the spreader) and the right joystick to the overview view (e.g., the aforementioned camera operations).

On the tablet, located between the joysticks, there were functions for example for changing the different camera views: in the four-view concept there was only a possibility to change the top-middle overview view while in the two-view concept it was possible to change both the left and right side camera views. In addition, the received task could be canceled during operation or finalized after operation from the tablet.

2.2 Participants

In total, six work-domain experts were recruited as participants for the evaluation study. Three of them had previous experience in remote crane operation. All subjects were familiar with the operation of different traditional container cranes: two of them had over ten years of experience of operating different types of industrial cranes, three of them had 1-5 years of experience, and one of them had 6-10 years of experience.

2.3 Test Methods

In order to evaluate how the originally defined UX goals and user requirements are fulfilled with the evaluated prototype, we used a combination of different methods. During a one evaluation session, the participant was first interviewed about his experience and opinions regarding crane operation. Then, the participant was introduced to the developed prototype system and asked to conduct different operational tasks with the two alternative concepts of the system.

The test tasks included container lifting and landing operations to and from road trucks in varying simulated conditions. The first task was for training purposes and included a very basic pick-up operation; its aim was to learn to use the controls and the simulator after a short introduction to them. To support the joystick operation, the participants received a piece of paper describing the function layouts of the joysticks.

The other operation tasks were more challenging than the first one, and included different disruptive factors, such as for example strong wind, nearly similarly colored container chassis as the container to be landed, other containers in the surrounding lanes, a truck driver walking in the loading zone, and a locked chassis pin. These tasks were conducted with both of the concepts, but not in the same order.

The two different concepts (the four- and the two-view concepts) were tested one at a time. The order of starting with the two-view or with the four-view concept was counterbalanced. Therefore, every other user started first with the two-view concept and every other with the four-view concept.

A short semi-structured interview was conducted after each operational task. In addition, two separate questionnaires were used to gather information: the first one about the user experience and the second one about the systems usability [8] of the concepts. The UX questionnaire consisted of twelve user experience statements that were scaled with a 5-point Likert scale. The UX questionnaire was filled in when the test participants had completed all the tasks with either of the concepts. Ultimately, the UX questionnaire was filled in regarding both of the concepts.

In the end of the test session, some general questions related to the concepts were asked before the participants were requested to select the concept that they preferred and that in their opinion had a better user experience. Finally, a customized systems usability (see e.g., [8]) questionnaire was filled in for the selected concept. The systems usability questionnaire included thirty-one statements that were also scaled with a five-point Likert scale. Due to space restrictions, neither of the abovementioned questionnaires is presented in detail in this paper.

The test leader asked the participants to think-aloud [9], if possible, while executing the operation tasks. The think-aloud protocol was utilized to make it easier for the researchers to understand how the participants actually experience the developed concept solutions. The evaluation sessions were video recorded to aid data analysis.

2.4 Analysis

The ultimate aim of the evaluations was to assess whether the chosen UX goals were fulfilled with the VR prototype version of the system. To do this, we utilized the Usability Case method, because we wanted to explore the suitability of the method for this kind of research. UC provides a systematic reasoning tool and reference for gathering data of the technology under design and for testing its usability in the targeted work [10]. The method applies a case-based reasoning approach, similar to the Safety Case method [11]. Throughout the development process, the UC method creates an accumulated and documented body of evidence that provides convincing and valid arguments of the degree of usability of a system for a given application in a given environment [5]. The main elements of UC are: 1) *claim(s)* (nine main claims of systems usability [8], of which three are related particularly to UX) that describe an attribute of the system in terms of usability (e.g., "User interface X is appropriate for task Y"), 2) *subclaim(s)* describing a subattribute of the system that contributes to the main claim (e.g., "X should work efficiently), 3) *argument(s)* that provides ground for analyzing the (sub)claims (e.g., "It is possible to quickly reach the desired result with X"), and 4) *evidence*, which is the data that provides either positive or negative proof for the argument(s) (e.g., task completion times in usability tests) [5].

In line with the UC method, the data gathered from our studies was carefully analyzed regarding each defined user requirement (i.e., a subclaim in UC) on whether positive or negative cumulative evidence was found about the fulfillment of each requirement. This fulfilment was based on the arguments derived from the evidence. On the basis of the fulfilment of different user requirements, it was possible to determine whether a certain UX goal (i.e., a claim in UC) is fulfilled or not. If most of the user requirements connected to a certain goal were met, then also the UX goal could be said to have been fulfilled. In addition to this kind of evidence-based reasoning, the UC method also provided us with data on the usability and UX of the concepts under evaluation. These results support the design work by providing feedback for future development.

3 Results

The results of our studies are presented in the following order: First, we present general user experience and usability related results that affected the chosen UX goals regarding both the four- and the two-view concepts. Then, we discuss which of the concepts the participants chose in the end of the test sessions and why. Finally, we discuss whether the defined UX goals were fulfilled and make hypotheses on what were the underlying reasons for these results.

3.1 Notes on General UX and Usability of the Concepts

Four-view Concept. In general, the participants felt that the information provided by the main display's four-view setup was appropriate and understandable: for example, the participants commented that the amount of presented camera views at once was suitable and most of the necessary information was available for the basic crane operations. However, some of the participants felt that for example information about possible fault conditions concerning the crane were missing from the current solution.

While performing the test tasks, the participants utilized most frequently the area and the spreader camera views. The spreader camera view was experienced to be useful especially at the beginning of a lifting task. However, when the spreader approached the container, it became more difficult to understand the position of the spreader in relation to the container in detail. In addition, the participants thought that the provided lane camera views did not support the beginning phase of the container pick-up operations, because the participants could not clearly comprehend the orientation of the provided views until the spreader was seen moving in the views.

Regarding the joystick functions in the four-view concept, the placement of some functions was not reported to support the operations very well. For example, the positions of the skew and trim functions were not optimal, since participants made frequent mistakes with them and reported to get emotionally frustrated with them. In addition, the position of the zoom was proposed to be placed together with the steering functions, i.e., to be designed into the right-hand joystick.

The overall nature of the results of the UX questionnaire statements related to sense of control with the four-view concept was positive. The participants felt that they were able to start, conduct, and stop the operations at their own pace. In addition, according to the interviews, the provided joysticks were experienced to be suitable for the remote operation of cranes and the feel of the joysticks to be robust enough. Also, the crane's reactions to the joystick movements was experienced to be appropriate.

Nevertheless, the UX goal feeling of presence did not get as much supportive results as sense of control. This was mostly due to the problems identified with the solutions aimed to fulfil requirements concerning the operation view. For example, the four-view concept's camera views were experienced to be too small for the participants to easily see everything that was necessary. In addition, combining two camera views together (in the lane cameras) received negative evidence; the participants had difficulties to orientate themselves with the combined camera views and perceive to which direction each of the cameras was shooting at.

The experience of safe operation with the four-view setup was reported to be negatively affected by the presentation layout of the operational parameters. For example, the grouping of the information was not experienced to be in line with a typical task flow of one operation.

Two-view Concept. The two-view setup in the main display was generally experienced to be clearer than the four-view concept according to the participants' thinking-out-loud comments and interviews. For example, the camera views were found to be big enough to spot relevant things from the object environment. Especially the area view was utilized a lot during the operations, because it offered a possibility to see better the spreader in relation to the container.

With the two-view concept the users felt that all the needed operational information was available and in a logical order (i.e., in line with a typical task flow of one operation). The participants for example mentioned that it was possible to perceive easily the status of the operation with one glance from this information.

The UX questionnaire results concerning statements related to sense of control with the two-view concept were positive, mostly due to the same reasons as they were with the four-view concept. In addition, these results showed that the participants felt that they were able to concentrate on a sufficient level on performing their operations with the two-view concept.

However, the UX goal feeling of presence received somewhat negative results from the tests. For example, the participants had difficulties to perceive the operation view provided through the different combined camera views. As with the four-view setup, especially the views of the combined camera views of spreader and lane cameras were experienced to be hard to understand what is seen from them. In addition, the camera views were not reported to support the comprehension of depth and different distances between objects in the loading zone very well.

Furthermore, the results regarding requirements connected to the provided camera views were fairly negative. Some of the participants commented that due to the placement of the camera views they were not able to see critical objects related to the task at hand through the camera views in the outmost truck lanes; for example, it was not possible to see easily all corners of the container and the truck's position. These results had a significant effect to the experience of safe operation UX goal.

3.2 Concept Selection

When asked at the end of the test session that which of the two concepts the participant preferred, four of the participants selected the two-view concept and two of them chose the four-view concept. Based on the participants' experience, the two-view concept was easier to understand: it was reported to be effortless to observe the loading zone through the big camera views and the provided operational information was said to be placed in a logical order. However, according to the participants, some of the joystick functionalities were placed better in the four-view concept than in the two-view concept.

In general, it can also be said that the results of the systems usability questionnaire were fairly positive regarding the both concepts. These results were further utilized in the analysis of fulfillment of the defined user requirements and UX goals described in the next section.

3.3 Fulfilment of User Requirements and UX Goals

Most of the user requirements were not fulfilled on a comprehensive level with neither the four- nor the two-view setups of the current prototype system. Especially the evidence related to the user requirements that were connected to the UX goals experience of safe operation and feeling of presence was mostly negative. Therefore, it can be said that these two goals were not fulfilled with the current versions of the ROS's two- and four-view concepts.

The experience of safe operation was affected for example by the fact that the participants were not able to form a clear picture of the situation in the loading zone when handling the container in the outmost truck lanes. Therefore, they needed to manually adjust the cameras a lot in order to gain a better view to the position of the truck and corners of the container. In addition to the aforementioned factors, the overview camera was not experienced to be sharp enough (when zoomed in) for the participants to be able to see whether the truck's chassis' pins are locked or unlocked when starting a lifting operation. An obvious danger to safety from this problem is that if the pins are locked when starting a container lifting operation, also the truck will be lifted to the air with the container.

The feeling of presence UX goal was negatively affected for example by the fact that some of the camera views (e.g., lane cameras) were difficult for the participants to understand and orientate themselves into. Furthermore, understanding distances between different objects in the loading zone was not experienced to be sufficient with the current camera views. In addition, some of the default zooming levels of the cameras were not very optimal for the conducted task in question and the participants had to do a lot of manual zooming. In Fig. 3, we provide an example of the used Usability Case-based reasoning regarding negative evidence for one requirement connected to the UX goal feeling of presence.

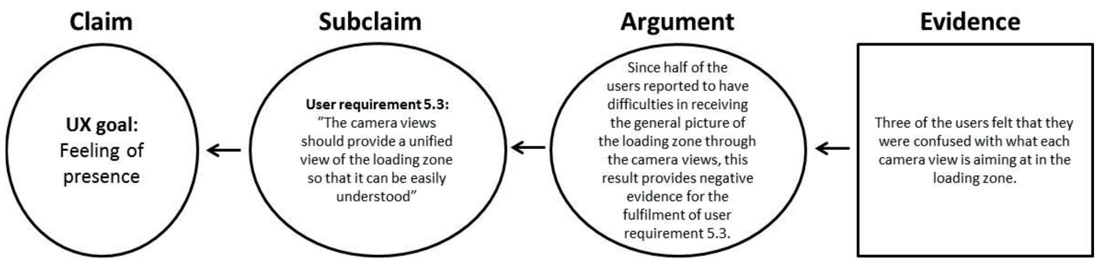


Fig. 3. Example of Usability Case based reasoning in our analysis.

The example of evidence in Fig. 3 was negative comments from three different participants while conducting the tasks with the ROS. In general, other than verbal evidence (the thinking-out-loud comments or the interview answers) provided by the participants were for example the results of the (UX and systems usability) questionnaires and task performance indicators. All this data was considered when creating the final Usability Case, which cannot be described here entirely due to its large size.

Regarding the sense of control UX goal, there was clear positive evidence in the end results from both of the concepts. For example, the utilized joysticks were felt to be robust enough and to control the crane with an appropriate feel of operation. Overall, the participants felt that they were able to master the crane's operations and concentrate on the task at hand. In addition, the possibilities to freely decide when to start and stop operating and to easily adjust the speed of operation with the joysticks were felt to be positive features. Therefore, it can be said that sense of control was achieved with both of the evaluated concepts.

4 Discussion

The results indicate that the evaluated concepts had both positive and negative aspects. The design of the final concept solution should be based on the positive aspects taken from both of the evaluated concepts. From the two-view concept, especially the placement of the operational data and size of the camera views should be adopted to the final concept. From the four-view concept, for example the layout of the joystick functions regarding the basic crane movements should be utilized.

In general, the results confirmed that providing real-time camera feeds for this kind of remote operation is essential. Visual validation of the situation in the object environment allows taking into consideration possible extra variables affecting the operation, such as weather conditions or debris on top of the container to be lifted up. Therefore, good quality camera views could support the experience of safe operation and feeling of presence goals with the final system.

The ecological validity of the prototype system also needs to be discussed as it may have had an effect to the UX goals. First, the fact that the operations with the system were not happening in reality, had an obvious effect on the participants' user experience and attitude towards the operations; if for example the people seen in the object environment would have been real human beings instead of virtual ones, the participants could have been more cautious with the operations. This fact had an obvious effect especially to the experience of safe operation UX goal.

Second, the virtual camera views cannot of course correspond to real camera views from the object environment. This had an obvious effect on the feeling of presence UX goal. However, it must be noted that some of the test participants thought that the virtual simulator was near equal to a real remote crane operation system, since the provided virtual camera views were implemented with such a good resolution. The simulator was also reported to provide a relatively precise feel of the operation, but did not for example have as much swaying of the container as it would have in real operations.

Third, the fact that in real life there are truck drivers with whom the operators communicate through the phone in case of problems affected the ecological validity of the conducted tasks. In addition, the participants conducted the tasks individually in a small room, which is not the case in real remote crane operation work. Therefore, as in real conditions the work is actually much more social than in our evaluation study, this had an obvious effect on the validity of the results of the studies.

5 Conclusions

The conducted study did not give an exact answer to the question, which one of the concepts should be selected for future development. Both concepts had positive factors that should be taken into account when designing the final system.

Different camera views provided essential information from the operating area. A decision concerning the amount of cameras in the loading zone and the camera views provided in the ROS needs to be made for the final concept to support safe crane operation. Another important factor is the size of the camera views in the main display. The two-view setup was experienced to have large enough views for the operation. A balance between the amount and size of the views presented in the user interface needs to be found. If the display space of a one monitor does not allow to present big enough camera views, then the possibility of two monitors needs to be considered.

To some extent, it was possible to evaluate the user experience of remotely operated crane operations with our virtual simulator even though the camera views were not real. However, the user experience of the system was not the same as if it was when operating in a real work environment. For example, the sounds, tones, or noises from the operating environment were not in the focus of the concept development or this evaluation study. In the final system's development, careful attention should be paid to the auditory information provided by the system from the object environment.

In general, as most of the user requirements related to the UX goals feeling of presence and experience of safe operation were not supported by the evidence from the evaluation studies, it can also be said that the originally defined main UX theme of 'hands-on remote operation experience' was not yet fulfilled with the current prototype system. In the future development, the requirements that were not met should be taken under careful investigation and answered with sufficient solutions. In this way, also the defined UX goals could be met better with the final system.

Nevertheless, the evidence from our study results supported the fulfillment of the UX goal sense of control for both of the concepts. Especially the feeling of the joystick operation and reactions of the crane were experienced to be appropriate and realistic. Support for aiming the spreader and the container to the correct position could enhance the sense of control even more in the future versions of the UI.

In the future development of the ROS, special attention should also be paid to the experience of fluent co-operation UX goal and different aspects related to it (e.g., the interaction between the co-workers and the truck drivers) as in the present study it was not possible to address this goal appropriately. Therefore, future studies with the system should include for example several test participants operating simultaneously with the system in order for the operational setting to be more realistic. To increase the ecological validity of the results, a more comprehensive study with a wider range of data inquiry methods could be carried out in a real control room setting with actual operators. This kind of a study could be conducted by adding some features of the proposed concept to the current, already implemented ROS solutions at some port and then evaluating whether the new features are useful and make the work more pleasant.

Methodologically, this paper has contributed to the discussion on how UX goals can be evaluated. According to the results, although the evaluated concepts were still in quite early stages of their design, the Usability Case method seemed to suit to this kind of UX goal evaluation with some modifications. Firstly, further work is needed especially on linking the arguments regarding the user requirements to the detailed design implications (for details see e.g., [3]) of the UX goals. Secondly, a scoring method for the evidence provided by study data should be implemented to the UC method in general, so that more emphasis could be placed on the data concerning the most critical parts of the evaluated product. Finally, it should be experimented whether other than the utilized data gathering methods could provide relevant data in constructing the Usability Case and studied how the method supports also later phases (than just the early-stage evaluation) of UX goal driven product development.

Acknowledgements. This research was carried out as part of the Finnish Metals and Engineering Competence Cluster (FIMECC)'s UXUS program. We would like to thank the participants of the evaluation sessions and our partners for the possibility to organize the evaluations.

References

1. Rogers, Y., Sharp, H., Preece, J.: Interaction Design: Beyond Human-Computer Interaction. John Wiley & Sons, Chichester (2011)
2. Karvonen, H., Koskinen, H., Haggrén, J.: Defining User Experience Goals for Future Concepts. A Case Study. In: Proc. NordiCHI2012 UX Goals Workshop, pp. 14–19 (2012)
3. Koskinen, H., Karvonen, H., Tokkonen, H.: User Experience Targets as Design Drivers: A Case Study on the Development of a Remote Crane Operation Station. In: Proc. ECCE 2013, article no. 25 (2013)
4. Hassenzahl, M.: Experience Design – Technology for All the Right Reasons. Morgan & Claypool (2010)
5. Liinasuo, M., Norros, L.: Usability Case - Integrating Usability Evaluations in Design. In: COST294-MAUSE Workshop, pp. 11–13 (2007)
6. Savioja, P., Liinasuo, M., Koskinen, H.: User experience: Does it matter in complex systems? Cognition, Technology & Work (2013) (online first)
7. Karvonen, H., Koskinen, H., Haggrén, J.: Enhancing the User Experience of the Crane Operator: Comparing Work Demands in Two Operational Settings. In: Proc. ECCE 2012, pp. 37–44 (2012)
8. Savioja, P., Norros, L.: Systems Usability Framework for Evaluating Tools in Safety-Critical Work. Cognition, Technology and Work 15(3), 1–21 (2013)
9. Bainbridge, L., Sanderson, P.: Verbal Protocol Analysis. In: Wilson, J., Corlett, E.N. (eds.) Evaluation of Human Work: A Practical Ergonomics Methodology, pp. 159–184. Taylor & Francis (1995)
10. Norros, L., Liinasuo, M., Savioja, P., Aaltonen, I.: Cope Technology enabled capacity for first responder. COPE project deliverable D2.3 (2010)
11. Bishop, P., Bloomfield, R.: A Methodology for Safety Case Development. In: Redmill, F., Anderson, T. (eds.) Industrial Perspectives of Safety-Critical Systems, pp. 194–203. Springer, London (1998)

IX

CONTEXT-SENSITIVE DISTRACTION WARNINGS–EFFECTS ON DRIVERS' VISUAL BEHAVIOR AND ACCEPTANCE

by

Tuomo Kujala, Hannu Karvonen, & Jakke Mäkelä

International Journal of Human-Computer Studies vol 90, 39–52

https://doi.org/10.1016/j.ijhcs.2016.03.003

CONTEXT-SENSITIVE DISTRACTION WARNINGS –

EFFECTS ON DRIVERS' VISUAL BEHAVIOR AND ACCEPTANCE

Tuomo Kujala*, Hannu Karvonen¶, Jakke Mäkelä*

*University of Jyväskylä

Department of Computer Science and Information Systems

P.O. Box 35

FI-40014 University of Jyväskylä

Finland

Corresponding author: Tuomo Kujala

tuomo.kujala@jyu.fi

jakke.makela@gmail.com

tel. +358 40 024 7392, fax +358 14 260 4400

¶VTT Technical Research Centre of Finland Ltd

P.O. Box 1000

FI-02044 VTT

Finland

hannu.karvonen@vtt.fi

Keywords: driver distraction; smart phone; warning system; situation awareness; acceptance; trust

Abstract

In this study, we investigated the effects of context-sensitive distraction warnings on drivers' in-car glance behaviors and acceptance. The studied prototype warning application functions on a smart phone. The novelty of the application is its proactive and context-sensitive approach to the adjustment of warning thresholds according to the estimated visual demands of the driving situation ahead. In our study, novice and experienced drivers conducted in-car tasks with a smart phone on a test track with and without the warnings. The application gave a warning if the driver's gaze was recognized to remain on the smart phone over a situation-specific threshold time, or if the driver was approaching a high-demand part of the track (an intersection or a tight curve). Glance metrics indicated a significant increasing effect of the warnings on glance time on road while multitasking. The effect varied between 5 to 30% increase depending on the in-car task. A text message reading task was the most visually demanding activity and indicated the greatest effect of the warnings on glance time on road. Driving experience did not have an effect on the efficiency of the warnings. The proposed gaze tracking with current smart phone technology proved to be highly unreliable in varying lighting conditions. However, the findings suggest that location-based proactive distraction warnings of high-demanding driving situations ahead could help all drivers in overcoming the inability to evaluate situational demands while interacting with complex in-car tasks and to place more attention on the road. Furthermore, survey results indicate that it is possible to achieve high levels of trust, perceived usefulness, and acceptance with these kinds of context-sensitive distraction warnings for drivers.

1 INTRODUCTION

Modern smart phones offer car drivers a lot of useful services on the road such as navigation, entertainment, communication, and information on nearby points-of-interest. However, a concern has been addressed lately on the increasing smart phone usage while driving and the related inattention towards the traffic environment (Fitch, Soccolich, Guo, F., et al., 2013; Klauer, Dingus, Neale, Sudweeks, and Ramsey, 2006).

From earlier research it is known that driver inattention is a major cause of safety-critical incidents in traffic. In a naturalistic driving study with one hundred car drivers (Klauer et al., 2006), it was concluded that almost 80 percent of all crashes and 65 percent of all near-crash situations involved visual inattention, i.e., the driver's eyes weren't on the road the moment before or at the moment of the incident.

As a cause of visual inattention by secondary activities in these safety-critical events, the use of a mobile device (mainly mobile phone) was by far the leading factor by at least 30% of the cases (Klauer et al., 2006). Another naturalistic driving study on the topic by Fitch et al. (2013) indicated that drivers engaging in visually complex tasks with their smart phones have a three-time higher safety-critical incident risk compared to drivers who pay attention to the road ahead.

Unfortunately, the most obvious solution to the problem, legislative measures, does not seem to work. For instance, in Finland a recent poll by the Finnish Road Safety Council revealed that over 30 percent of drivers admit texting while driving, despite of the fees on hand-held device usage and distracting in-car activities while driving (Jääskeläinen and Pöysti, 2014). This means that there is an urgent need for other, more effective means to mitigate the negative effects of driver distraction by mobile devices. Other possible approaches are, for instance, driver education and technological counter-measures. In order to provide efficient counter-measures, the priority should be on means that are widely accepted by the drivers (Donmez, Boyle, and Lee, 2007). In this paper, we study the efficiency and acceptability of context-sensitive distraction warnings that could serve this purpose.

2 DISTRACTION ALGORITHMS AND DRIVER ACCEPTANCE

Due to the increasing significance of driver distraction to traffic safety, a number of distraction detection algorithms and distraction warning systems are currently under development by car manufacturers (NHTSA, 2013a; Lee, Moeckli, Brown et al., 2013). These warning systems operate on the basis of distraction detection algorithms, i.e., algorithms that are meant to detect when the driver is distracted. However, there are basic conceptual difficulties in defining and operationalizing accurately what is distracted (inattentive) driving (Regan, Hallett, and Gordon, 2011). This places great challenges for the sensitivity and reliability of the algorithms in detecting distracted driving, and consequently, to drivers' acceptance of the distraction warnings.

Liang, Lee, and Yekhshatyan (2012) studied 24 different possible algorithms that could be used for detecting distraction and evaluated their ability to predict crash risk based on behavioral data collected in the 100 car study by Klauer et al. (2006). They concluded that the most sensitive indicator for crash risk seemed to be algorithms that measure instantaneous changes in off-road glance duration, that is, individual glance durations seem to matter. 1.5th power of glance duration, glance history, or glance location, did not significantly improve the sensitivity.

Even if the algorithms are highly valuable for indicating the general statistical link between off-road glance durations and crash risk, environmental and external situational factors (e.g., driving speed, road curvature and road type) were missing in all of the evaluated 24 algorithms (Liang et al., 2012). That is, one can argue that the severity of an off-road glance duration should be in a relationship with the visual demands of the driving situation, as suggested by the naturalistic driving study of Tivesten and Dozza (2014) as well as the 100 car study report by Klauer et al. (2006). Taking into account the situational visual demands of the driving task could further improve the sensitivity of the single glance algorithms.

The existing and proposed distraction warning systems and detection algorithms do not utilize context and driver data to the extent that could be possible with modern technology. Instead, the algorithms focus only on off-road glance durations and the direction of gaze (NHTSA, 2013a). Context-sensitivity of distraction warning systems could decrease substantially the high levels of false alarms experienced with the current systems (NHTSA, 2013a). In addition, context-sensitivity could improve the visibility of the system behavior by providing the driver a possibility to better associate the criticality of the warnings to the observable demands in the driving environment (e.g., an intersection ahead). All these factors should increase driver acceptance of these systems and make the systems more reliable. In addition, positive learning effects could be expected if the driver learns to associate the warnings to certain driving environments or situations observed ahead.

Like other available technical solutions to mitigate the negative effects of driver distraction, such as braking and lane-keeping assistants, most of the distraction warning systems today are reactive, that is, the systems react to observed distraction or its negative effects by counter-measures (e.g., Wege and Victor, 2014; You, Montes-de-Oca, Bao1 et al., 2012). This means often already a degraded driving performance.

Other, somewhat context-sensitive counter-measures act as workload managers, limiting the access of drivers to certain in-car services when the situational demands are considered to reach a certain level of high demand (Green, 2004). These kinds of forced solutions are rarely well accepted by the drivers. In addition, the high workload conditions are often recognized based on the high levels of activity by the driver (e.g., steering frequency, Green, 2004; Broström, Engström, Agnvall, and Markkula, 2006), whereas lack of sufficient attention on the driving task manifests often as low levels of activity compared to what the situational driving task demands would require (Regan et al., 2011).

The ideas about drivers themselves acting as dynamic workload managers and driver assistant systems for this purpose are relatively new (Donmez, Boyle, and Lee, 2008). A basic requirement for this kind of tactical behavior is that the driver is capable to evaluate the dynamic demands of each driving situation ahead. In-car tasks undermine this ability because it has been shown that drivers can have a low level of awareness of their own performance as well as the elements in the road environment while multitasking (e.g., Schömig and Metz, 2013; Young, Salmon, and Cornelissen, 2013; Horrey, Lesch, and Garabet, 2009). For instance, Young and Salmon (2012) have suggested that high levels of cognitive workload due to in-car task demands can have a negative effect on driver's situation awareness of the road environment, which could at least partially explain this inability. The study by Lee, Lee, and Boyle (2007) indicated that brief glances off road together with cognitive load are additive in their effects on drivers to miss safety-critical events in the driving environment.

In addition, even if the drivers would be aware of the situational driving demands, the most popular survival strategy in multitasking while driving seems to be "ASAP"; the in-car task is completed as soon as possible without considering the situational driving demands. Horrey and Lesch (2009) showed that although drivers seemed to be aware of the demands of the driving situation in their experiment, the drivers did not tend to postpone the presented secondary tasks even if they were given the chance. Based on the findings, the authors suggested that training drivers on tactical decisions and planning of timing in multitasking is worth considering. The effects of this type of training of tactical and strategic skills has been tested by Horrey et al. (2009), giving promising results. Another possibility is to provide real-time feedback for the drivers (Donmez, Boyle, and Lee, 2007), or both real-time and retrospective feedback (Roberts, Ghazizadeh, and Lee, 2012) on distracted behaviors. The study by Donmez, Boyle, and Lee (2010) indicated the positive effects of combined real-time and retrospective feedback on distracted driving behaviors among young high-risk drivers, in particular. Roberts, Ghazizadeh, and Lee (2012) suggested that systems providing immediate feedback on distracted behaviors are experienced in general as less pleasant and less easy to use than retrospective feedback systems. However, the specific implementation of the warnings can be argued to have a significant effect on the acceptability of the real-time warnings.

Instead of mere feedback, one possibility is to give the drivers proactive suggestions to postpone in-car tasks if the driving situation ahead is recognized as high demanding. A proactive and context-sensitive distraction warning system that would adjust warning thresholds according to the expected visual demands of the driving situation ahead and indicate these in real-time for the driver could in principle answer the issues raised by earlier research. In this paper, we study one possible implementation of such a prototype system called VisGuard ("Vision Guard", Kujala, 2013).

3 VISGUARD: PROTOTYPE FEATURES

In order to study the effects of context-sensitive distraction warnings on the drivers' visual behaviors and driver acceptance, we developed an Android-based mobile application called VisGuard (Kujala, 2013). The VisGuard prototype application displays the warnings on the smart phone that the driver is using while driving (see Figure 1). The application is intended to work proactively; warning the driver of the usage of the phone already before the driver enters a visually highly demanding situation.

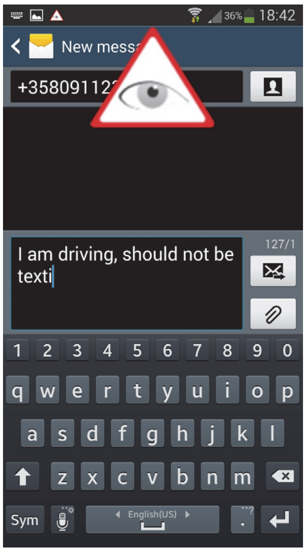


Figure 1. The warning icon on a smart phone screen.

The VisGuard prototype works as a background process in an Android smart phone. It constantly monitors the GPS, and uses a special-purpose map to estimate the visual demand level of driving at a given location on the road. The prototype was built using open source software and open data for the main tasks: Open Street Map for the visual demand map and Open CV software for gaze recognition. Near Field Communication (NFC) with a Radio-Frequency Identification (RFID) tag can be used to automatically activate the software once the driver places the device into a dashboard holder. For hand-held use, Android's activity recognition API is used to detect when the device is being used in a car. The software has been verified on Samsung Galaxy S3 and Note 2, and may work also on other high-end Android smart phones running Android 4.2 or later.

The visual demand algorithm behind the calculation of the adaptive situational warning time is based on a driving simulator study with visual occlusion and 97 drivers by Kujala, Mäkelä, Kotilainen, and Tokkonen (2016). The data allowed us to identify visually high-demanding driving scenarios in which visual distraction would be potentially particularly dangerous. It also allowed us to adjust the warning thresholds for in-car glance durations by situational and driver-specific factors.

The visual demand level is determined based on the experience level of the driver, the proximity of intersections, junctions and pedestrian crossings ahead, the winding of the road ahead, as well as the speed of the car. In addition, other forms of data, such as headway distance, traffic, weather, visibility, and device input data, may be be utilized in future implementations of the application.

The warning threshold is expressed here as a *Warning Time* (WT): a situation-specific threshold for a single in-car glance duration above which off-road glances are considered as risky. This varies from 0 to 2.0 seconds and corresponds to the time it takes to travel with the driver-selected speed the distance the 85^{th} percentile of drivers preferred to travel occluded (i.e., blinded) when fully concentrating on driving in the study of Kujala et al. (2016). The use of the 85^{th} percentile is based on a common standard in traffic engineering (TRB, 2003), assuming that 15 percent of drivers represent risky or unacceptable behaviors in traffic. The curvature of the road ahead is taken into account in our algorithm as OD = 85^{th} percentile OD - 9.1*W, where OD is the occlusion distance in meters and W Winding as meter/meter, a measure of how much the driver needs to turn the steering wheel while driving along the road, based on Kujala et al. (2016). The areas with 85^{th} percentile ODs of less than 5 meters (e.g., in intersections or tight curves) are considered to be visually high-demanding, and the WT is set to zero. When VisGuard determines that the threshold is exceeded or that the WT=0 (a highly demanding area ahead), it shows a warning icon (Figure 1) on the screen, which tells the driver to focus on the driving scene. For the areas of WT=0, the borders of the area are calculated based on vehicle speed and the corresponding stopping distance to the middle point of the area.

The application monitors the face of the driver in real-time, and aims to identify when the driver is looking at the phone. The gaze recognition is technically based on the phone's front camera, Open CV face recognition software, and the angle of the driver's face from the camera. For optimal gaze recognition for our experiment, the smart phone was placed on a dashboard holder so that the driver's face was not recognized when the driver was looking ahead on the road. It is known that people prefer to move their heads, instead of eyes only, if the target is more than 30 degrees away from the line of sight (Flannagan and Sivak, 1993). Therefore, a placement of the holder with at least a 30 degree angle to the driver was preferred for the testing purposes (as in Figure 2). The negative side effect of this placement is that it will recognize gaze to the device whenever the driver is looking at the right-hand field of view, but it was assumed that the driver ignores what is on the display when looking elsewhere in the driving environment. No sound or vibration was used by default, because of the possibility of false alarms and the consequent possibility to capture the driver's attention unnecessarily in these situations.

Figure 2. The phone holder installed on the dashboard air vent.

Grey circle symbol on the display indicated for the driver that gaze at the device is not recognized and the fulfillment level of the circle indicated the WT for the current situation (see Figure 3). This was intended to provide the driver a chance to check the WT with a very brief glance (less than 500 ms) before the gaze tracking catches the gaze in order to support the assessment of the expected visual demands of the situation ahead. Empty circle indicated a 2.0 s warning threshold and a half circle a 1.0 s threshold. VisGuard started a countdown as soon as a glance to the phone is identified. This was indicated by changing the color of the circle symbol to orange and showing a counter inside the circle symbol (the circle filling up, see Figure 3). In controlled laboratory settings the mean delay in gaze recognition for Samsung Galaxy S3 was near 500 ms (RMSE=521 ms, N=49 glances) and the delay was taken into account in the countdown. The orange symbol itself was intended to serve as a reminder of a potentially dangerous activity and to remind the driver to look back at the road before the orange circle was full. When a risky glance duration was observed by the application (i.e., the circle had filled up), the warning icon was displayed (see Figure 1), intended to signal the driver to pay attention to the road immediately. In situations with very high visual demands, such as approaching an intersection or a tight curve, the warning threshold was set to zero. In these situations, the warning icon was displayed immediately based on the GPS and map information, regardless of the driver looking at the phone or not, in order to minimize gaze time on the screen.

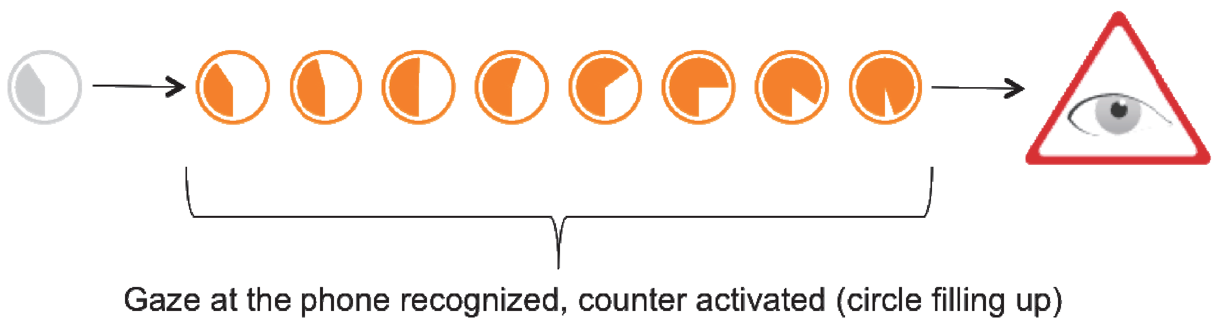


Figure 3. The user interface icons: the circle symbols and the warning icon, and their relative sizes.

The application does not restrict the usage of the smart phone in any way. The idea is to leave the final responsibility of safe driving to the driver, but the application is intended to help the driver in estimating the risks and demands of her or his behavior. The application is not yet available for the public.

4 METHOD

In order to test the effects of the VisGuard prototype on drivers' in-car glancing behaviors and to study how the drivers experience the application, we organized a test track experiment and a survey with 31 participants.

4.1 Design and Hypotheses

The mixed-model experimental design of 2 x 3 x 2 focused on the effects of the following independent variables; Warnings: Control/Warnings (within); In-Car Task Type: Calculator/Navigation/Text message (within); and Driving Experience: Novices/Experienced (between). In particular, we looked at in-car glance durations, their distributions, and the ratio of glance time on road of the total task duration.

The dependent variables were glance time on road (%), percentage of over-2-second in-car glances, and median in-car glance duration. In particular, the experiment tested whether the VisGuard warnings affect these safety-relevant visual behaviors. Glance time on road measures how much visual attention the driver devotes to the driving environment during an in-car task. Low percentages suggest that in-car tasks are performed as fast as possible, without regard for the situational driving conditions. Previously, for instance, Volvo has utilized this metric in testing their distraction warning systems (Wege and Victor, 2014). The ratio of in-car glances over 2 seconds to all in-car glances in percentage was chosen as another metric, because it is a standard verification criterion in test guidelines for in-vehicle electronic devices (NHTSA, 2013b) and it is known to have an association to elevated level of a safety-critical incident risk in real traffic (Liang et al., 2012). In a similar fashion, if the median in-car glance duration is high, it suggests a high number of long safety-critical in-car glances. According to Wierwille's (1993) general visual sampling model, drivers tend to keep off-road glance durations between on average 500 to 1600 ms in most traffic scenarios. Median glance durations for an in-car task above 1600 ms could suggest high visual in-car task demands that interfere with this general visual sampling tendency.

Furthermore, it is known that complexity of the in-car task increases the cognitive demands of the task, which can lead to cognitive capture of attention (e.g., Young and Salmon, 2012; Baumann, Petzoldt, Groenewoud, Hogema, and Krems, 2008; Blanco, Biever, Gallagher, and Dingus, 2006). We studied three different in-car tasks with varying interaction demands and complexity (see Table 1) in order to see if the in-car task demands interact with the warning efficiency.

In addition, we were interested if the warnings are more useful for novice drivers than for the experienced ones, as could be expected from the vast literature on risks related to novice drivers (e.g., Wikman, Nieminen, and Summala, 1998). Experienced drivers are expected to have a better situation awareness than the novice drivers already in the control condition (Underwood, Chapman, Bowden, and Crundall, 2002) as well as to have lower in-car glance durations in general (Wikman et al., 1998). The study by Donmez et al. (2010) suggests that the warnings could be beneficial for high-risk novice drivers, in particular. Based on the factorial design and the relevant literature, we formulated three hypotheses:

- H1. The warnings increase glance time on road and decrease individual in-car glance durations.
- H2. The warnings are more efficient for novice than experienced drivers.
    - The interaction Warnings x Driving experience should be significant and there should be greater increase in glance time on road as well as greater reduction

    - for Novices on the percentage of over-2-s in-car glances and median in-car glance durations than for Experienced from Control to Warnings conditions.
- H3. The efficiency of the warnings is in-car task-dependent.
    - This means that there should be significant interactions Warnings x In-Car Task Type on the dependent measures. If these are observed, more detailed analyses on the differences between the tasks are conducted.

In addition, the efficiency of the location-based (WT=0) warnings while multitasking will be further analyzed by taking the most demanding part of the test track under a closer analysis. The survey results should indicate the participants' experiences and in particular, the acceptance, of the warning system:

- H4. The participants experience the warnings as acceptable.
    - The means for the constructed experience factors differ to a positive direction from the midpoint of the scale, indicating positive general experiences towards the application.

4.2 Participants

In total 31 participants were recruited via university's student emailing lists. The requirements for the participants included experience on smart phones, a valid driving license, normal or corrected vision, and either less than 2,000 km (novice drivers) or more than 50,000 km (experienced drivers) of lifetime driving experience. Twenty of the participants were men (64.5%) and 11 were women (35.5%). The ages of the participants varied between 18 and 67 years with the mean age being 31.2 years (SD 12.1). The participants had had their driving license on average 12.3 years (SD 11.6). There were 10 novice (in or just out of driving school with less than 2,000 km of lifetime driving experience) and 21 experienced (with over 50,000 km of lifetime driving experience) drivers. Due to technical difficulties with the face recognition during the practice trials, one novice and one experienced driver did not drive the trials, but they were introduced to the system with a demonstration and they filled in the survey based on this experience (*N*=31 for the survey). Due to vibrations of the car and low levels of light during trials driven after sunset, the quality of the video image of these trials proved to be challenging to code the eye-movements reliably afterwards from the videos. This led to a sample size of *N*=24 (9 novices and 15 experienced drivers) for the glance analyses. All the participants were rewarded with a movie ticket, a mobile car holder, and a car charger.

4.3 Apparatus

The experiment was conducted on a driving practice track in Lievestuore, Finland during two consecutive weekends in early March 2014. The track was closed from other vehicles. The utilized car was a Volkswagen Golf GTI 2005 with direct shift automatic transmission to make the driving task as fluent and easy as possible (Figure 4).

Figure 4. The car on the Lievestuore track.

The driving paths on the track and the highly demanding parts of the track (WT=0) are illustrated in Figure 5. The most demanding parts of the track (circles in Figure 5) included three intersections as well as four tight bends with low visibility. The most demanding path on the track included a three-way intersection after an icy downhill slope, a stop sign and a tight (over 90 degrees) turn to right with low visibility. As an additional “warning sign”, there was a wrecked car on the right-hand side of the road, which marked the start of the WT=0 area in the intersection. The driving paths and directions were varied between the trials and participants in order to mitigate unwanted learning effects.

Figure 5. The driving paths on the track. The circles indicate high demand areas with WT=0, the intersection marked with stop a sign being the most demanding area.

Two Samsung Galaxy S3 smart phones (GT-I9305, Android 4.4.2) had the VisGuard application installed and these were alternated between the Warnings blocks. The dashboard holder was securely installed on the passenger side air vent, as seen in Figure 2. The angle of the phone towards the driver was adjustable, and was adjusted for each driver for optimal gaze recognition before each trial. In the other phone, the VisGuard warnings were set off, and in the other, they were set on. The phones had differing content (e.g., text messages), but the search targets were always set on the same locations. In addition, the smart phone with the warnings on was switched between days. These controls kept task complexity at a similar level across the conditions.

In contrast to possible real-world application settings, in the experiment, the same warning times for both novice and experienced drivers were used in order to avoid this confounding variable on the analyses of the effects of driving experience. Because of the fairly static and low visual demands of the practice driving track (excluding the areas with WT=0) and the low, but fairly constant driving speed (max 40 km/h), the WTs remained at almost static ca. 1.7 s levels on the relatively straight parts of the track.

A Sony HD video camera was used to capture driver's gaze and the screen of the smart phone in the holder. The camera was on a holder attached to the backseat of the car and operated by a research assistant who was recording the driver's gaze via a rearview mirror. An extra-large rearview mirror placed above the standard rearview mirror was used in order to keep the driver's eye constantly visible in the video image. Other equipment included a laptop for filling in the two (pre- and post-experiment) questionnaires online.

4.4 Procedure

At the arrival to the driving track, the participants filled in a consent form and a pre-study questionnaire for demographic data. After this phase, they completed four practice tasks with a Galaxy S3 smart phone while being stationary. The conducted tasks were similar to the tasks used later in the experiment. The participants were given the time they needed to get comfortable with driving on the track and the car with its controls. Two practice tasks without warnings were performed while driving; a task of using the calculator to count "89-56" and a task of searching and setting a destination (by typing and selecting a suggestion) to Tampere. Before the experiment and between each trial, a gaze recognition test was performed and the car phone holder adjusted, if needed.

In the trials, the in-car tasks started always on the home screen and the car was stationary at the starting position in the big crossings (in the middle of Figure 5) before the participant said to be ready to start the next trial. The participants were briefed to focus on the driving, to try to keep the speed near the assigned speed limit of 40 km/h if possible, and to try to keep the car on its lane. However, the participants were asked to adjust the speed according to the situation and their own feeling. In addition, the participants were instructed to drive as if they are driving in real traffic and to obey traffic regulations. For safety reasons, there was no other traffic on the track during the trials. However, as there was no fence on the track, there was always a possibility of other road users. Therefore, the participants were instructed to pay attention to the road. Even though the safety measures were strictly enforced, in three trials the risk of other road users was realized when hikers were crossing the track. However, no safety-critical incidents occurred during these events.

The participants were further instructed that there is always the possibility of unexpected events, as for example the road surface can be slippery. Therefore, they were instructed to take their time in completing the tasks with the smart phone. They were also told to use their own discretion in completing the tasks. However, they were instructed to keep the car in movement, if possible, when completing the smart phone tasks. They were instructed to take turns only according to the verbal instructions given by the experimenter sitting on the passenger seat. An important last advice was that if they would hear someone saying STOP, they should brake the car to a halt immediately.

The in-car tasks were selected to be equal with realistic in-car activities one could imagine drivers are willing to engage in while driving (see Table 1). The in-car tasks varied in many aspects, but the Gallery task was used as a practice task to reduce unwanted learning effects. It was the simplest task at least in terms of the number of interaction steps required as well as the number of possible interaction options per screen. The last Text message task was intended to be the most visually complex one, as it involved visual search of finding five targets in an unordered list among 13 items in total in a sentence-like text message. Blanco et al. (2006) as well as Kujala and Saariluoma (2011) have shown that finding a semantic target on a compressed text is a visually highly demanding task while driving. In addition, it included a small cognitive decision-making component of choosing if an item is a member of the target category (dairy product or fruit), whereas the other tasks included only a memory component of remembering the verbal target.

Table 1

In-Car Tasks (Trials) in the Order of Execution per Block (Top-Down).

Task type	Black phone	White phone	Task steps
Gallery	Find an image of a dog (with page-by-page scrolling)	Find an image of a sheep (with page-by-page scrolling)	1. Gallery 2. Browse (7 items)
Calculator	Use the calculator to count 851/742, and 269*358	Use the calculator to count 962*853, and 158/247	1. Calculator (on the Home screen) 2. (E.g.) 962*853= 3. (E.g.) 158/247=
Navigation	Search and set a destination to Helsinki, after which stop the navigation and search and set a destination to Oulu	Search and set a destination to Jyväskylä, after which stop the navigation and search and set a destination to Turku	1. Google Maps (on the Home screen) 2. Search 3. (E.g.) h 4. (E.g.) Helsinki 5. Car icon 6. Select a route 7. Start 8. Stop the guidance (and go to 2.)
Text message	Read a text message – read aloud what dairy products you have to buy from the grocery store	Read a text message – read aloud what fruits you have to buy from the grocery store	1. Messages (on the Home screen) 2. Select the first message in the menu 3. Read the targets aloud (13 items with five targets)

The tasks with the same phone (i.e., warnings off or on) were performed in a sequence in a block, but the orders of the Control and Warnings blocks were counter-balanced across the sample and level of driving experience. In addition, the exact contents of the task varied with the phone (black or white, see Table 1) and therefore, from the Control to Warnings blocks from day to day.

After the trials, the participants filled in a post-study questionnaire on their experiences, after which they were rewarded. A single experiment lasted on average one hour and 15 minutes.

4.5 Survey Design

Before the experiment, the participants filled in a web-based pre-study questionnaire, which included questions about the participants' background such as gender, age, and amount of years the participants had had a driving license. The results of the pre-study questionnaire were mostly used in analyzing the demographic data from the participants, which is presented in the "4.2 Participants" section of this paper.

After the experiment, the participants filled in a web-based post-study questionnaire, which included statements (i.e., questions or items) about the VisGuard application. Most of these statements were positive in nature (e.g., "The application was useful for me"), but the questionnaire included also some negative statements (e.g., "The application was annoying"). The answering options for the statements were in a 5-point rating scale with response options ranging from "strongly disagree" (1) to "strongly agree" (5). Response option 3 was "neither agree nor disagree". Each statement had also an additional text field for possible open comments regarding the chosen answering option. The post-study questionnaire items and the constructed factors based on these are presented in the "4.6 Analysis" section of this paper.

In detail, the post-study questionnaire included questions about different themes, which were based on previous literature. From research regarding trust in technology, the questionnaire had three items related to the application's trustworthiness (Lee and See, 2004; Dzindolet et al, 2003), perceived consistency (i.e., reliability) (Lee and See, 2004; Bisantz and Seong, 2001), and timeliness (Grandison and Sloman, 2000). In addition, two items were related to whether the participant would recommend the application (Jonsson, Harris and Clifford, 2008; Ghazizadeh, Lee and Boyle, 2012) and to application designers' benevolence (Lee and See, 2004; Mayer, Davis and Schoorman, 1995).

From technology acceptance research, two items were related to the application's and its warnings' perceived usefulness (Roberts, Ghazizadeh, and Lee, 2012; Pavlou, 2003;

Venkatesh and Davis, 2000) and another two on whether the participant would use the application after the study (Venkatesh and Davis, 2000). Also, four items were related to the general acceptance of the application and user satisfaction (Pavlou 2003; Venkatesh and Davis, 2000; Lewis, 1995; Van der Laan, Heino and De Waard, 1997). In contrast to perceived usefulness, four items were related to the perceived harmfulness (Bisantz and Seong, 2001) and annoyance (Van der Laan, Heino and De Waard, 1997; Weinstock, Oron-Gilad and Parmet, 2012) related to the application.

In addition, four items were related to the suitability of the application for its intended task (i.e., validity) (Lee and See, 2004; Venkatesh and Davis, 2000). The application developers' definition of the application's intended task was defined before these statements in the questionnaire. Finally, further four items related to the participants' experiences with the functioning of the circle symbol of the application. The questionnaire included also other items regarding, for example, the participants' opinions about the commercialization of the application, but the results of those items are not analyzed or reported here as they are not in the focus of this paper.

4.6 Analysis

Because of the expected unreliability in the glance data collected with VisGuard, two data reducers coded independently the in-car glance durations to the smart phone from video after the SAE-J2396 standard (SAE, 2000). Noldus Observer XT software was used for coding the glance data. Inter-rater reliability was assessed by calculating Cohen's Kappa and Intraclass Correlation Coefficient (ICC) for the glances of eight randomly selected trials the both data reducers had scored. Both metrics indicated high levels of inter-rater reliability (Kappa=0.874, 95% CIs [0.848, 0.900], N=160; ICC=.910, 95% CIs [.871, .937], N=122). For Kappa, there were 140 events during which both data reducers scored a glance within 500 ms of each other, and a total of 20 events when only one of them did.

Repeated measures ANOVA and paired samples t-tests were used for testing the hypotheses on the glance metrics. For each ANOVA, assumptions of sphericity were confirmed. If the assumption of sphericity was violated, degrees of freedom were adjusted with the Greenhouse-Geisser correction.

The received questionnaire answers were analyzed with exploratory factor analysis (Principal Axis Factoring) using IBM SPSS Statistics (Version 20). The used rotation method was Varimax with Kaiser Normalization. The purpose of the exploratory factor analysis was not to create a novel general scale but to reduce the data set by constructing the most suitable factors amongst the items for the purposes of the current experiment (H4). Twenty-five questionnaire items that were based on the previous research presented in "4.5 Survey Design", and which correlated at least with $r = .5$ with at least one other item were selected to the initial factor analysis.

The factor analysis was done five times, because some of the items originally thought to contribute for the acceptance-related factors had to be excluded from further analysis. On the first run, "I could recommend this application to my friends" had high cross loadings (> .40), "The intentions of the application's designers are good" had no loadings at all, and "I am satisfied with the application" had high cross loadings on the second run. On the third run, the items "The application supported my driving performance", "The application supports my safe driving", "The warnings support my safe driving", and "The circle symbol did not have a harmful effect on my driving performance", had still high cross-loadings, and were removed. On the fourth run, items "The warnings given by the application were annoying" and "The application increased my alertness in traffic" had high cross-loadings, and one of the factors had only two items with loadings over .40; "I was happy to use the application" and "I accepted the application as part of my driving activity during the study". All these items were removed before the fifth run. In addition, the item "The application works well in its intended task" was removed from the final solution, as it loaded only for a factor otherwise closely related to the circle symbol.

For the final solution with 13 items, the Kaiser-Meyer-Olkin measure of sampling adequacy was .65, that is, greater than the limit of acceptable (.50, Kaiser, 1974), and Bartlett's test of sphericity indicated high significance, $\chi^2(78) = 250.643$, $p<.001$, indicating that factor analysis is appropriate. All the communalities were over .40. Following the Kaiser criterion, only factors with Eigenvalues over 1.0 were selected. This resulted in the selection of four factors. Scree-plot indicated the last notable drop in the Eigenvalues after the fourth factor, after which there was a less steep decline. The rotation converged in 5 iterations. The factor loadings are presented in Table 2.

Table 2

Factor Loadings for Exploratory Factor Analysis with Varimax Rotation of the Questionnaire Items.

Item	Factor 1	Factor 2	Factor 3	Factor 4
The warnings given by the application appear when they are needed.	**.76**	.33	-.27	.12
The application works coherently and logically.	**.62**	.03	-.29	.21
The application is trustworthy.	**.63**	.39	.05	.29
After this study, I could use the application every day while driving.	.33	**.79**	-.21	.14
I could use this application after the test run.	.11	**.73**	-.14	.23
The warnings that the application gives are useful for me.	.24	**.80**	-.16	-.04
The application was useful for me.	.02	**.78**	-.10	-.09
The warnings had a harmful effect on my driving performance.	-.07	-.28	**.78**	-.01
The application had a harmful effect on my driving performance.	-.26	-.05	**.78**	-.14
The application was annoying.	-.12	-.18	**.77**	-.29
The circle symbol was not annoying. [1]	.02	-.08	-.12	**.78**
The circle symbol supports my safe driving.	.29	.17	-.25	**.76**
The circle symbol is useful for me.	.23	.09	-.05	**.69**

Notes [1]Reverse-coded values. Factor loadings > .40 are in boldface. *N*=31.

Factors were constructed from items that had a loading of over .40, but only if the Cronbach's alpha was over .70. The factors were finally selected and labeled according to themes identified based on the literature presented in "4.5 Survey Design". Table 3 presents the questionnaire items that contributed to the four different factors identified in the factor analysis. All the four factors can be interpreted as indicative of an underlying factor contributing to the acceptance of the application. The four factors explained 67.1% of the total variance of the items. Finally, means for the constructed factors were calculated by adding up the scores of each item in a factor and dividing the total by the number of items included in the factor. We hypothesized that the scale means for the constructed factors differed to a positive direction from the theoretical mean (or median) of 3.0 (one-sample t-test and one-sample Wilcoxon Signed Rank test), indicating a positive general experience on the latent factors (see H4).

Table 3

Factor Loading of Each Questionnaire Item and Cronbach's Alpha for Each Identified Factor.

Factor number and short label / Item (Factor long label) (% of total variance)	Factor Loading	α
Factor 1: Trust (Trust in the application) (13.4%)[1]		.79
The warnings given by the application appear when they are needed	.76	
The application works coherently and logically	.62	
The application is trustworthy	.63	
Factor 2: Usefulness (Usefulness of the application) (21.8%)[1]		.88
After this study, I could use the application every day while driving	.79	
I could use this application after the test run	.73	
The warnings that the application gives are useful for me	.80	
The application was useful for me	.78	
Factor 3: Harmfulness (Harmfulness / annoyingness of the application) (16.5%)[1]		.85
The warnings had a harmful effect on my driving performance	.78	
The application had a harmful effect on my driving performance	.78	
The application was annoying	.77	
Factor 4: Circle symbol (Functioning of the circle symbol) (15.4%)[1]		.82
The circle symbol was not annoying[2]	.78	
The circle symbol supports my safe driving	.69	
The circle symbol is useful for me	.78	

Notes [1]Rotation sum of squared loading (% of variance), [2] Reverse-coded values. *N*=31.

For all the statistical analyses the alpha level was set to .05. For multiple comparisons, Bonferroni correction was applied for the alpha level. Where applicable, partial eta-squared and Cohen's *d* were used as a measure of effect size.

5 RESULTS AND DISCUSSION

5.1 Glance Metrics

The glance metrics indicated varying levels of support for the hypotheses. The glance metrics relevant for the three hypotheses H1-H3 are presented in Table 4.

Table 4

Glance Metrics Relevant for the Hypotheses H1-H3 (N=24). Values are Means (Standard Error of Mean).

	Glance time on road (%)		Percentage of over-2-s in-car glances (%)		Median in-car glance duration (s)	
	Control	Warnings	Control	Warnings	Control	Warnings
Total	55.4 (2.7)	62.3 (2.3)	28.1 (4.1)	29.1 (4.4)	1.57 (.09)	1.58 (.09)
Calculator	60.5 (2.6)	63.7 (2.5)	24.9 (4.4)	22.4 (4.0)	1.49 (.09)	1.42 (.08)
Navigation	58.6 (2.8)	61.6 (2.6)	24.1 (3.7)	26.6 (4.2)	1.44 (.09)	1.52 (.09)
Text message	47.0 (3.6)	61.0 (2.7)	30.6 (4.7)	34.0 (5.7)	1.70 (.10)	1.75 (.11)
Novices	56.9 (4.2)	62.7 (3.6)	32.6 (6.3)	31.0 (6.7)	1.66 (.13)	1.59 (.14)
Experienced	53.9 (3.5)	61.8 (3.0)	23.5 (5.3)	27.3 (5.6)	1.49 (.11)	1.57 (.11)

A significant main effect of the VisGuard warnings was observed on glance time on road, which was increased on average by 12.5 percent, $F(1,20)=13.125$, $p=.002$ partial $\eta^2=.396$. However, no significant effects were found on percentage of over-2-second in-car glances ($p=.743$) or on median in-car glance durations ($p=.908$). Therefore, Hypothesis 1 was supported by the fact that the warnings increased glance time on road, but not supported by the absent effects on individual in-car glance durations.

There were no significant interaction effects of warnings and driving experience on glance time on road (*p*=.578), percentage of over-2-second in-car glances (*p*=.409), or median in-car glance duration (*p*=.301). This finding suggests the effects of the warnings are independent of the driving experience of the driver (i.e., Hypothesis 2 is rejected). In general, the novice drivers seemed to have a greater number of over-2-second in-car glances than the experienced (see Table 4), but the difference was not significant with this sample size (only 9 novice drivers). Due to the unequal group sizes, we tested the main effect of driving experience on in-car glance metrics in the control and experiment conditions by averaging over the tasks, and by testing the differences also with Welch's t-test that is more reliable when the two samples have unequal sample sizes. Still, we did not find significant differences between the driving experience groups with any of the in-car glance metrics (glance time on road, control: *p*=.566, experiment: *p*=.843; percentage of over-2-second in-car glances, control: *p*=.274, experiment: *p*=.670; median in-car glance duration, control: *p*=.308, experiment: *p*=.904).

There was a significant interaction effect of Warnings and In-Car Task Type on glance time on road ($F(2,28.198)=10.975$, $p=.001$, partial $\eta^2=.354$), but not on percentage of over-2-second in-car glances (*p*=.579), or on median in-car glance durations (*p*=.434). These findings give partial support for Hypothesis 3; the efficiency of the warnings on glance time on road seems to be in-car task-dependent. However, more detailed analyses are required in order to better understand the task type effects on the efficiency of the warnings.

All the in-car task types proved to be visually demanding because the mean percentages of over-2-second in-car glances were well above the 15 percent for the 85th percentile verification criteria for in-car tasks set by NHTSA (2013b) (Calculator: 25.7 [SEM=4.0], Navigation: 27.5 [SEM=3.9], Text Message: 32.6 [SEM=4.8]). The driving speeds were low (40 km/h) compared to the NHTSA (2013b) testing scenario (80 km/h), which may partly explain the high percentages. On the other hand, the visual sampling model of Wierwille (1993) suggests that drivers prefer to keep in-car glance durations well below 2 seconds in most traffic situations.

More detailed analysis revealed that there were significant differences between the in-car task types on glance time on road ($F(2,40)=7.485$, $p=.002$, partial $\eta^2=.272$), percentage of over-2-second in-car glances ($F(2,40)=3.333$, $p=.046$, partial $\eta^2=.143$), as well as median in-car glance durations ($F(2,40)=11.800$, $p<.001$, partial $\eta^2=.371$). The mean differences between the in-car task types are displayed in Table 5. Paired comparisons revealed, as expected, that the Text message task seemed to be the most visually complex task with the glance time on road as well as on the median in-car glance duration metrics. Besides the Text message task, the participants seemed to be able to follow the general visual sampling behavior suggested by Wierwille (1993); keeping the median in-car glance durations below 1.6 seconds (see Table 4).

Table 5

Mean Differences Between In-Car Task Types (N=24).

	Glance time on road (%)	Percentage of over-2-s in-car glances (%)	Median in-car glance duration (s)
Text message - Calculator	-8.9** (*p*=.001)	7.0 (*p*=.024)	.243*** (*p*<.001)
Text message - Navigation	-6.0 (*p*=.032)	5.2 (*p*=.051)	.204** (*p*=.001)
Navigation - Calculator	-2.8 (*p*=.200)	1.8 (*p*=.562)	.039 (*p*=.518)

Notes * $p < .017$ (Bonferrroni-adjusted alpha level for multiple comparisons), ** $p < .01$, *** $p <.001$

However, the greatest effect of the warnings on glance time on road was also seen in the visually most complex Text message task (see Figure 6). The percentual increases on glance time on road from Control to Warnings condition per task type were 5.3% for the Calculator ($t(23)=1.393$, $p=.177$), 5.1% for the Navigation ($t(23)=2.210$, $p=.037$, $d=.225$), and 29.7% for the Text message ($t(23)=4.407$, $p<.001$, $d=.929$). The increases on glance time on road were achieved without significant increases on individual in-car glance durations (Table 4). Because of confounding factors, we cannot say definitely that the visual complexity of the in-car task affected the efficiency of the warnings but it seems to be one plausible factor.

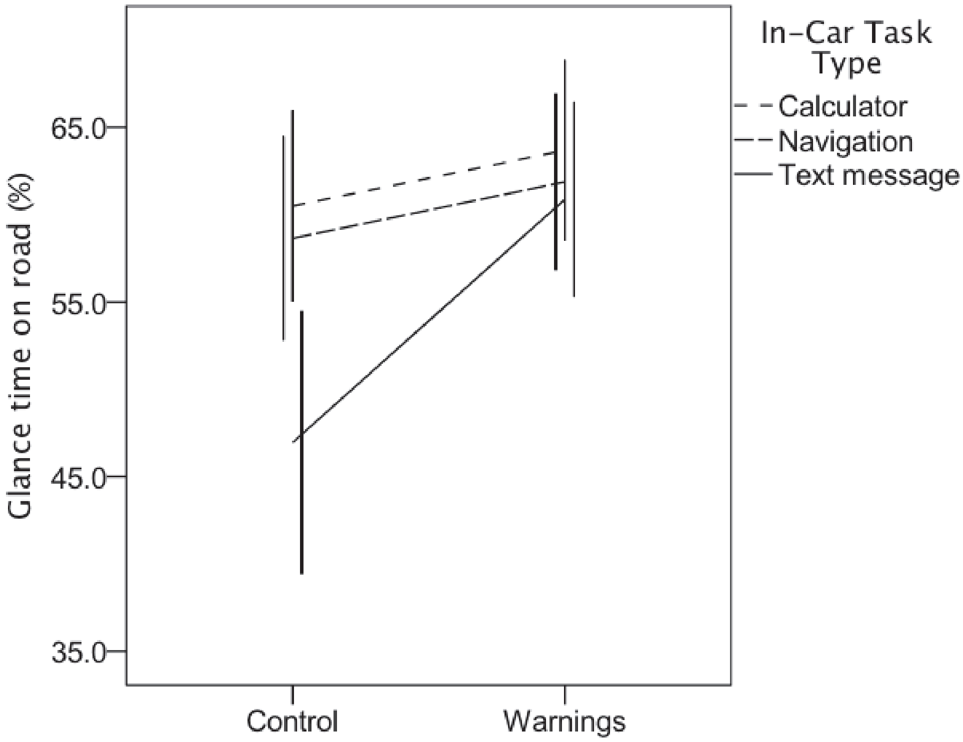


Figure 6. Mean glance time on road (%) by Warnings and In-Car Task Type (*N*=24). The bars represent 95% CIs.

Due to varying lighting conditions, we noted a number of technical difficulties in gaze tracking with the VisGuard application, which may at least partly explain the absent effects of the warnings on the individual in-car glance durations. Therefore, we wanted to see if the observed effects of warnings on glance time on road would be due to the location-based warnings (WT=0) that worked reliably based on the GPS signal, and thus, did not require the gaze tracking. We took the glance time on road in the most demanding intersection on the track with the stop sign (see Figure 5) and WT=0 under closer analysis. The driving directions and paths were systematically varied between the participants in order to mitigate unwanted order effects, but 14 participants drove the Calculator and Text message trials on exactly the same path that included the stop sign intersection for both Control and Warning conditions. The sample of the 14 participants was in balance for the trial orders. This enabled us to analyse if the general task type effect was visible in the glance time on road also in this highly demanding intersection. The glance time on road was coded frame-by-frame from the video material from a clearly visible landmark (the wrecked car on the side of the road) leaving the picture as marking the starting point of the WT=0 section to another clear landmark (a ploughing stick) leaving the picture as marking the end point of the section identically for both Control and Warnings conditions.

Figure 7 illustrates that, at a general level, the participants seemed to acknowledge that the intersection required more visual attention than driving in general and in the Calculator trials they were successful in devoting high levels of visual attention on the road environment. However, in the Text message trial in the Control condition, the mean glance time on road was only 71.3% compared to the 93.5% in the Calculator trial in the Control condition ($t(13)=2.971$, $p=.011$, $d=1.105$). The warnings succeeded in raising the percentage up to 90.1% (26.4% increase) in the Text message Warnings trials ($t(13)=2.761$, $p=.016$, $d=.838$. These findings give further support for the result that the Text message tasks were the most engaging ones, capturing participants' attention also in highly demanding situations and perhaps undermining their situation awareness. However, the location-based warnings (WT=0) seemed to increase participants' attention on road on these tasks, at least on this highly demanding point of the track.

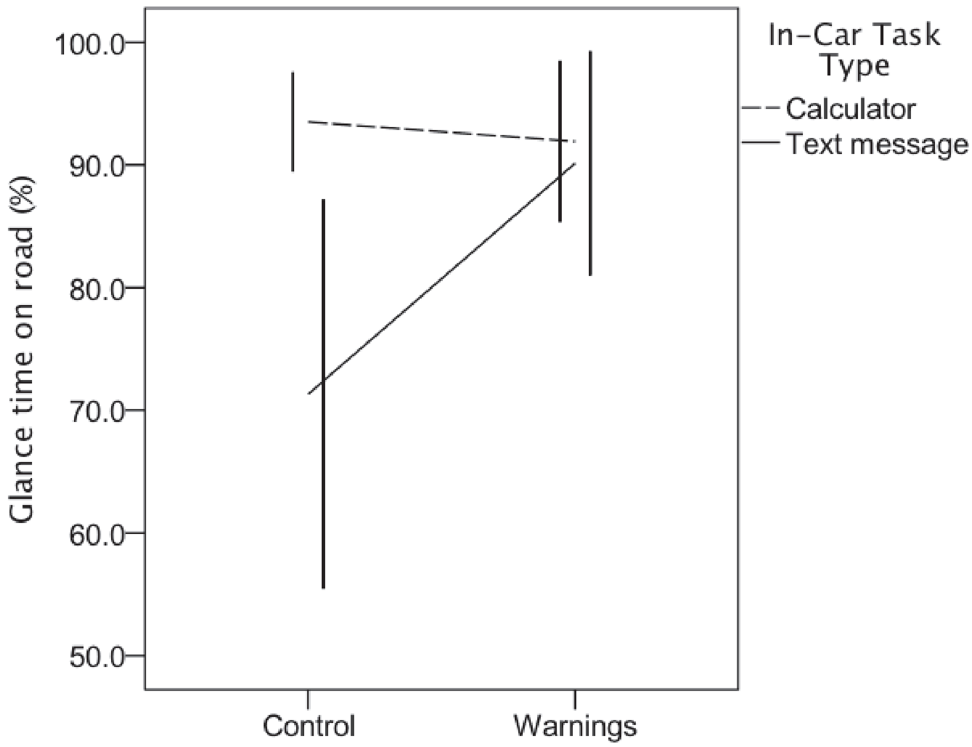


Figure 7. Glance time on road (%) by Warnings and In-Car Task Type in the most demanding intersection on the track with WT=0 (*N*=14). The bars represent 95% CIs.

Overall, the glance metrics indicated that the warnings increased glance time on road, and in particular in the visually most complex Text message trials, whereas no effects on individual in-car glance durations were found. Driving experience, or the lack of it, did not affect the efficiency of the warnings. The missing effects on in-car glance durations by the warnings can be attributed to the poor working of the gaze tracking in real driving environment with varying lighting conditions. However, the location-based warnings (WT=0) seemed to have a clear impact on the participants' glance time on road and in particular in the visually complex Text message tasks. This finding is well reflected in the analysis on the glance time on road in the most demanding intersection, as illustrated above, but also in the survey results on the participants' experiences towards the application.

5.2 Survey Results

Table 6 presents the means and standard deviations for the four different factors identified in the factor analysis as well as for the individual questionnaire items that contributed to the factors.

Table 6

Mean and Standard Deviation for Each Identified Factor and Contributing Questionnaire Item.

Factor number and short label (Factor long label) / Item	M	SD
Factor 1: Trust (Trust in the application)	**3.71**	**.68**
The warnings given by the application appear when they are needed	3.61	.96
The application works coherently and logically	3.94	.68
The application is trustworthy	3.58	.77
Factor 2: Usefulness (Usefulness of the application)	**3.45**	**.91**
After this study, I could use the application every day while driving	3.10	1.25
I could use this application after the test run	3.77	.88
The warnings that the application gives are useful for me	3.68	1.11
The application was useful for me	3.26	.97
Factor 3: Harmfulness (Harmfulness and annoyingness of the application)	**2.06**	**.92**
The warnings had a harmful effect on my driving performance	1.94	.93
The application had a harmful effect on my driving performance	2.06	1.15
The application was annoying	2.19	1.05
Factor 4: Circle symbol (Functioning of the circle symbol)	**3.23**	**1.02**
The circle symbol was not annoying[1]	3.55	1.15
The circle symbol supports my safe driving	3.06	1.21
The circle symbol is useful for me	3.06	1.21

Notes [1] Reverse-coded values. *N*=31.

As can be seen from Table 6, in general, the means of the identified factors indicated positive experiences of the participants towards the application. The participants' textual questionnaire comments regarding the application also supported the received mean values of the factors. For example, the rather high level of trust in the application (M = 3.71, SD = .68) was reflected in the following participant comment: "*The application had relevant warnings in dangerous situations. The detection of dangerous situations works well*". The usefulness of the application (M = 3.45, SD = .91) was commended, for example, in the following way: "*It is a useful application, which increases the safety of driving*".

The low level of experienced harmfulness and annoyingness of the application (M = 2.06, SD = .92) was supported by the following comment: "*The usage of the application did not have a harmful effect on my driving performance. Instead, I think it improved it*". Finally, Factor 4: Functioning of the circle symbol (M = 3.23, SD = 1.02) received a mean only slightly larger than 3.0. In line with this result, the participants commented the circle symbol to be "*rather unnoticeable*" and "*difficult to read*". Furthermore, one participant commented that "*The warning triangle works nicely, but the circle symbol does not – it is difficult to perceive*".

Hypothesis 4 that the scale means for the constructed factors differ to a positive direction from the theoretical mean of 3.0 was tested with one-sample t-test. In addition, the result of non-parametric one-sample Wilcoxon Signed Rank test (with 95% confidence level) was used to support the one-sample t-test result with each factor. The results of these analyses are described in Table 7 along with the short labels of the factors. Based on these results, the means differ statistically significantly from 3.0 on all the other factors, except on Factor 4 (Circle symbol). The reason for Factor 4 to not differ statistically significantly from 3.0 may be that the participants did not seem to experience the circle symbol to be neither very useful nor harmful. Rather, their attitudes were neutral towards the circle symbol. This was reflected in participant comments such as "*I think the circle symbol did not have an effect on my behavior*" and "*I focused more on the given task and on the driving [than on the circle symbol]*".

Table 7

Factors' One-sample Test Mean Differences, One-Sample t(30) values and Wilcoxon Signed Rank Test Z Values With Their Significance Levels.

Factor's short label	One-sample test mean difference	One-sample t(30) value	Wilcoxon Signed Rank Z value
Factor 1: Trust	0.71	5.84^{***}	4.05^{***}
Factor 2: Usefulness	0.45	2.78^{**}	2.28^{*}
Factor 3: Harmfulness	-0.94	-5.68^{***}	-4.05^{***}
Factor 4: Circle symbol	0.23	1.23	1.26

Notes * $p < .05$, ** $p < .01$, *** $p < .001$

With all the other factors (than Factor 4), the means of the factors were statistically significantly greater than 3.0, except with Factor 3 (Harmfulness), which was significantly lower than 3.0. This result is due to the fact that, in contrast to the other factors, all the statements regarding the application in Factor 3 were negative.

Pearson correlations between the factors were calculated (see Table 8) to establish the construct validity of the factors. Table 8 indicates that the correlations between the factors were otherwise statistically significant (at least on the .05 level), except that the factors "Usefulness of the application" or "Harmfulness and annoyingness of the application" did not correlate with the "Functioning of the circle symbol" factor. Based on this result — and also on the comments given by the participants regarding the circle symbol — it can be concluded that the circle symbol's functioning was not experienced to have importance over the general perceived usefulness (or harmfulness) of the application. Rather, especially the warnings given by the application were experienced to contribute to the perceived usefulness of the application, which became also evident in participant comments like the following one: "*Yes, the warning triangle was useful, and it could have been even more visible*".

Table 8

Correlation Matrix Presenting the Values (Pearson's r) for Inter-Correlations between the Identified Factors.

Factor short label	Trust	Usefulness	Harmfulness	Circle symbol
Trust	1			
Usefulness	0.52^{**}	1		
Harmfulness	-0.42^{*}	-0.38^{*}	1	
Circle symbol	0.43^{*}	0.20	-0.35	1

Notes * $p < .05$, ** $p < .01$

There were also other relevant items in the post-study questionnaire than the ones presented in Table 6. These individual questionnaire items are presented in Table 9 in a descending order by the mean value. The individual item means are well in line with the factor means, further supporting the notion that the application was well accepted by the drivers.

Table 9

Mean and Standard Deviation for Each Individual Questionnaire Item Left Out of the Factor Analysis. Sorted by Mean.

Item	M	SD
The intentions of the application's designers are good	4.58	.56
I accepted the application as part of my driving activity during the study	4.19	.70
The warnings support my safe driving	4.06	.77
I could recommend this application to my friends	4.03	.88
I was happy to use the application	3.94	.81
The application supports my safe driving	3.87	1.02
The application works well in its intended task	3.74	.82
The application increased my alertness in traffic	3.74	.82
I am satisfied with the application	3.55	.85
The application supported my driving performance	3.48	.96
The circle symbol did not have a harmful effect on my driving performance[1]	3.42	1.09
The warnings given by the application were annoying	2.16	1.04

Notes [1] Reverse-coded values. N=31.

The four individual items intended to measure the suitability of the application for its intended purpose got all median scores higher than 3.0 (one-sample Wilcoxon Signed Rank test); "The application works well in its intended task" ($Z = 3.77$, $p < .001$), "The application increased my alertness in traffic" ($Z = 3.77$, $p < .001$), "The application supports my safe driving" ($Z = 3.64$, $p < .001$), and "The warnings support my safe driving" ($Z = 4.44$, $p < .001$). The suitability was commented, for example, in the following way: "*If you use a mobile (phone) while driving, the application will increase the safety of driving*".

Based on the survey results, it can be said on a general level that the application was well accepted by the participants (H4 supported). The application was trusted by the participants as they felt that it works coherently and logically especially in demanding road locations (e.g., tight curves or intersections). The warnings of the application were experienced to be useful and the participants thought they could use the application also after the study. This attitude was also indicated in the survey results regarding the harmfulness and annoyingness of the application and its warning messages, which were experienced to be fairly low. The suitability of the application for its intended task (to support safe driving) was thought to be high, especially in demanding road situations according to the participants. However, the survey results suggest that the functioning of the circle symbol needs reconsideration in order for the application to support safe and distraction-free driving even better.

6 GENERAL DISCUSSION

In this paper, we have studied the effects of a context-sensitive distraction warning application on in-car glance behaviors and subjective experiences of drivers who conducted smart phone tasks on a test track while driving. We have also studied the moderating effects of driving experience and in-car task type on the efficiency of the warnings. The general goal of the study was to better understand the effects of the proactive and context-sensitive distraction warnings on drivers' visual behaviors and acceptance.

The glance metrics indicated a significant increasing effect of the VisGuard warnings on the glance time on road while multitasking with the smart phone. The average increase was 12.5 percent of the total in-car task time. Volvo has reported even 37 percent increases on glance time on road with their Visual Distraction Alert (VDA) systems (Wege and Victor, 2014). However, Volvo's studies have been conducted in driving simulators in highly controlled settings, whereas our experiment was done in a real car with varying environmental conditions, although on a closed track. Our experiences clearly indicated that gaze tracking with current smart phones outside the laboratory is highly unreliable due to varying lighting conditions. Besides the variable environmental lighting conditions, we found the gaze tracking based on face recognition, head pose, and phone holder position in the passenger side air vent problematic. The method requires that the driver turns his/her head sufficiently towards the phone for optimal gaze recognition. In addition, the holder position makes it challenging to utilize peripheral vision for driving while looking at the phone.

The in-car task type had a clear effect on the efficiency of the warnings on the glance time on road. The maximum increase of 29.7% on glance time on road was observed with the visually most complex Text message tasks. The findings are in line with previous research, suggesting that the increased demands of in-car tasks decrease drivers' ability to appropriately evaluate the dynamic visual demands of the driving situation (Young and Salmon, 2012; Baumann et al., 2008; Blanco et al., 2006). According to the data, the context-sensitive distraction warnings can help drivers in overcoming this inability and to place more attention on the road.

However, the in-car tasks differed also in many other aspects than the measured visual demand and these confounding effects should be considered when drawing conclusions. For instance, the Text message reading task included the lowest number of manual interactions (only two inputs to open the message) and the warning icon was displayed always closer to the visual targets (i.e., the message) than in the other tasks. The other tasks' input elements, in particular, were often located on the lower part of the smart phone display (N.B. the icon did not cover the targets in any task). Due to the latter aspect, the warning icon could have been noticed more easily in the Text message trials than in the others. The Text message trials were also always the last trials within a block. However, the data suggests that the visual complexity was significantly higher and the glance time on road considerably lower for the Text message tasks in the Control condition (at least for the most demanding section of the track) than for the other task types. It seems the participants were able to prioritize the driving task at a sufficient level in the Control conditions in the other trials, but not in the Text message trial. The warnings seemed to increase the glance time on road to a comparable level for the Text message task compared to the other tasks. Therefore, it is hard to see how the other (listed) task properties than the visual or cognitive complexity could explain the observed interaction between the warnings and the in-car task type on the glance time on road. The text message task represented the only in-car task in which there was a clear cognitive component together with the brief off-road glances, an unfavorable combination from the viewpoint of missing safety-critical events in the road environment (Lee et al., 2007).

Against our expectations, there were no significant interaction effects of warnings and driving experience on any of the glance metrics. The finding suggests the effects of the warnings are independent of the driving experience of the driver. In general, the novice drivers seemed to have a greater number of risky over-2-second in-car glances in general than the experienced drivers (as in Wikman et al., 1998), but the difference did not become significant with our sample size. In the study by Donmez et al. (2010) their distraction feedback had a significant effect on the glance behavior of the high-risk drivers only. The study suggests there are individual differences in visual sampling behaviors among young, inexperienced drivers, in a similar fashion as one can expect differences among more experienced drivers. Therefore, the novice drivers in our study should probably not be regarded as a homogenous group. However, we did not analyse their glancing behaviors in order to form subgroups due to the small group size. In addition, it should be noted that even if our data does not support differences between novice and experienced drivers, the group sizes were small and unequal, and this finding should be interpreted with caution.

To summarize, the glance metrics indicated that the warnings significantly increased glance time on road, and in particular in the visually most complex Text message trials, whereas no effects on individual in-car glance durations were found. The missing effects on in-car glance durations can be attributed to the poor working of the gaze tracking in a real driving environment with varying lighting conditions. On the other hand, the location-based warnings on the demanding parts of the track (curves, intersections) ahead seemed to have a clear impact on the participants' glance time on road and in the Text message tasks in particular. This finding is well reflected in the analysis of the Text message trials in the most demanding intersection, where glance time on road increased by 26.4% due to the warnings (from 71.3% to 90.1%). The level of driving experience did not seem to affect the efficiency of the warnings.

The survey results indicated a rather high level of trust in the application. This can be a positive result, if the trust is at an appropriate level (i.e., matches the capability of the system), as discussed, for instance, by Lee and See (2004). In addition, the results indicate a high level of perceived usefulness, and low level of perceived harmfulness of the application. All of these factors can be seen to contribute to the general acceptability of the application according to the previous literature as the results of our factor analysis are in line with the Technology Acceptance Model (TAM) and its extensions. From TAM (Venkatesh and Davis, 1996), perceived usefulness (PU) was evaluated especially with the items in "Factor 2:

Usefulness of the application" (Table 6). In our survey, it was not reasonable to measure TAM's perceived ease of use (PEOU), as the VisGuard application did not have or require user input. From a TAM extension labelled Automation Acceptance Model (AAM) by Ghazizadeh et al. (2012), AAM's "trust" is similar to items in our "Factor 1: Trust in the application". In addition, the high mean scores on the individual items on suitability, such as "The application works well in its intended task" and "The application increased my alertness in traffic", are well in line with the factor means and support the importance of the suitability of the application for its intended purpose for technology acceptance (Ghazizadeh et al., 2012). Therefore, we see that the results of our study contribute to the current discussion of what factors affect technology acceptance in general and the acceptance of distraction warning systems in particular (Roberts, Ghazizadeh, and Lee, 2012). However, the interpretation of the results of the exploratory factor analysis should be done with care, as the ratio of cases per items (31 / 13) was low.

Drivers' experiences towards the warning system were significantly more positive than towards the real-time feedback system studied by Roberts, Ghazizadeh, and Lee (2012). We suggest at least three plausible reasons for this finding; 1) the proactive context-sensitivity of the warnings, 2) people in general seem to acknowledge that the use of mobile devices while driving is always a distraction (Jääskeläinen and Pöysti, 2014), and 3) the warnings on the screen of the smart phone were more subtle than the flashing LED + auditory warnings in Roberts, Ghazizadeh, and Lee (2012).

The functioning of the circle symbol, intended to display the remaining warning time threshold, was not thought to be useful by the participants. One plausible explanation for these experiences is the larger than expected delays in gaze tracking. The individual comments also suggest that the circle caused only additional visual load. The warnings, on the other hand, were experienced as highly useful, which suggests that the circle symbols could be removed and the warnings kept as the only icon for guidance. In this study, we expected challenges with the gaze tracking, and therefore, we did not utilize warning sounds, vibrations, or blinking in order to avoid attracting (instead of detracting) the attention of the driver towards the phone in a high-demand situation. However, the effects of other modalities than visual only for the efficiency and acceptability of the warnings should be studied (Smith, Clegg, Heggestad, and Hopp-Levine, 2009).

Existing and suggested distraction warning systems have typically high false positive rates, which can undermine the acceptability of these systems by the drivers (Lee et al., 2013; NHTSA, 2013a). The false positive rate of the VisGuard system was low in the experiment due to the larger than expected recognition delays in the gaze tracking. Consequently, this must have also led to lower true positive rate than intended for the single glance duration warnings (WT~1.7 s), which may partly explain the absent effects on the in-car glance durations. For the location-based warnings of demanding road conditions ahead (WT=0), however, the true positive rate was 100%, as the warnings were displayed based reliably on the GPS signal and map data, regardless of whether the driver's gaze was recognized or not.

Our study had also some shortcomings that should be taken into account in further research. The experience of the participants with the application was fairly short. The acceptability of the application in daily use as well as possibility of long-term negative behavioral adaptation (e.g., Rajaonah, Tricot, Anceaux, and Millot, 2008) should be studied. The warning application could increase the use of smart phones while driving due to false sense of security, undermining its positive effects. According to the theory of task difficulty homeostasis by Fuller (2005), a driver has a preferred range of driving task difficulty that s/he is prepared to accept and prefers to maintain. Via its warnings, the application could improve the driver's ability to calibrate the perceived situational task demands to her/his capability while multitasking. On the other hand, if the driver's experienced risks of multitasking, informing the estimates of task difficulty, are reduced by the application, the driver could be motivated to increase multitasking behind the wheel. The long-term effects of providing drivers feedback of and support for multitasking behaviors in order to develop their tactical visual off-road sampling and task prioritization skills via a distraction warning system should be carefully studied.

We did not measure directly participants' situation awareness in this experiment but the increase in glance time on road in the most demanding part of the track achieved by the location-based warning could suggest that the warning helped the drivers to recognize the demanding section also with the most demanding in-car task, as they did while conducting the other, easier in-car tasks. Increased attention on road increases the possibilities to detect task-critical events on the road environment compared to a situation where the driver's eyes are off road. However, more glance time on road does not necessarily mean higher situation awareness, and the effects of the warnings on drivers' situation awareness should be more carefully studied in future work. Furthermore, the warning icon was the same for both glance-duration and location-based warnings, and it remains unclear if the drivers were able to correctly associate the location-based warnings to the increased visual demands ahead for each situation.

Due to the unreliable gaze detection, the current study focused mainly on the location-based warnings of the highly demanding parts of the track (intersections, tight bends). This together with the fairly static visual demands of the track on the straights allowed us to test and prove the concept of context-sensitive warnings in a controlled setup, but did not allow us to test the adaptability of the glance duration warning time in more dynamic settings. However, the proactiveness and context-sensitivity of the distraction warnings got significant support from both objective and subjective data.

Whereas in-car user interface design (Lee, Forlizzi, and Hudson, 2005) and testing tools (Kujala and Salvucci, 2015), driver education (Horrey et al., 2009), reactive in-car driver assistance (e.g., lane-keeping assistants) and feedback systems (Donmez et al., 2007), as well as legislative and governmental regulations (NHTSA, 2013b) may help in reducing the negative effects of driver distraction by in-car activities, there is additional demand for fast and cost-effective counter-measures that can be easily deployed by a driver. According to our study, mobile applications aimed to supervise the use of the smart phone while driving and aiding the driver to place more attention on road seems to be a one viable and acceptable option.

Acknowledgments

The authors would like to thank Hilkka Grahn, Niina Holmstedt, Päivi Kinnunen, Jussi Jokinen, Timo Tokkonen, Marko Salmela, Riitta Jokela, Ilkka Kotilainen, and Anton Valojää for their invaluable help in conducting the study. This research was supported by DisGuard, an innovation and research project funded by TEKES (Finnish Funding Agency for Technology and Innovation) and University of Jyväskylä. Jyvässeudun Ajoharjoittelurata Oy provided the test track for the experiment.

References

Baumann, M.R.K., Petzoldt, T., Groenewoud, C., Hogema, J., Krems, J.F., 2008. The effects of cognitive tasks on predicting events in traffic. In *Proceedings of the European Conference on HCD for Intelligent Transport Systems 2008* (pp. 3-11).

Bisantz, A.M., Seong, Y., 2001. Assessment of operator trust in and utilization of automated decision-aids under different framing conditions. *International Journal of Industrial Ergonomics*, *28*, 85-97.

Blanco, M., Biever, W.J., Gallagher, J.P., Dingus, T.A., 2006. The impact of secondary task cognitive processing demand on driving performance. *Accid. Anal. Prevent. 38*, 895-906.

Broström, R., Engström, J., Agnvall, A, Markkula, G., 2006. Towards the next generation intelligent driver information system (IDIS): The Volvo car interaction manager concept. *Proceedings of the 13th ITS World Congress.*

Donmez, B., Boyle, L., Lee, J.D., 2008. Designing feedback to mitigate distraction. In M.A. Regan, J.D. Lee, and K.L. Young (Eds.), *Driver Distraction: Theory, Effects, and Mitigation* (pp. 519-532). Boca Raton, FL: CRC Press.

Donmez, B., Boyle, L., Lee, J.D., 2007. Safety implications of providing real-time feedback to distracted drivers. *Accid. Anal. Prev., 39*, 581-590.

Donmez, B., Boyle, L., Lee, J.D., 2010. Differences in off-road glances: The effects on younger drivers' performance. *Journal of Transportation Engineering, 136*, 403-409.

Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., Beck, H.P., 2003. The role of trust in automation reliance. *International Journal of Human-Computer Studies*, *58*, 697-718.

Fitch, G.A., Soccolich, S.A., Guo, F., McClafferty, J., Fang, Y., Olson, R.L., Perez, M.A., Hanowski, R.J., Hankey, J.M., Dingus, T.A., 2013. *The Impact of Hand-Held And*

Hands-Free Cell Phone Use on Driving Performance and Safety-Critical Event Risk (Report No. DOT HS 811 757). Washington, DC: National Highway Traffic Safety Administration.

Flannagan, M., Sivak, M., 1993. Indirect vision systems. In Peacock B. & Karwowski W. (Eds.), *Automotive Ergonomics* (pp. 205-217). Washington DC: Taylor & Francis.

Fuller, R., 2005. Towards a general theory of driver behaviour. *Accident Analysis and Prevention*, *37*, 461-472.

Ghazizadeh, M., Lee, J.D., Boyle, L.N., 2012. Extending the Technology Acceptance Model to assess automation. *Cognition, Technology & Work*, *14*, 39-49.

Green, P., 2004. *Driver Distraction, Telematics Design, and Workload Managers: Safety Issues and Solutions* (SAE Paper No. 2004-21-0022).

Grandison, T., Sloman, M., 2000. A survey of trust in internet applications. *Communications Surveys & Tutorials, IEEE*, *3*, 2-16.

Horrey, W.J., Lesch, M.F., 2009. Driver-initiated distractions: Examining strategic adaptation for in-vehicle task initiation. *Acc. Anal. and Prev, 41*, 115-122.

Horrey, W.J., Lesch, M.F., Garabet, A., 2009. Dissociation between driving performance and drivers' subjective estimates of performance and workload in dual-task conditions. *Journal of Safety Res., 40*, 7-12.

Jonsson, I.M., Harris, H., Nass, C., 2008. How accurate must an in-car information system be? Consequences of accurate and inaccurate information in cars. In *Proceedings of the 26th SIGCHI Conference on Human Factors in Computing Systems (CHI'08)* (pp. 1665-1674). ACM.

Jääskeläinen, P., Pöysti L., 2014. Tarkkaamattomuus tieliikenteen turvallisuusongelmana – suomalaisten käsityksiä (Driver distraction and inattention as a safety issue in road traffic – Finnish views). *Liikenneturvan selvityksiä 7/2014.* Helsinki: Finnish Road Safety Council.

Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika, 39*, 401-415.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J., 2006. *The Impact of Driver Inattention On Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data* (Report No. DOT HS 810 594). Washington, DC: National Highway Traffic Safety Administration.

Kujala, T., 2013. *VisGuard – Distraction Prevention System.* Oral presentation at the 3rd International Conference on Driver Distraction and Inattention, Gothenburg, Sweden. Abstract retrieved from http://www.chalmers.se/hosted/ddi2013-en/program/program-papers

Kujala, T., Mäkelä, J., Kotilainen, I., Tokkonen, T. 2016. The attentional demand of automobile driving revisited – Occlusion distance as a function of task-relevant event density in realistic driving scenarios. *Human Factors*, *58*(1), 163-180.

Kujala, T., Saariluoma, P., 2011. Measuring distraction at the levels of tactical and strategic control – The limits of capacity-based measures for revealing unsafe time-sharing models. *Advances in Human-Computer Interaction,* vol. 2011, Article ID 594353, 13 pages.

Kujala, T., Salvucci, D.D., 2015. Modeling visual sampling on in-car displays: The challenge of predicting safety-critical lapses of control. *International Journal of Human-Computer Studies, 79*. 66-78.

Lee, J., Forlizzi, J., Hudson, S.E., 2005. Studying the effectiveness of MOVE: a contextually optimized in-vehicle navigation system. In *Proceedings of the 23th SIGCHI Conference on Human Factors in Computing Systems (CHI'05)* (pp. 1665-1674). ACM.

Lee, Y. Lee, J.D., Boyle, L.N., 2007. The interaction of cognitive load and attention-directing cues in driving. *Human Factors, 51*, 271-280.

Lee, J.D., Moeckli, J., Brown, T.L., Roberts, S.C., Schwarz, C., Yekhshatyan, L., Nadler, E.,

Liang, Y., Victor, T., Marshall, D., Davis, C., 2013. *Distraction Detection and Mitigation Through Driver Feedback.* (Report No. DOT HS 811 547A). Washington, DC: National Highway Traffic Safety Administration.

Lee, J.D., See, K.A., 2004. Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*, 50-80.

Lewis, J.R., 1995. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, *7*, 57-78.

Liang, Y., Lee, J.D., Yekhshatyan, L., 2012. How dangerous is looking away from the road? Algorithms predict crash risk from glance patterns in naturalistic driving. *Human Factors, 54*, 1104-1116.

Mayer, R.C., Davis, J.H., Schoorman, F.D., 1995. An integrative model of organizational trust. *Academy of management review*, *20*, 709-734.

National Highway Traffic Safety Administration, 2013a. *Distraction Detection Algorithm Evaluation* (Report No. DOT HS 811 548). Washington, DC: National Highway Traffic Safety Administration.

National Highway Traffic Safety Administration, 2013b. *Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices* (NHTSA-2010-0053). Washington, DC: National Highway Traffic Safety Administration.

Pavlou, P.A., 2003. Consumer acceptance of electronic commerce: integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, *7*, 101-134.

Rajaonah, B., Tricot, N.J., Anceaux, F., Millot, P., 2008. The role of intervening variables in driver-ACC cooperation. *Int. J. Hum.-Comput. Stud., 66*, 185-197.

Regan, M.A., Hallett, C., Gordon, C.P., 2011. Driver distraction and driver inattention: Definition, relationship and taxonomy. *Acc. Anal. & Prev., 43*, 1771-1781.

Schömig, N., Metz, B., 2013. Three levels of situation awareness in driving with secondary tasks. *Safety Science, 56,* 44-51.

Roberts, S.C., Ghazizadeh, M., Lee, J.D., 2012. Warn me now or inform me later: Drivers' acceptance of real-time and post-drive distraction mitigation systems. *International Journal of Human-Computer Studies*, *70*, 967-979.

Smith, C.A.P., Clegg, B.A., Heggestad, E.D., Hopp-Levine, P.J., 2009. Interruption management: A comparison of auditory and tactile cues for both alerting and orienting. *International Journal of Human-Computer Studies*, *67*, 777-786.

Society of Automotive Engineers, 2000. *SAE-J2396 Definitions and Experimental Measures Related to the Specification of Driver Visual Behavior Using Video Based Techniques.* Warrendale, PA: SAE.

Tivesten, E., Dozza, M., 2014. Driving context and visual-manual phone tasks influence glance behavior in naturalistic driving. *Transportation Research Part F, 26*, 258-272.

Transportation Research Board, 2003. *Design Speed, Operating Speed, and Posted Speed Practices* (NCHRP Report 504).

Underwood, G., Chapman, P., Bowden, K., Crundall, D., 2002. Visual search while driving: skill and awareness during inspection of the scene. *Transportation Research Part F, 5*, 87-97.

Van Der Laan, J.D., Heino, A., De Waard, D., 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies*, *5*, 1-10.

Venkatesh, V., Davis, F.D., 2000. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Man. Sci. 46*, 186-204.

Wege, C., Victor, T., 2014. Safer attention. *Vision Zero International*. 28-31.

Weinstock, A., Oron-Gilad, T., Parmet, Y., 2012. The effect of system aesthetics on trust, cooperation, satisfaction and annoyance in an imperfect automated system. *Work*, *41*, 258-265.

Wierwille, W.W., 1993. An initial model of visual sampling of in-car displays and controls. In A.G. Gale, I.D. Brown, C.M. Haslegrave, H.W. Kruysse, & S.P. Taylor (Eds.), *Vision in vehicles IV* (pp. 271-279). Amsterdam: Elsevier.

Wikman, A.S., Nieminen, T., Summala, H., 1998. Driving experience and time-sharing during in-car tasks on roads of different width. *Ergonomics, 41*, 358-372

You, C.-W., Montes-de-Oca, M., Bao, T.J., Lane, N.D., Lu H., Cardone, G., Torresani, L., Campbell, A.T., 2012. CarSafe: A driver safety app. In *Proceedings of UbiComp'12.*

Young, K.L., Salmon, P.M., 2012. Examining the relationship between driver distraction and driving errors: A discussion of theory, studies and methods. *Safety Science, 50*, 165-174.

Young, K.L., Salmon, P.M., Cornelissen, M., 2013. Missing links? The effects of distraction on driver situation awareness. *Safety Science*, *56.* 36-43.

Vitae

Tuomo Kujala is a Post-Doctoral Researcher at the Department of Computer Science and Information Technology in the University of Jyväskylä. He obtained his PhD degree in Cognitive Science in 2010 at the University of Jyväskylä, Finland.

Hannu Karvonen is a Research Scientist at VTT Technical Research Centre of Finland Ltd where he works in the Human Factors in Complex Systems team. He is also currently a PhD student in Cognitive Science at the University of Jyväskylä, Finland.

Jakke Mäkelä obtained his PhLic degree in Theoretical Physics in 1998 and PhD degree in Physics in 2010, both at the University of Helsinki, Finland. He is currently a Project Researcher at the Department of Computer Science and Information Technology in the University of Jyväskylä.