VISIONS • SCIENCE • TECHNOLOGY • RESEARCH HIGHLIGHTS

Dissertation

57

Evaluating systems usability in complex work

Development of a systemic usability concept to benefit control room design

Paula Savioja

VTT

VTT SCIENCE 57

Evaluating systems usability in complex work

Development of a systemic usability concept to benefit control room design

Paula Savioja

A doctoral dissertation completed for the degree of Doctor of Science in Technology to be defended, with permission of Aalto University School of Science, at a public examination held at lecture hall T2, of the school on the 25th of June 2014 at 12 noon.

ISBN 978-951-38-8145-0 (Soft back ed.)
ISBN 978-951-38-8146-7 (URL: http://www.vtt.fi/publications/index.jsp)

VTT Science 57

ISSN-L 2242-119X
ISSN 2242-119X (Print)
ISSN 2242-1203 (Online)

JULKAISIJA – UTGIVARE – PUBLISHER

VTT
PL 1000 (Tekniikantie 4 A, Espoo)
02044 VTT
Puh. 020 722 111, faksi 020 722 7001

VTT
PB 1000 (Teknikvägen 4 A, Esbo)
FI-02044 VTT
Tfn. +358 20 722 111, telefax +358 20 722 7001

VTT Technical Research Centre of Finland
P.O. Box 1000 (Tekniikantie 4 A, Espoo)
FI-02044 VTT, Finland
Tel. +358 20 722 111, fax +358 20 722 7001

Cover picture: Fortum

Kopijyvä Oy, Kuopio 2014

Evaluating systems usability in complex work
Development of a systemic usability concept to benefit control room design

Systeemikäytettävyyden arviointi monimutkaisessa työssä. Systeemisen käytettävyyskäsitteen kehittäminen valvomosuunnittelun tueksi. **Paula Savioja.**
Espoo 2014. VTT Science 57. 168 p. + app. 116 p.

Abstract

The design of industrial control rooms assumes fulfilling the goals of production, safety, and human well-being. Control rooms and the user interfaces within them should enable the effective and efficient conduct of work in all the operating conditions which can be foreseen during specification and design. At the same time, the user interfaces should enable the control of the process in unprecedented and totally unexpected situations while simultaneously maintaining the safety of the process. During the design or modification of the control room, the potentiality of the emerging solution to fulfil the objectives is assessed by conducting empirical evaluations. This dissertation presents the development of an evaluation methodology which enables developing the control room towards meeting these objectives.

Industrial process control constitutes a socio-technical system in which people and technologies have multiple and sometimes overlapping roles. In order to meet the demands of maintaining safety in all situations, the socio-technical system should have built-in capability of dealing with the unexpected. The control room, and the user interfaces within it, are an integral part of the socio-technical system. Thus, they have a role in construing and maintaining the safety of the system. The concept of systems usability (SU) is introduced in the dissertation to evaluate the systemic effects of control room solutions. SU is a human-centred quality attribute of user interfaces and control rooms attributed to technology, but the value in the use of the technology is evidenced in the success of the activity in which the technology is used. Thus, the research makes sense of the significance of the individual technological solutions in, and for, the entirety of an activity system.

Systems usability means that a tool in an activity serves the functions of 1) an instrument, 2) a psychological tool, and 3) a communicative tool. The meaning of each function in the specific domain is contextually defined. Furthermore, the quality of the tool can be assessed utilising different perspectives on the usage activity: performance, way of acting, and user experience. By combining the functions of the tool and perspectives on activity, a systemic framework for developing contextual indicators for a good control room is construed.

Utilising the concept of SU in the control room requires a model-based evaluation approach. This means that the general contextual work demands are considered in defining the reference of a successful process control activity. In addition, the scenarios utilised in the evaluation are modelled also taking the general work demands into account. A specific scenario modelling method, functional situation modelling (FSM), is presented in this dissertation. FSM combines a functional and

a chronological view of the activity of an operating crew in a particular situation. By making explicit the connection between required operating actions and critical domain functions, the model lays the ground for analysing operating activity from the point of view of maintaining the critical functions.

On aspect of control room evaluation is identifying whether the operating practices are attuned to maintaining safety in all situations. For this purpose, a data analysis method is presented in this dissertation. The analysis is based on identifying how operators identify and interpret signs depicted in the control room. Signs are, for example, information technological representations of process information, but they may also be the activities of other crew members. If the interpretation of signs has an identifiable global safety-related aspect, it may be concluded that the operating practice, even though situated, is also attuned to the general functions of the work which must always be maintained.

A particular viewpoint in evaluating SU is provided by analysing user experiences (UX) which emerge in the complex work. For this purpose, a UX questionnaire was developed within the SU evaluation framework. The questionnaire is based on UX indicators which reflect experiences of appropriateness concerning the three tool functions (instrument, psychological, communicative). The importance of UX as a measure of germinating, not yet existing tool appropriateness, is presented in the dissertation.

The contribution this dissertation makes is in the intersection of the research fields of human factors and ergonomics (HF/E) and usability engineering. The theoretical foundations of the research are in activity theory and cognitive systems engineering. The empirical work has been conducted by following the control room modernization efforts of Finnish nuclear power plants, during which evaluations have been carried out.

Keywords control room evaluation, human factors, systems usability, control room design, nuclear power plant

Systeemikäytettävyyden arviointi monimutkaisessa työssä
Systeemisen käytettävyyskäsitteen kehittäminen valvomosuunnittelun tueksi

Evaluating systems usability in complex work. Development of a systemic usability concept to benefit control room design. **Paula Savioja.**
Espoo 2014. VTT Science 57. 168 s. + liitt. 116 s.

Tiivistelmä

Teollisuudessa käytettävien valvomoiden suunnittelu tähtää tuotannon, turvallisuuden ja ihmisten hyvinvoinnin tavoitteiden saavuttamiseen. Valvomoiden ja niissä käytettävien käyttöliittymien tulisi olla sellaisia, että ne mahdollistavat tuloksellisen ja tehokkaan työskentelyn kaikissa suunnitteluvaiheessa ennakoitavissa olevissa tuotannollisissa tilanteissa. Samalla käyttöliittymien tulisi mahdollistaa prosessinhallinta myös ennalta odottamattomissa tilanteissa turvallisuustavoitteista tinkimättä. Kehittyvää valvomoratkaisua arvioidaan valvomon suunnittelu- ja muutosprojekteissa tekemällä empiirisiä kokeita. Tämä väitöstyö käsittelee sellaisen valvomoarviointimenettelyn kehittämistä, joka mahdollistaa valvomon kehittämisen aikaisen arvioinnin edellä mainittujen yleisten tavoitteiden suhteen.

Teollinen prosessinhallinta muodostaa sosioteknisen järjestelmän, jossa ihmisillä ja teknologioilla on useita, usein päällekkäisiä rooleja. Jotta sosiotekninen järjestelmä voisi saavuttaa turvallisuustavoitteet kaikissa tilanteissa, siinä tulisi olla sisäänrakennettua kyvykkyyttä käsitellä epävarmuuksia. Valvomo ja sen sisältämät käyttöliittymät ovat keskeinen osa sosioteknistä järjestelmää. Siten niiden rooli turvallisuuden rakentumisessa ja ylläpidossa on olennainen. Tässä väitöstyössä esitellään systeemikäytettävyyden käsite, jonka avulla voidaan eritellä valvomoratkaisujen systeemisiä vaikutuksia suunnittelun aikaisissa valvomoarvioinneissa. Systeemikäytettävyys on valvomoiden käyttöliittymien ihmiskeskeinen laatutekijä ja sinällään teknologian ominaisuus, mutta teknologian käyttöarvo liittyy aina toimintaan, jossa sitä käytetään. Systeemikäytettävyyden käsitteen kautta tämä tutkimus käsittelee siis teknologioiden merkitystä osana toimintajärjestelmää ja sen tavoitteiden saavuttamista.

Systeemikäytettävyys erittelee työvälineen funktioita toimintajärjestelmässä toiminnan teorian pohjalta: Työväline on 1) instrumentti, 2) psykologinen väline ja 3) kommunikatiivinen väline. Jokaisen funktion merkitys tietyssä toiminnassa määritellään asiayhteydessään. Työvälineen laatua voidaan tarkastella kolmesta eri näkökulmasta toimintaan: suoritus, toimintatapa ja kokemus. Yhdistämällä työvälineen funktiot ja tarkastelunäkökulmat saadaan kehikko kontekstuaalisten hyvän valvomon indikaattoreiden määrittämiseen.

Systeemikäytettävyyden käsitteen käyttäminen valvomoarvioinnissa vaatii mallipohjaista arviointia, jossa työn yleiset vaatimukset ymmärretään ja otetaan huomioon arviointiskenaarioiden suunnittelussa. Tässä väitöstyössä on kehitetty erityinen skenaariomallinnustekniikka: funktionaalinen tilannemallinnus (FSM, Functional situation model). FSM yhdistää funktionaalisen ja kronologisen näkökulman ope-

raattorivuoron toimintaan tietyssä tilanteessa. Tekemällä eksplisiittiseksi yhteydet vaadittujen operaatioiden ja prosessin kriittisten funktioiden välillä malli luo perustan operaattoritoiminnan analyysille prosessin kriittisten funktioiden näkökulmasta.

Arvioinnissa on kiinnitettävä huomiota kehittyviin operaattorityön käytäntöihin ja siihen, että ne yleisesti tähtäävät turvallisuuden varmistamiseen kaikissa tilanteissa. Tätä tarkoitusta varten kehitetty operaattoritoiminnan analyysimenetelmä esitellään tässä väitöstyössä. Analyysi perustuu siihen, miten operaattorivuorot havaitsevat ja tulkitsevat valvomossa esiintyviä prosessista kertovia merkkejä. Merkit voivat olla informaatioteknologisia, mutta ne voivat olla myös esimerkiksi vuoron muiden jäsenten tekoja. Jos merkkien tulkinnassa pystytään tunnistamaan globaaleja yleisiä piirteitä, voidaan päätellä, että toiminta suuntautuu yleisten tavoitteiden, esimerkiksi turvallisuuden, ylläpitämiseen. Samalla tunnistetaan myös välittömien prosessinhallintaan liittyvien tekojen toteutuminen tai toteutumatta jääminen.

Valvomoarvioinneissa on myös syytä tarkastella monimutkaisessa työssä ilmeneviä käyttäjäkokemuksia. Ydinvoimalaitoksen operaattoreiden käyttäjäkokemuksia tutkittiin kolmessa eri teknologiaolosuhteessa. Käyttäjäkokemuksia tutkittiin kyselymenetelmän avulla, joka on kehitetty systeemikäytettävyyskehikon pohjalta. Työssä esitellään käyttäjäkokemusindikaattorit kaikille kolmelle välinefunktiolle (instrumentti, psykologinen väline, kommunikatiivinen väline) ja eritellään kolmen eri tutkimuksen tuloksia. Työssä tarkastellaan myös käyttäjäkokemuksen merkitystä kehittyvän, ei vielä valmiin työvälineen arvioinnissa.

Väitöstyön kontribuutio on inhimillisten tekijöiden ja ergonomian (human factors and ergonomics) ja käytettävyyssuunnittelun (usability engineering) välialueella. Työn teoreettinen perusta on toiminnan teoriassa ja kognitiivisessa järjestelmäsuunnittelussa (cognitive systems engineering). Empiirinen osuus on tehty seuraamalla suomalaisten ydinvoimalaitosten valvomokehityshankkeita, joiden yhteydessä erilaisia systeemikäytettävyysarviointeja on suoritettu.

Avainsanat control room evaluation, human factors, systems usability, control room design, nuclear power plant

Preface

I became interested in usability of industrial applications during my studies for master’s degree in the Helsinki University of Technology, department of Automation and Systems Engineering. During that time, I was fortunate to work three summers as an automation mechanic; first at the Salmisaari Power Plant of Helsingin Energia, and then at Kaukas Paper Mill of UPM Kymmene. These time periods were exciting: I had the opportunity to immerse in the reality of the industrial plants and experience the everyday life, the workarounds and sometimes clumsy tools but also the fantastic knowhow and skills of the people working at the plants. I believe this was the true starting point of the work which is now presented in this dissertation. The official research process stretches in time from 2003 when I enrolled as a doctoral candidate in the former Helsinki University of Technology, to this date, spring of 2014 when the final phases of my doctoral work are taking place at the Aalto University School of Science. Naturally, during these years there have been periods in which the research has been more intensive along with periods during which I have been occupied with other things in life. Nevertheless, today is the day to think back and thank everybody who have supported me in this endeavour.

First and for most, I want to express my gratitude and my most sincere appreciation to my thesis advisor Research Professor, Docent Leena Norros. We have worked together over 10 years on the different aspects of human factors in complex systems. The time has been full of interesting events, ups and downs, but all and all, it has been exciting and a lot of fun. I want to thank you for introducing me to Activity Theory and other theoretical philosophical background relevant in the dissertation. Your passion for research is truly inspirational, and without your excellent conceptual innovations and continuous support this work would not have been possible to complete.

Professor Marko Nieminen has been the supervising professor of this work. I want to thank him for support during all phases of the doctoral studies. Professor Nieminen can be accredited with being one of the pioneers of usability in Finland and his enthusiasm for advancing the whole field has also had positive effect on my possibility to carry out research in the industry.

I have been honoured to have Professor Kari Kuutti from the University of Oulu and Professor Greg Jamieson from the University of Toronto as the official preliminary examiners of my dissertation. Your kind remarks and encouragement have

strengthened my trust in the findings of the research and helped me in clarifying the strongest points of my work. I am grateful and looking forward to having Professor John M. Flach from the Wright State University as the opponent in the public defence of my dissertation in the near future. I also want to express my gratitude to the tens of anonymous reviewers during the process of writing the dissertation papers. The constructive comments have been extremely useful in construing the argumentation for the empirical findings and theoretical developments presented in the dissertation.

Writing the articles which constitute the dissertation has been collaboration in the most profound sense of the word. I want to thank all my co-authors in addition to already mentioned Leena Norros: Leena Salo, Jari Laarni, Marja Liinasuo, Hanna Koskinen, Hannu Karvonen and Iina Aaltonen, thank you for the joint work and fruitful discussions during the writing processes. Furthermore, I want to thank my other former and current team members Paula Laitio and Mikael Wahlström for good collaboration and new viewpoints, and all the rest of the current team Maiju Aikala, Susanna Aromaa, Pertti Broas, Göran Granholm, Juhani Heinilä, Kaj Helin, Eija Kaasinen, Jaakko Karjalainen, Tiina Kymäläinen, Simo-Pekka Leino, Juhani Viitaniemi, and Antti Väätänen, for the support an encouragement in the final phases of the dissertation work.

I have been fortunate to work as a full time research scientist at VTT during the whole course of my doctoral process. I want to thank Research Manager Olli Ventä for hiring me, and thus enabling a career in research in the first place. Technology Manager Jari Hämäläinen made it possible for me to concentrate on finalising the articles and the dissertation from time to time in 2012–2013 which was invaluable. I want to also mention my current manager Head or Research Area Riikka Virkkunen and thank her for the support in the final phases of the dissertation work. During my career in VTT I have worked under four different team leaders Teemu Tommila, Leena Norros, Maaria Nuutinen and Jari Laarni. They have all been excellent bosses and done everything in their power to enable conducting long term scientific work in the rigorous operative environment of VTT. I thank them all for that. A special thank you belongs to Jari L for commenting and proofreading one of the late versions of the dissertation.

The research which constitutes this doctoral work has been conducted in the empirical context of control room modernisations of the Finnish nuclear power plants. Only the practical questions related to control room modifications show the importance of this research work. Therefore, both Finnish nuclear power operators Fortum and Teollisuuden Voima (TVO) and the key personnel within the organizations must be mentioned. Late Esko Rinttilä from Fortum was always favourable towards control room development related research, and his great efforts made it possible to conduct such research in Finland in the first place. Ville Nurmilaukas, Mika Lehtonen, and Leena Salo have since continued successfully the control room development work at Fortum and I want to thank them and the whole control room team of Fortum for great collaboration. Furthermore, I want to thank simulator instructor team of the Loviisa plant lead by Pekka Kettunen for invaluable support in the course of the empirical work. Most importantly I must thank all the operating crews of the Loviisa plant who have taken part in the studies conducted by the VTT Human factors team in 2008–2012. Also Teollisuuden Voima (TVO) must be thanked for the opportunity to conduct simulator studies in the training simulator. Olli Hoikkala, Sanna Haapala and Mauri Viitasalo have shown support for our research and on their part enabled the participation of TVO. I want to thank also the simulator instructor team of TVO lead by Matti Rantakari and similarly all the operating crews who took part in the experiments in 2009 conducted by the VTT Human Factors team. In addition, I want to thank Harri Heimbürger, Heimo Takala and Jukka Kupila from the Finnish Radiation and Nuclear Safety Authority (STUK) for their support for control room related research work in the context of the Finnish nuclear safety research.

The international nuclear safety research community has been a favourable environment for developing methods for control room assessment. I want to thank numerous colleagues from EdF, OECD Halden Reactor Project, and University of Toronto for sharing ideas and experiences concerning control room modifications during the years. Furthermore, I have been fortunate to take part in two HCI related cost actions that have been led by Dr. Effie Law. It has been great, from time to time, to detach from the nuclear safety research and immerse in HCI related discussions, because usability is, in the end, my scholarly home. During the late stages of my research, the Finnish HCI research community working together in FIMECC UXUS-program has been important. I want to especially thank Research Manager Hannu Paunonen from Metso Automation for insightful discussions.

The funding bodies must also be thanked. Most of the research was carried out within the Finnish National Research Programs on Nuclear Safety (SAFIR, SAFIR 2010, and SAFIR 2014). Within the three programs, the research has also been funded by the OECD Halden Reactor Project, Fortum Oyj, Teollisuuden Voima Oyj, and VTT. Additionally, I have received direct funding (in the form of work hours) from VTT and a personal grant from the Fortum Foundation.

I am greatly indebted to my family, friends, and relatives for the interest and support for my doctoral work. I thank your encouragement during the process and not asking too many questions about the level of completeness. Especially the efforts of Hannele, Riikka, Ossi, Liisa and Maka in the form of endless supply of childcare must be mentioned.

I want to dedicate this dissertation to the ones who are the closest. Tuomas and Elmo you are an inspiration to me. You have been patient when I have needed the time, but also given me reasons to focus on other aspects of life when needed. I thank you for being there and for the love and support.

Otaniemi, 30.4.2014 (Wappuaatto)

Paula Savioja

Academic dissertation

Supervisor	Professor Marko Nieminen Department of Computer Science and Engineering Aalto University School of Science Finland
Thesis advisor	Research Professor Leena Norros Systems Engineering VTT Technical Research Centre of Finland Finland
Pre-examiners	Professor Kari Kuutti Department of Information Processing University of Oulu Finland Professor Greg Jamieson Cognitive Engineering Laboratory University of Toronto Canada
Opponent	Professor John M. Flach Department of Biomedical, Industrial & Human Factors Engineering Wright State University USA

List of publications

This thesis is based on the following original publications which are referred to in the text as Papers I–VI. The publications are reproduced with kind permission from the publishers.

I Savioja, P. and Norros, L. (2013) Systems Usability Framework for Evaluating Tools in Safety-Critical Work. Cognition, Technology and Work, 15, pp. 255–275.

II Savioja, P., Aaltonen, I., Karvonen, H., Koskinen, H., Laarni, J., Liinasuo, M., Norros, L. and Salo, L. (2012) Systems Usability Concerns in Hybrid Control Rooms. In: Proceedings of the 8th International Topical Meeting on Nuclear Plant Instrumentation and Control and Human-Machine Interface Technologies. San Diego, USA.

III Savioja, P and Norros, L. (2008) Systems Usability – Promoting Core-Task Oriented Work Practices in Complex Work. In: Law, E., Hvannberg, E. and Cockton, G. (eds). Maturing Usability: Quality in Software, Interaction, and Value. Springer, London. Pp. 123–143.

IV Savioja, P., Norros, L. and Salo, L. (2012) Functional Situation Models in Analyses of Operating Practices in Complex Work. In: Proceedings of ECCE 2012 Re-Thinking Cognition. Pp. 58–64. Edinburgh, United Kingdom.

V Savioja, P., Norros, L., Salo, L. and Aaltonen, I. (2014) Identifying Resilience in Proceduralised Accident Management Activity of NPP Operating Crews. Safety Science, 68, pp. 258–274.

VI Savioja, P., Liinasuo, M. and Koskinen, H. User experience: does it matter in complex systems? Accepted for publication 2013 in Cognition, Technology and Work. Springer. 44 p.

Author’s contributions

Paper I “Systems Usability Framework for Evaluating Tools in Safety-Critical Work” presents the systems usability framework which had been developed for the so-called baseline studies of the control rooms of Finnish nuclear power plants. The paper is methodological, but also presents some findings from the baseline studies. Construing the evaluation framework was a joint effort between Savioja and Norros. The studies were conducted collectively by the VTT team. Savioja conducted the secondary data analyses concerning systems usability. Savioja is responsible for the structure an all sections of the paper. Savioja is the main author of the paper.

Paper II “Systems Usability Concerns in Hybrid Control Rooms” presents in greater detail the results of the baseline studies of the Finnish NPPs. Conducting the studies was a collective effort by the list of authors. Savioja had a leading role in scenario development and modelling. She was responsible for the design of the process tracing interview method and the user experience questionnaire. She contributed to the design and analyses of the orientation interviews. She also contributed to all the data collection, analysis and interpretation of results. Savioja conducted the secondary data analyses concerning systems usability. Savioja is the main author of the paper.

Paper III “Systems Usability – Promoting Core-Task Oriented Work Practices in Complex Work” presents the systems usability concept and the related evaluation process which was construed based on studies conducted in the initial phases of the dissertation work jointly by Savioja and Norros. Savioja is responsible for the structure of the paper. Savioja is the main author of the paper.

Paper IV “Functional Situation Models in Analyses of Operating Practices in Complex Work” presents the functional modelling method the development of which was a joint effort between Savioja, Norros, and Salo. The graphical presentation of the situation model was developed jointly by Savioja and Salo. Savioja conducted the literature review for the paper alone. She is responsible for the structure and conclusions of the paper. Savioja is the main author of the paper.

Paper V "Identifying Resilience in Proceduralised Accident Management Activity" presents yet another analysis conducted on the data gathered in the baseline studies. Savioja conducted the initial analyses together with Salo and Aaltonen. The interpretation of the results and identification of functions in which operating crews' practices differed from each other was carried out by Savioja and Norros, Savioja is responsible for the structure and all sections of the paper apart from the theoretical basis of the analysis method. Savioja is the main author of the paper.

Paper VI "User Experience: Does It Matter in Complex Systems?" presents the application of a user experience questionnaire in three different studies in an NPP context. Savioja is responsible of the development of the UX indicators within the three SU functions. Savioja designed the initial version of the questionnaire and it was developed further by Savioja and Koskinen. Conducting the studies was a collective effort. Interpretation of the results was conducted by Savioja and Liinasuo. Savioja combined the results of the three studies and is responsible for the literature review presented in the paper. Savioja is the main author of the paper.

Contents

1. Introduction

Digital technologies are everywhere in the modern world. They have become embedded in our natural habitat and the practices of our daily lives in ways that are to a point unobservable by us, when the technology functions appropriately. The possible unreliability of digital technology may in most cases only be a nuisance, and quick fixes such as rebooting the systems are immediately available. But in the case of safety-critical systems the situation is different: Digital technologies affect the functioning of complex systems just as profoundly as they do everyday lives, but the impacts of unreliability are more severe. Digital technologies change the nature of accidents [Leveson 2011a]. Consequently, our ways of engineering safety into the systems need to be compatible with the potential new forms of hazards (ibid).

The research presented in this dissertation deals with methodologies of human factors and ergonomics which are utilised in the development of socio-technical systems. These methods are applied in engineering systems which aim to be as safe as possible in human use. A central concept is *usability*, which is treated as a quality attribute of user interfaces (UI) within complex socio-technical systems. Such UIs are utilised by professional operators, for example in a control room (CR), to operate safety-critical industrial processes. The whole control room constitutes a mediating technology element in the overall socio-technical system aiming to achieve the general objectives of production, safety, and well-being [Vicente 1999] of humans. In this work, usability as a concept is attributed to technology, but the value in use of a particular technological solution is evidenced in how well the whole socio-technical system succeeds in meeting its objectives. For this purpose, to connect the features of technology to the overall performance of the socio-technical system, the concept of *systems usability* has been developed. This principle of the connection of the technology to the appropriate functioning of the production system has been the flywheel of the development of the evaluation approach presented in the dissertation. The interest is not in technology per se, but rather in the capability of the technology to serve the variety of functions relevant for the viability of an activity system.

The practical empirical context of the work is control room development in the nuclear power domain. Control rooms are important for plant safety, as operator performance is related to plant performance [NEA 2004, Abu-Khader 2009, Jang, Park & Seong 2012, Park, Cho 2010]. When new digital technologies are introduced in a nuclear power plant (NPP) main control rooms (MCR), the effects on the operating work and safety need to be analysed.

An NPP constitutes a complex socio-technical safety-critical system as characterised from multiple perspectives by for instance Vicente [1999] and Saurin & Gonzalez [2013]: complexity is evident in the large number and diversity of the dynamically interacting elements of the system and in the behaviour of the system which contains both unanticipated variability and self-organised adaptations. Safety-criticality is obvious. In an accident condition, the possibility of the release of radioactive material into the environment may be increased. Nuclear accidents such as the Three Mile Island accident, the Chernobyl disaster, and the Fukushima Daiichi disaster have shown that accidents can take place in an NPP, and that the main control room, as a scene of the interplay of operators, procedures, and control room systems, plays a role in accidents and their management [Le Bot 2004]. The balancing between human and technical reliability [Papin 2011] is important in learning from the accidents so as to support development of the whole field [Rempe et al. 2012]. Poor design of control rooms is one of the latent conditions acknowledged by modern safety research to lead to production problems [Li, Powell & Horberry 2012] and threats to system safety and ultimately to accidents [Stanton et al. 2010], and usability has been identified as one of the key issues of development [Hamilton et al. 2013] within the socio-technical system of NPP operations.

1.1 Background and motivation for research

The research presented in this dissertation is motivated by the extensive modifications taking place in NPP control rooms in the form digitalisation of technologies.

The design phase of the life cycle of safety-critical systems is crucial for safety. According to combined accident and incident data from the aviation, railway, and nuclear industries, about 50% of reported accidents and incidents have a root cause in design [Kinnersley, Roelen 2007]. In a review of reported modification-related nuclear events, i.e. near misses, incidents, and accidents, it was found out that 41% of the events could be traced back to a failure in the design phase of the modification [Zerger, Maqua & Wattrelos 2013]. This means that, in order to improve safety, it is absolutely essential to focus on the design phase of systems.

The technical development in the nuclear industry is slow in comparison to other industries [Abu-Khader 2009]. Where in conventional industries the digitalisation of control room systems started widely already in the 1980s, the situation in NPPs is different; most control rooms are still in the era of analogue UI techniques [Silva Junior, Borges & Carvalho 2012]. In non-nuclear industries, a contemporary control room contains many types of digital systems which have been designed for various aspects of process control work [Suomen Automaatioseura ry 2010]. Due to progress in other industries, information concerning the design of control room systems can be found in the literature [Stanton et al. 2010, see e.g. Hollnagel 2003] and international standards which are comprehensively introduced by Stanton et al. [2010]. However, even though the digitalisation of information may open up new possibilities for the development of new representations of process information, the possibilities have not been fully exploited even in conventional industries. Often the main form of information representation is still based on the process and instrumentation (P&I) diagram complemented with trend and alarm displays [Braseth et al. 2009, see also Jamieson 2007]. This means that the prevailing representations still partly derive from the era of analogue user interfaces. This is understandable considering the nature of the investment needed in control room modifications and especially the development of new representations, although cost benefits of the human factors effort have been found to exist [Bruseberg 2008]. Despite the conservatism of the design solution in control rooms, users are known to use the control room systems for various purposes in addition to the obvious monitoring and conducting operations, for instance learning, development work, information sharing, and knowledge management [Paunonen 1997]. This means, that also in the context of complex work, users appropriate technologies for purposes for which they were not originally designed [Dourish 2003]. From the point of view of design of new control rooms and process information representations, this means that there may be room to develop even further the ways in which organizations can benefit from the extensive process monitoring and control conducted by modern automation systems. The thorough investigation of the usages of control room systems may shed light on these new opportunities.

In the nuclear industry today, new plants are being built and old ones modernized due to the obsolescence of the original technologies, and thus the digital era is also reaching NPPs. This may open possibilities for new representations such as ecological interface design [Burns, Hajdukiewicz 2004], multilevel flow modelling [Lind 1999], function-oriented design [Pirus 2004], information-rich design, task-based design [Braseth et al. 2009], and other non-conventional interfaces [Aghina et al. 2012]. Simultaneously, there are also inevitable challenges involved in control room upgrades. From an analysis and evaluation perspective, it is notable that human activity in a computer-based control room, as with any other digital technology, is highly cognitive and thus not so easily observable by an analyst [Meister 1995]. So-called secondary tasks of managing the interface will become more dominant than in the current control rooms. Because of this, the operators' performance in process control work per se may be threatened as attention is shifted to interface operations. In addition, under high mental workload, interface operations may become difficult to conduct [O'Hara et al. 2002]. Also collaboration within the crew, and particularly the shared understanding of the situational demands, may become more difficult to maintain [Roth et al. 2010]. Therefore, the probability of failures such as missing information necessary for operation or decision-making may be increased and this may further affect operational safety.

Evaluation of the safety implications of human system interaction is extremely complicated [Cacciabue 2010]. As it is not possible to evaluate all possible usage situations, a relevant research approach should be developed which enables generalisation of the results in a way that is meaningful in the domain. In addition to being accurate as such, i.e. concentrating on the meaningful aspects of human performance, the evaluation should be such that it supports the design of control room systems. Human factors methods are not always like this, as they do not link the human performance measures to the design features of the tested systems [Norros 2013]. Actually, rather the opposite, values such as independence of design, are typically emphasised in human factors testing to avoid biases. Nevertheless, in modernization of control rooms the changes impacting operational work by the new designs must be monitored because of the possible safety implications. This can be interpreted as management of change in operating work, which is one of the most important ways to support human performance in the evolving socio-technical system [Nuutinen 2005a]. In addition to supporting the design and management of change, the evaluation approach should be systematic, traceable and reflective about the scope of the change [Cacciabue 2010].

Therefore, in the prevailing situation of technology transformation in the nuclear industry, the motivation for this research is to identify and develop a control room system evaluation approach and methods which both support design efforts and enable assessment of the safety implications of the new systems.

1.2 Research object and goals

As stated above, this dissertation concerns the evaluation of control room tools from the point of view of suitability for operating work. The object of the research is thus *evaluation of tools in complex work* (Figure 1).

Evaluation as such cannot be treated in isolation of its purpose and the practical context in which it is carried out. The purpose of the evaluation dictates in large part the methodology and approach utilised in the evaluation. In the control room digitalization context, the purpose is at least twofold: design of tools needs to be supported and identification of safety implications is necessary to inform the acceptance of tools. These two purposes affect the choice of methods. The interaction is two-way: the way the evaluation is conducted affects the usefulness of the results for the intended purposes.

Starting from a different viewpoint, evaluation is also affected by the nature of the operating work in which the tools are to be used and the conceptualisation of what constitutes a good tool in the particular work. The nature of operating work affects the choice of aspects of the work to be treated in the evaluation. For example, in particularly risky tasks it may be exactly the aspects known to be related to the identified risks which need to be studied carefully in the evaluation. The conceptualisation of a good tool is important because it has implications, for example, for the measures of success utilised in the evaluation which constitutes the core of the validity of the evaluation procedures.

The object and main goals of this research concern the evaluation of tools, but, as stated above, in order to treat the evaluation in a contextually meaningful way, other topics also need to be investigated. These are: nature and characteristics of operating work, conceptualisations of good tools, and the design and implications of tools. These topics expand the object of the research accordingly (Figure 1).

The main goal of the research is to investigate, test, and develop research methods which enable evaluating NPP control room solutions in a manner which is both relevant for design and considers the safety implications of the new solutions. For this purpose, it is necessary to acquire a profound understanding of the characteristics of operating work and also to investigate and develop the conceptualization of a good tool.

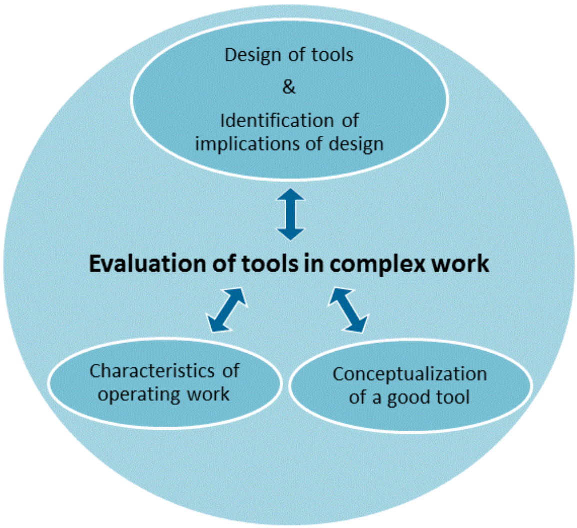

Figure 1. The object of the research is the evaluation of tools, which supports the design of tools and is based on an understanding of the characteristics of operating work and a conceptualisation of what constitutes a good tool in the specific context.

1.3 Research themes

In what follows, I develop the goals of this research in three separate themes which correspond to three of the four objects of the research presented above (Figure 1). The characteristics of operating work are not treated as a separate research theme due to the amount of existing research on this topic. Instead, literature concerning operating work is reviewed in Chapter 2. The remaining objects of research are investigated both through the literature and empirically under three themes, labelled 1) Conceptualization of a good tool in complex work, 2) Methods of usability evaluation of control room UIs, and 3) Implications for the design of control rooms.

The research themes are further specified and focused through the analyses of relevant research in Chapter 2 of the dissertation, and the detailed research question are presented accordingly in Chapter 3 (p. 68).

1.3.1 Theme 1: Conceptualization of a good tool in complex work

In evaluating the appropriateness of control rooms and control room systems, it is evident that it must be clear what constitutes a good control room. Within the first research theme in this dissertation, I want to emphasize that conceptualising the characteristics of an appropriate control room is a central precondition of any evaluation work, especially in support of design. The characteristics of the good control room set the objective for the design and thus the reference for the evaluation. The choice of concepts is vital as it largely dictates what kind of data is collected, the analyses conducted and thus what kind of results may be achieved.

The starting point in the conceptual research theme is the concept of usability. As a concept, usability addresses the quality[1] of the system *in use*. It is quite obvious that a good control room is such that the operating crew of a plant is able to use it in order to fulfil the goals of the work. In the usability paradigm, it is specifically the outcomes of use which determine the appropriateness of a system for its purpose [Cockton 2006]. Usability as a concept has a definition written into an international standard such as: usability is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [ISO 9241-11 1998]. The definition implies that the outcomes of use can be measured in terms of the effectiveness, efficiency, and satisfaction mentioned. In the context of control rooms, this would mean that the outcome of use, measured, for example, as the success of the process control work, would be the standard to which the appropriateness of the system is compared. The standard definition of usability emphasizes contextuality of use, it implies that the measures, e.g. efficiency, are always to be defined contextually. The goal of the dissertation is to explore what the context of complex sociotechnical systems, i.e. nuclear power production, means for the contextualisation of the indicators of usability.

Usability research has over the past decade found new perspectives on quality in use by developing the concept of User Experience (UX). UX research emphasises the overwhelming profoundness with which technologies have transformed people's everyday lives. Research in UX is typically conducted concerning consumer products. Within the dissertation work, my goal is to explore the meaning of the concept of UX in the context of complex sociotechnical systems.

Therefore, the goal of the research theme concerning conceptualization of a good tool is to explore the meaning of two central concepts; usability and UX in the context of complex sociotechnical systems.

[1] Quality is a heavily loaded concept in engineering science, and e.g. Cockton [2006] prefers to use the terms value or worth. This makes a distinction with the quality engineering paradigm.

1.3.2 Theme 2: Methods of usability evaluation of control room UIs

The second main research theme addresses the methodological implications of the conceptualisations of usability and UX in the context of complex systems. The particular interest is in the methods of evaluation which have a key role in providing feedback to design.

Probably a countless number of different kinds of usability evaluation methods can be found in the literature, but the standard methods are not as such directly applicable to the evaluation of control rooms [Han, Yang & Im 2007]. Usability evaluations may be divided into inspections and empirical user tests [Riihiaho 2000]. The scope of this research covers only the empirical user tests which truly test the system in use. In a usability evaluation, the specification of the "context of use" is often the starting point [ISO 9241 - 210 2010]. The context of use is divided into user, tasks, equipment, and environment. In complex sociotechnical systems, the characteristics of the users are known, as the operators of an NPP constitute a group of whom their education and other background information is known. The concept of task is problematic. In controlling the NPP process, the operators deal with an inherently open system, the state of which is effected by external influences uncontrollable from within the system (e.g. extreme weather conditions), system-internal characteristics, and the activity of the operators. This means that tasks, as sequences of actions, cannot be pre-determined entirely. Therefore, the concept of task is not suitable as such for evaluation purposes, and methodical conceptual new ideas are needed. In an NPP the equipment and environment are both characterised by complexity to the point that make it impossible to analyse them completely. Therefore, the specification of context of use, the starting point of evaluation needs to be developed in order to carry out an evaluation in NPP control rooms. Also, methods of data analyses need to be developed when conducting evaluations within complex systems in order to achieve results which are relevant from the points of view of design and safety.

The goal of the second research theme is to identify and test different kinds of evaluation methods in order to develop procedures of evaluation which both support control room design and enable taking the safety perspective into account in evaluations.

1.3.3 Theme 3: Implications for the design of control rooms

The third research theme addresses the empirical reality of NPP control rooms as they are now and the directions in which they are developed in the context of digitalisation.

As the practical empirical research in this dissertation is conducted following the modernization efforts of the MCRs of Finnish NPPs, it should also produce knowledge concerning the current and developing control room solutions. This means that it should be possible to analyse and identify both the problems and benefits of the solutions. The modernization efforts will benefit from these results because, during the long process of modernization, it is also possible to channel design efforts in an already ongoing project. The results achieved in the third empirical theme will constitute the practical value of the research.

The goal of the third research theme is to identify the problems and benefits of the current and developing control room solutions.

1.4 Research process

This subsection concerns the research approach and proceeding of the research in meeting the research goals.

1.4.1 Research approach

As presented above, the motivation for this research is in the practical problem of evaluating the appropriateness of NPP control rooms. I have broken down the practical problem into three intertwined research themes as described in the previous section. The general aim of this research is to produce *prescriptive knowledge,* i.e. new knowledge that is field problem-driven and solution-oriented, and deals with certain organizational problems. This is a distinction of *descriptive knowledge* creation which is often theory-driven and focuses on describing and explaining phenomena in terms of independent variables. [Van Aken 2005]

Van Aken [van Aken 2004] distinguishes three categories of scientific discipline: 1) the *formal* sciences, such as philosophy and mathematics, 2) the *explanatory* sciences, such as the natural sciences and major areas of the social sciences and 3) the *design* sciences, such as the engineering sciences, medical science and modern psychotherapy. Design Science approach solves some of the problems [Bennet, Flach 2011] with the traditional dichotomy of dividing research into basic and applied. The research presented in this dissertation, within the themes presented above, can be interpreted within the frame of design science. The notion of design science is mainly utilised in the field of information systems (IS) research [van Aken 2004, March, Smith 1995, Hevner et al. 2004] and in the engineering sciences, in which it may also be called constructivist research [Crnkovic 2010].

Design science is about developing knowledge for the design and realization of artefacts [van Aken 2004]. Design science consists of two design processes which are *building* and *evaluating* [March, Smith 1995]. Within the frame of my research, I am *building* an evaluation approach and *evaluating* it by applying it in real cases. The building part starts within the existing evaluation methodologies and a review of development needs within them. After identifying this “research gap”, I will build an evaluation approach which will be utilised in real world evaluations within NPP control rooms. The evaluation part concerns the applicability and benefits of the developed approach. Based on analytical evaluations, I will improve the approach in the course of the progress of the work, and develop it further to be applied in improved form in the next evaluations.

As research outputs, design science produces constructs, models, methods, and instantiations [March, Smith 1995]. Constructs provide the language in which problems and solutions are defined and communicated [Hevner et al. 2004, citing Schön 1983]. Models use constructs to represent a real-world situation and represent the connection between problem and solution, enabling exploration of the effects of a design decision and changes in the real world [Hevner et al. 2004]. In the frame of my research, I have defined the first research theme within which I am exploring concepts to reflect the appropriateness of control rooms and finally building a model of systems usability. Methods define processes [Hevner et al. 2004]. Within the second research theme I am building and evaluating a process of NPP control room evaluation. Finally, instantiations prove that constructs, models, and/or methods can be implemented in a working system [Hevner et al. 2004]. In my research, the instantiations are the actual evaluations which are conducted so as to benefit control room design and to evaluate the appropriateness of developed control room solutions.

According to this interpretation, the research approach in which I am seeking to address the research goals, falls under the broad umbrella of design science.

1.4.2 Proceeding of research

The chronological process of the research is presented in Chapter 4 which presents overview of the empirical work but a schema of the process of the work (Figure 2) is presented here as it is related to the research approach.

The research proceeds through iterative steps in which evaluation methods are first built, based on theory and my own experiences in other fields. After instantiating methods in practical evaluations, reflections of the appropriateness of the methods are conducted in order to improve the methods. In the improvements theories and experiences will again be taken advantage of.

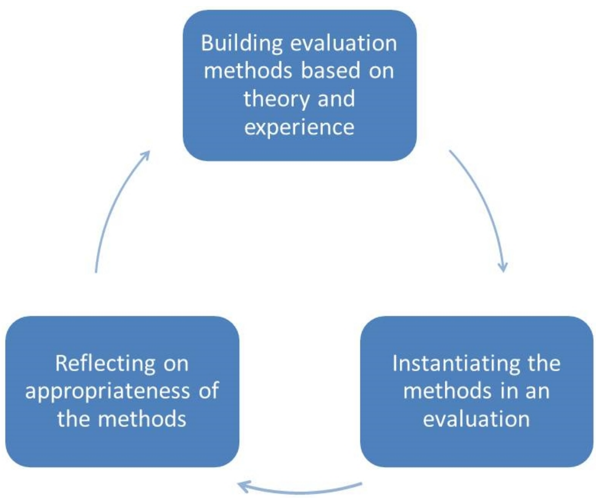

Figure 2. Schematic proceeding of the research.

1.5 Relevant fields of research

The research presented in this dissertation is at the intersection of several scientific disciplines which all have their own distinctive research approaches and methodologies. In what follows, I will (very) briefly introduce the fields of research which have been inspirational for the development of the control room evaluation approach. Also, I discuss briefly the typical research approaches of the fields and make the connection to the approach used in the dissertation.

The most obvious research field related to the topic of this dissertation is the Anglo-American tradition of human factors and/or ergonomics[2] (HF/E). The International Ergonomics Association defines HF/E as "*the scientific discipline concerned with the understanding of the interactions among humans and other elements of a system, and the profession that applies theoretical principles, data and methods to design in order to optimize human wellbeing and overall system performance*" [IEA 2010]. Traditionally HF/E, at least in the domain of NPP, is based on positivistic research approaches[3]. This means that values such as objectivism in the absolute sense [Megill 1997] are emphasised. This kind of research typically utilises controlled experiments and statistical analyses of quantitative data as methods for making conclusions concerning the object of research, for instance the control room. Such approaches are powerful when there is a need to prove, for example, the superiority of one solution over another. The downside may be that the abstractions and simplifications, which the conduct of controlled experiments requires, may in some cases exclude the complexities of the real world from the experiment. Also design implications may be difficult to deduce [Norros 2013]. This is a challenge, as it is absolutely necessary to make the connection between the design and the performance in order to utilise the results for the benefit of novel digital systems for process control. Therefore, the human factors approach alone is not a sufficient research approach for the dissertation work.

The French tradition of ergonomics poses a different approach to developing human-technology systems. The starting point is the concept of activity (*activité*) and the research methodologies based on contextual analyses of activity in real situations and conceptual developments [Daniellou 2005]. The researcher (an ergonomist), as an observer of activity, is not only concerned about *what* people do but also about *how* the activity is conducted by the individuals. The question of how helps the researcher to understand the intrinsic qualities of the activity, "how this activity is constructed by a given operator as a response to a given context" (ibid). The research approach emphasises the researchers' profound understanding of the intrinsic structures of activity and his/her interpretation of the in-situ ob-

---

[2] The terms "human factors" and "ergonomics" are used interchangeably.

[3] This statement is elaborated in Section 2.3.1 (p. 53) concerning different interpretations of the concept of safety.

servations as prerequisites for conducting interventions in order to develop the activity.

The cultural historical theory of activity (CHAT), or activity theory (AT) is in short rooted in the Soviet psychology of Leontjev [Leontjev 1978] and Vygotsky [Vygotsky 1978] is one of the most influential theoretical underpinnings in the development of the unique research approach of our own research group: the Core-task analysis (CTA) which is presented comprehensively in a book by Norros entitled "*Acting Under Uncertainty*" [2004]. Consequently, the CTA methodology is the framework within which I am working in each of the research themes of this dissertation.

Human-Computer Interaction (HCI) research and Usability Engineering [Nielsen 1993] together form the next important field of research which is related to the research themes of this dissertation. HCI is an interdisciplinary research area which aims to improve the understanding of the relationship between people and technology in order to improve design [Sears, Jacko 2008]. Usability engineering addresses specifically the engineering process of "making usable products". Closely related fields are also computer-human interaction (CHI), Computer-supported collaborative work (CSCW) and Information Systems (IS) [Grudin 2008]. The strength of HCI, usability and related approaches is their inherent orientation to design in which the starting point and a crucial driver is the human use of the systems. Also, in practice this viewpoint complements the traditional HF/E approach. HCI as a discipline has been able to bridge between university and society, and between scientific acceptability and practical relevance [Kuutti 2007] and, therefore, it is relevant in finding approaches to the practical issue of evaluating control room systems appropriateness.

The research work presented in this dissertation also contributes to the scientific discussion in the frame of cognitive systems engineering (CSE). CSE is a field of research rooted in the late 1970s and early 1980s when it was realised that changes are needed in human-machine-system design because of the increasing level of automation and the resulting changes in human operators' work [Hollnagel, Woods 1983]. CSE is based on the early works of Jens Rasmussen, for instance the distinction of the skill-, rule- , and knowledge-based behaviours [Rasmussen 1983] and the abstraction hierarchy [Rasmussen 1985]. Important developers of the field are Erik Hollnagel [e.g. 2006], Morten Lind [e.g. 1999, 2003] and Kim Vicente [e.g. 1999] among others. In a recent paper, Upton et al. [2010] presented an interesting characterisation of the different research approaches which have been reported in studies positioned under the umbrella of CSE. The research approaches can roughly be divided into *empirical-positivist* and *empirical-hermeneutic* according to their epistemological stance reflected in the way of conducting analyses of behaviour of socio-technical systems [ibid]. The empirical-positivist approaches consider human behaviour in some way a priori determined by the constraints and goals of the work system. Upton et al. [2010] include in these approaches cognitive work analysis (CWA) [Vicente 1999], abstraction hierarchy [Rasmussen 1985], and multilevel flow modelling (MFM) [Lind 1999]. The empirical hermeneutic approaches on the other hand, "examine work practices to derive behavioural patterns that can be used to describe system functionality. It views human behaviour as a product of interpretation, which requires an understanding of social and historic context. This can be considered a *post-modern* perspective as interpretation, or the establishment of meaning, requires the analyst to observe and explain the actions of workers." [Upton et al. 2010]. Upton et al. (ibid) include approaches such as distributed cognition [Hutchins 1995] and activity theory in the empirical-hermeneutic approaches, and further claim that in CSE the positivist approaches have mainly been used in analyses of causal systems (e.g. chemical reactors), whereas the hermeneutic approach is used in analyses of so-called intentional systems (e.g. emergency dispatch) in which the human actor brings profound variability to the system behaviour. The conclusion of the authors is however, that the dichotomy of the approaches is unnecessary and a mixed model approach including characteristics of both could be beneficial. The benefits of utilising complimentary analyses methods have also been identified by Jamieson et al. [Jamieson et al. 2007] who claim that in eliciting information requirements for UI design it is beneficial to utilise analysis approaches of different origins.

My personal stance concerning the research approach coincides with that of previously cited authors coming from the discipline of CSE. I believe a mixed model approach which applies methods from both positivistic and hermeneutic approaches provides a multitude of perspectives concerning the phenomena under investigation. I believe it is important to make sure that a practical problem, for instance the appropriateness of the control room solution, is examined from multiple viewpoints, which in practice means that data is collected utilising methods originating in different research approaches. In the end, it is the task of the researchers and the subject matter experts to triangulate the meaningful conclusions from the data which will provide help in solving the practical problems to which answers are sought in the research. Therefore, the empirical research work conducted for the dissertation in the form of different studies is conducted utilising a mixed model approach. And thus, the methodology I am proposing for the evaluation of control rooms within complex socio-technical systems represents the same research approach.

1.6 Structure of dissertation

The structure of this dissertation is as follows.

Section 1 is an introductory section which describes the practical needs for development of control room evaluation methods. In Section 1, the research goals and research themes are also presented along with the research approach and related process. Also, related fields of research are introduced.

Section 2 presents the related research as it can be found in the international scientific literature. In Chapter 2 the research gap to be filled is identified.

Chapter 3 outlines the detailed research questions which have been formulated based on the research themes and the initial goals, and specified based on the review of existing related research.

Chapter 4 gives an overview of the empirical work conducted for the dissertation. The empirical work is presented in the light of individual studies conducted and the materials and methods utilised in the studies.

Chapter 5 presents the main results of the research work. The detailed results are evident in the papers of the dissertation. Answers to the research questions are given as a summary to Chapter 5.

Chapter 6 is the discussion in which the results are discussed in the light of a contribution to the practice of conducting control room usability evaluations and the theory concerning quality of tools in use. The chapter ends in a conclusive subsection (6.6).

2. Related research

This chapter presents and discusses the related research relevant for the dissertation. The section is divided into 4 subsections in the following manner.

First, in Subsection 2.1, I review empirical studies the results of which are *descriptive characterizations of operating work* carried out in control rooms. Together, the studies cover a wide range of domains, but the emphasis in the review is on the nuclear industry. The motivation to review literature concerning operating work in general was to discover what, according to the empirical findings, constitutes operating work in the context of control rooms. In the review of the descriptive studies, I concentrate on the identified inherent characteristics of the operating work and the specific discoveries relating to characteristics which actually make the work complex and demanding apart from the obvious possibility of hazards. In this section there is also a slight historical undertone, as I present the studies conducted in 1980s and 1990s roughly in chronological order.

In the second subsection (2.2), I review empirical studies concerning *evaluation of control rooms* both in NPPs and a few other relevant domains. Evaluation is the main object of the research presented in this dissertation (see Figure 1). For the NPP domain I have tried to comprehensively include the control room evaluation and validation studies conducted after the year 2000 and which have been published in international scientific journals[4]. Even though NPP operating work is known to be more demanding than work in other process control domains [Bobko et al. 1998], some seminal works in other domains are also included which have relevance to the research questions of this dissertation. In reviewing the articles, I sought answers to the question: How is operating work treated in the empirical studies in which control room systems are evaluated and whether one can identify an explicit or implicit conceptualisation of what constitutes a good control room in the reports of the studies. Subsection 2.2.3 contains a discussive summary on the findings (p. 51).

---

[4] In order to not have to evaluate the quality of the presented evaluation study, I have only included first-hand empirical evaluations from journal level publications. This demarcation is surely reflected in the number of studies.

In Subsection 2.3 I consider the threads of development taking place in the fields of human factors and ergonomics (HF/E) and human-computer interaction (HCI) which are both relevant and provide inspiration for the development of the evaluation approach called for in the research goals of the dissertation work.

In the synthetic last section (2.4), I summarize the related research in the light of the research themes in order to detail the final research questions in the chapter following.

2.1 Characteristics of operating work

This subsection presents studies in which the nature and general characteristics of operating work have been investigated. The motivation to present the studies and their findings is twofold: firstly, to demonstrate how the quest for understanding complexities involved in controlling dynamic phenomena has evolved over time, and secondly, to present the current understanding of what makes operating work demanding.

Development in the field of human factors in the NPP domain made a step advancement in the early 1980s after the Three Mile Island accident, the cause of which had been declared to be related to human performance in severe conditions. Then, it was realised that human factors issues, such as human operators' ability to conduct the operating work in all plant conditions, needed to be examined in control room design [Maddox, Brickey 1982] and methods were sought in order to study the activities of the operators [Leplat, Hoc 1981].

A central theme in empirical studies then, was identification of human error mechanisms [Reason 1990] based on the prevailing information processing theory of human cognition. Looking back on this, it seems that in the early studies the NPP operating work was treated somewhat mechanistically. For example, it can be interpreted that work was treated as consisting of intentional actions consisting of planning and execution phases continuously following each other, because the errors were considered to take place in one or the other of the phases. This is, of course, a natural implication of the theory base utilised in the studies. [Cacciabue 2010]

During the 1980s, based on the results of empirical studies, important contingencies were nevertheless innovated and implemented for operating plants, such as emergency operating procedures (EOPs) [see e.g. Filippi 2006 for the French approach to procedure design] and expert systems [see e.g. Hoc 1989], to aid operating crews' decision making.

In Finland, the human factors research efforts were directed towards supporting PSA studies of the operational plants, designing operator training, and defining the cognitive content of operational tasks [Wahlström, Norros & Reiman 1992].

2.1.1 Situated activity shaped by the environment

Towards the end of the 1980s and early 1990s the underlying assumptions concerning human action and communication started to receive more and more critique, an example of which is Lucy Suchman's seminal book *Plans and Situated Actions* [1987] which, through the results of anthropological ethnographic studies among users of "intelligent machines", proves the situated nature of human actions which constitutes a different understanding of the basic principles regulating human activity from that utilised in empirical studies of the time. Instead of an execution of prescribed plans, the theory of situated action understands human action as being based on continuously developing plans, the adjustment of which is based on the feedback from the environment. The findings of Suchman also brought understanding to questions relevant for understanding the complexities of operating work in the nuclear domain. For example, it could be concluded that writing a procedure cannot be the only remedy for a known difficult operative situation, because the mere existence of a procedure does not guarantee a successful operating activity in a situation.

The ethnographic approach in research, aiming for understanding the real characteristics of everyday working life spread to the field of control room work also. Heath and Luff [1992] studied the collaborative operating practices, in situ, in the London underground line control rooms, and demonstrated how "sociological and naturalistic analyses of work practice can inform the design of a tool". The research, conducted with ethnographic methods, was ground breaking in that it changed the view of the operating work by revealing the tacit work practices not described in procedures. It also demonstrated the everyday challenges and complexities with which the operators are engaged, for example, the ways in which operators monitor and follow each other's conduct and systematically distribute information among the shift. With practical examples, Heath and Luff made a step contribution to the understanding of how work flows actually emerge in situations in ways that cannot be fully described in advance. A similar approach and findings in an NPP domain are reported by Theureau et al. [Theureau et al. 2001].

Also in studies conducted with an ethnographic approach, Hutchins [1995] made a novel conceptualisation by framing the distributed nature of human cognitive activity. By studying the actual practices of ship navigation both in modern and ancient vessels, he discovered the profound connections of human cognitive activity with the resources in the surrounding environment. The theory of distributed cognition is relevant for control room modernizations, because the control room constitutes a comprehensive environment for the operating activity which may be totally transformed in the modernization project.

Core-task analysis (CTA), a work analysis method developed by Norros and colleagues [Norros 2004, see also Klemola, Norros 2002, Klemola, Norros 2001, Nuutinen, Norros 2001, Nuutinen, Norros 2009, Norros et al. 2012] also constitutes an ecological approach to the analysis of human conduct. CTA takes the characteristics of the environment as a starting point. In CTA, the complexity, dynamism and uncertainty features of the respective domain are modelled in order to explain the general demands placed on the operators. These demands, the core-task demands, are those requirements concerning operating work which the operators must always take into account in order to meet the general objectives of work.

The main content of process control work can be described in terms of primary tasks, the most important of which are monitoring and control [Roth, O'Hara 2002]. The complexity of the primary tasks is exemplified by the study of Li et al. [2006] who say that in monitoring an open system (a hydro power plant) the human controller’s work is a mixture of reactive and proactive control. The processes of monitoring for failures and intervening may be characterised as reactive. However, monitoring also involves a proactive search for situations that might become problematic if not handled early. When detected, such problems lead to discussion, decision-making and planning of future action, which often take the form of changing the initial conditions of the process. This shows how tightly the demands of the operating work in situ are coupled to a controlled process situation and regulated by the collaboration and general problem-solving in the work.

The research work emphasises the situated nature of human activity in general, and the role of the environment in shaping operators activity show that dynamic environment in which the operating work takes place, for example the controlled process, is a crucial determinant of the complexity of the work.

2.1.2 Involving multiple levels of cognition

Ethnographic studies have also been conducted in NPP Control rooms Mumaw et al. [2000] report a series of cognitive field studies which were conducted during normal operations in two traditional NPP MCRs. The results of the studies emphasise the *problem solving nature* of the operating work: “In their daily work operators need to identify and pursue relevant findings against a noisy background of the control room UIs and exploit proactive strategies in making important information more salient.” Operators’ monitoring activity was evidenced to be highly knowledge-driven. Vicente et al. [2001] continued the in-situ observations of operators’ work in a more modern control room, and discovered that the same strategies are also utilised in the modern control rooms, but that the strategies may be implemented with different observable behaviours. In addition to verifying the problem-solving demands of operating work, this result also demonstrates the behaviour-shaping effect that tools in use may have.

Although the studies conducted with ethnographic research approaches delivered a great deal of information concerning the everyday work of NPP operators, simulator studies of accident situations are considered necessary for understanding the nature of work in extreme conditions which may pose the greatest threats to safety. Perhaps the research approaches developed for studies of normal work had influenced a comprehensive accident situation study which was conducted by Roth et al. [1994]. The aim of the study was to gain knowledge concerning the role of “higher-level” cognitive activities such as diagnosis and situation assessment in proceduralized conditions. The results of the study did indeed prove the existence of these cognitive functions by demonstrating the mechanisms through which operators may conduct successful process control in the phases of scenarios which are not fully addressed by the procedures. Later, decision making and situation assessment became labelled as non-technical skills (NTS) in emergency management in an NPP among others such as communication, teamwork, and stress management [Crichton, Flin 2004]. These empirical studies together contributed to the understanding that operators work on multiple levels of cognition simultaneously, a An issue which had actually been discussed already by Rasmussen [1983], but the practical examples and implications of which had not been fully adopted in the industry.

Operators' decision-making is an important feature of operating work which has been studied extensively. Crichton et al. [Crichton, Flin & McGeorge 2005] studied on-scene incident commanders' (OIC) decision-making in an emergency situation in the nuclear domain. By conducting a card-sorting experiment, the authors gained an understanding of the complex situational factors which play a role in nuclear OICs' decision making. It was discovered that decision-making is influenced by such factors as: availability of procedures, degree of uncertainty, responsible agent in the situation, and advice being sought from others. This result shows the complexities involved in managing emergency situations; procedures may or may not be available; there is always uncertainty in information available on the situation; there are several responsible agents involved; and sometimes advice can be acquired from outside the immediate systems, which means that system boundaries are not always clear cut.

In a study concerning shift supervisors' decision-making, Carvalho et al. [2005] found that decision-making in managing micro-incidents under normal operating conditions, is primarily based on naturalistic strategies. This means that the supervisor's decision-making can be characterised as being based on pattern recognition, tacit knowledge, and condition-action rules. These are decision-making strategies which may not always be supported by the available tools which are often based on a rational model of human decision-making. In a study by Hayes [Hayes 2011], research in three domains (a chemical plant, a nuclear power station and an air navigation service provider), a situational approach in decision-making taken by experienced operating personnel was also identified. When abnormal situations arise, the operational managers focus on the status of safety barriers rather than consider risks from first principles and develop self-imposed situation-specific limits for decision-making. Shattuck and Miller [2006] illustrate by analyses of decision-making in ship navigation "how decisions made are often a result of the interaction between a variety of technological and human agents and how errors introduced into the complex system can propagate through it in unintended ways". Again, the complexity and naturalistic way of making operative decisions is demonstrated.

2.1.3 Un-describable in operating procedures

In complex socio-technical systems, some tools are intended to become contingencies against failures and they can also be referred to as safety barriers [Kecklund et al. 1996]. They may be social, technical, and/or organizational measures which aim to prevent or stop an undesired consequence [Basnyat et al. 2007/6]. The most important safety barriers are the operating procedures. The complexity of conducting proceduralized work has been demonstrated and operationalized by Park [2009].

Even though the whole industry of nuclear power production is heavily proceduralized, professionals may still sometimes engage in unsafe behaviours [Choudhry, Fang 2008] for various reasons. For example, in some domains the utilisation of procedures, due to historical reasons, can be considered as against what Norros [2004 citing MacIntyre 1984] calls the ethos or “the internal good of practice” as e.g. Knudsen [2009] observed with written procedures and seamanship. Even more often, the reason for deviating from the procedures may be in the gap between the procedures and practice [Pentland, Feldman 2008]. Generally, it can be stated that designed contingencies and decision-making aids for problematic situations aim to help operators in the work the nature of which is problem–solving, but in order to be valuable they need to be designed to correctly suit the practices of operating work.

A large study in a digitalised control room was conducted within the research facilities of Electricité de France (EdF). The results concerning utilisation of procedures are reported by Filippi [2006]. She concludes that managing accidental scenarios with EOPs requires several competencies from the operators such as “score reading”, which refers to understanding the instructions and the way to implement them, managing attention resources, building expectancies concerning the future of the process, of the procedure and of the other operators’ procedure. As a result of the studies, the level of control in the computerized procedures was lowered in order to better reflect the realism of the level on which human activity can be prescribed.

Dien [1998] describes an application of intelligent procedure which is neither violence (negligence) of – nor strictest possible adherence to – procedure, but can be characterized as adapting the given procedure to the situation. This means that operators themselves must realise when “the procedure prescription diverts from reality”.

By carefully analysing operating work of different crews in simulated accident situations, Furniss [Furniss et al. 2011] identified multiple strategies which operating crews had developed so as to recover from deviations and adjust to environmental variations. This result shows the adaptive power which operating crews possess and take advantage of when required by the situation, and the courses of action which are needed in the situation but not possible to prescribe in procedures.

The studies concerning operating with procedures clearly demonstrate that utilising procedures is not straightforward and un-problematical as such, but that it demands special skills and knowhow from the operators.

2.1.4 Collaborativeness

Operating work in an NPP is conducted in teams and is thus collaborative, and this constitutes an important characteristic of the work. The ergonomic field study conducted by Carvalho [2006], based on analyses of communications during micro-incidents in normal operations, shows how operators use verbal exchanges to produce continuous, redundant, and recursive interactions to successfully construct and maintain individual and mutual awareness, which is paramount to achieving system stability and safety. Such continuous interactions enable the operators to prevent, detect and reverse system errors or flaws by anticipation or regulation. Farrington-Darby & Wilson [2009] emphasize the significance of team work by suggesting that ‘‘the social’’ should not merely be regarded as a situational variable, albeit a very powerful one, but as a key focus of study in control room development. in its own right In a study by Kontogiannis and Kossiavelou [1999], efficient teams were observed to adopt several strategic adaptations such as thinking ahead of adverse events in less busy periods, building contingency steps in their plans to cope with later events and errors, and switching to more analytical strategies when time pressure is not so great. There are also several other empirical and analytical considerations concerning collaboration issues in operating work: characteristics of collaboration [Luff, Heath 2000], team skills [O’Connor et al. 2008, Kim, Byun & Lee 2011]; team situation awareness [Stanton et al. 2009]; common frame of reference [Hoc, Carlier 2002], and diversities in decision making and strategies of teams [Patrick, James & Ahmed 2006].

2.1.5 Co-operation with intelligent tools

The role of tools in operating work is an interesting research topic within the scope of this dissertation. The main tool in operators’ work is notably the user interface to the automation system residing in the control room. Paunonen [Paunonen 1997] investigated the roles in which process operators use the automation and control system. He found out that, in addition to the self-evident role of monitoring process information and conducting operations of planned tasks (the primary tasks [O'Hara et al. 2002]), the automation system is also used for the management of abnormal situations, learning and process development, for instance optimization, communication, and knowledge transfer. Creative uses of tools for awareness and monitoring are also reported by Luff et al. [2000].

These findings together show that the role of control room systems is actually wider than the immediate process control needs, and thus control room modernizations have more profound effects on the operating work than might be imagined if only monitoring and operations were considered.

Furthermore, the relationship between the human operator and the automation system may nowadays be characterised as co-operation [Hoc 2001a]. Automation systems have become automatic and intelligent to the extent to which many operations are carried out completely autonomously. These increased capabilities have changed the relationship between the human and the machine, and in order to keep the human operator "in the loop" new ways of communicating about the autonomous activities of automation need to be developed.

2.1.6 Requiring of psychological resources

Professional operators are process experts of their respective domain. Nuutinen [2005b] studied the development of expertise of NPP operators, and discusses how shifting the focus of research to the emotional side of human actions and cognition creates new perspectives on the problem of how to support the human operator in the control of rare disturbances. She concludes that the key challenge and motive for the trainees in the development of expertise is not only to achieve an adequate degree of competence but also to construct confidence in being able to cope with the potential disturbance situations. This *expert identity* is an important psychological resource in the work.

It is also evident that operating work is psychologically demanding in inducing stress and workload in the operators [Stanton et al. 2010, Suomen Automaatioseura ry 2010, see e.g. Wickens, Hollands 2000]. Studies concerning these very much studied psychophysiological phenomena are not reviewed here.

Expert identity may be affected by a control room modernization in which the tools are profoundly modified, because the skills in using the tools may be shattered and must be re-constructed in order to maintain sufficient confidence in one's own skills.

2.1.7 Embedded in a social and cultural context

The operating activity in an NPP is embedded in the social and cultural context of the whole industry and more specifically of the respective NPP. Studies of organizational decision making [Carroll, Hatakenaka & Rudolph 2006] safety culture and safety climate [Carvalho, Santos & Vidal 2006, Reiman, Oedewald 2007/8, Hahn, Murphy 2008, Martínez-Córcoles et al. 2011], and operating activity [Theureau et al. 2002] have stressed that the level of safety of production is not only observable through possible negative outcomes such as accidents and disasters but is also evidenced in the people's conceptualisations and appreciations concerning their own work and values of the organization.

2.1.8 Summary of characteristics of operating work

To summarize the descriptive studies of operating work in an industrial plant, and more specifically in the safety-critical context of an NPP, it can be stated that operating work can be broken down and thus viewed from very many different perspectives and on different abstraction levels. Operating work can be characterised as a complex problem-solving, knowledge-driven activity which involves multiple levels of the actors' cognition simultaneously. It is tightly coupled in the operative situations and with the current available operating tools. This means that the work is complex and only a portion of it can be described in procedures. One of the key characteristic is collaboration. The successfulness of operating work is affected by operators' expert identity and emotions, and the wider cultural context of the organization and the whole industry.

My background assumption is that it is not uncommon that the complexity of the human operators' work and the diversity of viewpoints from which it can be investigated, is not considered to the full extent when technological advancements and modernizations of systems are carried out[5]. In the next subsection, empirical studies concerning evaluation of tools are reviewed in detail.

2.2 Empirical studies in control room evaluation

I this section I review empirical studies concerning the evaluation of control room systems and user interfaces utilised in process control work in an NPP and some other contexts. In reviewing these studies, I was especially interested in whether the diversity and complexity of operating work, discovered in the studies referred to in the previous subsection (Section 2.1), is reflected in methods of evaluating the successfulness of the new operating tools. The aim was to find how control room evaluations are being conducted in practice within the industry and academia.

2.2.1 Empirical studies in the NPP domain

Empirical studies concerning the evaluation of control rooms published in international scientific journals were reviewed. The scope of the studies was variant: Some investigations concerned the totality of a control room and some only a particular system or a feature of a control room.

5 For example, in our own study concerning subway automatisation [Karvonen et al. 2011] the role of the train driver not considered to the full extent.

2.2.1.1 Studies concerning the totality of the control room

Studies in which the object of investigation was *the full control room* were reported by Ha et al. [2007], Chuang & Chou [2008], Hwang et al. [2009b], dos Santos et al. [2009], Gatto et al. [2013], and Jang et al. [2013b].

Ha et al. [2007] describe the human performance parameters utilised in the validation of an APR (advanced power reactor) 1400 plant to determine whether the main control room user interface design acceptably supports safe operation. Plant performance, personnel task performance, situation awareness, workload, teamwork, and anthropometric and physiological factors were utilised as performance indicators. The authors make a distinction between *product* and *process* measures. Product measures refer to the plant level outcome of human activity whereas process measures indicate qualities of human activity in order to achieve plant performance. Hierarchical task analysis, based on information in the procedures, was utilised as a method to identify the relevant human actions needed in the evaluation scenarios.

Chuang & Chou [2008] give a report of the extensive HFE V&V tests which were being conducted for the Lungmen NPP in Taiwan. The report does not reveal the methodical details of the studies, but describes, on a general level, a series of tests. In the first round of testing, HFE experts acted as participants (operators) in the tests. In the later rounds the prospective operators of the new plant were brought in to act as operators in the empirical tests. The information concerning the appropriateness of the new VDUs was collected using the concepts of usability, monitoring and detection, situation assessment and awareness, workload, and communication and teamwork. The authors describe a sequence of tests which were conducted in simulator settings and during which the progress of the important HFE parameters were followed from one test to another. Overall, the authors describe that the efforts needed in design and V&V are tremendous. The approach used in the evaluation, based on the indicators utilised, seeks to understand the connection of the HSI designs to the operating performance.

Hwang et al. [2009a] studied the main control room of NPP4 in Taiwan prior to the commercial use of the plant in order to ensure safety and the efficiency of operations. The study was conducted after the operators' three-week training period. In the study, task analyses and structured interviews with the operators were conducted. Altogether, 14 display-related and four control-related problems were identified in the new control room. Operators selected the most important identified problem by selecting which problem required most immediate attention[6]. In the study, corrective measures were also identified. In the published article about the

[6] The top three problems were 1) Inaccuracy of water level indication on a particular screen 2) Poor perceivability of a decimal point in the dry-well pressure monitor and 3) Incomprehensibility of alarm abbreviations. The first problem relates to the algorithm utilised in calculating the value which is presented to the operators, whereas the latter two problems are related to interface design. All three problems are quite severe and may potentially lead to erroneous interpretation of the information by the operators. [Hwang et al. 2009b]

study there were no reflections concerning the suitability and benefits of the utilised methodology. It can be concluded that, in this study, the operators' expert opinions concerning the new system were utilised as the main data. The problems which were identified in the main control room were addressed from the point of view of latent human errors, meaning that, in the interviews, the operators discussed the features of the new control room from the point of view of error possibilities.

Dos Santos et al. [dos Santos, Isaac José Antonio Luquetti et al. 2009] used an HF questionnaire to investigate a design of an NPP control desk of a research nuclear reactor. The actual aim of the article is to publish validation results of the developed questionnaire, but simultaneously the evaluation study through which the questionnaire's validity is demonstrated is presented. The human factors questionnaire consisted of fifty questions about panel layout, panel labelling, information displays, controls and alarms. Experts (designers, HF experts, and operators) completing the questionnaire scored the questions on a conformance scale and on importance weighting from the point of view of safety. In this study, the effects of the UI design on the operating activity were not explicitly considered, but the expert users completing the questionnaire may have operating practice-related background assumptions affecting their ratings.

Jang et al. [Jang et al. 2013b] studied human error probabilities in advanced digital control rooms with a simulator study. As an operational environment, a compact simulator was used. The participants were students majoring in nuclear engineering. The error possibilities were analysed prior to the experiment and checklists for recording the errors were prepared. In the check lists, the observers categorized the errors that occurred during the conduct of the scenario. In the methodology, the users did not comment on their own performance, but the reason for the error was analysed by the observer.

Gatto et al. [Gatto et al. 2013] utilised a virtual reality model to assess the physical layout of a control room. A loss of coolant accident (LOCA) was postulated to be used in the evaluation in a totality of three different operating conditions. The respective EOPs were utilised as task model. The simulation was conducted using avatars playing the roles of the supervisor, primary operator and secondary operator according to a script. Operating time was utilised as an indicator of the successfulness of the evaluated design.

2.2.1.2 Studies concerning alarm systems

Empirical studies concerning *alarm systems* were reported by Norros and Nuutinen [2005], Hwang et al. [2008a], Huang et al. [2007], Jang et al. [2013a], Huang et al. [2006], and Lin et al. [2010a].

Norros & Nuutinen [Norros, Nuutinen 2005] studied the usability of a safety information and alarm panel (SIAP) utilised in an NPP MCR. The aim of the study was to investigate whether the SIAP was an appropriate tool for NPP operations in emergency situations. In the experiment, process performance and operator practices were utilised as indicators of the success of the new technology. Operator practices were operationalized as specific behavioural indicators identified in the data. The experiment was conducted as a series of simulator tests in which professional operators acted in their operative roles. The data collected and the completed analyses allowed the authors to make conclusions concerning the effects that the new system will have on the development of the working practices of the operating crews.

Hwang et al. [2008a] studied an advanced alarm system by engaging university students and staff member in the task of conducting NPP process control-related isolated operating tasks. The aim of the study was to find out about the benefits of the proposed pre-alarm concept. The variables used to study and demonstrate the benefits were task performance, workload, and team situation awareness. Within these evaluation conditions, the authors discovered that use of the pre-alarm system reduced the occurrence of “real alarms” and also operators’ work load but there was no effect on the team situation awareness. In the paper, there are no reflections concerning the validity of the results in the light of the methods utilised.

Huang et al. [Huang et al. 2007] studied the interaction of operators and alarm systems in order to inform the alarm reset design of a new plant. In the study, two experimental conditions were compared: a manual reset of alarms and an auto-reset of alarms. Participants were experienced operators (experts) and operators in initial training (novices). Operation time, situation awareness (subjective and objective), task load, and subjective ratings were used as indicators of operator performance. In addition, suggestions for improvement were gathered from the participants. Statistically significant results in comparing the two automation modes were achieved for task completion time and effort in favour of the automatic reset. In addition, minor differences between the expert and novice users were identified.

Jang et al. [Jang et al. 2013a] conducted an alarm reduction method validation study. Eight subjects took part in the experiment, and cognitive performance in the form of time, accuracy, and situation awareness were utilised as indicators of success. In the results there were no differences in cognitive performance time between the new alarm reduction method and the control condition. Nor were there any differences in accuracy between the conditions. Differences in favour of the proposed new method were found in the situation awareness measures.

Huang et al. [Huang et al. 2006] conducted a study to evaluate the impact of the auto-reset alarm system on the plant performance, operators’ preference and task load. The background assumption was that an automatic alarm reset has the benefit of reducing the operator’s workload by reducing the number of alarms, but at the same time it may also have negative effects. In the test, the auto reset of alarms was compared to the manual reset of alarms. 30 students took part in the actual test. Indicators of success were task performance time, end-of-task subjective rating (including NASA TLX and comprehension test), and end-of-experiment subjective rating concerning personal preference. The results of the study show that there were no differences in task performance between the two conditions. Nor were there differences in the subjective ratings. In the end-of-experiment subjective ratings, the participants preferred the auto reset for multitasking and effectiveness, whereas manual reset was considered better for alarm handling, monitoring, and decision-making.

2.2.1.3 Studies concerning display design concepts

A significant proportion of the studies that could be identified in the literature search was dedicated to demonstrating the benefits of particular *design concepts* such as ecological interface design (EID). In these studies, the system under investigation as such was not the focus, but the manner in which the UI had been designed and the design principles utilised in the design process, were intended to be the actual objects of evaluation. Such studies were conducted by Naito et al. [1995], Riera [Riera 2001], Lau et al. [2008], Burns et al. [2008], Kim et al. [2012], and Yang et al. [2011].

Almost twenty years ago, Naito et al. [1995] conducted a study to validate EID-based display designs. The main purpose of the validation test was to confirm the validity of the MMS to support the operators during plant transients. Professional operators took part in the test. After utilising the new system in selected scenarios, the operators and the system designers conducted discussions to identify the information balance (supply and demand) during the scenarios. In addition, the operators were interviewed and asked to complete questionnaires. The results of the test indicate the positive features of the displays tested and support for the operators' supervisory control tasks.

Riera [Riera 2001] evaluated an interface of a supervisory system for a nuclear fuel reprocessing system. The system had been designed based on analytical models of human operator's monitoring and control tasks developed by Rasmussen [Rasmussen 1983], Sheridan [Sheridan 1992], and Hoc [Hoc 1996]. The aim of the design had been to develop tools which promote the active behaviour of the human operator. In the evaluation, seven professional operators took part in the test the approach of which was mainly qualitative. All the operators conducted 4 simulations containing a system failure. The scenarios were of two types: failures difficult to detect and explain and failures less difficult to evaluate. Four different UI compositions were tested, ranging from the current system (mimic display) progressing towards the most advanced by adding newly developed information presentation features to the composition. In the scenarios the operators could choose which displays they utilised and indication of success of a design was the extent to which operators chose to use it. The results of the study show that in terms of usage the advanced displays proved successful: Some operators chose to use them practically exclusively. The results also show that operators created innovative ways to utilise the systems which had not been anticipated by the system designers.

An EID experiment was conducted at the HAMMLAB facility at the Halden Reactor Project in Norway in 2006[7]. The results are reported in multiple articles and those describing the methods in evaluation are by Lau et al. [2008] and Burns et al. [2008]. In the experiment, the operator task performance support of an EID-based design of a BWR secondary side display system was investigated. In the study, EID displays were compared to traditional mimic displays and advanced mimic displays in a simulator setting in which six operating crews conducted operating activity in 6 scenarios. Three of the scenarios were procedure-guided and three “knowledge-based”, characterised by unanticipated failures. During the simulator run a process expert scored operator task performance with scenario-specific indicators relating to crew performance in detection, inference, action, and teamwork and communication. The index gained by this method for crew performance reflects the discrepancy between operator performance and predefined optimal solutions to scenarios. Workload was measured utilising a subjective rating of the task-related difficulties operators are experiencing during the scenario. The results of the study show that both of the advanced display types (EID and advanced mimic) induced lower increases in workload in the transition between detection and mitigation phases of the scenario. Also EID was marked as advantageous in the detection phase of the knowledge-based scenarios when actual task performance was utilised as a performance indicator. The success of producing evidence in support of EID was greatly dependent on the success in defining the scenario-dependent performance indicators. Also, the process expert judging crew performance with the pre-defined indicators is in the key position to enable successfulness of the experiment. The authors discuss the limitations of the study which reflect the general difficulties in evaluating control room systems from the human factors point of view such as the hybridity of the real industrial solutions, the effort demanded in conducting studies, and limited time on training the operators on the features of the new system.

Kim et al. [2012] completed an empirical study concerning an EID-based user interface designed for monitoring the primary side processes of an NPP. The authors explicitly state that it was the effectiveness of the EID display system that was the question in the study. In the methodology, three experienced operators were trained in the specificities of EID in order to guarantee sufficient level of expertise. Situation awareness methods were utilised to gather data on the effectiveness of the displays developed. In addition, time and accuracy were measured. In the experimental condition, using an EID display accompanied by traditional mimic displays was compared to using only mimic displays. Some additional, contextually meaningless, tasks were included in order to simulate the normal working conditions of the operators in an NPP. The results of the study indicate that there were no differences in performance time, performance accuracy or objective situa-

[7] VTT took part in the experiment and was responsible for the qualitative data collection and analyses. The results of the VTT research exploring the cross-effects of situation, operating crew’s way of acting, and display concept are reported by Norros et al. [2009].

tion awareness between the two test conditions. With subjective situation-awareness measures, the EID condition produced better results. The authors point out that they could not evaluate how the new display system would affect the operating activity on a more general level.

Yang et al. [Yang et al. 2011] studied operators' signal-detection performance in VDU monitoring tasks. Two different display modes, mixed mode and consistent mode, were compared in an experiment in which thirteen students acted as operators and conducted detect-and-hit operations with the VDUs in order to simulate a monitoring task. Performance indicators utilised in the experiment were: frequency of miss, reaction time, situation awareness, subjective performance and subjective visual fatigue. In the study it was found that mixed mode VDUs resulted in a higher frequency of misses and longer reaction times, and SA was higher for consistent mode VDUs. Also, subjective performance was higher with the inconsistent mode. The results speak for aiming for consistency in VDU design.

2.2.1.4 Studies concerning automation concepts

Studies concerning automation systems and human-automation collaboration have been reported by Skjerve and Skraaning [2004], Jou et al. [2009b], and Lin et al. [2010a, 2010b]

Skjerve and Skraaning [Skjerve, Jr. 2004] studied the quality of human-automation cooperation in an experimental setting. The experiments were performed in a full-scale nuclear power plant simulator using licensed operators as subjects. The quality of human-automation cooperation was assessed from subjective operator judgements. The experiments demonstrated a clear improvement in human-automation cooperation quality when the observability of the automatic system's activity was increased. The relationship between human-automation cooperation quality and the effectiveness of the joint human–machine system's performance was also explored, but no clear results were found.

Jou et al. [Jou et al. 2009b] report a study in which the mental workload was studied in different automation mode conditions. Subjects of the study were students. The aim of the study was to gather information for appropriate automation design. Task completion time, reaction time, heart rate, work load and error rates were utilised as indicators of the success of automation design. The result found two statistical differences: subjects performed more slowly in a reactor shutdown task with low level automation than with high level automation, and the error rate was lower with low automation. The validity of the results is not discussed in the paper. Recommendations concerning automating certain tasks are made based on the study.

Lin et al. [Lin, Yenn & Yang 2010b] studied the human error types and occurrences of two different automatic modes in order to gain knowledge about appropriate human automation function allocation to inform automation design. The research method was a questionnaire survey in which the professional users could make claims about the probability of errors. Before completing the questionnaire, the participants utilised a system having two different modes of automation for 5 hours. The study results emphasize the benefits of automatic mode from the point of view of reducing human errors. The limitations of the study are not contemplated by the authors, but it is clear that the questionnaire method and the scope of the system studied are insufficient for producing comprehensive information concerning the phenomena studied.

Lin et al. [Lin, Yenn & Yang 2010a] studied the effects of different levels of automation (LOA) on human performance. Human performance was measured as situation-awareness and subjective ratings such as NASA TLX. The experimental results indicated that blended decision-making (LOA 6) generated the lowest mental workload. Furthermore, the SA results indicated a better SA at intermediate LOA and poorer SA at low LOA and full automation.

2.2.1.5 Studies concerning operating procedures UI features

Empirical evaluation of computerised procedures has been reported by Yang et al. [2012] and Huang & Hwang [2009].

Yang et al. [2012] studied the effects of computer-based procedures (CBP) and paper-based procedures (PBP) to human performance. The authors acknowledge that computerizing procedures may have an impact on the roles, responsibilities and interactions of the crew members. Yet, the study only concerns work load and situation-awareness as indicators of human performance. In the study, non-professionals conducted pre-defined tasks with CBPs and PBPs. The result of the study is that CBPs allow faster task completion, lower cognitive load and higher situation awareness. Based on these results, the authors recommend utilising CBPs in NPP MCRs.

Huang & Hwang [2009] studied the effects of computerised procedures and team size on operating performance. 30 students took part in the experiment and performed several simulator runs with the developed teamwork system. Operation time, errors, detected system errors and subjective measures of teamwork skills were utilised as indicators of team work successfulness. In the results, the teamwork system, using the computerized procedure interface resulted in higher operating performance (i.e. shorter operation time and fewer errors) than using the paper procedure interface. Automation bias effect was also identified, as the teams did not succeed in detecting system related failures.

Particular features concerning digital control rooms were considered by Al Harbi et al. [2013] who studied touchscreens. In the study, the effects of soft control on human error probabilities were investigated. More precisely, a particular touch screen technology was studied in emergency management conditions. To measure the success of the touchscreen, the operating errors were calculated for the measure of task completeness. In addition, selected physiological measures were utilised (electroencephalogram (EEG), electrocardiogram, and skin temperature). Participants in the study were students, and it was evidenced that the students with little prior experience of the system produced more errors with the touch screen system than with the control condition. This result was interpreted such that novice operators would be prone to errors with touch screen technology.

2.2.1.6 Summary of empirical studies in NPP control rooms

As a summary of the studies reviewed, it can be observed that all together 15 types of measures of successfulness of control room design had been utilised (Table 1) in the 22 empirical studies reviewed. The most common measures were task performance in the form of time and errors, situation awareness and work-load. The studies concerning totality of control rooms typically utilised a greater variety of measures than studies focusing on particular solutions.

Table 1. Summary of the measures utilised in empirical control room studies in an NPP context. Number in the table refers to the number of studies in which the particular measure was utilised.

Object of evaluation Type of measure	Control room	Alarm system	Display concept	Automation concept	Other (procedures, UI features)	**sum**
plant performance	1	2				**3**
task performance, time	1	2	2	1	2	**8**
task performance, errors	3	2	3	1	2	**11**
situation awareness	2	3	2	1	1	**9**
workload/task load	2	3	1	2	1	**9**
teamwork and communication	2	-	1	-	1	**4**
anthropometric measures	1	-		-	-	**1**
physiological measures	1	-	1	1	1	**4**
usability	1	-	-	-	-	**1**
expert opinion concerning error probability	1	-	-	-	-	**1**
expert opinion concerning safety	1	-	-	-	-	**1**
subjective preference	-	2	-	1		**3**
effect on work practices	-	1	-	-	-	**1**
subjective experience of information balance	-	-	1	-	-	**1**
extent of usage	-	-	1	-	-	**1**
variety of measures	**12**	**7**	**8**	**6**	**6**	

In quite a large number of studies, university students acted as participants, tasks were equated with procedures, and sometimes artificial tasks were also utilised. All these factors, together with the utilised measures, may increase the control required in the experiments, but at the same time they simplify the work of the professional operators to an extent which may be interpreted to lessen the ecological validity [Hoc 2001b] of the studies. The decrease in ecological validity means that the conditions utilised in the study become so different from the conditions of real work that results may lose their meaning. This is a problem because the real complexities of the work may not then be addressed to the extent necessary for development of the systems.

2.2.2 Some empirical studies in other relevant safety-critical fields

In this subsection I have included studies which I have found to represent some aspects of the complexity of real operating work. Review like this cannot be comprehensive in any industrial domain; instead I have focused on studies which have seemed promising in the light of the research themes of this dissertation.

In other industrial domains, apart from NPP, the control room systems started to become digitalised already throughout the 1980s. The prevailing way to present information to the operators in the UIs has been, and still to a large extent is, to utilise different kinds of formats basing on process design. The most typical format is the process mimic display which is a dynamic representation concerning the process status based on the process and instrumentation (P&I) diagram. Jamieson [2007] describes the situation in the petrochemical industry: "The display scheme is based primarily on a process mimic graphic that places some setting and flow values in their physical context. Trending capability and tabular alarm summary pages are provided so as to support the mimic displays. Very little higher order information is made available, and designers rarely employ human factors standards for interface design. In most cases, experienced operators designed these displays based on reviews of piping and instrumentation diagrams and drawing from their prior experience as users." Other display formats do exist however, and they include, for example, task-based displays, automation information displays, functional displays, event lists, sequence displays, etc. [Suomen Automaatioseura ry 2010].

Some empirical studies concerning control room evaluation in other relevant fields were also included in the literature review in order to provide an understanding of the general situation concerning evaluation studies. Reports on methodically interesting and innovative studies which concern empirical evaluation of control room user interfaces in fields that can be characterised as complex work were reviewed[8]. In the review, special emphasis was laid on the way of constellating the operating work in the experiments and discussions concerning suitability and ecological validity of the utilised evaluation methodology.

[8] Some domains I have excluded from the literature survey purely to keep the scope manageable. These are e.g. the medical domain, aviation and military applications. Although they all constitute safety critical work domains, the work environments are not control rooms per se as in the NPP.

Jamieson [2007] conducted an empirical study in the petrochemical industry in which three display types were compared, one of which was an EID display. The study aimed to make a step improvement in ecological validity in comparison to the previous empirical evaluations conducted for the EID concept as an actual full fidelity simulator was used and professional operators took part in the evaluation. In addition, familiar proceduralized events were utilised in the scenarios. The scenarios were analysed from the point of view of how familiar they were to the operators and from the point of view of how anticipated they were, i.e. whether a procedure had been developed for the event. In the study, fault detection and diagnosis times, diagnosis accuracy and control action counts were utilised as indicators of performance. What is more, plant performance was considered utilising measures of material throughput. In the results of the study, the two UIs presenting physical and functional information resulted in significant advances in support of performance in terms of completion time and efficiency. No main effect was found on the level of plant performance.

Li et al. [2006] conducted an evaluation of hydropower system control displays. The task of the hydropower operators consists of monitoring and analysing vast amounts of data from diverse domains and on different time frames so as to control the energy source, to generate electricity, and to meet economic dispatch targets while keeping an eye on transmission status and constraints, and on changes in the market. The demands on the controller are exacerbated by the fact that hydropower companies often serve peak load, which is highly driven by dynamic market forces. In the evaluation study, the authors state that the key concern was the development of scenarios so that they would be indicative of the advances that the new displays would bring to the operating work. The scenarios were developed analytically so as to provide assistance in known risk situations. In the experiment conducted, the functional displays were compared to the current displays utilised in the domain. Performance was measured in the following areas: (1) capturing the problem solving time frame as intended, (2) supporting situation awareness, (3) promoting trust and self-confidence. The results of the study state that evidence could be found to support the development of the new system in the form of operators' comments and questionnaire results. Interesting comments were made by the participants concerning their view of the limitations of the study.

> *[The participants] commented that because the evaluation time was so compressed it was hard for them to ''set up their direction'' [of control], incorporate the new displays into their routines, and fully test their potential advantages. The participants felt they needed more time and greater realism to fully evaluate the value of new displays. In addition, many participants commented that they also needed to test the displays on the equivalent of a normal day, rather than a highly abnormal day, to see how the displays indicate where they are and where they are going.*

Morineau et al. [2009] utilised a similar approach with Li et al. [Li et al. 2006] in that they developed the tasks in the evaluation to fit the designed features of the system, in their case a tide prediction card. The actual study was conducted in an experimental setting in which two groups of participants (experts and novices) utilised three different tide card formats in order to extract tidal information. Participants' time and accuracy in answering 8 different tidal information-related questions were measured. In addition, participants were asked for their personal preference. The results of the study indicated that the novel tidal information format was fastest to use and also resulted in higher scores for preference. The authors claim that the novelty of the format did not hinder its acceptability as might have been expected.

Blandford and Wong [2004] conducted an ethnographic study in situation awareness in an emergency medical dispatching centre. The study reported in the paper involved a number of different methods of investigation, to provide complementary information to yield a broad understanding of the nature of the work of EMD operators. The methods utilised were based on e.g. contextual inquiry [Beyer, Holtzblatt 1998] and the critical decision method [Hoffman, Crandall & Shadbolt 1998] but performance data was also gathered. Findings from the study were presented at a meeting with the control centre managers, supervisors and the EMD operators themselves, so as to verify the results and to correct any misinterpretation of the data. This constellation of research methods makes the approach participatory, emphasising the active engagement of the participants in the research.

A socio-technical systems approach was utilised in a study reported by Stanton and Ashleigh [2000] in which the effects of major changes within the company on team performance were investigated in an energy distribution domain. Along with the commercialisation of the domain new technologies had also been adopted, and the impacts on teamwork needed to be studied. Several indicators of teamwork had been identified, based on previous literature and these were divided into three rough categories: 1) inputs meaning prerequisites of teamwork such as individual abilities and the organisational context, 2) process meaning ways of actually conducting teamwork, and 3) output meaning measurable outcomes of teamwork. These factors were assessed by a combination of direct observation, questionnaires and performance ratings by experts. The study conducted was very comprehensive and took several viewpoints on the investigated phenomena reflected in the choice of methods and measures. The authors note that "research of this nature is always opportunistic and it is difficult to control the variables to the same degree as laboratory studies." Nevertheless, the effects of time of year, stage of team development, and structure of team configuration were identified on team work. Based on the results, the authors were able to identify mechanisms in team behaviour which underlay the differences between the different control centres.

Van Laar and Deshe [2002] studied control room display types for design purposes in laboratory settings. The benefit of utilising visual layering in control room UI design was the central question in their study. Three design formats were tested for search time, errors and user preference. All participants were new to the investigated subject matter and had no previous experience with any of the systems tested. The error rates did not differ between the conditions but the search times were shorter for the layered display. User preference, investigated in verbal feedbacks, was also in the favour of the visually layered user interface format.

A case study concerning control room design and evaluation in steel manufacturing was conducted by Han et al. [2007]. A human factors engineering process is described which consists of initial surveys, defining requirements, development of guidelines, evaluation of current technologies, development of design guidance and rules, design of interfaces and prototypes and evaluation of prototypes and redesigning. Operator tasks were taken as they are described in the operating manuals, and problems in the current UIs were investigated utilising heuristics defined based on literature. In testing the prototypes, the operators' preference was utilised as a success factor. Several design options were presented to the operators who then expressed their preference of the options.

Palviainen and Väänänen-Vainio-Mattila [2009] have conducted studies in machinery automation. The motivation for the work was to develop the usability of the human-machine systems, and the specific view point was that of user experience UX. The results of the study state that increasing support in problematic situations could enhance UX. This means that the support a technology is capable of providing to the user in problematic situations could be a relevant UX factor. Palviainen and Väänänen-Vainio-Mattila (ibid) also state that it is quite challenging to take into account factors related to users' needs of development and self-fulfilment while maintaining the safety, effectiveness, and compliance to regulations required in a safety-critical domain. Thus, the results seem to put the "hard measures" in contradiction with the softer values related to UXs. It can be argued that this may not be a correct interpretation of the concept of UX in safety-critical work. The UXs of the professional users could also be interpreted as concerning the appropriateness of the work practices and tools from the point of view of promoting the general goals of activity such as safety and effectiveness. It is hard to believe that a positive UX could emerge in work that puts for instance the safety of the people conducting the work in danger.

Tools for the emergency response commander were investigated by Norros et al. [2011]. The specific domain of emergency response was firefighting, which is a task quite difficult for the external observer to gather data about. Thus, an augmented research method was developed in which a user of the evaluated technology followed a live rehearsal of complex accident management, and evaluated the potential of the new tool in comparison with the effectiveness of the management of the on-going situation which he could fully observe visually and in audio. The indirect measurement of the performance of the new tool proved successful in that several development ideas could be identified together with the participants of the study.

Obrist et al. [2011] studied user experiences in a semiconductor factory using a probing method. The aim of the research was to allow for empathic interaction design which would be informed of the subjective operator perspective. In the study, stress, usability/ergonomics, emotion and social aspects were discovered to be factors related to UX.

As a summary of the empirical studies in other industrial domains apart from NPP, it seems that the evaluation approaches are not so dependent on the experimental evaluation approach, or at least that other kinds of approaches are also utilised and innovated.

2.2.3 Discussion of the methods utilised in the empirical studies

The descriptive studies of operating work provided a pluralistic and complex view of the work conducted in real industrial control rooms (Section 2.1 p. 31). However, this multi-perspective view of the work is not accounted for in the published studies reporting empirical evaluation of control room tools in an NPP context. The methodologies utilised in evaluation studies in other industrial domains have more inclination towards taking the real world complexities into account but still studies concentrating profoundly on the impacts of new tools in different aspects of work are rare. Instead, the most common measures utilised in control room and user interface-related evaluation studies are task performance in the form of time and errors, situation awareness and workload.

The prevalence of the concept of situation awareness as an indicator of good operator performance can be considered to be surprising taking into consideration the fact that the definition of the concept is not fully agreed upon by the researchers working with complex systems [Stanton, Chambers & Piggot 2001, Salmon et al. 2006, Lau, Jamieson & Skraaning 2012, Vaitkunas-Kalita, Landry & Yoo 2011]. Stanton et al. [2001] present three most cited models and definitions of situational awareness and the background theories. They are: 1) The three level model by Endsley and Kiris [1995] based on an information processing model of human cognition, 2) The interactive sub-systems model by Bedny & Meister [1999] based on activity theory, and 3) The perceptual cycle model by Smith & Hancock [1995] based on ecological psychology. All of the reviewed empirical studies of control room evaluation utilised the definition and model of situation awareness proposed by Endsley and Kiris [1995]. This shows how all the studies are based on the theory of information processing psychology and do not consider the later developments concerning human cognition and activities.

In the studies addressing the totality of a control room, the largest variety of measures were: 12 (Table 1). The reported evaluation approaches which utilise a mixture of measures comply with the advice of previous CSCW research [Ross, Ramage & Rogers 1995, e.g. Neale, Carroll & Rosson 2004] which advocates pluralism and a multi-method evaluation approach including observation among other methods, to be used in evaluations when the context is such that multiple users use the system in a work setting. Nevertheless, the observational aspects of evaluation were not emphasised in any of the studies.

In quite a number of the studies the operating task was, equated with operating procedures. This is a problem because the gap between procedures and practice is known to exist [Theureau et al. 2001, Pentland, Feldman 2008, Dien 1998]. It may also be problematical because a procedure is always a generalisation and often even a simplification of the real operating task. This means that equating the operating task with the procedure also changes the nature of the analyses conducted in a more general and simplified direction and in this process the complexities of the real work may easily become neglected.

The most common indicators (performance, situation awareness and workload) are all quantitative in nature and can be interpreted as reflecting a tradition of experimentalism in which the human performance is considered to be the “dependent variable” by which the effect of the independent variable is investigated. In this evaluation approach, the actual operating activity is treated as a black box [Kirlik 2012] as only the defined measures are considered as interesting indications of performance. It can also be stated that values such as the control of conditions, the independency of evaluators and the power of statistics are emphasised in these approaches. The main disadvantage with this approach, according to Kirlik [2012] citing Dewey [Dewey 1896], is that behaviour is, “at one and the same time, both dependent and independent variable as our actions alter our perceptions even more regularly than our perceptions alter our actions”. This criticism means that actions and perceptions actually constitute an inseparable whole, and the effort of measuring them separately is futile in the sense that it lessens the ecological validity of the study to the extent to which the results are no longer relevant from the point of view of development of the system. It may be possible to address these problems in evaluation by applying some of the notions of the later understandings of human cognition and action such as theories already presented of situated action and distributed cognition. The challenge is that operationalization of the concepts of the theories is not straightforward and perhaps because of this such approaches have not become more prevalent. The real in-situ complexities are difficult to operationalize. Thus, this aspect of the evaluation methodology is related to the definition of unit of analyses in the evaluation: What is it that should be measured in the evaluation of control rooms?

A similar concern for the adequacy of the evaluation method in terms of the indicative power of the utilised measures has been brought up by Randell et al. [2010] in the healthcare domain. They argue that, since healthcare applications are set in a complex context which involves users (personnel), patients, and the ever evolving practice of medicine, the evaluation methods should be adjusted. The tool evaluation methods which concentrate on clinical outcome measures or which attempt to meet the “gold standard” of randomized controlled trials or which aim for quantitative results by exploiting laboratory user test do not reveal the complexities with which the components of the system – technological, clinical, social, organizational, and professional – interact together to produce health care for society [Randell, Wilson & Fitzpatrick 2010]. Concerns have also been expressed with regard to the adequacy of methods used in control room technology evaluations in the nuclear power production domain [Braarud, Skraaning jr 2006, O'Hara 1999], yet, without explicit questioning of the unit of analysis used.

Another problem with the reported approaches in the empirical studies is that they are seldom able to consider the developmental nature of human activity. When operators act as participants in a test of a new control room system, they are, of course, trained. But the time it takes to adopt a new system into practice is definitely more lengthy than the training period of the test[9]. The same problem has been identified by Dix et al. [Dix, Ramduny & Wilkinson 1998], who claim that “interaction in the large” is not considered enough in HCI evaluations. This view-point of time scale in human system interaction would be very important in a study of a new control room, because the operating activity develops simultaneously with the tools [see Béguin, Rabardel 2000 for instrument genesis]. In the literature review, I did not find any studies which would have considered the longitudinal nature of the process of adopting new tools into practice. Thus, it may be possible that some of the new design ideas are being rejected too lightly, as the participants do not have enough experience of the systems to develop practices of use.

Based on the review of published empirical evaluation studies of control room technologies in nuclear and other industrial domains, I conclude that the evaluation methodologies tend to simplify and generalize the operating work to the extent which may put the relevance of the results in question. In a control room study, the evaluation framework should preferably be such that the complexities of everyday work of operating crews in NPPs can also be addressed.

2.3 Towards a comprehensive approach in control room evaluation

In this section I present lines of recent developments taking place in HF/E and other related fields which are relevant from the point of view of the research goals of the dissertation. I am specifically interested in developments which address the issues related to the development needs of the evaluation approach i.e. understanding the safety implications of systems in use, taking into account the context of operating work and relevance to design.

2.3.1 Extended view of safety

In the specific context of this dissertation work, nuclear power production, safety is probably the most important objective of HF/E. Although HF/E generally aims for universal goals such as effectiveness, efficiency and reliability, safety is the one specific aspect within the nuclear domain which unquestionably needs attention from HF/E. The conceptualisation of how nuclear safety can be addressed from the point of view of HF/E has evolved over the years [Hollnagel 2008]. Nowadays, the scientific consensus is that safety is definitely a much wider concept than freedom from accidents or singular component breakdowns [Besnard, Hollnagel 2012], and thus in the HF/E sense it is more than freedom from operative errors. I call this new elaboration of safety an extended view of safety to make a distinction

9 In our own previous study [Salo, Savioja 2006], we identified periods ranging between 2 and 18 months for adopting new control room technologies into practice.

to the previously utilised rather narrow views on what constitutes safety in NPP operations.

Nevertheless, human error has been and still is, one of the best known concepts utilised to explain the causes of accidents in safety-critical domains [Leveson 2011a]. For some reason, a human error is a very easy concept to embed in our everyday thinking and language. And in the scientific pursuit of its time, the characterization of different types of errors by James Reason [1990] definitely increased our general understanding of errors by demonstrating how they stem from normal cognitive processes. But the focus of human factors analysis has expanded from identification of errors into a more comprehensive analysis of human activities. The significance of the concept of human error as an indicator of safety has been argued against for instance by Besnard and Greathead [Besnard, Greathead 2003], who claim that, for instance violations do not always cause accidents per se. The authors claim that the case is actually the opposite; in a dynamic environment violations and deviations from procedures may just as well be necessary adaptations which actually cause favourable outcomes.

The evolution towards modern understanding concerning safety has taken place through analyses of accidents. By analysing an accident or disaster, an unarguably un-safe situation, the characteristic qualities of the absence of safety can be understood, and then through an analytical contemplation of safety can be understood and characterized[10]. An example of an accident investigation contributing to the understanding of safety, is a study conducted by Furuta et al. [2000] presenting an investigation of the Tokai-mura accident which took place in 1999 in Ibaraki Prefecture in Japan. The motivation for the investigation was that the accident had in the initial analysis been hastily declared to be caused by a worker's unsafe action deviating from the approved procedure, i.e. human error. Nevertheless, in the thorough human factors investigation carried out by the authors, several managerial, organisational and institutional causes were identified for the accident. As the main factors contributing to the accident the authors claim the degradation in the safety culture and safety regulation. This investigation shows how an unfortunate action which at first sight seems like a simple human error, is actually influenced by the whole organisational and institutional context in which it takes place. Similarly analyses of previous accidents and disasters which have even in judicial process been deemed to be human errors or violations, are presented in a popular book by Kim Vicente [2004] in which the wider organizational, social, and societal context is actually proved to be linked to the propagation of events. In an analysis of 13 piloting accidents, Nuutinen and Norros [2001] found that over time, the societal demands of piloting activity had changed so that the prevailing practices and the prerequisites no longer fully corresponded with current demands. In the study by Nuutinen and Norros, the quest for understanding the profound societal context in which the activity takes place leads to the discovery of

[10] The role and significance of hindsight in foresight can be argued though; see e.g. MacKay & McKiernan [2004].

a discrepancy between the current needs of the activity and the prevailing practices and tools. This is an example of a drift which may take place in the operational phase of complex systems [Dekker 2011]. Dekker suggests that the accidents may take place because, over time, the practices of operations may inconspicuously drift closer and closer to the limit between safe and unsafe, and one day the margin has disappeared and disaster may follow.

The modern understanding is that safety is a systemic concept. Dekker and Nyce [2012] state that one of the driving forces of recent safety research has been an understanding that successes and breakdown in safety-critical systems should not be attributed solely to component breakage or human failures, but rather should be seen as connected to the complexity and dynamics of the activity itself. Nemeth [2012] states that “the whole” is what really matters, because, even though local problems may be easier to fix, they seldom improve the overall performance and actually they can make things worse. Also Dekker et al. [2011] argue for a systemic safety concept in their critique of the traditional philosophical-historical and ideological bases for linear thinking about failure in complex systems. Systems thinking in safety has been especially advocated by Nancy Leveson [2011a, see e.g. 2011b], by introducing the new accident model STAMP which is based on systems theory instead of reliability theory, and demonstrates the systemic mechanisms through which accidents occur in complex sociotechnical systems. Rollenhagen [2011] describes how a systemic approach has been utilised in Sweden for NPP event investigation in practice: Systemicity is achieved by utilising three levels in the analyses of events: level 1 deals with the basic event sequence and utilises the concept of barriers; level 2 identifies the intermediate level causes and conditions in the event, and level 3 deals with the safety management system and its practices.

In all of the above-cited theoretical contemplations, there is the similar idea, namely that safety is maintained somewhere in the organization’s way of conducting its operations. This can be concluded based on the results of the accident analyses, which claim that the reasons for an accident can be found in the organizational and societal context of the activity. The organization’s ability to function safely in a dynamic natural and societal environment can be called resilience. The concept of resilience is defined in the preface of a recent handbook by Erik Hollnagel and colleagues [2011] as follows: “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”. This definition of resilience emphasizes the ability of the system (e.g. the organization) to continuously adapt its functioning in order to meet the operative objectives (e.g. production and safety). The logic is quite easy to endorse[11]. Adjustments in functioning are necessary on all levels of human activity if/because the environment in which we live in is ever-changing.

[11] Sometimes improvement in resilience does not necessarily improve safety however [Morel, Amalberti & Chauvin 2009].

The operationalization of the concept of resilience is more tricky. It is not clear, especially on the level of control room activities, what kinds of adjustments should be looked for when one is trying to identify resilience. The research by Back et al. [2008] and Furniss et al. [2011] has tried to find answers to these questions. Together they identify a set of so-called resilience markers which are behavioural indices of adaptations in operating activities.

Studies of anaesthetists' practices conducted by Norros and Klemola [Klemola, Norros 2001, Norros, Klemola 1999, Norros 2005] adopted the semiotic model of habit introduced by Peirce [1958] as an operationalization of operating practice. The observed behaviours of the anaesthetists were graded according to different levels of interpretative quality of reasoning in action which the practitioners expressed in their reactions to environment (i.e. the patient and the on-going operation). An interpretative relationship was characterised by operative or communicative acts that demonstrated active attention to the specific features of the situation, attempts to construct a cumulative interpretation of the patient's physiological state, and/or pro-activeness with regard to the on-going anaesthesia process. A reactive response was characterised by operations and communications that did not demonstrate active epistemic interest in the situation, i.e. perceiving or searching information is scarce and no interpretation was expressed, and the actor appeared to be driven by the events. It was further assumed that a confirmative response would express an intermediate quality between interpretative and reactive responses. Confirmative response was characterised by the interpretation of a situation in the form of identifying it to be as expected or falling into a known class of situations, and assuming certain responses. Specific features of the situation may be interesting for classifying the situation, but are not taken as a source of new knowledge of the patient's physiological characteristics or to inform of the effect of the used drugs, etc. Later, it has been contemplated by Norros that the characteristics of the interpretative practices answer to the needs expressed by the demands of resilience on the level of operating activity [Norros 2012].

From the point of view of a control room evaluation approach, the operationalization of resilience on the level of operator activity are extremely important if the evaluation approach is to be such that it does not rely on considering the procedures equal to operating activity. The aim should instead be to reveal the proneness to unsafe acts, as Kövers and Sonnemans formulate [2008]. The practice indicators could also respond to the need expressed by Julien [2008] in the call for evaluation approach which could also utilise study of normal situations as a source of information.

Based on the above cited theories, I conclude that the evaluation approach to support control room development should benefit from the modern understanding of safety as a systemic concept and especially resilience as its operational manifestation.

2.3.2 Addressing contextuality

As already stated, the operating activity always takes place in a particular context. The context can be examined from different perspectives: environmental, societal, organisational, physical, etc. For example, Räsänen and Nyce [2006] explore the role of social context within HCI and CSCW research. Falzon [2008] claims that designers do consider the users and their roles in the functioning of the system to be designed, but not necessarily in the context in which the users will have to operate. Within the scope of this dissertation I will concentrate on the immediate operational context: the controlled process. It is the most important contextual factor which has effects on the evaluation. The evaluation approach should be such that, depending on the controlled process, different aspects of the control room may be emphasised. For this purpose, the controlled process needs to be modelled and analysed as it constitutes the context in which the operating activity takes place.

The controlled process can have multiple different representations, the most common of which is probably the piping and instrumentation (P&I) diagram created in process design. This diagram portrays the piping of the process flows in connection to the process equipment utilised so as to control the process by the automated control system and the operators. The P&I diagram is an important starting point of representing the process context, because in it are portrayed material resources with which the process control is expected to be handled in all situations.

Cognitive work analysis (CWA) [Vicente 1999] expands the process analyses by adding several qualitatively different abstraction levels on top of the physical components portrayed in the P&I diagram. The abstraction levels were originally introduced by Rasmussen [Rasmussen 1985] as the abstraction hierarchy (AH), a five-level functional decomposition of the work domain. The abstraction hierarchy is a task-independent general representation of the work domain. It combines the physical function (e.g. the P&I diagram) upwards to the functional purposes of the domain and downwards to the physical form, utilising "means-ends" relationships. The result of the work domain analysis is an AH (or several AHs) portraying how the functional purposes of the domain are fulfilled by the means of abstract functions, i.e. the (natural) laws and principles governing the system, which again are fulfilled by generalised functions i.e. the processes of the system. At the next level down, is the physical function in which the physical components and equipment are associated with the generalised functions. The lowest level, the physical form, relates to the most concrete aspect of the domain, the shape, condition, layout and for instance coupling of the physical components in the respective plant. By connecting the different functional abstractions of the process, the AH represents the complex view of the process control work. It shows how operating work is not only about operating the individual physical components as often represented in the procedures, but is also about complex problem-solving and monitoring the status of the functions and purposes of the domain. The AH makes explicit the content of the knowledge-driven aspect of the operating work in the respective domain. Thus, it is an important contextual representation that may be exploited for the purposes of evaluation of control room tools.

Another functional modelling method, multilevel function modelling (MFM), was developed by Morten Lind [see e.g. Lind 1999]. MFM depicts the functional decomposition of the controlled process on different levels, and it can be used to analyse existing processes or design new ones. MFM is conducted with specific graphical languages in which a pre-set of icons represents the generalised functions of process control.

The context may be taken into account in analyses of operating crew performance as a performance-shaping factor (PSF) [Lee et al. 2011]. The authors present a categorisation of different PSFs and make it evident how the characteristics of the controlled process in the situation have a direct impact on operating crew performance. This connection between the context and the situation means that for the purposes of understanding successfulness of the operating activity the context needs to have also a situational representation.

For the purpose of understanding the demands on activity in particular situations, specific models of situations are needed. Norros [2004] presents a situational manifestation of the work domain which emphasizes the role of available information in the situation within the propagating flow of process events. Flach et al. [Flach, Mulder & van Paassen 2004] present a theory-based depiction of how Rasmussen's levels of abstraction could be used to describe resources within situations. Petersen [2004] has also developed modelling in order to understand control situations for the purpose of design of human-machine systems. Later Bjørkli et al. [2007] have added a level of strategy to Petersen's model based on empirical work in a maritime environment. Possible context-dependent deviations from optimal task flow can be modelled with the approach introduced by Paternò and Santoro [2006]. Situational models of contextual design [Beyer, Holtzblatt 1998] emphasize the role of the physical environment as an enabler of fluent task flow of users.

Both AH and MFM are significant modelling methods because they depict the connection of the lower level physical components to the higher level purposes of the domain. The purposes are the meanings which should be conveyed in good user interface designs. But the problem of using AH or MFM directly in control room evaluation is that the models are general in nature, and evaluation is carried out with an operational scenario in which certain operating activity is conducted so as to control the process. Thus situational models are also needed. The above-mentioned situational manifestations of the operative context are important in emphasizing some aspects of the operating task, but none of them is directly suitable for the evaluation of control room technologies as such. I conclude that the evaluation approach should include a contextual, situational description of the operating tasks against which the operating activity in a test could be compared.

2.3.3 Design orientation in human factors

In theory, human factors is a science of design by definition [Dul et al. 2012]. But in practice the intervention of the human factors specialists is often limited to evaluation [Falzon 2008] and more design-oriented human factors approaches are needed [Bas 2007, Healey, Cacciabue & Berman 2013]. In my view, this issue was reflected in the published control room studies reviewed in the previous section: Not very many studies described the actual benefits to design acquired in the evaluations[12].

In order to understand how a human factors method could be advantageous to design, one must consider what actually constitutes design in the field of complex socio-technical systems. Design is an inherently humanistic and pluralistic activity [Bardzell 2010]. Falzon [2008] characterises design as a problem-solving activity during which the problem is defined simultaneous to its resolution. This means that by designing, e.g. sketching, prototyping, or simulating, the original design problem is understood more and more profoundly and thus it may become reframed. This phrasing seems to bring to the fore an essential problem related to the design of such complex systems as NPPs. The scale and complexity of the object of design are so enormous that in the beginning it is practically impossible to comprehend the involved couplings and related design problems, and thus the original problem needs to be defined over and over again during the design. In my personal view, the HF evaluation approach should be such that it could be informative in the issue of re-phrasing the design problem and thus providing new perspectives on design.

The insight of Falzon [2008] concerning the nature of design also reveals an important aspect of design which may make the “collaboration” of design and human factors activities inherently difficult: They represent very different scientific paradigms. Human factors (often) rely on an idea of science which emphasizes rationality, independency and control, whereas design as an activity is different. There are ingenious design outcomes which are not based on rationality, and in the design of which the designers have become highly involved in the usage activity. Design may produce excellent outcomes and at same time be participative, iterative and sometimes chaotic. In order to support a design process like this, the human factors evaluation methods may need to adapt some of its core values just as well as design activity may need to move even more towards formalised processes.

Béguin [2003] sees design as a mutual learning process between users and designers. The invention of new uses can then be seen as a continuation of the design process [Falzon 2008, also Walker et al. 2009]. Even function allocation

[12] The reason for this may partly be that the published studies are often validation studies in which the particular purpose is to convey the independence of evaluation from the design, and the focus of the study is on the accept side of the pendulum between accepting and improving design [Braarud, Skraaning jr 2006].

between human and automation may become altered in use [Wright, Dearden & Fields 2000, Fuld 2000] and designers themselves may be involved in the continuous development of the system during use [Alby, Zucchermaglio 2006]. Béguin and Rabardel [Béguin, Rabardel 2000] point out that design is not really "ready" until it has been adopted into a meaningful use by someone. The design outcome becomes an instrument for an activity in a process labelled instrument genesis. In my view, if instrument genesis were to be taken into account in the HF evaluation it would mean that the purpose of evaluation or validation would be to identify the potential for instrument genesis, both in the tool and in the people utilising it.

Falzon [2008] cites Daniellou [1992, 2004], in saying that the goal of the ergonomist when designing work systems is to build the "space within which the activity will be developed". Also, it is acknowledged by Falzon [2008] that Vicente's [1999] idea of boundaries or constraints of use seems to be related. According to this, the objective in design is to define envelopes of uses (rather than to specify uses). These contemplations can be interpreted to mean that the end result of ergonomic evaluation is not only the identification of the effects of tools on current human performance but also to evaluate the potential of the new tool in having an effect on practices of use and keeping them within the safety envelope.

It must be acknowledged that in design it is not possible to know comprehensively all the information needs of future users. Petersen and May [2006] suggest that scale transformations could help solve the problem. Their idea is that scale transformations are a key to understanding how operators tailor information presented in the control system in order to acquire support for higher level cognitive tasks. If an HF evaluation were to support the design of such a system which supports scale transformations it should be very profoundly grounded in the critical information of the domain, i.e. it should be able to connect the critical functions of the domain to the way operators utilise information concerning them in their activity.

Evaluations should be conducted at an early stage in the design process. A challenge in trying to assess the safety of early concepts is that there is often little detail on the procedures or controller working practices proposed for the concept. This amounts to a lack of a mature operational concept, one that is sufficiently detailed to allow safety theses (e.g. 'what would happen if…?') to be answered (other than – 'well, it depends how we operate it') or even asked. [Kirwan 2007] Professional end users, for instance operators, are an important resource in gathering safety assessment information on new concepts and procedures [Pasquini, Pozzi & McAuley 2008]. And it is possible that their expert views on the new systems could be better exploited in the evaluations of early design concepts.

Design is communication in many ways [McCarthy 2000]: communication with end users and other stakeholders and communication within the different design disciplines. To support the communication, boundary objects are needed in design [Norros, Salo 2009]. If one manages to come up with good boundary objects, they can be utilised for various purpose as is explained by Naikar [2006]. The role of boundary objects is related to human factors evaluation because the concrete outcomes of evaluation need to be such that they communicate the results among the different interest groups.

Safety as an important design goal has many manifestations [Boy, Schmitt 2013]. Earlier it was identified that resilience is an important characteristic of safe operating activity. Yet, an unsolved question is how resilience can be supported by design. Nemeth et al. [2011] present a practical example from the emergency healthcare domain. They argue that control does not increase resilience. If one is aiming for resilience, some requirements on control may have to be loosened. In the petroleum industry, Skjerve et al. [Skjerve et al. 2011] utilised a coaching approach for designing resilient collaboration with the employees of a new plant.

Of course the technical rational view on design should also be supported by human factors analyses and evaluations. This is because design flaws can be reduced by proper review and evaluation methods [J. Robert 2007] and utilising appropriate tools [Jou et al. 2009a]. Baxter and Sommerville [2011] describe what they have labelled socio-technical systems engineering (STSE), which draws from human-centered methods but also complies with technical design quality requirements. The role of evaluation in STSE is described as filling in the gaps that have not been addressed in the initial analyses and requirements definitions. Also through evaluation, the system will be domesticated to the organization in which it will be used. These ideas resemble ideas presented earlier concerning instrument genesis.

Evaluation within design should be such that it mediates between the current form of the outcome of design (i.e. the artefacts such as new displays) and future uses. In evaluation, the new artefacts are being put to use, so their future can be envisioned especially by professional expert users. As design continues, iterative and often incremental evaluation should be able to produce results in formats which enable improving the design.

2.3.4 Widening the concept of usability

HCI is a field of study interested in interaction between a computer and its human user. HCI is a natural source for inspiration in the development of control room evaluation, as modern control rooms consist of different computer systems. Successful adoptions of HCI methodologies to human factors issues have been reported e.g. by Wilson [1995] and Vink & van Eijk [2007]. Both reports present a study in which participative methods have successfully been used for the design of industrial tools.

Usability is a central concept in HCI. It stands for “The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” [ISO 9241-11 1998]. The ISO definition of usability represents quite a broad understanding of what usability is. First of all, it states that usability is a quality attribute of “a product” and thus specifies that usability resides deeper in the system than in the user interface. Secondly, it declares effectiveness as the first characteristic of a usable product. Effectiveness in achieving goals means that relevant functionalities must be provided by the system. This is an extension in comparison to Nielsen’s [1993] often utilized conceptualization, which differentiates usability from utility. Thirdly, the significance of the context of use is emphasized. This must be interpreted so that the usability of a control room may be very different from the usability of, for instance, some consumer products.

Vrazalic [2004] reviews the criticisms of the traditional definitions of usability originating in the 1980s and early 1990s and in her dissertation utilizes the notion of distributed usability. The main critique is that the definitions are too system-oriented and technocratic. They do not take into account the different kinds of contexts such as social, cultural, and physical. In addition, evolutionary aspects of use are not addressed. Distributed usability implies that usability is a property of interaction between users and technologies and must concern the totality of the surrounding information ecology.

Sørensen & Al-Taitoon [2008] outline a concept of organisational usability so as to demonstrate that a system performance may be observed also on the level of the organization by demonstrating how well technologies enable and support fulfilling the organization's goals. A system which may be usable on the level of individual users may not be that on the level of the organization. For example, this may be true if a system does not comply with the processes with which a business is assumed to operate.

During the new millennium, a paradigm shift has been claimed to have occurred within HCI with the so-called third wave HCI [Bødker 2006, Harrison, Sengers & Tatar 2011, Ylirisku et al. 2009] which emphasizes the general role and the socially constructed meanings conveyed by technologies in people's everyday lives [McCarthy, Wright 2004]. Within the third wave HCI, the user experience (UX) has become an important concept. Kaptelinin and Nardi [2012] even state that currently "the experience" is actually an emblem for the whole field of HCI. UX as an object of study in HCI emerged as a counterforce to the previously dominant work and task oriented HCI [Hassenzahl, Tractinsky 2006]. The emergence of the third wave HCI can be interpreted as a force aiming to overcome the previously too narrow conceptualization of usability and the fact that it did not take into account the totality of the aspects of human activities in conducting usability studies [Nelson, Buisine & Aoussat 2013].

In my view, the extended view of usability should also be reflected in the development of evaluation methods to support control room design and safety assessment.

2.3.5 Adopting activity theory in HCI

The third wave HCI has several theoretical underpinnings, one of which may be considered to be activity theory (AT). AT, which stems from 1930s Soviet psychology [Diaper, Lindgaard 2008], has been suggested as a promising theoretical framework for HCI already in the early 1990s [Kuutti 1991, Bannon 1991], and AT-inspired design approaches [Kaptelinin, Nardi 2012, Gay, Hembrooke 2004, Hyysalo 2002, Kaptelinin, Nardi 2006] have since been developed [Law, Sun 2012] and advocated [Norman 2006]. Although, it must be noted that the adoption of the theory in practical research work has not always been straightforward [Diaper, Lindgaard 2008].

The potential significance of AT, from the point of view of HCI and usability, has been comprehensively outlined and justified in a recent book by Kaptelinin and Nardi [2012]. In what follows I present the key concepts of AT which I find relevant from the perspective of control room evaluation.

One of the fundamentals in AT is the concept of mediation. For example, Gay and Hembrooke [2004] discuss the bidirectional nature of tool mediation, i.e. how perceptions, motivations, culture and actions shape the tool and simultaneously are also shaped by the tool (pp. 5–6). Mediation is a relevant concept in analyses of control room work because, in a modern industrial control room, there are practically no means for direct interaction with the process. All the interaction between the operators and the process takes place via the medium of automation and the control system. Thus, the capability of the control room to mediate all the relevant process information to the operators and vice versa can be considered to be a success factor of control room systems.

In AT, activity is understood as historically and culturally developed. Hence, the central aim in the analysis is to ascertain the current state of affairs but also its historical roots and possible trends toward which development is proceeding. The approach suits the needs of control room evaluation well in a state in which hybrid technologies have been implemented and more profound modernizations are being designed. In order to understand whether the development of tools is proceeding in a good direction, the wider historical context of tools must be understood. This seems valid as transition processes are known to be long in safety-critical domains [Ans, Tricot 2009].

Internalisation in AT signifies the process through which the external activities (i.e. manipulations of real world objects) become adopted to the extent that they can be performed internally. Internally performed actions are for instance mental simulations, considerations between alternative plans and other mental exercises. The capability to perform internal activities signifies a profound capability to perform external activates. In analyses, internal and external activities should not be separated because they form a unified system in which the transformation, from external into internal and again to external, is continuous. In considering the process of internalization in the case of control room modernization, this means that the internalised functions lose their external counterparts as the representations and modes of interaction are transformed in the control room. This means also a shattering of the internalised functions, as they are tightly connected to the external ones. Therefore, even though the controlled process will remain roughly unchanged in the control room and automation modernization, the operators' skills and capabilities may be under threat.

The significant contribution of AT is to consider the design and use of technology in a socio-cultural context. As a result, the unit of analysis of tool usage is a multi-layered system of activity. Usage of tools is seen to be constructed as an interaction between users of different user groups and designers [see e.g. Gay and Hembrooke 2004, pp. 2–14]. The detailed means for the evaluation of tools is something that the earlier appliers of AT in HCI have not proposed even though evaluation has been identified as an important function of an activity-cantered design [Gay and Hembrooke 2004, p. 12].

As already stated, the concept of UX has become important for the third wave HCI. Even though there have been numerous practical examples in analysing, evaluating and even designing for UX, presented at HCI-related scientific conferences and in journals [Bargas-Avila, Hornb\aek 2011], the basic concept of UX still remains debated, bearing many definitions in the community [Roto et al. 2011]. There have also been claims that the published practical descriptions of the studies lack the theoretical explanations and justifications based on which the research has been conducted [Kuutti 2010, Luojus 2010]. In this dissertation, I utilise the activity theoretical understanding of the concept of experience as a basic underpinning for understanding the significance of the operators' UX concerning the proposed new control room tools.

In AT[13], human-environment interaction is considered to take place as continuous embodied action-perception cycles. Perceptions of the environment, including the embedded technologies, are used in acting on the environment and accumulated in the experience of the actors. The qualities of experiences in action are interesting, because they reveal inherent features of action that cannot be perceived by observation from outside [Norros, Liinasuo & Hutton 2011]. Mediation is related to experience in a special way: when people learn to use particular tools and technologies successfully, new mediations are created into the complex activity and positive emotions emerge [Vygotsky 1978, Koski-Jännes 1999]. For example, when one learns to use a new tool, this may happen. From this point of view the general genuine positive emotions of operators could be considered as indications of comprehending and appreciating the potentiality of the new tool.

AT understands subjective experiences as profoundly embedded in the activity and actually determining courses of activity, as opposed to being merely a consequence of interaction, or epiphenomena of the actual defining processes of activity. From the research point of view, this stance of AT means that understanding experiences may help in making sense of the activities: Rather than trying to capture the experience that a particular technology produces, one should concentrate on *the experience that produces the activity* [Kaptelinin, Nardi 2012]. This is a justification of studying experiences in connection with control room evaluations. If it is possible to understand the experiences of the operators for instance in emergency situations, it may be possible to design tools which are better suited to the situational demands.

In AT, emotions are understood as direct indicators of the status of an activity as a whole [Kaptelinin, Nardi 2012]. From the interaction research point of view this means that the users' feelings concern the totality of the activity, and not

[13] And according to Norros [2004] also in philosophically oriented analyses of activity within pragmatism [Peirce 1998a, Dewey 1999, Määttänen 2009] and phenomenology [Kestenbaum 1977, Merleau-Ponty 1986].

merely the singular technological solutions within the activity. As emotions do not reveal the reasons for their occurrence, people may not always be aware of why they experience a certain emotion. This means that the experiences which cannot be thoroughly explained by the users may also be of significance from the point of view of the whole activity. The user may have a feeling concerning a tool that "there is something wrong here" but is not capable of attributing the feeling to a particular technological feature. Nevertheless, the feeling should be taken seriously because it may be an indication of the inappropriateness of the tool in the context of the particular activity.

Based on the theoretical contemplations presented above, the background assumption concerning the concept of UX within complex systems is concluded as follows. UX is a subjective phenomenon which is profoundly embedded in the activity as emphasised in AT. Similarly, the experiences of the users concern the status of the activity as a whole and can, therefore, be utilised to assess the general potential and promise that individual tools have in light of the objectives of the whole activity.

2.4 Synthesis of related research: development needs of control room evaluation approach

In the previous sub-chapters, I have reviewed the related research which I have found relevant and topical from the point of view of the research goals of my dissertation. The following synthesis of related research presents the research gap which I will address with the empirical work of the dissertation.

First, I reviewed literature concerning operating work in process control industries and in particular in NPPs. As a result, a pluralistic notion of the work of was construed and sources of complexity in the work were discussed. A particular point was made concerning operating procedures: Human work in a control room covers a significantly wider spectrum of activities than can be described in the operating procedures.

In the second subsection, I reviewed literature concerning empirical evaluation of control room technologies both in NPP and other industries. I found that in most cases, especially in NPP evaluations, the work of the operators is treated solely as it is described in procedures. This is a clear deficiency in the prevailing evaluation methods. Almost all the evaluation utilised quantitative measures such as time, number of errors, task load, and situation awareness as the main information sources concerning the successfulness of the evaluated solution. In my view the objective measures mentioned are important indications of the quality of control room solutions, but the challenge is that they do not cover the so-called extended view of safety. They do not address the *potential* because they are not sensitive to the delicate features of the activity that do not pose an immediate threat but are the possible seeds from which the operating practices start to develop into a non-optimal direction. Therefore, in my view, the spectrum of measures utilised in control room evaluations should be widened (for developments in this direction, see also Schraagen et al. [Schraagen et al. 2008]). Specific features that should be addressed in evaluation are for instance the effects that new technologies have on working practices, the development of professional skills, the development of operating work and the effects on collaborative aspects of work.

In the following subsection I reviewed recent theoretical development in HF/E and other important fields of science. Based on the review several requirements for control room evaluation methods can be identified.

In the development of an approach for evaluating control rooms the extended view of safety should be taken into account. Safety should be treated as a systemic concept. It should be reflected in the measures based on which the successfulness of design solutions is concluded. The collection of measures utilised in evaluation should be such that the capacity of operators to act in a resilient manner may be identified. This need calls for a methodology which enables the identification of resilient operating practices. As a prerequisite for developing measures that reflect resilience, a sole characterization of what constitutes resilient operation practice is needed.

Evaluation methodology should be contextual so that it takes into account the particular environment for which the new tools are designed. This means thorough analyses of the controlled process as it constitutes the immediate surroundings for operating activity. The evaluation approach should include a contextual, situational description of the operating tasks against which the observed real operating activity can be compared. A functional and situational task model is needed in order to understand the meanings that should be conveyed to the operators in the situation.

It is evident that HF/E activities should be more beneficial to actual design activity than the current evaluation methods allow. In order for the evaluation to benefit design, it should be conducted in such a manner that the design implication of the results can be made from the evaluation results. This means that usage activity cannot be treated as "a black box". Instead, careful attention should be paid to the ways in which operators utilise information technological tools. The question of how people use tools reveals the internal mechanisms of activity which are important for understanding which features of the new design solutions the operators are capable of exploiting and which need more development.

To utilise the concept of usability in an evaluation of appropriateness of control rooms seems adequate, as usability particularly considers systems in use. Nevertheless, the concept should be understood broadly, and defined contextually, so that specific characteristics of safety-critical work can be addressed in evaluations.

The strength of the concept of UX in HCI is that it allows the researcher or designer to address the meanings that new technologies have in peoples' felt life. Through this understanding, it may be possible to design technologies that support human roles and responsibilities in profound ways. NPP control rooms are typically designed from conservative starting points. Solutions which have been demonstrated to work adequately are seldom changed. The principle of conservatism arises from the concern of maintaining safety. Up to 75% of design effort may be dedicated to detailed design, verification, validation, and documentation [Eckert & Clarkson 2005] thus leaving the innovative aspects of design for less attention.

The conservative philosophy is well justified, but the problem is that it impedes general development and thus is not in line with the modern understanding of safety culture which promotes continuous development as an attitude and a practice. In order for design activity to continuously develop, it must be aware of developments taking place in the neighbouring fields of science, such as UX in HCI. Therefore, the potential value of concept of UX in the safety-critical domains should also be evaluated for understanding its meaning in light of the development of the field.

Activity theory provides conceptual tools which are meaningful in the context of control room modernization. Therefore, the potential inputs from AT to address the expressed development needs of control room evaluation should be further explored.

3. Final research questions

In this chapter, I present the final research questions which have been specified according to the review of the related research presented in the previous chapter. The research questions are presented under the research themes the goals of which were identified in introduction (Section 1.2, p. 20).

3.1 Theme 1: Conceptualisation of a good tool in complex work

The first conceptual research theme addresses the issue of characterising what constitutes a good control room system in use. The research under the conceptual theme aims at producing constructs and models as proposed by the research approach of design science.

As presented earlier, the goal of the conceptual research theme is to explore the meaning of two central concepts' usability and UX in the context of complex sociotechnical systems. In the light of the research gap presented in the previous chapter, I have formulated the following two research questions for the conceptual theme:

> *RQ1: Conceptually, how should usability be addressed and understood in the context of complex safety-critical work?*
>
> *RQ2: What is the significance of the concept of user experience (UX) in the development of tools for safety-critical work?*

3.2 Theme 2: Methods of usability evaluation of control room UIs

The methodical research theme addresses the issue of conducting usability evaluations of UIs within the control rooms of complex socio-technical systems. The research under methodical theme aims at producing methods as proposed by the research approach of design science.

As presented earlier, the goal of the methodical research theme is to identify and test different kinds of evaluation methods in order to develop procedures of evaluation which both support control room design and enable taking the safety perspective into account in evaluations. Based on the research gap presented in previous chapter, the emphasis in the methodical theme is threefold and condensed in the following research questions:

> *RQ4: What kind of task model would enable the identification of important features of operating activity?*

> *RQ5: What kind of analysis method enables addressing the safety implications of particular design solutions in the empirical data of process control activity?*

> *RQ3: What kind of evaluation process reveals aspects of control room quality in use that are relevant both for design and understanding the implications of new technologies for operating work?*

3.3 Theme 3: Implications for the design of control rooms

The empirical research theme addresses the practical issues of control room design that are relevant from the point of view of control room modernisations. The research under the practical empirical theme aims at producing instantiations as proposed by the research approach of design science.

As presented earlier, the goal of the empirical theme is to identify problems and benefits of the current and developing control room solutions. In the light of the research gap presented earlier, I am interested in the possible benefits and challenges the digitalization of control room technologies induces for operating activity. Therefore, I have formulated the following research question:

> *RQ6: What kinds of problems exist in the current control rooms that are critical from the point of view of control room modernizations and design of digital control rooms?*

4. Overview of the empirical work

This chapter presents the research work, both empirical and analytical, conducted to complete the dissertation. First the individual studies are presented (Sections 4.1–4.4) and in the summary (4.5) the contribution of each study to the development of conceptualisation of systems usability and the related evaluation approach is described.

The individual studies presented in the following subsections were all conducted in the nuclear industry between the years 2003 and 2012 in order to support modernization efforts of the control rooms of the Finnish NPPs.

4.1 Initial studies

This research work started in line with the development and start of the Finnish nuclear safety research program SAFIR. Automation and control rooms had been identified as one of the key areas in which there were research needs which could benefit advances in nuclear safety. At that time, both Finnish nuclear companies were in the initial phases of planning automation modernisation strategies. A project called "Interaction approach to the development of control rooms" (IDEC) was launched as part of the SAFIR programme, and the research into the usability of NPP control room systems started.

Among the first research contributions was a literature study concerning usability evaluation of complex systems [Norros, Savioja 2004]. The literature study covers both scientific literature concerning usability evaluation and human activity in safety critical industries, and an overview of the relevant international standards[14] concerning usability evaluation and the development of control rooms in the nuclear domain. One of the most important findings of the literature review was a problem related to the definition of acceptance criteria to be used in control room validation studies. It was clearly stated in the literature and standards that in a validation test there needs to be accepted criteria based on which the system is

[14] Since the time of the literature review in question, the amount and coverage of industrial standards has increased. Also, some of the standards reviewed then have later been revised.

either accepted or judged to need further development efforts. But how to define these criteria, in other words, how to define what is a good control room, was not explicit in the literature studied. Further findings of the literature study showed that, even though the industrial standards and the literature describe relevant attributes that need to be addressed in a usability evaluation, they mainly attribute usability to the quality of the actual user interface and do not include the wider context of technology usage in the analyses. In our own conclusions we emphasized the role of tools in shaping human operators' work practices.

Simultaneously, in the first phase of the research work, empirical studies were conducted at the training simulators[15] of Fortum's Loviisa NPP and Teollisuuden Voima's (TVO) Olkiluoto NPP. In Loviisa three operating crews each conducted three scenarios, and in Olkiluoto four crews conducted two scenarios. Evaluation methodologies which were initially laid out for these tests were based on state-of-the-art usability testing methods identified in the literature, and earlier work conducted in the VTT research group in evaluating medical tools and NPP control room systems [Norros 2004, Klemola, Norros 2002, Norros, Nuutinen 2005, Norros, Klemola 1999]. In the previous studies, the scope of evaluation had been narrower than the whole control room. Completing the first comprehensive NPP control room studies made evident the practical challenges of conducting comprehensive control room usability evaluations. Approaching the problem with the methods that had previously been applied in smaller scale own studies or in consumer domains was not practical or sufficient. Conducting a task analysis to make sense of what the crew was supposed to do in a situation proved nearly impossible, let alone observing whether the crews were actually acting according to the prescribed tasks. It was clear that new ways of representing the performance reference of an operating crew were needed in addition to other methodological developments already identified in the literature review. This work lead into developing and outlining the modelling efforts required in comprehensive usability evaluation which was published in the Enlarged Halden Programme Group Meeting in 2004 [Norros & Savioja 2004].

The VTT group took part in the so-called EID (ecological interface design) experiment which was conducted within the Halden reactor project in 2005. At this point the methodical development had advanced and work practices of the crews were included in the evaluation of new innovative user interface ideas developed by Canadian and Norwegian colleagues. The results of EID study concerning the relation of innovative interface concept and operating crew work practices were published in a book addressing the display support for situation awareness in process control work on a more general level [Norros et al. 2009].

---

[15] The training simulator is a high fidelity replica of the actual NPP control room, and it is used for the annual training of the operating crews. As the actual control room is not accessible for research purposes, the training simulator is typically used for studies concerning operating activity and control room studies. All the empirical work of this dissertation is physically conducted at a simulator.

From the point of view of this dissertation the initial studies resulted in completion of Paper III, which outlines the activity theoretical assumptions concerning the role of tools in complex work, and thus makes the first suggestion of the concept of systems usability. Paper III also describes the process of systems usability evaluation to be utilised in the control room context. This process was first developed based on standard usability evaluation methods and then developed based on the identified needs of the empirical evaluation case.

4.2 Baseline studies at the Finnish NPPs

The second phase in the research aiming for the dissertation was started after the initial outlining of the systems usability evaluation had been completed. At this point both Finnish power companies had completed the initial stages of the control room modernization projects. Fortum had modernized the control rod monitoring and manoeuvring systems, and TVO had modernized the turbine control systems. At this stage it was decided that comprehensive baseline studies should be conducted in order to gather evidence concerning the user interfaces and work practices in the control rooms of the time.

In both plants, the evaluation process outlined in Paper III of this dissertation was utilised, and the process on a general level proved feasible. Conducting the studies provided an opportunity to reflect in practice on the theoretical underpinnings of the concept of systems usability which had also initially been outlined in Paper III. These reflections led to a more detailed operationalization of the concept which was then utilised in the analyses of the data collected in the reference tests. The conceptual developments of systems usability are reported in Paper I, which presents the systems usability framework for analyses of human technology interaction. Paper II presents the more practical results of the reference tests of Loviisa an Olkiluoto NPPs.

4.2.1 Reference tests at the Loviisa NPP

The reference tests in Loviisa power plant were quite comprehensive, as all 12 operating crews consisting of 3–6 operators (including the trainees within the crews) took part and conducted three different scenarios: a loss of coolant accident (LOCA), a primary secondary leak (PRISE), and an electrical bus bar failure (BB). Therefore, the total number of participants was 42.

Data collection efforts were extensive: all the operators were interviewed individually concerning personal orientation towards work, all the simulator runs were video-recorded with overview and head-mounted cameras, task load information was gathered after each run and process tracing interviews were conducted as group interviews after each simulator run. In addition, subjective evaluations concerning control room systems usability were gathered from all the participants using a questionnaire method, and a process expert evaluated crew performance with a contextually defined evaluation scheme. The simulated scenarios were accident and incident situations, as one motivation for the plant was to gather data concerning recently renewed emergency operating procedures (EOP).

The outline of the concept of systems usability assumes that not only the directly measurable human performance is evident when considering appropriateness of tools, but also the *way operators act* and the way in which *operators experience* the suitability of the new technology are of importance. For the purpose of analysing ways of acting, a new model of operating activity was developed. This model, the functional situation model (FSM) was developed and formalised utilising earlier models presented by Norros [2004] as an inspiration and a starting point. The FSM technique is presented in Paper IV. For the analyses of operators' experiences, a questionnaire was developed which follows the principles of systems usability by applying the three tool functions and their experiential dimensions. The application of the questionnaire in the baseline studies is reported in Paper VI[16].

During the progress of the research presented in this dissertation, the theoretical concept of resilience [Hollnagel, Woods & Leveson 2006, Hollnagel et al. 2011] became prevalent in the scientific discussion of safety-critical systems. Thus, the data gathered in the baseline studies of the Loviisa plant was revisited with a special emphasis on identifying resilience in the proceduralized accident management activity. The results of the detailed level analyses of operating activity are reported in Paper V.

4.2.2 Reference tests at the Olkiluoto NPP

The reference tests at the Olkiluoto plant concerned half of the operating crews (6 crews) consisting of 3–4 operators each. The total number of participants was 21. Each crew took part in three different scenarios: a failure in a decay heat removal system, an ejector failure, and an automation failure in a pre-heater line. The scenarios conducted at Olkiluoto were of incident category because a particular motivation for the plant was to study situations which have the possibility to escalate if not acted upon promptly.

The data collection concerning the control room was extensive: the operators were interviewed individually both concerning personal orientation to work (prior to the simulator run) and process tracing (post-simulator run), simulator runs were video-recorded with overview and head-mounted cameras, task loads were measured, and subjective evaluations of the control room systems usability were made. Also, expert evaluation of crew performance was gathered with a contextually defined evaluation scheme.

In the Olkiluoto reference tests there were two significant differences in comparison to the Loviisa tests: 1) The FSMs were generated by the simulator trainers. This proved effective and shows that a scenario modelling method, the FSM, is adoptable in practice, 2) The process tracing interviews were conducted individ-

[16] Paper VI also reports UXs in an operating experience study and in a validation study.

ually for each operator. The reason for this was that, in the Loviisa tests, we had experienced that sometimes there was too much dominance by one participant in the group interview. Therefore, we were not able to use the process tracing interview data to analyse the crews' conceptions concerning past events. For this reason, we developed the evaluation method into an individual interview.

In the Olkiluoto reference tests, we identified important issues concerning the utilisation of a new digital automation system user interface. In the scenario concerning automation failure, only one crew out of the six was able to solve the situation successfully and avoid an unnecessary partial reactor scram. This same crew differed from other crews concerning one operating practice-related indicator: collaboration. These findings lead into further research questions concerning human-automation collaboration, a research line which has been continued in an ongoing research project concerning automation awareness [Laitio, Savioja & Lappalainen 2013, Laitio 2013].

The results of the Olkiluoto reference tests are reported in Paper I and II of this dissertation.

4.3 Operating experience review at the Loviisa NPP

An operating experience review (OER) was conducted at the Loviisa NPP in order to gather evidence concerning a particular digital UI solution implemented in the first stage of the modernization. The OER was conducted while the design of modernization phase 2 was ongoing. At that time, the technology platform for the major modifications to take place at the end of stage 2 had been chosen to be a particular touch screen display already in use in one subsystem. Thus, in order to guide the design further, information concerning the usage of this existing system was needed. In addition, the system had had some problems in the validation tests in the first stage of the modernization, when it was taken into use. Some modification had been conducted in the positioning of the system screen in the operating desk. Thus the motivation was also to monitor whether further design modifications were necessary.

The particular system is utilized by shift supervisors and reactor operators in order to monitor and control the status of the control rod manoeuvring system. The system had been in use in the MCR for about three years, and it is important for maintaining the safety objectives of the plant. The system comprises a touch screen and a manual rotary switch. In the control room, the system is located in two separate places; there is one UI in the reactor operators' desk which is installed roughly at the same angle as the operating desk and one in the shift supervisors' desk as a normal screen installed in a vertical position.

The systems usability questionnaire, developed originally for reference tests, was slightly modified so as to fit the particular aims of study B. There were both modifications and additional statements, especially concerning the instrumental function questions. A small observation and interview study was conducted in order to detail the necessary modifications, prior to distributing the questionnaire to the operators.

All licensed operators of the plant were sent the questionnaire in an electronic format. Participation was voluntary during shift and the response rate was 57% (n=39).

Results of the OER concerning UXs are reported in Paper VI of this dissertation.

4.4 Prevalidation study at the Loviisa NPP

During the design of stage 2 of the modernization of the control room of the Loviisa NPP, a comprehensive series of so-called sub-system validations (SSV) were launched. The concept of SSV was an innovation jointly achieved by Fortum control room designers and the VTT research team to guide and direct the control room design with independent evaluations during the design process. The overall aim of the SSVs was to gather evidence of control room systems usability in order to both accept mature design solutions and support design. The overall validation of control room transformations will take place in a series of sub-system validations concerning individual systems and related operating concepts and finally in an integrated validation. Even though they are referred to as validations, sub-system validations also have a design orientation, as modifications to the overall design of the new system can still be conducted before freezing the specific sub-system design. Three SSVs constitute a comprehensive testing of the systems comparable in scope to the reference tests. Only the part concerning UXs in SSV1 is reported as part of this dissertation work.

The particular SSV study concerned a new long-term accident management system user interface based on the same touch screen technology was investigated in the OER study. The system will be utilised by all three crew members for the tasks of long-term accident management.

It was important to study the UXs in the situation as the operating concept of the new system is different from the existing one in the plant in a few important ways. First of all, the system provides process feedback to the operators only from the parameters which utilise safety classified, very robust and reliable, instruments and systems. This means that, in comparison to the current control room, there is less process feedback. Simultaneously, the feedback which is given is more reliable in nature. Secondly, the new system includes the EOPs in digital format on the displays. This is a change which may have an effect on the UXs, because the operators are quite used to the paper-based working with the EOPs. Thirdly, the operating interface is designed based on the respective EOP. This has resulted in a task-based UI which contains only the controls and monitors prescribed in the EOP. These changes in the operating concept aim at harmonizing operating practices and increasing the efficiency of operations in an accident management situation.

Nine operators took part in the experiment. The operators had varying levels of expertise.

The questionnaire utilised was again slightly modified to suit the particular aims of the SSV. Prior to completing the questionnaire, the operators had first passed a two-day validation training organised by the design organisation, and then on the test day had conducted three accident scenarios in which they had utilised the new system in nearly full fidelity development simulator conditions.

As the number of participants was quite small, no statistical analyses were conducted. Thus, the results should be considered as preliminary reflections of the UX of the new user interface and operation concept. The UX-related results of the SSV study are reported in Paper VI of this dissertation.

4.5 Summary of methods and materials

The methods and materials of the entirety of the dissertation work are summarised below (Table 2).

Table 2. Summary of the methods and materials in empirical work.

Study	**Initial studies**	**Baseline evaluation study in Loviisa NPP**	**Baseline evaluation study in Olkiluoto NPP**	**Operating experience review in Loviisa NPP**	**Sub-system validation concerning modernized MCR of Loviisa NPP**
Practical purpose of the study	Develop initial version of evaluation process	Create a baseline reference to be used in comparisons with the new control room Study adoption of new EOPs Identify UI problems	Create a baseline reference to be used in comparisons with the new control room Study effects of new digital UIs in MCR Identify UI problems	Monitor adoption process of new UI Formulate a new baseline concerning the particular system	Validate control room subsystems Identify design improvements
Scientific purpose of the study	Explore methodical choices	Develop the evaluation methods	Develop evaluation methods Study effects of new digital UIs in MCR	Explore concept of UX in NPP context	Explore concept of UX in NPP context
# of participants	--	42 operators	21 operators	39 operators	9 operators
Research process	Literature study Outlining initial evaluation process and developing it in two empirical evaluation cases	Situation modelling Data collection Data analyses Assessment	Situation modelling jointly with trainers Data collection Data analyses Assessment	Small scale observations and interview in order to develop the questionnaire Distribution of the questionnaire electronically	Questionnaire study (embedded within a larger sub-system validation project)

Facilities utilised	Training simulator, test facilities	Training simulator	Training simulator	Normal main control room of the plant	Development simulator
Interview materials	Designer interviews	Orientation interviews Process tracing group interviews	Orientation interviews Individual process tracing interviews	Interview concerning the studied UI	--
Observation materials	Crew activities in different accident scenarios	Crew activity in 3 scenarios: LOCA Electric bus bar failure Primary-secondary leak	Crew activity in 3 scenarios: Failure in decay heat removal system An ejector failure An automation failure in pre-heater line	Observation of crew activity in one scenario	--
Question-naire mate-rials	--	Task load index Expert evaluation Systems Usability Questionnaire	Task load index Expert evaluation Systems Usability Questionnaire	Systems Usability Questionnaire	Systems Usability Questionnaire
Data analyses	Qualitative	Qualitative and quantitative	Qualitative and quantitative	Qualitative and quantitative	Quantitative

Result reported in	Paper III	Paper I, Paper II, Paper IV, Paper V, Paper VI	Paper I, Paper II, Paper IV	Paper VI	Paper VI
New questions which emerged in the study	Further development of functional situation models Development of systems usability questionnaire	Development of process tracing interview method	Further research questions concerning operator-automation collaboration	Further development of UX research method	Longitudinal assessment of developing control room concept
Contribution to the systems usability concept and evaluation approach	Initial testing of different methods Outline of the initial evaluation process	Systems usability of a hybrid control room Functional situation models Resilience in proceduralised operating activity UX in complex work	Systems usability of a hybrid control room Functional situation models	Understanding the role of UX in evaluations	Understanding the role of UX in evaluations Understanding the challenges in evaluating less mature technologies

5. Results

This chapter gives an overview of the key results acquired in the research work comprising this dissertation. The full results are detailed in the papers attached at the end of the dissertation. This section is organised according to the research questions presented first as research goals in the introduction (p. 20) and as final research questions in Chapter 3 (p. 68). This chapter is divided into three main parts: Subsection 5.1 addresses the conceptual research questions concerning the notion of usability in the domain of complex work and the role of UX in development of tools for complex work. The details of these results are reported in Papers I and VI. Subsection 5.2 deals with the methodical research question concerning practical ways of studying the usability of control rooms. The detailed results of Subsection 5.2 are reported in Papers III, IV, V. Section 5.3 presents the practical empirical findings concerning identified usability problems. The detailed results are mainly presented in Paper II.

5.1 Characterising good usability in the domain of complex systems

This research deals with the quality in use of control rooms used by professional operating crews in order to control complex safety-critical industrial production processes. The fundamental starting point for understanding the appropriateness of a control room begins by investigating what is, in general, the purpose of the tools humans use in their lives. The magnitude of change in transferring to digital tools [Ha et al. 2007] is so great that the role of the tools needs to be considered "from the beginning" by theorizing about the role of tools in human activity. How do tools contribute to human activities? In this question, the theoretical contemplations of cultural historical theory of activity (CHAT) were utilised, as the role of tools in activity has been thoroughly explained in the paradigm. Thus, CHAT has been the main theoretical inspiration for the concept of systems usability the derivation of which is presented next.

5.1.1 Concept of Systems Usability

In the section concerning related research (Section 2.4) it was identified that in order to characterise “the good” of tools within complex socio-technical systems, the approach needs to be systemic. I have interpreted the quest for a systemic approach by including multiple viewpoints and abstraction levels in the constellation of the concept of systems usability.

5.1.1.1 The functions a tool shall serve in an activity

In CHAT, or in short, activity theory (AT), the role of tools in activity has been thoroughly examined. One of the founding notions of AT is *mediation*: In an activity, the relationships between constituents of the system are mediated by historically and culturally developed means, for example artefacts. This means, that the relationship between the subject and the object she is acting on is mediated by tools and instruments. This is a very fundamental characteristic of human activity; it is often stated that it is precisely mediation that has made Homo sapiens into such a successful species [Kaptelinin, Nardi 2012].

In the context of process control work, mediation means considering the subject as the individual operator and/or the entire operating crew, the object as the controlled process, and the main tool the automation and control system UIs which reside in the control room (Figure 3).

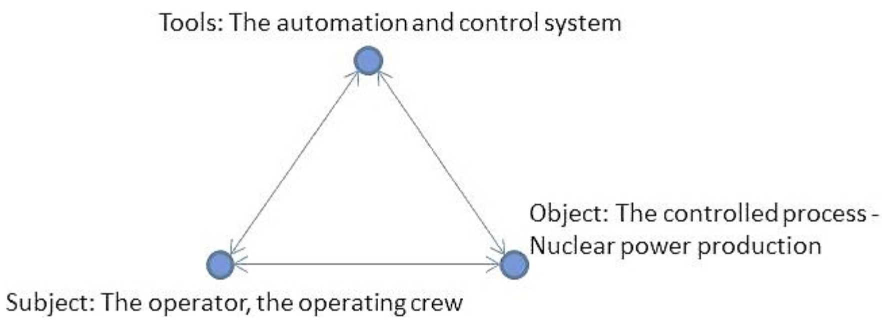

Figure 3. Mediating role of tools. In process control work the subject can be an individual operator or the entire operating crew, the object is the controlled process, and the main tool is the automation and control system.

From this notion of the mediating role of tools, it can be derived that the purpose of a particular tool, the automation system and in this case in particular the UI in the control room, is to provide the object to the operators in a way that enables appropriate acting on the object.

Vygotsky [1978] elaborated mediation by making a distinction between two different functions of tools in an activity: *instrument* and a *psychological* tool. Instrument refers to the tool’s ability to produce the intended effect in the environment, whereas the psychological function refers to the comprehension of the tool's potential by the human and the tool's capability to function as an external control of human action.

The distinction, in the context of process control work, means the following. As an instrument, the automation system allows the operators to operate the process effectively. This means that operators are able to monitor and operate the relevant process components and sub processes to the extent that the defined work tasks and responsibilities can be accomplished. The operators are able to have an effect on the process, which in this case constitutes the environment. Consequently, as a psychological tool, the automation system enables the psychological processes relevant in the operators work such as learning and remembering. By using the automation system, the operators learn how the process actually functions and for instance what are the interconnections within the process that are relevant for obtaining the objectives. In more theoretical terms, the automation system provides an external means to control one's own behaviour; it provides feedback and enables reflection on one's own activity, which are basic prerequisites for development.

Fairly recently, Georg Rückriem [2003] proposed a third general function for a tool in an activity: that of communication. The idea is based on media theory and the new aspects which the digital tools bring into tool functions in modern society [Rückriem 2009]. The communicative function of a tool denotes to the social aspects of using a tool.

In the process control context of an NPP, the communicative functions emphasize the role of the community which uses the same tool for the activity. The control room is the nerve centre of the whole plant, in which all the important information concerning the process is interpreted and acted upon by the operating crew. Within the community of an NPP, it is evident that the usage of particular features of the tool conveys meanings. For example, in a conventional analogue control room even the physical location of the operator communicates to the other members of the same crew the tasks that are being conducted by that particular operator. The same could be stated more obviously concerning the use of particular emergency-related systems and tools. In more general terms, the communicative function of a tool enables the construction of shared understanding concerning the work, which is a prerequisite for meaningful collaboration.

As a summary concerning the functions of a tool in an activity system, it is concluded that a technological tool always serves the three distinct, yet intertwined functions in the activity (Figure 4).

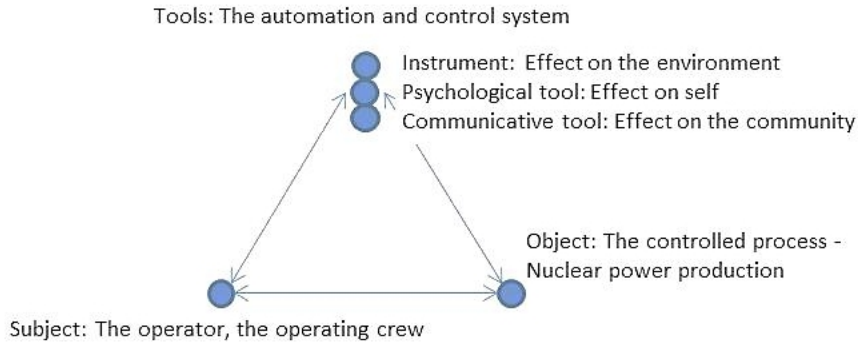

Figure 4. The functions of a tool in mediation between the subject and the object are threefold. They include instrumental function, psychological functions, and communicative function.

A central thesis in this dissertation is that in the systemic analysis of the usability of the tool, all three functions of a tool should be considered. From the point of view of conceptualising usability in the domain of complex socio-technical systems, this means that a usable system is such, that it fulfils the demands of all three functions.

5.1.1.2 Contextualisation of tool functions

Another central concept in AT, which is relevant for considering the purpose of tools in the activity system, is *object-orientedness*. According to this principle, the human activity is always directed to the parts of the environment that provide the possibilities to maintain and develop human existence. The material (or conceptual) entity towards which the activity is directed, forms the *object of activity*.

In the context of considering an automation system, UI as a tool in process control work, this means that the object of the activity is the controlled process. In an NPP, the object of the operating crew is the power production process which produces electricity by releasing and converting the energy stored in the nuclear fuel.

According to AT in general, the environmental possibilities and the expected outcomes that may be reached while acting on the object, are the central determinants in structuring the activity [Norros 2004]. Thus, in the NPP context, this means that the operating activity is structured according to the possibilities that the process design allows for acting. This is a very fundamental principle, and it describes the level of depth at which the human interaction is actually determined: The process design already dictates some of the ways in which the human may interact with the object [Papin, Quellien 2006].

To be successful in their activity it is necessary that actors take into account the actual and real possibilities and constraints of the domain in fulfilling the objectives of the activity. Norros [Norros 2004] has developed a framework for identifying and modelling the possibilities and constraints of the domain and the demands that they set on appropriate activity. The formative modelling approach of the human role in obtaining the objectives of safety-critical systems is called core-task analysis. A central concept is the core task, which denotes the main content of the work. In core-task modelling, the domain characteristics are analysed with a specific conceptual tool which elicits the dynamism, complexity, and uncertainty features of the object. Based on the features of the object, the demands on the actors are deduced that relate to the actors' skills, knowledge and collaboration[17]. The elaboration of the core-task demands in the particular context provides knowledge about what specifically are the features of the object that the tool is supposed to mediate in order for the workers to fulfil their core-task demands. With this method, the object of activity is analysed to provide the contextual characterisation of the tool functions.

The second thesis of this dissertation is that, in the systemic analyses of control room UIs' usability, the object of activity will be analysed so as to contextualize the functions of the tool in the particular activity. Core-task demands represent the contextually meaningful content of the work, and thus provide the basis for identifying appropriateness of activity also from the point of view of tools.

By applying the two theses presented above, a definition of systemic usability, i.e. systems usability in the context of complex sociotechnical systems, is construed:

> *Systems usability (SU) denotes the capability of the technology to fulfil the instrumental, psychological, and communicative functions of the tool in the activity and the capability of the technology to support the fulfilment of core-task demands in work.*

5.1.1.3 Perspectives to analysis of usage activity

With reference to the above definition of SU, it is evident that the evaluation of tools is possible only by following and analysing the actual activity in which the tool under evaluation is being used. The successfulness of the activity in fulfilling the core-task demands is the evidence of SU, along with the support for the individual tool functions. This means that, in order to identify the SU of a tool, empirical data concerning the usage activity has to be gathered[18]. This subsection deals with the approach utilised in the analysis of activity and details which perspectives on activity are relevant for evaluating the SU of tools. Below, three perspectives on the analysis of activity are introduced and justified: performance, way of acting, and experience.

Analysis of *performance* is the first perspective on the activity. Even in practice, it is necessary first to identify the sequence of actions and operations the actors accomplish, and the outcomes they produce in conducting the work. This is also

[17] For elaboration of core-task demands in nuclear power production, see Norros (2004).

[18] This is, of course, in line with the whole usability and user-centred design approach to technology development.

the natural way to start the analysis of usage activity. The performance aspect of activity answers the questions of what was done in a situation, and with what outcome. The analysis of performance sequence corresponds to those accomplished in most human factors and end-user studies, in which the attempt is to describe the quality of the tool by utilising indicators such as task completeness, errors in use, or time spent in completing the task. Thus, the reference in the performance analysis of activity is the outcome achieved in the activity. Outcome is something that is directly observable and even measurable by the external observers of the activity.

In safety-critical and complex domains, the outcome is, however, not sufficient as the only measure of appropriateness of tools [Norros 2004]. Analysis of *the way of acting* provides the second perspective on analysis of activity. Way of acting refers to the answers to the question of *how* the performance outcome was achieved [ibid]. This aspect is important due to the fact that numerous for example technical, training-related, or organizational (e.g. procedures) barriers, have been designed to neutralize the effect of possible performance variance on the outcome. In other words, performance outcome as a measure is not sensitive to variation in the tools. Well-trained operating crews often reach good outcomes even with poor tools. Yet, the actors themselves and the trainers report that there are differences in the *ways of* accomplishing the work, and sometimes one way may be superior to another by better meeting the general demands of the activity. Thus, the indicators concerning way of acting are brought to the centre of the analysis of activity, in order to identify the differences and their significance to safety and other objectives of activity. Hence, in addition to outcome, the way of achieving the outcome should be analysed. The reference for the quality of way of acting is the orientation to the core-task [Norros 2004] because it assumes that the operators are attuned to the general safety-relevant characteristics of the domain on top of fulfilling the immediate process control tasks. This orientation in activity creates the potential for resilience. From the point of view of evaluating control room tools, this means that it must be identified in the empirical data, that with the tool the operating crew is able to maintain the ways of acting which enable obtaining the objectives of the whole system.

The third perspective on the analysis of activity is *user experience.* Because of the complexity of the work of the operators and the expertise needed in it, it is not possible analytically to grasp all the possible implications that a tool will have in an activity. This is particularly difficult for an external observer who is not an expert on operations in the domain[19]. Therefore, the expertise of the professional operators concerning the appropriateness of the tools must be exploited. This is labelled user experience, because in this aspect of usage activity, the interest is in the feelings and emotions which the tool evokes in the users. These experiences,

[19] For example in the NPP context operator training takes several years, during which the production process and the tools in the control room become profoundly internalised by the operator candidates.

which have arisen in the real usage of the system, are an important indication concerning the development potential of the tool [Norros, Liinasuo & Hutton 2011]. Also, in AT, emotions are understood as direct indicators of the status of an activity as a whole [Kaptelinin, Nardi 2012]. Hence, analyses of experiences provide yet another perspective on the analysis of usage activity which reveals aspects concerning the suitability and potential of the tool, experienced by professional users, on the level of the totality of the activity. The overbearing quality of user experience is that the user feels that the technology has the potential to develop into a meaningful tool for the activity [i.e. instrument genesis in Béguin, Rabardel 2000], and benefits the interaction with the object of activity.

The three perspectives on activity add to the systemic nature of the approach of considering the systems usability of tools because they provide different viewpoints of and abstraction levels to the analysis.

In some way, the performance outcome is the primary perspective in comparison to other two perspectives. If outcomes are not acceptable, e.g. tasks cannot be completed with the tool, it is evident that the tool is not acceptable, and it is not worthwhile to more thoroughly investigate the ways of acting and experiences which are often more labour-intensive to analyse. But, even in this case, the other perspectives may provide explanatory power which enables the analysts to comprehend the mechanisms in usage that contribute to the poor performance. These analyses may be crucial in understanding how the tool should be developed to better support performance.

Activity can be analysed for *external* and *internal* good [Norros 2004 citing, MacIntyre 1984]. The three perspectives can be described as a continuum from the external to the internal good of activity (Figure 5). In a safety-critical domain which is regulated by laws and acts, and which is societally significant, it is necessary to demonstrate the quality of the system by utilising externally defined and observable outcome measures. At the same time, it must be acknowledged that the strict requirements on the external measures make them such, that they are not always sensitive, and even more importantly, they are not always easy to interpret because the mechanisms of usage that produce the outcome are hidden in the complex couplings within the socio-technical system. Therefore, also internal indications of appropriateness of activity are needed.

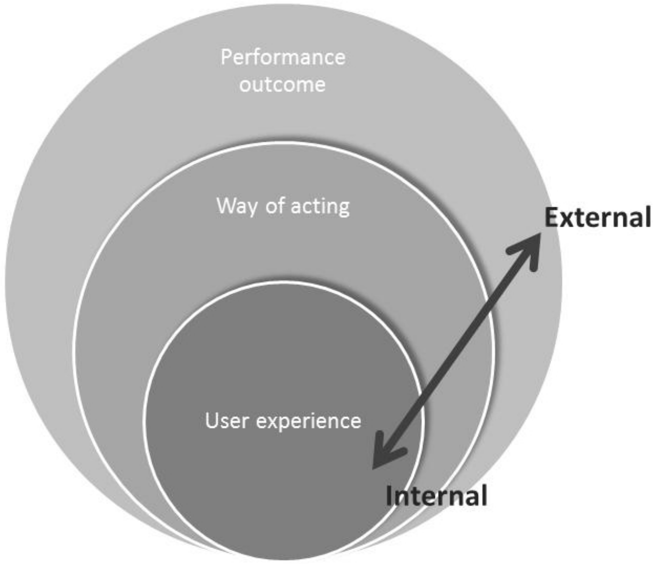

Figure 5. Perspectives on the analyses of activity: Performance outcome, way of acting, and user experience.

The definition of systems usability is completed by adding the three perspectives on activity:

> *Systems usability denotes the capability of the technology to fulfil the instrumental, psychological, and communicative functions of the tool in the activity, and to support the fulfilment of core-task demands in the work. Systems usability is evidenced in technology usage in appropriate performance outcome, way of acting and user experience.*

5.1.2 Framework for identifying SU indicators

The evaluation of SU necessitates defining concrete indicators that express SU. By combining the principles concerning the tool and the activity (functions of tool and perspectives to activity), a two-dimensional plane (Figure 6) can be construed. This plane is a frame which describes the general classes of SU indicators. The framework is a conceptual tool which aids in finding relevant measures to be used in comprehensive evaluation.

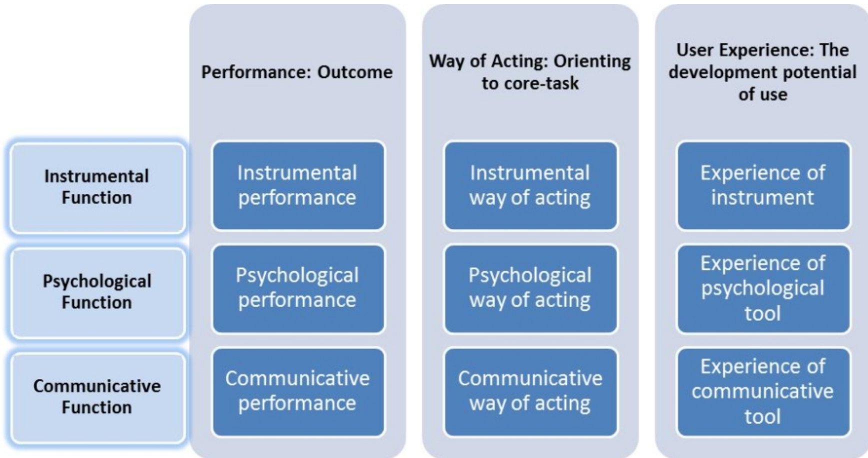

Figure 6. Nine categories of systems usability indicators.

In what follows, a description is given of the measures of systems usability in the context of control room evaluation. At the same time a characterization of appropriate process control work utilising the three perspectives to activity is made. An example of a set of measures that were utilised in a concrete evaluation is presented below (Figure 7).

Measures of performance: Measures of performance describe the outcomes of work and reflect the first perspective on activity. The measures aim at being objective, and characterize the work as it can be seen from the outside by an external observer. In control room evaluation, instrumental outcome refers to the operators being able to carry out their tasks with the tool. This means that the tasks are completed in a manner that does not endanger safety or the production- related goals of the activity. Psychological performance refers to the measures which capture the cognitive performance of the operators such as cognitive load. Communicative performance refers to the collaborative aspects of process control activity, for example the extent to which process information is communicated out loud by the members of the operating crew.

Measures of way of acting: On an instrumental level, the way of acting means that the tool helps the users to focus on relevant phenomena in the process; it aids in focusing on the tasks that are most crucial and relevant in the given situation. This is manifested in a way of acting that can be characterized as focusing on core issues and effective prioritization in work. On a psychological level, the way of acting means that the practices of use are such that they strive for understanding of the process situation, thus being able to anticipate and being in control. On a psychological level, this way of acting requires profound conceptual knowledge and operative schemes about the process, its conformance with the natural laws and the controlling automation system. The tool should be such that it enables and supports the development of these psychological capabilities. On a communicative level, the way of acting refers to the features of teamwork and a shared understanding of the prerequisites for good performance. A good way of acting in the communicative form is such that it enables the crew to share the object. This means that each individual is able to see their own influence on the joint object. In order to support a communicative way of acting, the tool should support good teamwork and shared awareness of the process status.

Measures of user experience: Instrumental experience means that users feel that the tool works well; there are no unnecessary complexities in using the tool. Feelings of achievement also belong to this category. In the psychological function, the user experience means that the user is confident in using the tool and the tool is embodied to the extent that the usage feels effortless and natural. In the communicative function, the user feels that they can trust the tool in the same way one can trust another operator within an operating crew. The tool is trusted to communicate all the required process information in a way that is comprehensible to the users. The tool improves transparency and anticipation within a team and also improves shared understanding of the situation. In a wider perspective it also mediates the values of good practice.

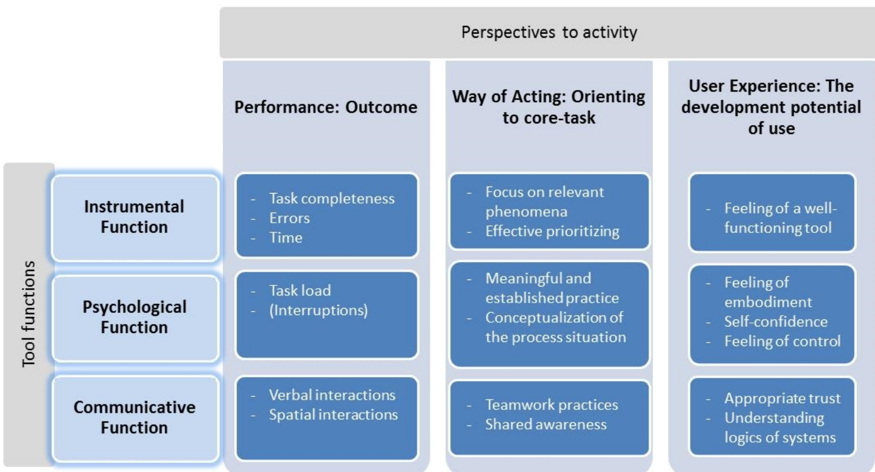

Figure 7. An example of applying systems usability framework to identify measures: The indicators utilised in the base line studies in the Finnish NPPs.

In what has gone before, the notion of systems usability and theory based justifications were described. They constitute one of the main results of the dissertation. SU emphasizes the systemic and mediating role of information technology in an activity system. The concept refers to a system's overall meaningful role in an activity, which can be demonstrated and assessed by defining a contextually meaningful set of SU indicators. The point of utilising SU in control room development and evaluation is to ensure systemic consideration of the tool. This means that the activity is considered from different perspectives, which enables inclusion of the global, societally defined purpose and objective, which are the reasons for the activity. The SU framework for developing indicators covers the functions of the tool and the perspectives to activity to ensure the necessary scope and depth in terms of abstraction levels in the analyses of tool appropriateness.

5.1.3 Significance of UX in designing tools for complex work

The other conceptual research question regarded the role of the concept of user experience (UX) in the development of tools for complex safety-critical work. As indicated above, the concept of experience is one perspective from which an activity is analysed (Figure 7) in the evaluation of the SU. The exploration of the concept of UX was conducted in the NPP domain, and the results of empirical work are reported in Paper VI of this dissertation.

In principle, the design of new control room UIs in the nuclear industry is based on a conservative strategy. Conservatism in design, if pondered in particular from the point of view of human factors and design of interaction technologies, means that the new control room systems in NPPs apply the same design philosophies and are based on the same design principles as the existing ones. This approach to design precludes the taking into account the modern understanding of human cognition and technology mediated activity cited in the related research section of the dissertation, such as the theory of situated action [Suchman 1987], embodiment [Dourish 2001], distributedness of cognition [Hutchins 1995], and the socially constructed and driven meanings conveyed by the technology [McCarthy & Wright 2004]. Thus, even if evolutionary improvements in human system interfaces are being conducted in the nuclear domain, they are mainly based on dogmas and “industry standards”. This conservative approach to design imperatively hampers the development of the whole field, because the requirements of a modern understanding of design as pluralistic and inherently humanistic [Bardzell 2010] are not fulfilled.

”The human factor” is, of course, reckoned to exist and actually to have quite a significant effect on overall system safety in the nuclear domain. Even though the advances in understanding the foundations of human behaviour cited above have shown that the human role in achieving the objectives of activity could be perceived more widely than seeing the human as a risk factor, the contemplation of only the possible negative consequence of human action is still quite dominant in a safety-critical industry. Constituting an opposite approach, the objectives advocated by resilience engineering [Hollnagel, Woods & Leveson 2006] aim at understanding also the positive role of human beings in construing safety. In considering the role of UX in the design of new systems, the idea of utilising UX as an early indication of the successfulness of a new design solution was developed.

Inspired by the resilience engineering paradigm, it is also possible to address the design of new systems from the point of view of positive variation. The question may be formulated: What makes design activity resilient? In order to be resilient, design should be sensitive in its practices to different kinds of signals concerning the successfulness of new solutions. Relying solely on dogmas and norms surely standardizes designs and lessens variability, but maybe at the same time it may hamper general development because the emergence of positive variability is also restrained by the standardization. In the three studies presented in Paper VI of this dissertation, the motivation was to investigate whether, by utilising UX as an indicator of system appropriateness, some conclusions could be made about the successfulness of designs which vary in maturity. In the studies, UX was utilised in addition to the standards, guidelines, and existing solutions which are prevailing design drivers today. It was claimed that the UXs of the potential users could be utilised as anticipatory indicators of the effects that the new system will have on the usage activity, because they are exactly such; users try out and test new tools, and general feelings concerning the tools emerge. The usage activities of the new tools are not yet mature enough that it would be meaningful to use indicators such as errors, performance accuracy or time. Therefore, the following and taking into account of the UXs emerging in different test situations during the design process is considered as a characteristic of a resilient design activity aimed at foreseeing the impacts of new technologies already in the design phase. It was proposed that, in resilient design activity, the UXs of the individual users concerning the new developing tool could be utilised as early signals indicating the potentiality of the new tool. The professional users, who have years of experience of working in the domain, may have an intuitive feeling concerning the general appropriateness of the proposed new tools for the activity. Also, a notion of promise has been utilised [Norros, Liinasuo & Hutton 2011] to indicate the added value that only the professional end-users could experience concerning the technology. It may only be a gut feeling of an operator that a proposed tool is, or is not, suitable for the particular work, but in the safety-critical settings, the feeling should be taken seriously and explored thoroughly to find its sources and significance.

Based on the explorative studies concerning UXs in a complex NPP context, I conclude that significance of the concept of UX is in informing design activity of the potentiality of the new tool very early when it is not yet feasible to use performance measures. A design activity which is able to take advantage of this kind of anticipatory indicator concerning future usage can be considered resilient, because it addresses the issues which do not yet exist but contains inclinations of developments which are relevant for safety.

5.2 Procedures for identifying the SU of the control room

This subsection provides the results related to the methodical research questions concerning practical ways of assessing the SU of a particular tool. In the review of existing approaches, development needs were identified in 1) analyses of the domain and the related implications for evaluation approach, 2) modelling of operating tasks, 3) analyses of operating practices to identify resilience. These specific methodological questions complement the results of the conceptual theoretical elaborations concerning SU, as it is expected that systemic evaluation may need detailing of the evaluation methods.

5.2.1 Evaluation process to reveal systems usability

In order to build a methodology for evaluating the appropriateness of control rooms, the human factors evaluation methods and the state of the art usability evaluation were reviewed and some development needs were identified (Section 2.4). As a result, an evaluation process which addresses the identified short comings in the existing methods was outlined (Figure 8). The approach aims at an evaluation process which is able to assess the effects of the technology on safety, and to understand the mechanisms of interaction on a level which enables identification of directions towards which the tool may drive the operating practices in the long run. The potential hazards may not be evident on the level of outcome, but the possible early signs of degradation in the sociotechnical system are made explicit in the evaluation. The process was labelled contextual assessment of systems usability (CASU), and it was first introduced in Paper III of this dissertation. In the CASU evaluation approach, the empirical-positivist approaches, as Upton et al. [Upton et al. 2010] term them, arising from, for instance, cognitive psychology, and the empirical-hermeneutic approaches deriving from a post-modern understanding of human conduct have been interwoven. The approach implements in practice the theoretical concept of SU introduced above. The CASU method has been used in studies of nuclear power plant control room modifications.

The CASU approach to evaluation consists of four separate phases – Modelling, Data collection, Data analysis, and Assessment – which are all briefly described below. The main additions of the CASU approach to other prevailing evaluation approaches are in the detailed analyses of the work domain and the operating situations, and the operationalization of concepts of practice and user experience.

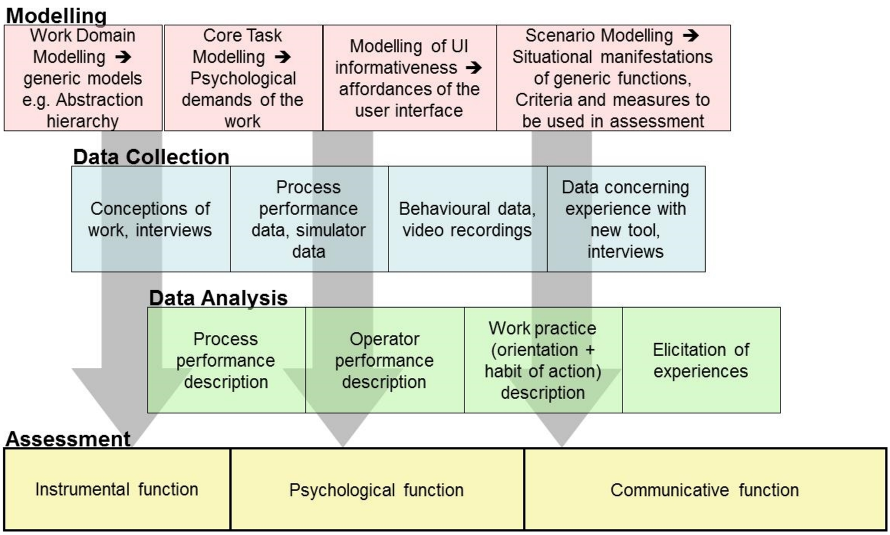

Figure 8. The process for contextual assessment of systems usability.

5.2.1.1 Modelling

The first phase – the modelling phase – outlines the basis for the evaluation by analysing in depth the context of use of the control room. This phase in particular is in concurrence with the activity theoretical thinking concerning the object-orientedness of activity, as the models elicit information concerning the object. The models act as internal references on which the data analyses are based.

The modelling includes analyses of the of the work domain. This means that the intrinsic physical laws governing the production process are identified, and the technical process components which realise the processes are mapped. For this purpose, an analysis method originally developed by Jens Rasmussen [Rasmussen 1986], the abstraction hierarchy (AH), may be utilised. An abstraction hierarchy provides a general view of the controlled process. It reveals, in means-ends relationships, how the process components are capable of realising the objectives of production.

In order to be successful in their activity, it is necessary that actors take into account the possibilities and constraints of the domain (expressed, for example, in the AH). So as to elicit this information, the modelling method developed by Norros [Norros 2004], the already briefly mentioned core-task modelling, has been utilised.

In order to conduct an evaluation of systems usability, it is necessary also to analyse what type of information and possibilities of acting the system under evaluation provides to the operators. This modelling is called modelling of the UI informativeness and it produces as an outcome the intended affordances in the system under evaluation. [Norros et al. 2009]

The previously mentioned domain and work modelling methods are general in nature, but as evaluation must be conducted in real usage, a model of the particular situation utilised in evaluation must be created. In the situation, it is possible to describe in more detail the functions of the domain which are in question in this situation, and the expected operator actions to respond to the requirements of the domain. To elicit this information, a task analysis method, the functional situation model (FSM) is used, which is presented in more detail below in Subsection 5.2.2 (p. 96).

All the modelling methods together produce the collection of contextual models which can be used as a reference for comparing the empirical data collected later. Also, the evaluation measures and criteria can be defined based on these models. In the modelling phase, the main functions of the domain and how they are realised by the technical system, what is the core task of the human operators of the system, what are the affordances in the UI with which the work is supposed to be conducted, and finally in the particular operating situations selected for the evaluation, how are all the above manifested.

5.2.1.2 Data collection

The second phase in the evaluation process is data collection. The aim of this phase is to collect empirical data which is relevant from the point of view of the SU. This means that the data should be such that fulfilment of the three functions and the core task can be evidenced.

In practice, there needs to be at least objectively and independently gathered data in order to judge performance, personal conceptualizations and observations of behaviour so as to reveal ways of acting, and subjective considerations so as to reveal experiences.

Performance data can be categorised to concern either operator (human) performance or process performance. Human performance refers to measures such as task completeness, time and errors. For these, there are approved methods, the predictive power of which has been studied [Hwang et al. 2008b]. Such measures are for instance the NASA Task load index [Hart, Staveland 1988], some physiological indices, task error rates. Process performance data is based on important process parameters. It can for instance be analysed that in the particular scenario the primary circuit pressure is an important parameter, and its behaviour is indicative of fulfilment of some of the goals identified in the modelling phase. Thus, data concerning primary circuit pressure should be collected. Typically, process simulators utilised in the test situations store this information in the log files. Expert evaluations basing on the FSM were also conducted in the empirical studies of research work so as to gather data concerning performance. In addition, process tracing interviews were conducted post scenario. In these interviews the operators revealed their understanding concerning the process situation and the tools utilised in order to control the process.

Data concerning ways of acting is somewhat more complicated. An interview method was developed (based on earlier work [Klemola, Norros 2001, Norros, Klemola 1999]) within the CASU framework to reveal operators' personal conceptualisations and valuations concerning their work. It is difficult to establish causal relationships, but there is little doubt that attitudes and behaviour are somehow linked [Stanton, Ashleigh 2000]. The method developed, the orientation interview, is a short interview in which, based on answers to a few indicative questions, the operators' personal epistemic attitude towards their work can be assessed[20]. In addition to conceptions, the way of acting is evidenced in the operators' actual corporeal practices of conducting process control work. This data is best captured by video-recording the activity of the crew. In the studies, head-mounted cameras were utilised to enable the operators to view the work and to enable detailed level analyses of information usage. In addition, an overview camera recording provided information concerning operators' physical locations, movements, and communi-

[20] The orientation interview method has been reported in detail by Norros et al. [Norros, Liinasuo & Savioja 2013], but the results are also discussed in articles 1 and 2 of this dissertation.

cation and collaboration. Additionally, the operators' own viewpoint on the conducted process control activity was acquired by conducting process tracing interviews.

In order to gather data concerning the operators' experiences, a questionnaire method was developed, in which the three functions of a tool were operationalized for the benefit of understanding experiences concerning them. This work concerning UX evaluation is reported in Paper VI of this dissertation. In addition to the questionnaire method, UXs can also be analysed based on the interview data, especially the process tracing interview.

5.2.1.3 Analysis

The third phase is the analysis phase. Analysis aims at making sense of all the data gathered concerning the SU of the control room systems under evaluation. In the analysis it is important to maintain the systemic approach. This means acquiring different perspectives on the collected data.

As the data collection really produces an abundance of data, it was found useful to start making sense of it by first creating a chronological view of the scenarios. In this, the information from the simulator logs is combined with the video-recorded observations of the activity of the crew. All the communications, movements, process operators, and critical parameter behaviours are transcribed in a single file. From the results of this analysis it is possible to identify the process, and the human performance, by comparing the results to the reference identified in the modelling phase.

In the analysis of practices, the results are based on two different data sets: the operators' conceptions concerning the work and the crews' ways of acting in process control. As main data, the orientation interview results and the observations of crew activity are utilised. The FSM serves as a fundamental reference in the analyses based on the ways of acting which can be valued. In analyses of activities special attention is in addition paid to ways the crew is utilising the information technological resources. This can be considered 1) an indication of the feasibility of the systems and 2) an indication of the person's intentions in activity [see Edmondson, Beale 2008 for "projected cognition"]. The methodology of analysing operating practices is presented in more detail in Section 5.2.3 (p. 102) and in Paper V of this dissertation. In Paper V the differences identified in practices between the crews in observational data are also characterised.

The questionnaire-based method utilising the analysis of user experiences is presented in Paper VI of this dissertation. UXs can also be identified from the interview data, especially the process tracing interview.

5.2.1.4 Assessment

The final phase of the evaluation process is the assessment. In the assessment, both positive and negative issues identified in the data analyses are gathered together. The assessment is made by combining three points of view: performance results, results concerning the tools' ability to promote appropriate work practices, and results concerning operators' experiences. In either of the three viewpoints there may be both positive and negative issues that are identified.

By connecting the identified issues to the three general tool functions, it is possible to discuss their significance and the possible design changes they necessitate. For example, it is evident that a significant problem in the instrumental function most probably may directly point to a design implication, whereas a problem concerning psychological function may require more thorough investigation of the possible reasons for the problem.

If the evaluation is a validation study, it is necessary to make a statement concerning the acceptance of the system. To achieve a final statement, methodical triangulation is needed. In this process the question of the appropriateness of the control room design is approached from different perspectives and in the light of different kinds of evidence. As the sheer amount of evidence may be colossal, ways of organizing and structuring evidence are needed so that traceability can be assured.

5.2.2 Scenario model for identifying operating practices

This subsection describes, in more detail, the task-analysis method developed specifically so as to provide an evaluation basis for analysis of operating practices[21]. A more detailed elaboration of scenario modelling is presented above as part of the modelling phase of the general process of evaluating systems usability of UIs in complex work.

As stated earlier, the analysis of operating practice answers the question of *how* operating work is being conducted. This is a distinction from the more typical question of *what* actions are being conducted. The question of how is important because it reveals the *meaning* that the particular action has for the actor. By analysing meanings it is possible to generalise the results, because people have a tendency to act in a way they find meaningful. For example, an action which is considered important or critical by the actor, is more likely to be conducted with particular care as opposed to something that is not considered critical. This feature in operating activity, the meaning that particular actions have for the actor, does not become evident if only "what" type questions are asked in the analyses.

In the domain of process control work generally the meaning of an action is reflected in its significance from the process control point of view. In order to understand what the situational meaning of each action is, a task model is needed which connects the actions to the critical functions of the domain. For this purpose, we present the FSM approach. A general schema of an FSM is presented below (Figure 9).

[21] "Way of acting" and "practice" are used synonymously here.

An FSM combines two viewpoints to the modelled situation; chronological and functional. These two dimensions define a two-dimensional plane in which the most important operator actions and process phenomena are mapped for the purpose of understanding both operating performance and practice.

5.2.2.1 Chronological view

The chronological view is the most obvious way of analysing the operating activity. In the chronological view the scenario is divided roughly into different phases in which the operating actions have a specific goal. In the FSM the chronological phases are labelled according to the general goals: Detection, Mitigation of effect, Diagnosis, and Stabilising the process state. Although the distinction is made between the phases, it must be acknowledged that in real activity all the goals are all present simultaneously. But, for the purpose of making sense of the process situation, the distinction is nevertheless made in FSM modelling. Also, in a different circumstance the phases might have different goals, and thus different labels for the phases could be utilised.

Detection: In the model, detection phase denotes that the crew identifies some process events requiring operator actions. The process information presented to the operators is typically alarm information and notifications. Simultaneously, monitoring of all process information is conducted by the crew. In the FSM, the detection phase is concerned with the *information* that is available in order for the operators to understand the deviations in the process. In modelling the detection phase, it is also important to explain the initial event(s) in the scenario. The most important alarms informing the operators are included in the model. The implications of the initial event on the process state are depicted at parameter and overall process level.

Mitigation: The line between mitigation and detection is not always distinct, as the operating activity actually constitutes a continuous cycle of monitoring and acting. Thus, the exact point in time when the mitigation phase “starts” is not so important, but we do want to make a distinction that some operating actions are tuned more towards mitigating the situation than perceiving information about the situation. It is also typical that automatic functions handle some of the actions in this phase. In this case, important operator actions are to monitor and confirm the fulfilment of the automatic functions. In the mitigation phase of an FSM, the operating actions that mitigate the process situation are mapped under the specific process information and the initial events of the detection phase to which they are connected.

Diagnosis: As the ultimate operating goal in an accident situation is to bring the process into a safe stable state, diagnosis related actions are required from the operating crew. It is important to realise what the process situation is in order to identify the required actions. In the diagnosis phase, these actions are depicted in the model under the specific parameters of detection phase to which they are related.

Stabilisation phase refers to the operating activities which aim at bringing the process into a safe and stable state. These actions are also connected to the relevant initial events.

5.2.2.2 Functional view

A functioning sociotechnical system has the objectives of production, safety, and health [Vicente 1999]. These objectives form the basis of the functional view of the FSM. These objectives are depicted as the three "lanes" below (Figure 9). The aim of the functional view is to identify, on a situational level, which process functions are the means to maintain the three overall objectives, and how these means translate into actual operating actions and operations.

The most important items in the functional view are the critical functions of the process which are endangered in the specific situation. Critical functions refer to the domain-specific core functionalities of the process, which enable achievement of the general objectives. Typically, in a complex situation there are critical functions related to all of the above objectives which are endangered. For example, in the nuclear domain the critical functions originally explained by Corcoran et al. [1981] may be summarized as Norros [Norros 2004] writes: Reactivity, Core heat removal, Coolant inventory, Pressure control, Heat transfer from the primary to the secondary circuit, Containment pressure control, Containment integrity, and Power for emergency systems.

Operating activity, on a high level, is oriented towards maintaining the critical functions of the process. Thus, the operating actions required in the situation can be collated under general means to respond to the endangerment of the critical functions. This is depicted in the FSM by presenting a specific functional level in the model "functional means to respond". This level makes the connection between the individual operating actions and the critical functions. The relationship with critical function level upward and with individual actions downward is of the type means-ends. The means-end relationship resembles the one utilised in the abstraction hierarchy AH [see also Lind 2003, Rasmussen 1986].

5.2.2.3 Connecting chronological and functional views

By connecting the individual level actions of the crew in depicted chronological view with the functional view to the situation, the meaning of each action in the wider context of the scenario is made explicit. This is the core of an FSM which enables the analysis of operating activity on the level of way of acting.

This notion is based on an idea that making this connection in action is actually what enables operating practice to support *resilient* activity in a situation. If and when the operating crew is tuned towards the critical functions of the domain, they always take them into account in their actions and follow how the critical functions behave and are affected during the course of complex process control activity. Making the critical functions the object of attention in the activity is a way of remaining vigilant also for the unexpected features of the situation and thus enabling resilient activity. Resilience in this sense means that the activity is able to respond to external events that may not have been rehearsed or pre-planned. In order for the operating crew to be able to handle the unexpected situation when it takes place, they must always be tuned to the critical functions, even in a situation which is well rehearsed or planned. The FSM makes explicit which critical function is targeted with which action. Thus, in following when the crew takes the particular action, it is possible to analyse whether they make the connection to the function. This connection may be verbal (in communication) or there may be evidence for instance in the direction of gaze data, whether the critical function-relevant parameters are followed. Observing this, enables analysis of the meaning processing [Bennet, Flach 2011] which takes place in the situation by the operating crew.

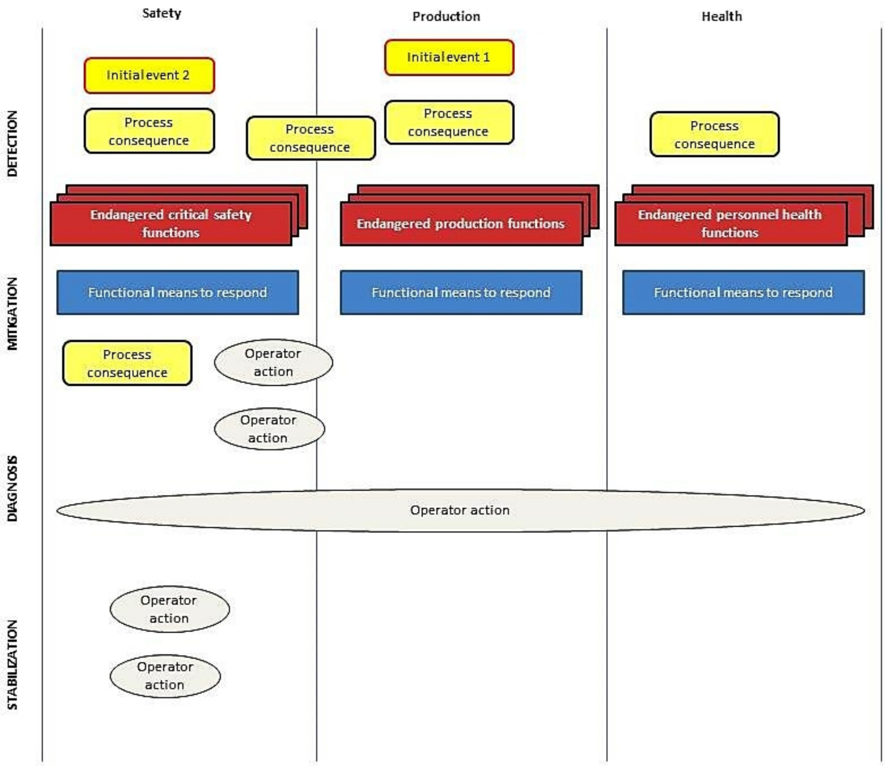

Figure 9. The general schema of an FSM. In modelling a scenario with an FSM, the boxes describing process events and information (yellow), critical functions (red), and operating crew actions (grey) are depicted as connected to the critical functions (read: safety, production, health) and the operative means to mitigate the functions (blue). The description is organised according to the chronological phases (detection, mitigation, diagnosing, and stabilization) according to the scenario in question.

5.2.2.4 Example: FSM for LOCA in a PWR plant

The following is an explanation of the FSM depicted below (Figure 10). The model was developed for the analyses of LOCA scenario of the Loviisa NPP baseline studies.

In the particular LOCA the initial event is a leak in one of the primary loops. The leak is of the size 50 kg/s. The process consequence visible to the operating crew is that the pressurizer level drops dramatically. Also some alarms go off, but they are not depicted in the model. Another initial event was also included in the scenario. This was a failure in a plant protection signal which governs the automatic protection systems of the plant. The failure was such that the signal was initiated, but it failed to launch the implications.

A critical function which is primarily endangered in the LOCA is that of primary circuit mass balance. This function is supposed to maintain that there is always enough mass (cooling water) in the primary circuit in order to cool the reactor core. In addition, heat transfer is endangered, as the coolant is leaking in the containment building and is not capable of transferring the heat to the secondary circuit, although there is no real failure in the systems related to heat transfer.

The failure in the plant protection signal (modelled as a secondary initial event) endangers the critical function of containment integrity. When the automatic plant protection system is not functioning correctly, the containment is not isolated properly.

The critical function of emergency power was also considered to be endangered, because power systems are crucial in any accident situation in order to enable the functioning of safety systems which require electricity.

Production-related critical functions were not considered in the LOCA scenario, as it is an accident, and operating activities are mainly directed towards ensuring safety.

A personnel well-being-related function was identified as in a LOCA situation the whole plant is evacuated for safety reasons.

Several means to mitigate the endangered safety functions were identified. In order to maintain mass balance in the primary circuit, it is necessary to take the auxiliary feed water systems into use. This happens via automatic plant protection functions, and the operating crews' responsibility is to monitor and detect that safety injection pumps are starting. Heat transfer is mitigated by concentrating on primary circuit cooling. At the outset, the important operating task is to assure that the automatic scram has functioned correctly. This can be done either by the shift supervisor's decision or based on an emergency operating procedure. Containment integration is mitigated by means of isolation and emergency power by manually assuring the start-up of the diesels. Personnel safety is mitigated by means of evacuation.

In the diagnosis and stabilisation phases, the relevant operating crew actions were picked from the emergency operating procedure and depicted under the critical function to which they mostly refer.

In the LOCA FSM the process events are depicted as yellow boxes. The initial events have a red line. The endangered critical functions are marked as red boxes and the relating operative means are depicted below as blue boxes. The operating crews' actions are marked as light grey boxes. In the text it is always marked whether the action is the duty of the turbine operator (T), reactor operator (R) or the shift supervisor (VP).

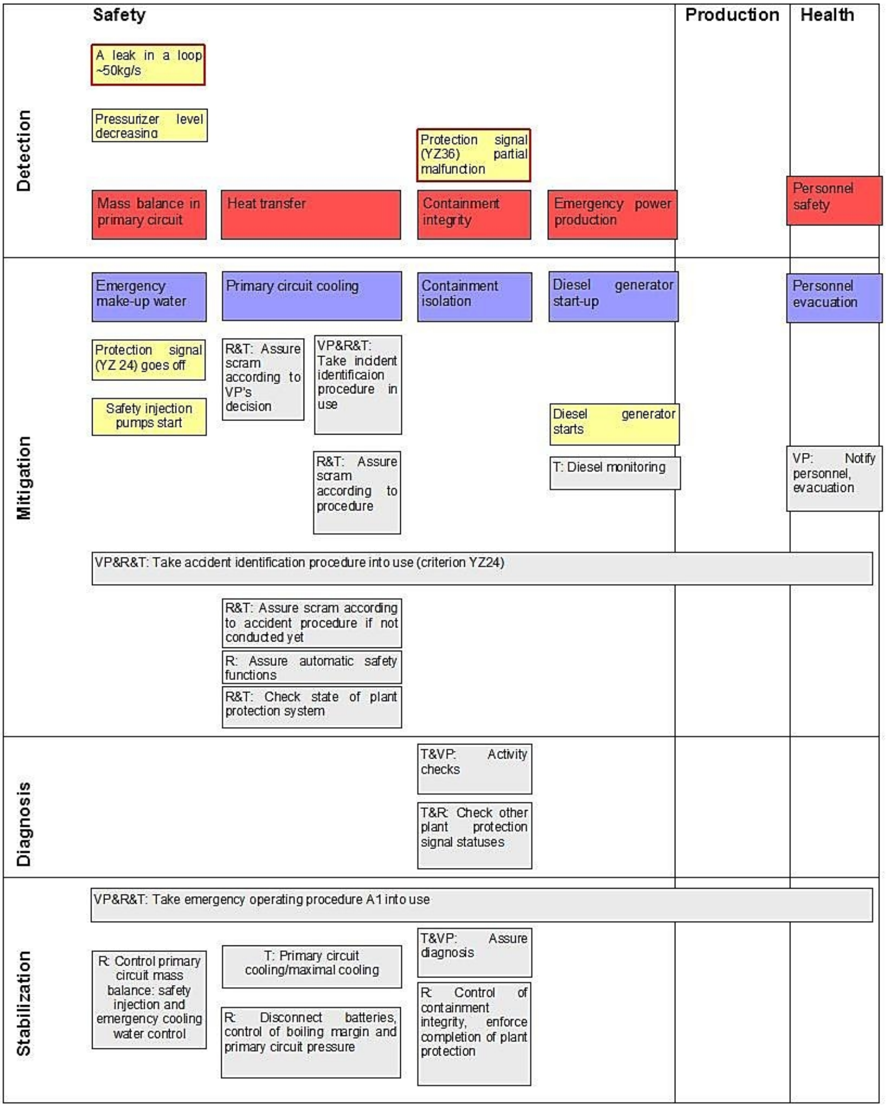

Figure 10. An FSM of a particular LOCA situation in a PWR NPP.

5.2.3 Analysis of operating activity for evaluating ways of acting

This subsection details the analysis approach utilised in identifying the potential for system level resilience in the observational data of operating crew activity in accident management. The detailed level results are presented in Paper V of this dissertation. The empirical results: the characterization of practices supporting resilience are reported later in Subsection 5.3.2 (p. 115).

The LOCA scenario of Loviisa NPP baseline studies was carefully analysed in order to understand the operating crews' ways of acting. Simultaneously, an analysis method was developed while the actual operating practices were investigated.

The data utilised in the analyses was the video-recorded behavioural data of each crew's activities in the scenario. Each scenario was video recorded with 5 cameras: Each of the three operators carried a head-mounted camera which recorded operations, directions of gaze, and communications. In addition, there were two overview cameras and an audio recorder which recorded as widely as possible the whole activity taking place in the control room.

The analyses were conducted on two different levels within the activity: action and habit of action[22]. In practice this means that, first the activity was considered on the level of action. Answers to the question "What do operating crews do in the situations?" were sought. In this analysis, a detailed description of all operations, communications, movements etc. was made[23]. Next, the level of analysis was raised to the level of habit of action. On this level the question guiding the analysis is "How do crews act in these situations?" This means that qualitative differences in the ways the crews acted were first identified and then classified. In the third phase of the analysis the categorisations identified on the level of habits were brought back to the level of action, and an understanding of explanations of actions with regard to system resilience was sought (Figure 11).

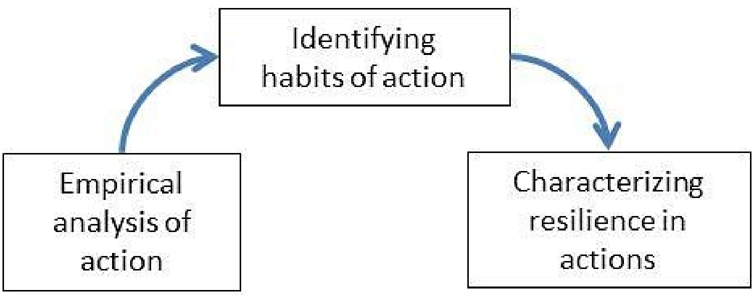

Figure 11. Process of identifying resilience by analysing first actions and then habits of action.

[22] Habit of action is the operationalization of operating practice.
[23] This also enables analyses of performance.

5.2.3.1 Habit of action as the unit of analyses

The idea of using habit of action as the unit of analysis and operationalization for the concept of practice comes from Norros [2004]. The general idea is that *habit* is the notion which expresses an organism's way of meeting the changing and sometimes unexpected features of the environment. By standardizing features of own behaviour into habits, the organism is prepared to face the dynamic world. Without habits, every situation would be treated as new, and acting in a situation would always be as difficult as the first time. The reasoning that a particular behaviour expresses a habit, and would appear again in corresponding situations, draws on Georg Henrik von Wright's idea concerning behavioural inference of reasons [von Wright 1998]. It states that, if a person understands the meaning of a particular sign as a reason for acting, this means that ordinarily the person responds to the sign accordingly, unless they have overruling reasons against the action (see also Norros [2004]).

Philosophers Charles Sanders Peirce and John Dewey saw habit as a fundamental principle of human thinking and action [Norros 2004 citing, Peirce 1998b, Peirce 1998a, Dewey 2002]. Human actors relate to the possibilities of the environment by continuous action-perception cycles, via which the outcomes of action are observed by themselves. Acting and observing and interpreting the effects of acting are inseparable. In the process, guesses concerning the world and possible explanations are made and tested. The fundamental role of habit is to enable interpreting the cues of the environment and anticipating the effects of one's own actions. Habits are repeated as they are meaningful.

Peirce proposed that habits have a semiotic structure [Norros 2004, citing Peirce 1958]. Habit expresses the principle of human thought, which allows a thing to be in some way substituted or represented by another thing, i.e. representing real world objects with crafted ones. For example, in a control room, the real process phenomena is presented utilising different kinds of signs. Signs, the form of which is heavily dependent on the technological medium applied, are used to represent objects, and understanding of this relationship becomes evident in an action, thought, emotion, or another act or behaviour labelled as interpretant by Peirce (Figure 12)[24].

[24] See also the presentation of Peirce's model by Bennet and Flach [2011, p.18].

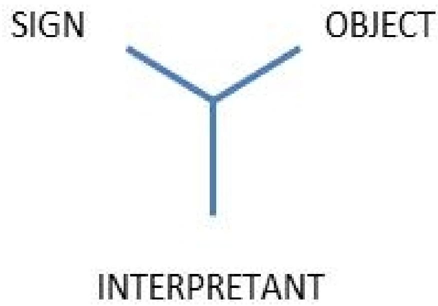

Figure 12. The semiotic structure of habit [Peirce 1958].

Exploiting the former theoretical ideas, Norros [2004] proposed that the semiotic concept of habit could be used in an empirical analysis of action so as to identify what meaning the observed actions of subjects have in a certain domain and in the particular studied situations. In this work I demonstrate that by utilising the semiotic notion of habits it is possible to find differences in operating crews' ways of handling proceduralized emergency situations which may be significant from the point of view of reaching the general goals of process control.

Using the above-described apparatus in the analysis of accident management practices in simulated LOCA, indicators of habits were constructed from the observed behaviour of the operating crews. In the LOCA study the intention was to characterise nuclear power plant operators' ways of acting and using procedures, in order to see whether it was possible to identify inclinations towards behaviours which could become unsafe over time or in a particularly demanding situation. A set of indicators of habits of action were constructed which characterise NPP operators' actions and the ways procedures were incorporated in the actions. The construction of the set of indicators is explained in the next section.

5.2.3.2 Identifying habits of action in the data

First in the analysis of practices based on FSM and analyses of the operating practice of one crew, the whole scenario was divided into meaningful episodes. Altogether, there were three separate episodes which were especially meaningful from the point of view of objectives in the situation. In the second phase of the analysis, the operating activity of each crew in these episodes was analysed utilising the concept of habit.

Identification of habits was conducted for each episode and for each crew separately. The treatment was carried out first for all crews in episode one, then for all crews in the next episode, etc.

In the study of habits, in actual operation, the perception action cycles via which the operating crews conducted the proceduralised tasks in the situation were investigated. The three interconnected components of habit – sign, interpretant, and object – were utilised. The interpretant was considered to be the element of habit which is directly observable in the behaviour of the actor. It was assumed that the acting of a person is driven by interpreting the signs in the environment. The sign was considered to be any perceivable element or feature in the environment. It could be an action (e.g. a communication) of another person, or it could be technology-mediated sound or a visual feature. The idea or object to which the sign can be connected via the interpreting behaviour reveals the objective which is sought by the actor, in other words, the way the crew acts based on particular signs reveals what they consider the situation to demand from them.

In the identification of habits, what people did (interpretant) was first studied. In this step the key actions (verbalisations, operations, movements etc.) of each crew member in the episode were written out in text format. The next step was to look at the very moments prior to the defining actions of the episode, and identify the signs based on which the particular actions were carried out. The signs were for instance previous actions, process information, information in the procedure, sounds, communications, etc. This information was detected from the video recordings and, for example, directions of gaze were utilised. In the last step, we deduced the objective in the habit by analytically comparing the interpretant, the sign and the demands of the process situation described in the FSM.

As this analysis of perception-action cycles was carried out for the entirety of the data, it was discovered that there was definable variation in the micro-level activities of the crews. All the crews conducted the process control actions complied with the procedure, but the ways of conducting the tasks were different. In particular, the signs based on which the actions were carried out. As these differences were further analysed, process control tasks became identifiable in which the crews' habits of action differed from each other. If, based on the analysis, there were significantly different perception action cycles in an episode concerning some particular issue, it was analysed whether it constituted a process control task. In episode 1 we identified four different tasks, in episode 2 two tasks, and in episode 3 one task. The tasks and the respective habits are presented in Section 5.3.2 (p. 115) of the dissertation.

5.2.3.3 Grading of habits with respect to resilience producing mechanisms

In the analyses of the habits it became evident that distinctively different habits exist to fulfil the identified tasks of proceduralized accident management activity. This variation occurs not only between the crews in different episodes, but also within the crews continuing process control activity from one episode to another.

In the grading of habits the aim was to identify the habits which could be justified as adding to the system level resilience in the activity. For example, an alarm as a sign might be interpreted by seeking other process indications confirming the particular alarm, which could be a resilience-producing habit of action because it is oriented to the process information and making thorough sense of the situation. Whereas, the same alarm could also be interpreted by conducting exactly and only the action the alarm indicates. This kind of habit of reacting to alarms can be argued to have less capacity to produce system level resilience, because it takes the alarm information provided for granted and does not aim for contextual interpretation of the situation, which is a prerequisite for identifying whether the predefined procedure actually fits the situation.

The grading of habits was carried out on a scale ranging from interpretative to reactive, interpretative being the one which is reflects most profoundly developed human-environment relationship. This scale of grading habits was originally introduced by Norros [2004]. As explained earlier, the basic distinction between reactive and reflective relationship with the environment is made, and a confirmative relationship is assumed between these two. In this study the scale was predefined, but the qualitative characteristics of each class were grounded in the data. An interpretative habit is such that behaviours can be identified which point in the direction of questioning the observed phenomena, building expectations of future events, and one's own interpretation of the situational demands. At the other end of the scale is a reactive habit. A reactive habit is such that it reflects passivity and a lack of expectations concerning the situation. No indications of own interpretations and reflections are identified in the data. Weakness of reactive habit is that it is only able to react to situations; there are no anticipatory aspects. A confirmative habit is such that neither reactive nor interpretative characteristics can be identified. A confirmative habit can be described as taking the situation for granted, acting in a pre-defined way, and over-emphasizing rules and procedures.

As an end result of all the analyses, it became possible to formulate results such as: Crew A fulfils a function z with a habit y. It also became possible to give a contextual description of the function z and the habit y. These results, the functions, characterisation of the interpretative, confirmative and reactive habits in them, and the distribution of the habits among the crews are reported in below in Subsection 5.3.2 (p. 115).

5.2.4 Procedures for analysing UXs in complex work

The significance of the concept of UX in NPP operating work was addressed by operationalizing the concept of UX within the SU framework for each function of a tool in an activity (instrumental, psychological, and communicative). Based on the operationalization a questionnaire was developed, which was utilised in three different studies to elicit information concerning the potential of the evaluated systems.

Below are also depicted the specific indicators which were utilised in the questionnaire studies (Table 3 – Table 5). Many of the statements were modified during the studies to fit the purposes of each study. Nevertheless, there was an attempt to keep the indicator levels constant.

5.2.4.1 Operationalization of UX for instrumental function

The instrumental function refers to the tool's ability to have an effect on *the environment* for which it has been designed, e.g. a control system user interface allows the user to manipulate process components. Instrumentality or instrumental value is often discussed in UX research also, but the point is usually to contrast UX with usability [Palviainen, Väänänen-Vainio-Mattila 2009]. In a systems usability approach, based on the notion of experience in AT, I do not separate UX from the instrumental role of tools. The experiences of the actor concern the whole activity, including the instrumental value of the tool. Especially in work settings, it is the fulfilment of the instrumental function of the tool which is important for the emergence of positive user experiences. Probably everybody can relate to the positive emotions associated with “getting the job done” or “finishing a project”. This is perhaps a difference in comparison with “main stream” UX research which considers UX almost as a counterforce to the earlier usability research which emphasized the more utilitarian values in product usability development. The instrumental function of the tool is concerned with the tool’s ability to allow the user to have an effect on the environment. If the control room functions well in instrumental function, the operators are able to use the system to conduct the process control actions they see adequate in the situations. They are also able to monitor the fact that the intended process effects are reached. Through analytical contemplation, including the principles identified in relevant theories, applying notions of dialogical empathy [Wright, McCarthy 2010], and exploiting one’s own experiences of several development projects in the NPP industry, five indicators of good UXs were developed for the instrumental function of a MCR under the general heading of: Feelings of appropriate functionality.

The first indicator (I1) addresses the feeling of suitability of the tool for controlling the process. The operators must have a feeling that the control room systems are actually capable of conducting the process operations that are indicated. This is quite a basic requirement for the system, and it reflects the AT principle that emotions concern the status of the whole activity [Kaptelinin, Nardi 2012].

The second indicator (I2) addresses the feeling that control system interfaces are consistent with what is required from the operating crew in the light of their operating procedure. The operators must feel that the control room systems support them in achieving the objectives of the work. This general “feeling of support” was also identified by Palviainen and Väänänen-Vainio-Mattila [2009] as a relevant UX factor in work settings.

The third indicator (I3) addresses the feeling of ease of use, and effortlessness in monitoring and controlling the process through the control room systems. Even though control room operating is work and can be very demanding in certain situations, the general feeling of the operators should be that utilizing the control room systems is effort-consuming only to the extent that is necessary.

The fourth indicator (I4) addresses the feelings that the operators have about the quality of the control room systems. In order for the tool to function in instrumental function, the operators must have accumulated experiences which indicate that there are no recurring or obvious problems in the control room systems. This indicator reflects the same phenomena that Obrist et al. [Obrist et al. 2011] referred to as usability/ergonomics.

The fifth indicator (I5) is concerned with the feelings that the physical form of the user interfaces induce in the operators. In a well-functioning control room the operators should have the feeling that the physical input devices “obey” nicely. The operators should have the feeling that the systems behave appropriately at their hands and fingertips. This is a kind of feeling of appropriate embodiment [Dourish 2001].

The indicators for good user experience of the instrumental function of the control room are not a complete set, but address control room functioning from multiple perspectives, for example appropriateness of work and embodiment and on different abstraction levels, thus fulfilling the general objective of a systemic approach.

Table 3. The statements used in the questionnaires. I1–I5 indicators of instrumental UX A, B, C refer to the separate studies and the exact format of the statement that was utilized in the particular study. The number after the letter, e.g. B2, refers to the modifications made to the statement in the course of the three studies. If there is no number, the statement was utilized in exactly the same way in all studies. If there are numbers, it means that the statement was modified for the purposes of the later studies. Some of the indicators were utilized for more than one statement, e.g., I1 had three statements in study B (I1_B1, I1_B2, I1_B3). In the modifications, the statement was contextualized for the study, but the main content of the indicator was kept constant.

Instrumental function UX: Feelings of appropriate functionality	
I1	(A1, B1, C1) In my experience, the control room systems function as they should
	(B2, C2) Touch screen response feels appropriate to me
	(B3) The system functions in a way which corresponds to my comprehension of how it should function
I2	(A1) According to my experience, the control room systems support operators in achieving the goals of work
	(B2) The layout of the pages supports operators in their tasks
I3	(A1,B1) I experience that following the process from the control room systems is easy
I4	(A1, B1, C1) There are no specific problems in the control room systems that I would encounter repeatedly
	(B2, C2) In my experience, touch screen technology is appropriate for the control of safety critical systems
I5	(A1, B1, C1) The control room systems allow a good sense of control in the fingertips
	(B2) The rotary switch allows a good sense of control in the fingertips

5.2.4.2 Operationalization of UX for the psychological function

The psychological function refers to the effect which the tool has on *the user*. In order to be able to use the tool, the user must have an understanding of the mechanisms with which the tool has an effect on the environment. In usage, the user develops ideas, mental models and schemas about the tool as such. Even further, in the psychological function, the tool shapes the user's understanding of the environment, for instance the controlled process and the physical phenomena taking place in the process. Thus, the tool affects the user's conceptualization of both the tool itself and the controlled process. The psychological function of a tool is of utmost importance in the activity because it enables *reflection* about the activity, which is a pre-requisite for learning and development. When the users' understanding concerning the tool and the environment develop, possibilities emerge for the whole activity to develop and thus to respond to external changes in the environment, which is a pre-condition for resilience. If the tool functions well in psychological role the operators are able to create and maintain personal interpretations and models concerning the controlled process and the tool. This interpretation enables appropriate situational awareness, which in turn is a pre-requisite for adequate activity. The following indicators of experiencing the psychological function of a tool were generated under the general heading: Feelings of suitability for self.

The first psychological experience indicator (P1) addresses the feelings that the operators have accumulated over time about learning to use the control room systems. A certain level of effortlessness should be connected to the learning experiences if the tool is felt to be supportive of the objectives of the whole activity [Kaptelinin, Nardi 2012].

The second indicator (P2) addresses the experiences which both the external appearance and the functional aspect of the tool induce in the operator users. In a well-functioning psychological control room system the users should have a feeling that the tool suits their expert identity [Nuutinen 2006] and professional self-esteem. The tool must appear to the users as the tool of a qualified professional operator.

The third indicator (P3) addresses the feeling of control emerging in the user in operating activity. If the psychological function is in good shape, the operators have accumulated experiences which allow them to gain a developing understanding of the complex interrelations within the controlled system. 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 [Hollnagel et al. 2011]. A feeling of control has also been identified as a relevant UX factor in a study in process control domain [Paunonen, Oksanen 2011].

The fourth indicator (P4) addresses the negative experiences possibly emerging if the psychological function is not fulfilled by the tool. Probably the most identifiable negative feeling is the frustration experienced in usage. Thus, lack of frustration is a positive user experience indicator. Similar indicators were identified in the study by Obrist et al. [Obrist et al. 2011] which they labelled stress and emotions.

The fifth indicator (P5) addresses the feelings of self and agency which emerge in using the control room systems. In a psychologically well-functioning tool the users have an appropriate self-confidence as regards their own skills. This indicator is also related to professional self-esteem, the development of which the tool should promote and enable, and is aligned with basic AT notions of agency [Kaptelinin, Nardi 2012].

The indicators concerning the psychological functioning of the tool have a resemblance to the UX indicators utilized in developing and evaluating consumer products and systems, but the work context is integrated by utilizing work-relevant concepts and related terminology.

Table 4. The statements used in the questionnaires. P1–P5 indicators of psychological UX. Letters A, B, C refer to the separate studies and the exact format of the statement that was utilized in the particular study. The number after the letter, e.g. B2, refers to the modifications made to the statement in the course of the three studies. If there is no number, the statement was utilized in exactly the same way in all studies. If there are numbers, it means that the statement was modified for the purposes of the later studies. Some of the indicators were utilized for more than one statement, e.g. P1 had two statements in study B (P1_B1, P1_B2). In the modifications the statement was contextualized for the study, but the main content of the indicator was kept constant.

Psychological function UX: Feelings of suitability for self	
P1	(A1, B1) Learning to use the control room systems feels effortless
	(B2) I have developed personalized routines concerning the operation of the system
	(C3) The system promotes formation of appropriate routines
P2	(A, B, C) The control room systems have a convincing look and feel and seem to be the tools of a professional user
P3	(A, B, C) In my opinion I can control the process through the control room user interfaces
P4	(A, B, C) Utilizing the control room systems does not frustrate me
P5	(A, B, C) I master the usage of the control room systems

5.2.4.3 Operationalization of UX for the communicative function

The communicative function refers to the effect that the tool has within *the community* of the users. Use of particular tools can be interpreted as communication. For example, by selecting a certain tool for a task, a person communicates something. Choosing a specific cooling method in an accident situation reveals how severe a situation is considered to be. The communicative function entails that by using the shared tools the community adopts and shares vocabulary, knowledge structures, and even values concerning the work. Thus, the tools shape and construct the user community. Most importantly, the communicative function emphasizes how meanings are conveyed by the tool in the community. If the tool functions well in the communicative function, it communicates what is valued in the community. And these values are approved by the user community. For example, the tool must convey to the operators what is considered to be the important phenomena in the process by the wider plant organization. As the design of the control room fulfils the requirements of the safety department of the NPP, the control room constitutes a view to what is considered safe operating of the plant. The control room design should also convey an appropriate level of trust in the technical systems. The following indicators of experiencing the communicative function of a tool were generated under the general heading: Feelings concerning the joint capabilities of the system.

The first communicative experience indicator (C1) addresses the feelings the operators have concerning a specific communicative feature of the control room system. That is the feature of communicating about the available resources. If the communicative function is supported correctly, the operators should have a feeling that the availability of resources is communicated on a level that exceeds the immediate situational needs. This is the prerequisite for achieving resilience, and is often referred to as building slack resources into the system [Hollnagel et al. 2011]. If the operators can feel that there are slack resources, it means that level of communication is such that help can be obtained from the system more widely than the immediate situational needs indicate.

The second indicator (C2) addresses the feelings towards collaboration with automation. A commonly used metaphor concerning operating a highly automated plant is that of collaboration of humans and automation. The indicator reflects positive feelings concerning collaboration with automation, and it deals with control room systems' support in finding operating paths in a situation in which it is unclear what to do. This indicates a feeling that a human and automation can jointly function with a superior performance in comparison to either one functioning alone. This aspect was also found relevant in the study by Palviainen and Väänänen-Vainio-Mattila [2009].

The third indicator (C3) relates to the feelings towards the EOPs' communicative features. If the EOPs function well in the communicative role, the operators should have a feeling that the EOPs help in understanding what is happening in the process. The EOPs should be experienced to communicate, not only what should be done next, but what is important about the process to monitor and control. This feeling is relevant from the resilience point of view.

The fourth indicator (C4) is also concerned with the feelings concerning specifically EOPs. It is the feeling of understanding the logics of the EOPs. The EOP should communicate to the operators why particular actions are required, and the operators should have a feeling that they understand and can accept the reasoning behind the logic of the EOPs. The fifth indicator (C5) addresses the feeling of trust concerning the control room systems. It is quite important in a safety critical domain that the operators do not either over- or under-trust [Lee, See 2004] the tools. The trust is related to the communicative function of the tool because trustworthiness and reliability are qualities which the tool should communicate to the users about itself.

Table 5. The statements used in the questionnaires. C1–C5 indicators of communicative UX. A, B, C refer to the separate studies and the exact format of the statement as utilized in the particular study. The number after the letter, e.g. B2, refers to the modifications made to the statement in the course of the three studies. If there is no number, the statement was utilized in exactly the same way in all studies

Communicative function UX: Feelings concerning the joint capabilities of the system	
C1	(A1) The control room systems present the operating resources available to operators
	(B2, C2) In my experience, the displays communicate the essential in formation
C2	(A) The control room systems help in finding the right solution in a situation in which it is unclear what to do next
C3	(A1) The emergency operating procedure helps me understand the process
	(B2, C2) In my experience, the system supports communication within the crew
C4	(A1) I understand why the emergency operating procedure directs me to a particular path of operations
	(B2) I am aware of the internal design restrictions of the system
C5	(A, B, C) According to my experience, I can trust the control room systems

5.3 Identified systems usability issues in control rooms

This subsection addresses the systems usability problems identified in the empirical work concerning the systems usability of the current hybrid control rooms of NPPs. The detailed descriptions of the studies and of the findings are presented in Papers I, II, V, and VI of this dissertation.

5.3.1 SU Findings in the hybrid control rooms

As reported earlier, baseline studies were conducted in both Finnish NPPs concerning their main control rooms[25] (MCR) prior to major modifications to the critical safety systems. The plants are referred to as plant A and plant B in this research.

[25] The actual data collection was conducted on a training simulator.

Although both of the MCRs were mainly based on analogue UI technology, they can be characterised as hybrid: Plant A had a digital control rod manoeuvring and monitoring system, and plant B had digitalised the turbine control systems. Although both MCRs studied constitute a unique case, some conclusions can be drawn concerning hybrid control rooms on a general level. In the studies the operators apparently conducted the required activities well, and no real safety-threatening findings were made. This result was as expected because this is the control room where the normal daily work is carried out, and the operating records are quite good. Nevertheless, some problematic issues were also identified in the studies, which may be traced back to the hybridity of the control room.

The study of systems usability proved beneficial, as the problems caused by hybridity did not fully become evident in the performance outcomes of operating activity, but were rather visible on the level of operating practices in which diversions were identified. Especially processes of learning were considered difficult in the hybrid control rooms. Also, consistency concerning different features of the control room easily becomes compromised within a hybrid control room. As a solution to these problems a more continuous monitoring (evaluation) of the prevailing operating practices in connection to the development of training programs was proposed. In addition the practice point of view should also be exploited in the design of new control room features.

5.3.1.1 Instrumental function in hybrid control rooms

In the instrumental function, both of the evaluated control rooms worked quite well. This means that the operating crews were able to effectively carry out the control operations considered necessary in the simulated accident and incident situations. This issue is featured in relatively good operator performance results obtained in both plants from an expert judgment method.

Specific problems in instrumental function were as follows: according to the expert evaluator, the operating crews in plant B had problems in the detection phase of the scenario. Also, the use of procedures received relatively low performance scores. Both of these tasks are heavily affected by the tool that is used. Also, in plant B it was seen that turbine operators, utilizing the modernized interface, were more satisfied with the tool. This is a hybridity-related concern because modernizing the control room only partly, may cause such polarization within operating crews.

Generally the results from systems usability questionnaire were also on quite a good level. When the statements concerning instrumental function were analysed separately for both plants, it was found that the percentage distribution of answers was such that in plant A 35% of the statements were responded to with completely agreement and in plant B 30%. Some problematic issues concerning instrumental function were identified, however. The lowest scores within instrumental function were received for feedback on the digital touch screen solution (in plant A). The operators stated that it is not always clear whether a command has been implemented. Another problematic issue was the possibility of errors especially in using the digital interface. The recovery from errors was in addition experienced as problematical. Also, the operators claimed that the systems in the control room do not help in focusing on the relevant phenomena when support for prioritization is needed. These issues were common in both plants, and, as both constitute a hybrid solution, it may be one of the causes of the problems.

5.3.1.2 Psychological function in hybrid control rooms

The psychological function was also supported relatively well, but it was the most problematical of the three general tool functions. In the questionnaire data, only 25% in plant A and 28% in plant B completely agreed with the positive statements concerning psychological function. The learnability of the control room systems was considered by the operators in both plants not to be on an optimal level. Even though the control room is a complex system, and thus learning its use cannot be expected to be effortless, the operators made statements that, in fact, it is confusing to have different information presentation principles and notations within the systems of the same control room. The learnability is affected by hybridity, as it induces more learning challenges in the form of multiple different systems utilizing different interfaces and operating logics.

It was evident in both plants that the psychological function of the control room was not fulfilled as successfully as the instrumental function. This is an indication that hybridity complicates the operating work. This means that more careful attention should be paid both to consistency throughout the different systems (in overall operation logic) and to the training processes. In a well-functioning psychological tool, there would not be inconsistencies in the operating practices between the crews, because the functionalities would be such that they could be exploited by all the crews to the same extent (at least roughly). In the current solutions there were tendencies in the operational practices to differ between the crews. This is a weakness on the overall level, because in the long run it may also develop into variation in the performance outcome.

In plant B, where there were more digital user interfaces in use, the understanding of the automation information provided in the interface was not on an optimal level. Only one crew was able to solve an automation-related problem in scenario 3, and was thus able to avoid the production loss related to an automatic scram[26]. In the analysis it was found out that the problem was related to the psychological function, because the other crews' abilities to utilize the system were not on the same level with the system capabilities. This is related to hybridity as it is evidence of polarization of the skill levels between the different operating crews.

In plant A there were differences on the micro-level in the crews' operating practices in proceduralised scenarios. We considered this also to be a problem in

[26] A reactor scram also always increases the risk level, thus it can also be regarded as impacting on safety.

the psychological function of the tool. Not all crews had been able to create interpretative ways of acting which would be manifested in the crews' mitigating actions concerning the core safety functions of the plant. This micro-level analysis of operating practices was not conducted in plant B.

A positive issue concerning psychological function was the new state-based EOPs in plant A. They truly seemed to lessen the task load of the users, and as such, support the psychological function of the totality of the control room.

5.3.1.3 Communicative function in the hybrid control rooms

Overall, the communicative function lies somewhere between the psychological and the instrumental in terms of successfulness. In plant A, 29% of the questionnaire answers were completely agreed upon and in plant B, 44%. Most problematic communicative features, according to the questionnaire data, concerned the interpretation of the audio information in the control room, operator support in unclear situations, and the EOP's ability to help in finding new solutions. The fact that operator orientations lay mostly in the confirmatory category was also considered to be a problem of the communicative features of the control room. The tool should more firmly support an operating attitude which identifies the connection of the sometimes mundane process control work to the higher level objectives of ensuring nuclear safety. The differences in collaboration practices (movements, communication, time spent together) show weakness in the communicative function of the tool – not all crews are able to collaborate effectively and efficiently in the control room.

The communicative function is reflected in the ways operating crews collaborate in the control room, and also in the ways the control room systems manage to convey meaningful information for operations. Based on the findings in the reported experiments, the communicative functions of both control rooms studied could be improved. The needs for improvement were evident in the results of the orientation interview results of both plants. The dominant orientation was confirmative, which reflects an attitude towards work according to which process control work is confirmatory in nature. This means that the benefit of hybridity, the variety of information presentation formats and operating interfaces has not been exploited by the operators. But rather, it works against the development of the interpretative orientations.

5.3.2 Variability in proceduralized activity in simulated LOCA

In the detailed analyses of operating practices in plant A, differences between the operating crews were identified. The method of analysis was explained previously, in Subsection 5.2.3 (p. 102).

Altogether, in the analysis of the crews' habits of action in the LOCA, six different process control tasks were identified, for which it was possible to grade the on a scale from interpretative through confirmative to reactive. The six tasks were:

Information usage (in two different episodes), Process situation interpretation, Dealing with automation, Decision making, Communication, and Leadership. These functions resemble, for example, the macro-cognitive functions described in naturalistic decision-making [Schraagen et al. 2008] and non-technical skills described by in previous literature [see e.g. Fletcher et al. 2004]. But the fact that, in the study, the differences in operating practices within these specific functions emerged in the micro-level examination of proceduralized activity is convincing evidence for the argument that, even in following procedures, human actors may act in different ways. A qualitative characterisation of the identified habits of actions is presented below (Table 6).

Characteristic of the interpretative habits is that the crew activity is profoundly connected to the process situation. On a general level, this characteristic enables resilience, because it becomes possible to observe the possible deviation from the expected process phenomena, which is a fundamental pre-requisite for dealing with unexpected situations. This means that, in order to create resilience in the system, process information produced by the automation system should be widely used; both redundancy and diversity in the types of information to be used are important. Cross-checking is utilised for gathering re-assurance [see also Patterson et al. 2007]. The interpretations made concerning the process status aim for deep functional understanding of the situation and the possible consequences for the other parts of the process. Dealing with automation reflects an appropriate division of labour. Human agency is reflected in the decision-making. The group work can be characterised as active and dialogical. All of these features of habits are such that they may help the operating crew to survive unexpected situations which have not been considered during the design of the system.

The confirmative habits are such that they also result in good outcomes of activity in most cases, but, in comparison to the interpretative habits, the confirmative habits are less tuned to understanding the process phenomena on a profound functional level and are more concerned with the events that are taking place. Thus, controlling the events with the pre-defined measures (procedures and practices) becomes the focus and objective of activity. This type of habit may easily fall into utilising a pre-categorised typology of events, which is not an effective strategy to manage the unexpected events and thus does not add to the system level resilience.

The reactive habits may also produce a relatively good end result in the situations in which the process behaves in a way that has been fully anticipated in the design of the system. The pitfall of this habit is that there is little potential for creating new ways of handling the work for some reason if it happens to be required at one point in time.

Table 6. Characterizations of interpretative, confirmative and reactive operating practices identified in the LOCA scenario.

	Information usage	Interpretation of situation	Dealing with automation	Decision-making	Communication	Leadership
Interpretative	Variety of sources, redundancy and diversity in information sources, dialogue in interpretation of information	Interpretation by considering functional meaning of process events	Human assures the automatic functions. Shared responsibility of human and automation.	SS decides to scram the process. Human as an active, present agent in decision making.	Dialogue concerning process status in the situation. Diverse and redundant information communicated. Reflects creation of joint awareness.	Active engagement of each operator in all the decision points. Transparency in contemplation enables one to spot false conceptions.
Confirmative	Multiple source but taken for granted	Identify the process events based on an existing typology of possible events, e.g. a leak	Automation functioning is observed but not acted on. Reliance on the pre-defined roles of human and automation	Scram is conducted paced by the procedure. Actions are controlled by the procedure	Statements made aloud concerning process parameters. Reflects confirmation of own interpretations.	The end result of the decision making process is stated and confirmed by all the operators
Reactive	Variation in information sources not sufficient	Identify that something is going on; no strive to understand or label the situation	Automation information is taken for granted, reflects total reliance on automation	Not identified in the data	Process state is not explicitly mentioned. Transfer of support system information.	No real collaboration. SS announces the next steps.

The summary of the different crews' habits in the different episodes with regard to the emerged tasks is presented below (Table 7). Most of the crews had constructed operating habits that vary on the scale from reactive to interpretative during the scenario. Almost all the crews had habit characteristics from all three classes. Nevertheless, some inclinations towards different habit profiles can be identified. In what follows an assumption has been made that, if the crew has 6 or more indicators of a certain habit, there is a strong inclination towards the respective habit. Secondly, if the majority of indicators is in a certain category, this is the dominant habit characteristic of the crew. Thirdly, if different habits are represented equally, then often the habit profile is diffuse. According to this analysis, crew J is quite strongly reactive. Crews A and D are dominantly reactive. Crew B is strongly confirmative, as is also crew G. Dominantly confirmative are crews F and H. Crews C, K, and L are quite diffuse between confirmative and interpretative, and crews E and I are dominantly and strongly interpretative respectively.

The point of the analysis was not to make comparisons between the operating crews, or to be able to say which crew has constructed the superior habits, but rather to have an idea of the variation and dominant characteristics of the habits of action of the operating crews within the one plant. At the top level 7 separate functions[27] were analysed altogether for each of the 12 crews, covering 84 separate habit characteristics in total. Out of those 27 were interpretative, which yields 32.1%, 34 confirmative, which yields 40.5%, and 23 were reactive which yields 27.4 % (Figure 13). This result shows that confirmative way of acting is the prevailing habit of the operating crews.

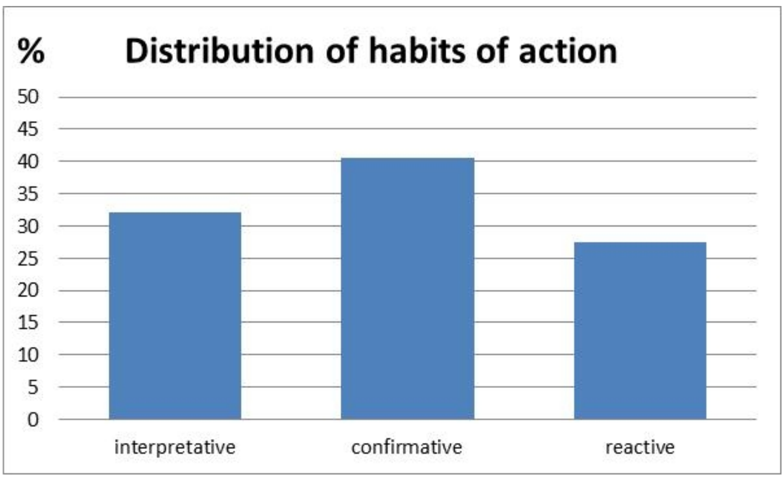

Figure 13. Distribution of habits of action in a simulated LOCA scenario. Confirmative habit of action is the prevailing one.

[27] Out of which 6 were different: information usage was analysed for two episodes.

Table 7. Summary of the different crews' habits in the different episodes with regard to the emerged functions.

Crew	EPISODE 1				EPISODE 2		EPISODE 3	Habit profile
	Info use	Sit. intrp.	Deal. w. autom.	Decis. making	Info use	Comm.	Leadership	
A								dominantly reactive
B								dominantly confirmative
C								diffuse btw. confirmative and interpretative
D								diffuse btw. confirmative and reactive
E								dominantly interpretative
F								dominantly confirmative
G								dominantly confirmative
H								dominantly confirmative
I								strong interpretative
J								strong reactive
K								diffuse btw. confirmative and interpretative
L								diffuse between all classes

5.3.3 Findings concerning UX in three studies

The proposed operationalization concerning UX for the three functions of a tool in an activity system were utilised in three different studies. In each of the studies, the investigated system varied in scale and technological maturity. Study A concerned the current control room of a particular plant; study B concerned an individual system which had been in use for couple of years during the study, and study C concerned a totally new system still under design. In study A, the systems studied were a collection of digital and analogue UIs, whereas studies B and C concerned a touch-screen-based digital UI.

Study A, concerning the totality of the control room, elicited the most positive UXs of all the three studies. This was especially evident when only the instrumental function UXs were considered (Figure 14). This may well be a reflection of the familiarity, of the studied technology for the users. The maturity of the technology in study A room is high. It is based on quite concrete analogue interaction devices which enable very tangible interaction with the controlled process. The current control room is the tool through which the operators have learnt their profession, and presumably many of their conceptions concerning the controlled process are based on the process representations utilized in this current control room. It is also evident that all operators perform quite well with the tool they are currently using, and thus the instrumental experiences also become high. The control room in general feels like an appropriate tool to control the process. In study B, the technology had been in use for couple of years, and study C concerned a completely novel system. The main finding in this line of reasoning is that, when the familiarity and maturity of the technology are lower, the instrumental UXs do not seem to reach the same level as with the current tool, the familiarity and maturity of which are higher.

Another interpretation of this same result is that it is an indication of the touch-screen based technology tested in studies B and C on a more general level. The operators may feel that it does not provide as tangible a feedback of operations as the analogue controls in the current control room. And thus it does not feel as if it is an appropriate tool for the work as the analogue user interfaces.

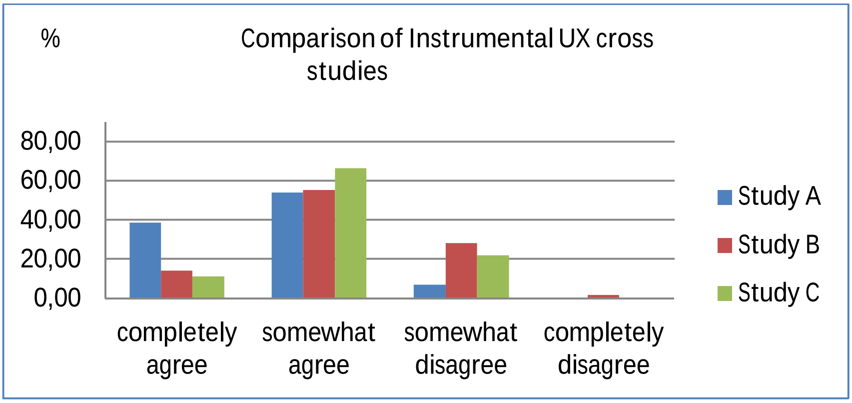

Figure 14. Comparison of instrumental UXs in three studies.

The fact that the psychological and communicative functions received lower scores in study C than in studies A and B (Figure 14) is well in line with the theory behind the functions of a tool. As the interface studied in study C was totally novel, the psychological and communicative functions have not yet developed to the same level as in the other two studies.

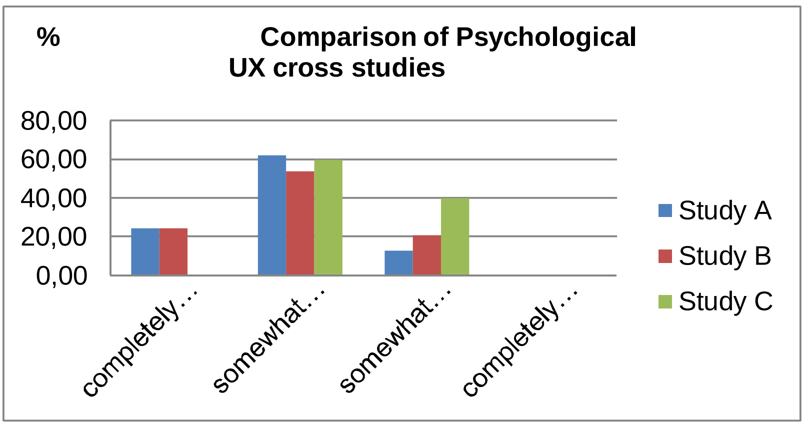

Figure 15. Comparison of the psychological UXs in three studies.

The individual indicator P3 (see p. 109 for description) concerning the feeling of control is quite interesting. It is the highest scoring psychological experience indicator in studies A and B (43% and 41% answers in the “completely agree” category respectively). But in study C its score was the lowest of all the psychological UX indicators, having 60% of answers in the “somewhat disagree” category. This is something that cannot be excluded in the contemplations concerning the operational concept of the systems studied in study C.

In the comparisons of communicative function UXs over the three studies there seems to be a similar difference in the answers to in the ones concerning instrumental function: In study A, there are more replies in the category “completely agree” than in studies B and C (Figure 16). Consequently, in the category “somewhat disagree”, the situation is reversed. This may be a reflection of how the communicative function of the tool is the slowest to develop. Perhaps it is more time-consuming to adopt the societal meaning of the new tool and develop trust in their capabilities than it is to “merely utilize” them in some familiar task.

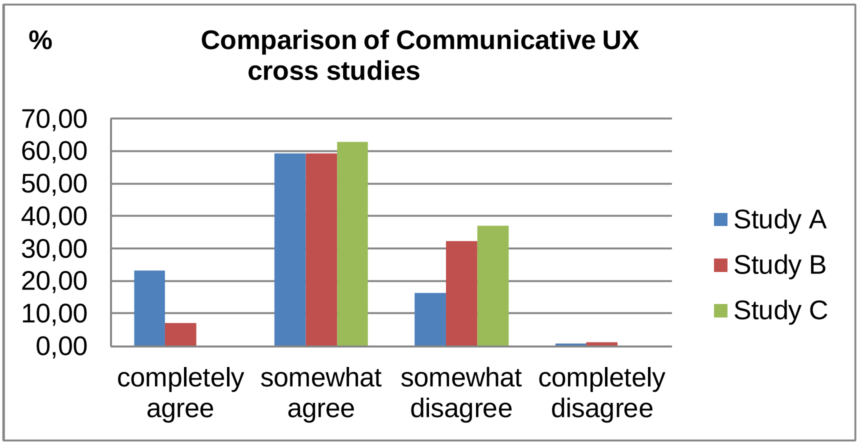

Figure 16. Comparison of communicative UXs in three studies.

The object of study varied between the three studies. Study A’s scope was the totality of the control room, whereas studies B and C focused on particular systems under development. This is probably a factor affecting the present results concerning UXs. The abstraction level at which the answers consider the object of study in study A is higher than in studies B and C. Considering the whole control room consisting of several different systems, it is probably easier to give a high score than when a particular system with particular features is in question.

Overall, the category “somewhat agree” is the most common answer in all three studies. Overall, 50–60% of the answers are in this category. On one hand, this can be interpreted to reflect the weakness of the questionnaire methodology. It is probably the easiest one of the choices to pick, if the system is such that it doesn’t feel perfect, but there are no specific features reflecting negative feelings either. And, as justifications for the answers were not required, it is not possible to trace back why operators choose this option so often. On the other hand, the expression may also just be felicitous and for this reason gather a large number of responses.

5.4 Summary of the results: Answers to the research questions

Previously in this chapter, the main findings of the studies comprising this dissertation work have been introduced. This last subsection of the chapter summarizes the results and provides concise answers to the research questions.

5.4.1 Theme 1: Characterising quality of tools in complex work

The research questions within the conceptual theme concerned the characterisation of usability in the context of complex process control work, and the role and significance of the concept of UX in the same context.

5.4.1.1 Systems usability outlines quality in use for complex UI

Based on theoretical contemplations concerning the functions of a tool in work, a model of systems usability was presented. Systems usability assumes that a tool in an activity system serves the functions of: 1) an instrument, 2) a psychological tool, and 3) a communicative tool. The meaning of each function in the specific domain (e.g. an NPP) is contextually defined and reflected in the core-task demands of the particular domain.

Furthermore, the quality of the tool can be assessed utilising different perspectives on the usage activity: performance, way of acting, and user experience. The three perspectives on activity constitute a set of viewpoints which enables an investigation of operating activity from multiple angles in order to fulfil the objectives of systemic analyses.

By combining the functions of the tool and perspectives to activity, a systemic framework for developing indicators for a good control room is construed. This framework depicts nine classes of SU indicators which revealing different aspects of quality of a UI utilised in complex work.

5.4.1.2 UXs indicate aspects of a tool's development potential

The significance of UX in complex work was addressed both analytically and empirically. Based on theoretical notions of concept experience in AT, the users experiences were hypothesised to concern the development potential of new tools, an issue which is otherwise hard to measure in empirical tests.

The idea was explored in three studies in which nuclear power plant operators' user experiences (UX) concerning tools were studied in different technology maturity phases. The method utilised in order to study UXs was a questionnaire, which was developed based on a systems usability framework. UX indicators were operationalized for all three tool functions (instrument, psychological, communicative), and the results obtained with the questionnaire method in three studies discussed. The importance of UX as a measure of germination, for an as yet non-existent tool, was explored.

The significance of the whole concept of UX in complex work is connected to practices of design in the safety-critical domain. Currently design activities in control room development rely heavily on existing solutions and other industry dogmas such as standards and guidelines to an extent which may hamper developments and natural progress in the outcomes of design. The conservative design approach has known advantages but also disadvantages. Basing evaluations on new designs only on strict performance measures may result in too early a rejection of innovative designs. Therefore, new indicators of potentiality of proposed solutions are needed. Formulation of the UX indicators for the three tool functions allowed a different perspective on the evaluation from that of the traditional task performance-based view. The results show that, in designing the work systems of the future, experience-related factors should also be treated as design drivers, as they may be considered signals concerning the potentiality of the tool, evaluated by the professional users.

For the design activity to be resilient, it should profoundly “address the potential”, i.e. consider not only actualised performances but also other possible outcomes. UXs of future users of the system could be used in resilient design activity for this purpose.

5.4.2 Theme 2: Methods of usability evaluation of control rooms

The research questions in the methodological theme concerned the development of practical methods for studying the quality in use of UIs within complex work. Several empirical studies were completed over the course of the research work for this dissertation. In the empirical studies, different evaluation methods were tested and developed. In what follows, the main characteristics of the evaluation methods are recapitulated.

5.4.2.1 Systematic and broad-scoped evaluation process

On the level of evaluation process, the whole approach to evaluating the systems usability of control rooms in safety-critical domains should be systematic, broad in scope, and based on relevant internal references.

Systematics may be achieved by utilising a process in which the line of reasoning concerning quality in use is explained to the extent which allows a critical review of the reasoning. Pre-requisite for this level of explication is the outlining of the evaluation bases, i.e. the reference which is used in evaluation as thoroughly as possible. In the evaluation process that I have proposed in this dissertation, the whole reasoning concerning the quality of the proposed tool is based on modelling both the general characteristics of the domain and its situational manifestations. In addition, the demands on operating work, i.e. the core-task demands, are used as a reference, as well as the informativeness of the proposed UI solution (the latter models are not included in the scope of the dissertation).

The models set requirements for data collection and analyses followed by the assessment of the systems usability of the proposed new tool.

The broad scope in evaluation is achieved by not drawing the lines of the evaluation too tightly around the system under evaluation. In the case of a complex work system, it is not possible always to understand all the relations between the system components in advance. Therefore, the scoping in evaluation should be sufficiently broad to allow emergence of the effects of new systems not envisioned in the design.

The individual methodical suggestions are presented in what follows.

5.4.2.2 Functional modelling of tasks

In the literature review part of this dissertation it was identified that, in evaluating the appropriateness of new control room solutions, the tasks of operators are most often treated as equivalent with operating procedures, despite the fact that procedures do not, as such, describe all aspects of work, and that what is more, operators also utilise systems for purposes other than the immediate process control needs. Therefore, a need was identified to develop a task modelling method which would allow evaluation of operators' ways of acting.

For this purpose, I have in this dissertation presented the functional situation modelling method (FSM). An FSM combines a functional and a chronological view of the activity of an operating crew in a particular situation. By making explicit the connection between required operating actions and critical functions, the model lays the ground for analysing operating activity from the point of view of maintaining the critical functions.

5.4.2.3 Semiotic analysis of operating practices

As an addition to the prevailing data analyses methods concerning task completeness and other outcome-related aspects of operating activity, I have presented a method for analysing operating practices.

The analysis of operating practices utilises the semiotic concept of habit as an operationalization of way of acting or operating practice. The semiotic model of habit combines the behaviour of the actor (interpretant) with the environmental sign which is the cue for the particular actions. By making this connection, it is possible to deduce the object of the action, which signifies the more global meaning of the particular action. This means that it is possible, in addition to performance, to study the implications that the studied UI solution have at a more global level on the operating activity. It is possible to examine whether the operating practices with the new tool have the potential to develop into a direction which allows connecting the individual actions to the more global objectives of the activity. This connection is an important signifier of resilient operating practice, because it allows perception and thus action also in a situation which does not comply with the pre-defined procedures of acting.

In the semiotic analysis of operating practices, resilience is treated as the general capability of the socio-technical system to survive and recover from both expected and unexpected situations. The operating practice which supports system-level resilience is characterised as interpretative. Interpretative practice aims for a profound connection to the controlled process via a continuous process of perception, interpretation and acting.

5.4.2.4 Evaluation of UXs of complex systems

An operationalization of the concept of UX was developed for the three tool functions. Based on the UX indicators, a questionnaire was formulated which was used in studying UXs of different control room systems.

The developed questionnaire was in each study contextualised according to the specific system under evaluation and the specific research questions in use. Yet on a general level, the indicators were kept the same. Collecting quantitative data concerning UXs enables following the development of UXs during the development of systems and thus if necessary aiming design efforts according to the results.

5.4.3 Theme 3: Implications for the design of control rooms

The practical empirical research questions concerned the identified benefits and drawbacks in the systems studied within the empirical research. Several empirical findings were indeed made in the course of research for the dissertation. In what follows I summarize the most important findings which have relevance from the point of view of modernizing control rooms in safety-critical industries.

5.4.3.1 Outstanding management of simulated accidents

NPP operating crews' level of skills in accident management is very high. During the course of my research I have followed the management of dozens of different simulated accident situations by almost all the operating crews of the operating plants. The operating crews' ability to manage the process is at a very high level. This alone signifies the outstanding skills which the professional operating crews possess for accident management in the situations included in training programmes.

This result does not mean that the control room systems may not contain problematic features. The situation actually tends more to the opposite: despite the obvious discrepancies in the systems utilised, the human operators are able to develop ways of working which are not disturbed by the designs which are not optimized for operating work.

Nor does this result mean that control rooms should not be developed further. On the contrary, I believe that, in a situation in which the external conditions developed further into the non-optimal direction, the discrepancies in control room solutions will play a more important role.

In the development of control rooms, the boundaries of the system under development should not be drawn too tightly. It is sometimes very hard to differentiate the effects of training, procedures, control room design and situational factors. Therefore, the development efforts should also have a broad scope.

5.4.3.2 Hybridity in control rooms complicates operating work

All the empirical studies conducted in the course of dissertation work took place in control rooms which can be described hybrid solutions. Even though fully digital control rooms are not yet in place in the nuclear industry, every control room has digital systems in use. Because of the safety regulations, the situation will also remain the same. Some controls will in fully modernised control rooms still be based on analogue user interfaces.

A particular challenge in the design of hybrid solutions is to maintain consistency throughout the control room solutions. Consistency is a common usability principle and it is not unfamiliar to the control room design standards either. But for some reason, in the operational environment of an operating NPP, it is difficult to maintain. Control rooms seem to be developed by adding on individual systems which serve important but often separate purposes, and from an operator's point of view the end result is a control room consisting of various different systems, which all have their separate operating logics and ways of interaction. This development approach may result in a situation where the operator has to learn and maintain skills in using several different systems on a daily basis.

The effects of hybridity on operating work were most evident in the so-called baseline reference tests conducted in both NPPs in Finland. The results showed that, on an instrumental level, both control rooms worked quite well. But in the psychological and communicative functions the results were not entirely unproblematical. Concerning psychological function, problems related to learnability were identified. It was also evident that not all operating crews had acquired skills to fully exploit new information contents provided in the digital systems. The question of learnability and particularly skill levels relates to both control room design and training. Many issues can certainly be improved through training, but utilising innovative presentations may make learning to use new systems less demanding.

5.4.3.3 Variation within proceduralized scenarios

Resilience in proceduralized accident management was studied utilising micro-level analyses of behavioural data. In the analyses, several functions were identified in which operating crews' practices differed from each other. This is an important practical finding. It means that, even though every crew acted according to the procedure, there were still differences. The crews fulfilled the demands of the procedure in different ways.

The differences in the operating crews' ways of acting were classified according to the resilience characteristics identifiable in the practices. Resilient information usage is such that a variety of information sources is utilised in such a manner that both redundant and diverse information sources are made use of. In the interpretation of information, the crew engages in active dialogue. In interpretation of the process situation, the crew considers the functional meaning of the individual process events. In dealing with high levels of automation, the human operator's role is to assure automatic functions. In decision-making, the human operators acknowledge the role of an active agent signified by its presence in the situation. In communication, the crew aims to create a shared understanding of the situational demands in the particular event. The leadership role of the supervisor is such that it enables active participation of all the relevant parties. Transparency in contemplations enables identifying possible false conceptions.

These resilient practices in an operative situation had been developed by the individual crews on top of fulfilling the immediate demands for acting expressed in the procedures. This means that the control room system, including the operating procedures, may be utilised in different ways by the operating crews. Therefore, the technical design of those systems is not the only determinant of crews' capabilities in operative situations. This is an issue which must be taken into account in design, evaluation, and training of the systems. In design, it should not be expected that all crews will utilise the systems in a similar manner. In evaluation, it should not be expected that the evaluated system will determine courses of action only. And in training, attention should be paid not only to introducing the features of the new systems but also to ways of using them to full capability.

6. Discussion

This chapter presents the discussion of results of the research work.

6.1 Practical implications

The results presented in the light of the three research themes have several practical implications. The literature reviewed in the related research section identified that many of the real-world complexities of operating work are not typically included in the evaluation studies of control room technologies. The prevailing evaluation approaches may produce results from which it may be difficult to deduce the design implications.

The evaluation approach presented here complements the prevailing approaches by including qualitative analyses of operating activity, which reveal the actual practices of usage, which can be utilised to improve design solutions. This practical implication is very important in today's operational context in which NPPs all over the world are designing digital information and control systems to be used in MCRs. My experience from conducting the empirical studies is that the immense possibilities for information presentation and human technology interaction are not yet fully adopted in the industry. Therefore, an evaluation approach which is able to dissect the actual corporeal practices of usage is of value for designing and innovating new ways utilise using the digital information which concerns the process of producing electricity through nuclear fission.

Another, possibly even more topical, practical implication of the research is the need for conducting independent evaluations of intended control room solutions. It has been agreed in the industry that, especially in validation type evaluation, it is important that evaluation is detached from the actual design [O'Hara, Higgins & Fleger 2012] o as to avoid possible biases in the results. But this independence poses several practical challenges. It is not possible to evaluate systems without a profound knowledge of the controlled system and the design solution. Without this knowledge it is not possible to design meaningful scenarios to be used in evaluation. The fact that the evaluation process proposed in this dissertation is based on several reference models solves this problem, because in the beginning of the evaluation process an evaluation internal reference is created. The models constitute this reference, and it can be discussed with the domain experts and designers without jeopardizing the independency of the actual evaluation and especially the drawing of conclusions.

Yet another practical implication of the proposed evaluation approach is that it changes the role of the participants, the operators (or people acting as operators), in comparison to the prevailing evaluation approaches. For the sake of standardization and control, it is typical in the prevailing evaluation approaches that participants are kept ignorant concerning many aspects of the experimental design. The objective then is to fulfil “the gold standards” of experimental design. But the question remains whether the design of technological systems is an area of engineering in which best possible results, i.e. designs, are achieved in this way? During the course of the empirical studies conducted for the dissertation, different approaches to engaging the participants in the evaluation process have been explored, borrowing from neighbouring fields of science such as usability engineering and participatory design. The methods which probe the participants to explain their own thought processes during the simulator runs are a step towards a more active role on the part of the operators in evaluation of the new control room systems. In the end, the operators are the ones who have the most profound understanding of the demands of the operating work. My belief is that this capability should be exploited to the full extent in evaluating new control room systems. Therefore, one of the practical implications of this research is a possible shift in the role of the operators as participants in evaluation studies. New perspectives may open up if operators have an active, participative role in the evaluations.

The need for evaluation methods has become more and more pressing in the nuclear industry. The general approach presented in this dissertation has been utilised in the so-called sub-system validations (SSV) conducted within the final stages of the control room design of the Loviisa NPP [Laarni et al. 2011]. These evaluations have been conducted by the VTT human factors team, utilising the approaches and principles described in this dissertation. In the SSVs we have combined the micro-level analyses of operating practices to the higher abstraction levels of assessing the implications of the new concept of operations for the general aspects of operating work. The results of the evaluations have provided significant design improvements for the forthcoming, mainly digital, control room.

The adoption of digital technologies in control rooms enables the fulfilment of the safety culture principle of “continuous development”, because the development of digital technologies is not as effort-consuming as the development of concrete panels and switches. For example, obvious design errors within the interface can be corrected online. If modifications are conducted in this manner, it means that monitoring of the effects of new systems should also be continuous. Having first conducted the baseline studies in the truly hybrid control rooms, then the validation studies in the developing new digital control rooms, we have established a knowledge base which enables monitoring and following the development of operating practices, which is a pre-requisite for noticing any drifts that are taking place within the control rooms on the level of operating practices.

6.2 Methodical implications

The research presented in the dissertation also has methodical implications beyond the practical issues presented in the previous subsection. The proposed methods are designed to fulfil the needs of methodical developments identified in the review of related literature.

The first of the methodical development needs was related to treating safety as a systemic concept in the evaluations of control room technologies. This is an important requirement, because, in the prevailing evaluation approaches, safety is addressed through congruence between observed behaviour and the operating procedures, i.e. assuming that strict compliance with procedures equates with safety. My personal view is that compliance with procedures is an important safety indicator, but it is not the only issue that matters. Safety is a far more complex phenomenon as has been suggested by several authors [e.g. Leveson 2011a, Besnard, Hollnagel 2012, Falzon 2008, Hollnagel 2013]. Therefore, in the dissertation work I have developed other, complimentary indicators to address the safety of the operating activity. These are the indicators concerning ways of acting and users' experiences of the new technologies.

Another requirement for methodological improvement identified that was based on the literature review was an emphasis on contextuality of evaluations. Not every evaluation of control room technologies may address the same issues. The scope of evaluation is dependent on the scope of the new design and its interrelations and expected effects on other systems, both horizontally and vertically, within the operative environment. In the evaluation process developed in the dissertation work, the scoping of the evaluation is made explicit by developing contextual models in the first phase of evaluation. The models depict operative requirements which are tested in the evaluations and thus make explicit the scope of evaluation. The particular development of the functional situation model enables positioning the designed evaluation scenarios into the context of the safety of the plant.

Yet another identified challenge of the reviewed evaluation approaches was that the link to the design features of the evaluated system was difficult to establish. A typical way of making the connection was to treat the tested design as the independent variable. In some cases the approach has been proved to work, but often the results are conflicting or otherwise difficult to interpret. One reason for this is that, within the complex sociotechnical system of NPP operations, total control in the experiments is very difficult to establish [O'Hara 1999, Gore et al. 2006], and thus the causalities are very difficult to trace, as the system includes very many dynamic components. Therefore, in the evaluation approach developed in the dissertation, the high level outcome-related indicators are complemented with micro-level analyses of operating activity. As a result of these analyses, it is possible to answer the question of *how* operators utilise the information provided by the technological tools in the process control activity. Through answering these questions, the link to design becomes more direct through an explanation of the individual reasoning people have developed.

An important methodological input from the dissertation work is the demonstration of how the semiotic model of habit can be utilised in the analyses of operating practice. The semiotic model connects human behaviour to its environmental cues and thus enables interpretation of the object towards which the activity is oriented. This is an important contribution, as the analyses of human technology interaction are often tuned more towards one or the other, human or technology. Utilising the habit concept enables analyses truly on the systems level. This is also a contribution to the discussion concerning Joint Cognitive Systems (JCS) [Norros, Salo 2009, Hollnagel, Woods 2005, Woods, Hollnagel 2006] in which the analyses and design of systemic patterns of behaviour has been called for.

The approach also allows identifying flexibility in the routines of operating crews, and the ways in which the operating crews in their activity always have to balance between the need for the standardization expressed in the procedures and the openness to change when prompted by external cues [Grote et al. 2009]. This would not be possible by utilising the procedures as a task model and only evaluating the compliancy aspects of procedure usage.

The introduction of operating practice as a unit of analysis also has implications on the mechanism of generalizing the results of the research. In prevailing methods, the generalization is conducted with statistical means. The assumption is that, if operators can cope sufficiently well in very difficult situations, they can do so also in seemingly less demanding situations. The challenge of this approach is to add complexity in the scenarios in a way which is meaningful and ecologically valid. This is not an easy challenge to overcome, because it is impossible to predict what kind of complexity will eventually put the real system into test. Therefore the standard solution is for instance to increase the demands by introducing more and more process events, or events that are unfamiliar and difficult to control and then measuring whether the operating crews are able to conduct the required actions in a time which is acceptable. This way operating activity is tested in conditions that are as difficult as possible. And if statistically significant differences can be identified between the experimental conditions, it can be concluded that one system is superior to another. The problem is that it is possible that the real world complexities are different from the ones that are utilised in the experiments. According to the modern understanding of how accidents propagate in complex systems, there are several contributing factors affecting the emergence of accidents on different “levels” of the sociotechnical system. The aim of utilising operating practice as a unit of analysis is to raise the level of analysis from the level of individual actions onto the level of practice, and on this level identify inclinations which may in very difficult situations lead into unfavourable outcomes. Practices (operationalized with the semiotic model of habit) reveal general patterns of behaviour which people have developed in work, and which they have found to coincide with and help in achieving the objectives of the activity. The practices manifest the core-task demands unfolding in real activities. When the operating practice is oriented to the core task demands in a normal operating situation, it can be assumed to be such also in an accident situation, because the personally meaningful practice reveals what the actor considers worthy in the situation.

The new methods developed in this research enable a profound understanding of the mechanisms of usage which produce the measurable outcomes which are typically paid attention to in human factors analyses of human technology interaction. The methods provide a complimentary approach to the human factors evaluation approach utilised in most published studies concerning control room systems. The proposed approach is particularly valuable in evaluations in which the aim is to improve the design of the evaluated systems. Improvement may be in terms of design of technologies, procedures of usage, and training.

6.3 Theoretical conceptual implications

The main conceptual contribution of the dissertation is the outlining of the concept of systems usability (SU) and introduction of the framework for developing contextual SU indicators. SU follows the two central notions of AT: mediation and object-orientedness. The concept of mediation is developed by analysing the role of tools in mediating between the subject and object in three distinct yet intertwined functions: instrument, psychological tool, and communicative tool. In each function, it is exactly the object of activity which is mediated to the actor.

The contemplation concerning functions of a tool is, I believe, generalizable also to the design and research of technologies for everyday use. An activity-theoretical approach to technology design has been proposed by several authors for many years [Kaptelinin, Nardi 2012, Diaper, Lindgaard 2008, Kuutti 1991, Gay, Hembrooke 2004, Kaptelinin, Nardi 2006, Norman 2006, Nardi 1996]. Yet, it has not become a prevalent or mainstream approach in HCI, the reason for which has been claimed to be the complexity of the approach. The multiple functions that technology may serve do not seem too complex an idea. I believe, that it could be useful also in the design of any technology to sometimes shift the viewpoint in design from the instrument to the psychological and communicative functions of technologies. This may open up new perspectives which enable innovativeness in design.

The notion of the communicative function of a tools is not customary for activity theory; it was only fairly recently proposed by Rückriem [Rückriem 2003], but it seems like a concept that has immediate value in interpreting the uses of technologies in modern society. By utilising different gadgets, people communicate their values and social status. Consider, for example, the meanings which different brands of mobile phones hold within the community of the users.

Another conceptual contribution of the dissertation concerns the role of user experiences (UX) in the context of serious work. It is not customary to study technological systems from the point of view of UX in a professional domain, as the concept is believed to be related to the commercial success of modern consumer products and services. Nevertheless, experiences also have a role in work life. It is undeniable that what people experience in work affects their performance in work. Therefore, UX, as it has been interpreted in the HCI community, may also play a role in development of tools to be used in work. This issue was explored in this dissertation by developing an operationalization of the concept for the different tool functions and utilising it in separate studies concerning control room systems of differing maturity. It was concluded that UX as a concept could be utilised in design in order to inform the design activity about the directions to which the professional future users of system foresee the development of the system going.

6.4 Validity of the research

Validity of research is a profound question rooted deeply in the philosophy of science of the particular field. Validity of research work can be addressed from several perspectives. A common distinction is to differentiate between the validity of the research process and the research results.

As presented in the Introduction (p. 24), the research conducted for the dissertation follows the general approach of design science. New constructs, models, methods and instantiations have been construed, used, and evaluated in a real-world context. The research process presented in this dissertation follows the process proposed for design science by Peffers et al. [2007], which includes six steps: problem identification and motivation, definition of the objectives for a solution, design and development, demonstration, evaluation, and communication. The problem identification and motivation are presented in Chapters 1 and 2 of this dissertation, namely the introduction and the related research. The definition of the objectives for a solution is presented in Section 2.4 concerning the research gap and in Chapter 3 in the formulation of the final research questions. The design and development of the solution is described in Chapter 4, the overview of the empirical work as well as in Chapter 5, which presents the results. The evaluation and communication of the results are presented in Subsection 5.4, which is the summary of the results, and in this final discussive chapter. The evaluation and communication are also presented in the individual papers which present the research work on a more detailed level. Congruence of the research process with that of the relevant field of research speaks in favour of the validity of the research process.

As a grand objective, all research is about creating new knowledge. Fairly recently, the traditional mechanisms through which scientific knowledge is produced have become accompanied and partly also challenged by new forms of knowledge creation [Gibbons et al. 1994, Novotny, Scott & Gibbons 2001]. Two distinct forms of knowledge have been identified: Mode 1 and Mode 2. The distinction between Mode 1 and Mode 2 knowledge refers to a general transformation in the practice of science which involves a profound shift in the epistemological underpinnings of what is considered scientific knowledge in the first place [Hessels, van Lente 2008]. Mode 2 knowledge production can be distinguished from traditional positivistic Mode 1, as van Aken [2005] explains: “Mode 1 knowledge production is purely academic and mono-disciplinary, while Mode 2 is multidisciplinary and aims at solving complex and relevant field problems.” In other words, Mode 2 knowledge is produced in the context of application by socially distributed transdisciplinary collaborations, and can be characterised for instance as reflexive [Hessels, van Lente 2008]. Several characteristics of my research fit to the envelope of Mode 2 knowledge production. The aim of the research has been to develop solutions to a complex real-world problem: the evaluation of control rooms in a way which would demonstrate the safety of the systems and be beneficial for the design activity. The research has been conducted in the real-world context of control room modernizations, and the resulting evaluation approach has been utilised in improving the designs.

The mission of design science is to develop knowledge that the professionals of the discipline in question can use so as to design solutions for their field problems [Van Aken 2005]. Therefore, the general validity of the results is demonstrated in their applicability by the professionals of the field. The research results presented in this dissertation have to some extent been applied by the professionals in the field. The conceptual results concerning the construct of systems usability have been utilised by, for example Karvonen et al. [2012] in formulating research concerning the UXs of crane operators. The methodical developments concerning evaluation approach and individual approach have been utilised in SSV tests concerning a new control room. In addition the FSM was successfully applied by the simulator trainers of TVO NPP in modelling the scenarios utilised in baseline studies. The practical empirical results have also to some extent been utilised in control room design.

6.5 Limitations and future work

Several ways exist in which the research conducted for the dissertation could be improved and the results developed further.

The main limitation of the results of the dissertation is the laboriousness of the developed evaluation methodology. It is evident that the concept of systems usability and the derivation of the SU indicators for all nine classes of indicators is a ponderous process which requires resources both in terms of time and knowledge concerning the domain. It is evident that the developed evaluation approach cannot be utilised without profound domain knowledge. Although this is a clear limitation of the results I personally also think that when dealing with complex phenomena such as sociotechnical safety-critical systems, the profound domain knowledge is an absolute necessity to even start understanding what the real development needs within the industry are.

The concept of systems usability could be certainly developed further. I particularly think that the communicative role of tools is something that needs further investigation because team performance analyses and assessment are still somewhat underdeveloped area [Palmqvist, Bergström & Henriqson 2012]. The distinction of different levels of maturity in team performance and the role of tools in either inhibiting or contributing to appropriate team performance is one research thread that could be followed.

Simultaneously a limitation of the results but also a future research topic is developing the evaluation approach to a more participatory direction. I feel that the voices of operators as professionals and experts in the field should be even stronger in the evaluation practices. Perhaps ideas from e.g. change laboratory [Engeström et al. 1996] concerning interventions and could be developed in order to support designers' and users' mutual understanding of each other's problem domains.

Another future research topic is development of the evaluation method to fit the needs of continuous monitoring of operational safety of plants. It is clear that operating NPPs need to monitor the development of operating practices and tools in order to spot possible drifts as timely as possible.

In addition, improved methods for collecting and evaluating operating experience are needed [O'Hara, Higgins & Brown 2008]. For this purpose involving the users more in the process could be beneficial. E.g. the probe method [Paay et al. 2009] could be something that could provide inspiration on how to involve actual users more profoundly in the evaluation and design of new systems.

The concept of UX in serious work was only quite superficially touched by the empirical work conducted within the research for the dissertation. In my view a method basing on more qualitative analyses of users experiences concerning technologies in work should be developed in order to gain more profound understanding of the significance of the concept.

6.6 Conclusion

Finally, in this dissertation I have presented research concerning evaluation of control rooms in nuclear industry. The research has been carried out applying the general approach of design science.

My main thesis is that by adopting a systemic UI quality concept, i.e. systems usability, it is possible to carry out evaluations within control room design which both enable the assessment of safety implications through the study of operating practices and are beneficial for identifying design improvements in the systems evaluated.

I have demonstrated this thesis through theoretical and empirical work. I have construed the concept of systems usability which denotes to the general functions a tool has in an activity system: instrument, psychological tool, and a communicative tool. In evaluating the appropriateness of tools the perspectives from which the activity may be analysed are: performance, way of acting, and user experience. The notion of systems usability and the related evaluation approach has been utilised in connection with the practical challenges involved in the modernization of control rooms in nuclear industry.

References

Abu-Khader, M.M. 2009. Recent advances in nuclear power: A review. Progress in Nuclear Energy, Vol. 51, No. 2, pp. 225–235. ISSN 0149-1970. doi: 10.1016/j.pnucene.2008.05.001. http://www.sciencedirect.com/science/article/pii/S0149197008000851.

Aghina, M.A.C., Mól, A.C.A., Jorge, C.A.F., Freitas, V.G.G., Pereira, C.M.N.A., Lapa, C.M.F., Nomiya, D.V., Santo, A.C.E., Landau, L. & Cunha, G.G. 2012. Non-conventional interfaces for human-system interaction in nuclear plants' virtual simulations. Progress in Nuclear Energy, Vol. 59, pp. 33–43. ISSN 0149-1970. doi: 10.1016/j.pnucene.2012.03.009". http://www.sciencedirect.com/science/article/pii/S0149197012000443.

Al Harbi, M.A.S., Kim, A.R., Jang, I., Seong, P.H., Shirouzu, S., Katayama, S. & Kang, H.G. 2013. Effects of soft control in the nuclear power plants emergency operation condition. Annals of Nuclear Energy, Vol. 54, pp. 184. ISSN 0306-4549. doi: 10.1016/j.anucene.2012.11.014. http://www.sciencedirect.com/science/article/pii/S0306454912004598.

Alby, F. & Zucchermaglio, C. 2006. 'Afterwards we can understand what went wrong, but now let's fix it': How Situated Work Practices Shape Group Decision Making. Organization Studies, Vol. 27, No. 7, pp. 943–966. doi: 10.1177/0170840606065703. http://oss.sagepub.com/cgi/content/abstract/27/7/943.

Ans, M. & Tricot, A. 2009. Information seeking in documents by pilots: Assessment of the reliability problems caused by the transition from paper to electronic. Safety Science, 11, Vol. 47, No. 9, pp. 1241–1247. ISSN 0925-7535. doi: 10.1016/j.ssci.2009.03.007.

Back, J., Furniss, D., Hildebrandt, M. & Blandford, A. 2008. Resilience Markers for Safer Systems and Organisations. In: Harrison, M.D. & Sujan, M. (eds.).SAFECOMP 2008, LNCS 5219. Berlin Heidelberg: Springer-Verlag. 99 p.

Bannon, L. 1991. From Human Factors to Human Actors. In: Anonymous Design at Work: Cooperative Design of Computer Systems. Mahwah: LEA. Pp. 27–44.

Bardzell, S. 2010. Feminist HCI: taking stock and outlining an agenda for design. Proceedings of the 28th international conference on Human factors in computing systems. Atlanta, Georgia, USA, New York, NY, USA: ACM. CHI '10. Pp. 1301–1310. ISBN 978-1-60558-929-9. http://doi.acm.org/10.1145/1753326.1753521.

Bargas-Avila, J.A. & Hornbaek, K. 2011. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Vancouver, BC, Canada, New York, NY, USA: ACM. CHI '11. Pp. 2689–2698. ISBN 978-1-4503-0228-9. http://doi.acm.org/10.1145/1978942.1979336.

Bas, d.M. 2007. Designing medical technology. Safety Science, 2, Vol. 45, No. 1–2, pp. 283–291. ISSN 0925-7535. doi: 10.1016/j.ssci.2006.08.003.

Basnyat, S., Palanque, P., Schupp, B. & Wright, P. 2007/6. Formal socio-technical barrier modelling for safety-critical interactive systems design. Safety Science, Vol. 45, No. 5, pp. 545–565.

Baxter, G. & Sommerville, I. 2011. Socio-technical systems: From design methods to systems engineering. Interacting with Computers, 1, Vol. 23, No. 1, pp. 4–17. ISSN 0953-5438. doi: DOI: 10.1016/j.intcom.2010.07.003.

Bedny, G. & Meister, D. 1999. Theory of Activity and Situation Awareness. International Journal of Cognitive Ergonomics, Vol. 3, No. 1, pp. 63–72.

Béguin, P. 2003. Design as a mutual learning process between users and designers. Interacting with Computers, Vol. 15, No. 5, pp. 709. ISSN 0953-5438. doi: 10.1016/S0953-5438(03)00060-2. http://www.sciencedirect.com/science/article/pii/S0953543803000602.

Béguin, P. & Rabardel, P. 2000. Designing for instrument-mediated activity. Scandinavian Journal of Information Systems, No. 12, pp. 173–190.

Bennett, K.B. & Flach, J.M. 2011. Display and Interface Design: Subtle Science, Exact Art. London: Taylor & Francis.

Besnard, D. & Greathead, D. 2003. A cognitive approach to safe violations. Cognition, Technology & Work, Vol. 5, No. 4, pp. 272–282. ISSN 1435-5558. doi: 10.1007/s10111-003-0131-1. http://dx.doi.org/10.1007/s10111-003-0131-1.

Besnard, D. & Hollnagel, E. 2012. I want to believe: some myths about the management of industrial safety. Cognition, Technology & Work, pp. 1–11. ISSN 1435-5558. doi: 10.1007/s10111-012-0237-4. http://dx.doi.org/10.1007/s10111-012-0237-4.

Beyer, H. & Holtzblatt, K. 1998. Contextual Design: Defining Customer-Centered Systems. San Francisco: Morgan Kaufman.

Bjørkli, C., Øvergård, K., Røed, B. & Hoff, T. 2007. Control situations in high-speed craft operation. Cognition, Technology & Work, Vol. 9, No. 2, pp. 67–80. ISSN 1435-5558. doi: 10.1007/s10111-006-0042-z.

Blandford, A. & Wong, B.L.W. 2004. Situation awareness in emergency medical dispatch. International Journal of Human-Computer Studies, Vol. 61, No. 4, pp. 421–452. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2003.12.012. http://www.sciencedirect.com/science/article/pii/S1071581904000102.

Bobko, N., Karpenko, A., Gerasimov, A. & Chernyuk, V. 1998. The mental performance of shiftworkers in nuclear and heat power plants of Ukraine. International Journal of Industrial Ergonomics, Vol. 21, No. 3–4, pp. 333–340. ISSN 0169-8141. doi: 10.1016/S0169-8141(97)00053-X. http://www.sciencedirect.com/science/article/pii/S016981419700053X.

Bødker, S. 2006. When Second Wave HCI Meets Third Wave Challenges. Oslo: ACM Press.

Boy, G.A. & Schmitt, K.A. 2013. Design for safety: A cognitive engineering approach to the control and management of nuclear power plants. Annals of Nuclear Energy, Vol. 52, pp. 125–136. ISSN 0306-4549. doi: 10.1016/j.anucene.2012.08.027. http://www.sciencedirect.com/science/article/pii/S0306454912003477.

Braarud, P.O. & Skraaning jr, G. 2006. Insights from a benchmark integrated validation of a modernized NPP control room: Performance measurement and the comparison to benchmark system. American Nuclear Society, NPIC&HMIT. Albuquerque, NM, November 12–16.

Braseth, A.O., Nihlwing, C., Svengren, H., Veland, O., Hurlen, L. & Kvalem, J. 2009. Lessons learned from Halden Project Research on Human System Interfaces. Nuclear Engineering and Technology, Vol. 41, No. 3, pp. 215.

Bruseberg, A. 2008. Presenting the value of Human Factors Integration: guidance, arguments and evidence. Cognition, Technology & Work, Vol. 10, No. 3, pp. 181–189. ISSN 1435-5558. doi: 10.1007/s10111-007-0100-1. http://dx.doi.org/10.1007/s10111-007-0100-1.

Burns, C.M. & Hajdukiewicz, J.R. 2004. EID: Ecological Interface Design. Boca Raton, FL: CRC Press.

Burns, C.M., Skraaning Jr., G., Jamieson, G.A., Lau, N., Kwok, J., Welch, R. & Andresen, G. 2008. Evaluation of Ecological Interface Design for Nuclear Process Control: Situation Awareness Effects. Human factors, Vol. 50, No. 4, pp. 663–679. ISSN 0018-7208. http://search.ebscohost.com/login.aspx?direct=true&db=asr&AN=34165992&site=ehost-live.

Cacciabue, P.C. 2010. Dynamic reliability and human factors for safety assessment of technological systems: a modern science rooted in the origin of mankind. Cognition, Technology & Work, Vol. 12, No. 2, pp. 119–131. ISSN 1435-5558. doi: 10.1007/s10111-010-0145-4. http://dx.doi.org/10.1007/s10111-010-0145-4.

Carroll, J.S., Hatakenaka, S. & Rudolph, J.W. 2006. Naturalistic Decision Making and Organizational Learning in Nuclear Power Plants: Negotiating Meaning Between Managers and Problem Investigation Teams. Organization Studies, Vol. 27, No. 7, pp. 1037–1057. doi: 10.1177/0170840606065709. http://oss.sagepub.com/cgi/content/abstract/27/7/1037.

Carvalho, P.V.R., Santos, I.L.d. & Vidal, M.C.R. 2006. Safety implications of cultural and cognitive issues in nuclear power plant operation. Applied Ergonomics, Vol. 37, No. 2, pp. 211. ISSN 0003-6870. doi: 10.1016/j.apergo.2005.03.004. http://www.sciencedirect.com/science/article/pii/S0003687005000724.

Carvalho, P.V.R., Santos, I.L.d. & Vidal, M.C.R. 2005. Nuclear power plant shift supervisor's decision making during microincidents. International Journal of Industrial Ergonomics, Vol. 35, No. 7, pp. 619–644. ISSN 0169-8141. doi: 10.1016/j.ergon.2005.01.010. http://www.sciencedirect.com/science/article/pii/S0169814105000272.

Choudhry, R.M. & Fang, D. 2008. Why operatives engage in unsafe work behavior: Investigating factors on construction sites. Safety Science, 4, Vol. 46, No. 4, pp. 566–584. ISSN 0925-7535. doi: 10.1016/j.ssci.2007.06.027.

Chuang, C. & Chou, H. 2008. Design Development and Implementation of the Human-System Interface for Lungmen Nuclear Project. IEEE Transactions on Nuclear Science, Vol. 55, No. 5, pp. 2654–2661. ISSN 0018-9499. doi: 10.1109/TNS.2008.2003977.

Cockton, G. 2006. Designing worth is worth designing. Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles. Oslo, Norway, New York, NY, USA: ACM. NordiCHI '06. Pp. 165–174. ISBN 1-59593-325-5. http://doi.acm.org/10.1145/1182475.1182493.

Corcoran, W.R., Porter, N.J., Church, J.F., Cross, M.T. & Quinn, W.M. 1981. The critical safety functions and plant operation. Nuclear Technology, Vol. 55, No. 12, pp. 690–712.

Crichton, M.T. & Flin, R. 2004. Identifying and training non-technical skills of nuclear emergency response teams. Annals of Nuclear Energy, Vol. 31, No. 12, pp. 1317–1330. ISSN 0306-4549. doi: 10.1016/j.anucene.2004.03.011. http://www.sciencedirect.com/science/article/pii/S0306454904000611.

Crichton, M.T., Flin, R. & McGeorge, P. 2005. Decision making by on-scene incident commanders in nuclear emergencies. Cognition, Technology & Work, Vol. 7, No. 3, pp. 156–166. ISSN 1435-5558. doi: 10.1007/s10111-005-0183-5. http://dx.doi.org/10.1007/s10111-005-0183-5.

Crnkovic, G. 2010. Constructive Research and Info-computational Knowledge Generation. Vol. 314, pp. 359–380. ISSN 978-3-642-15222-1. doi: 10.1007/978-3-642-15223-8_20. http://dx.doi.org/10.1007/978-3-642-15223-8_20.

Daniellou, F. 2005. The French-speaking ergonomists' approach to work activity: cross-influences of field intervention and conceptual models. Theoretical Issues in Ergonomics Science, 09, Vol. 6, No. 5, pp. 409–427. ISSN 1463922X. doi: 10.1080/14639220500078252. http://search.ebscohost.com/login.aspx?direct=true&db=asr&AN=17552360&site=ehost-live.

Daniellou, F. 2004. L’ergonomie dans la conduite de projets de conception de systèmes de travail. In: Anonymous Ergonomie. Paris: PUF.

Daniellou, F. 1992. Le statut de la pratique et des connaissances en ergonomie. Habilitation Thesis ed. Université de Toulouse-Le Mirail.

De Carvalho, P.V.R. 2006. Ergonomic field studies in a nuclear power plant control room. Progress in Nuclear Energy, 1, Vol. 48, No. 1, pp. 51–69. ISSN 0149-1970. doi: DOI: 10.1016/j.pnucene.2005.04.001.

Dekker, S. 2011. Drift into Failure – From Hunting Broken Components to Understanding Complex Systems. Burlington, VT: Ashgate.

Dekker, S., Cilliers, P. & Hofmeyr, J. 2011. The complexity of failure: Implications of complexity theory for safety investigations. Safety Science, 7, Vol. 49, No. 6, pp. 939–945. ISSN 0925-7535. doi: 10.1016/j.ssci.2011.01.008.

Dekker, S. & Nyce, J.M. 2012. Cognitive engineering and the moral theology and witchcraft of cause. Cognition, Technology & Work, Vol. 14, No. 3, pp. 207–212. ISSN 1435-5558. doi: 10.1007/s10111-011-0203-6. http://dx.doi.org/10.1007/s10111-011-0203-6.

Dewey, J. 2002. Human nature and conduct. Toronto: Dover Publications, Inc.

Dewey, J. 1999. The Quest for Certainty. A Study of the Relation of Knowledge and Action (Finnish translation). Helsinki: Gaudeamus.

Dewey, J. 1896. The reflex arc concept in psychology. Psychological Review, Vol. 3, No. 357, pp. 357–370.

Diaper, D. & Lindgaard, G. 2008. West meets East: Adapting Activity Theory for \HCI\ & \CSCW\ applications? Interacting with Computers, Vol. 20, No. 2, pp. 240–246. ISSN 0953-5438. doi: 10.1016/j.intcom.2007.11.006. http://www.sciencedirect.com/science/article/pii/S0953543807000872.

Dien, Y. 1998. Safety and application of procedures, or how do “they” have to use operating procedures in nuclear power plants. Safety Science, Vol. 29, No. 3, pp. 179–187.

Dix, A., Ramduny, D. & Wilkinson, J. 1998. Interaction in the large. Interacting with Computers, Vol. 11, No. 1, pp. 9–32. ISSN 0953-5438. doi: 10.1016/S0953-5438(98)00031-9. http://www.sciencedirect.com/science/article/pii/S0953543898000319.

dos Santos,Isaac José Antonio Luquetti, Grecco, C.H.d.S., Mol, A.C.A. & Carvalho, P.V.R. 2009. The use of questionnaire and virtual reality in the verification of the human factors issues in the design of nuclear control desk. International Journal of Industrial Ergonomics, Vol. 39, No. 1, pp. 159–166. ISSN 0169-8141. doi: 10.1016/j.ergon.2008.08.005. http://www.sciencedirect.com/science/article/pii/S0169814108001376.

Dourish, P. 2003. The Appropriation of Interactive Technologies: Some Lessons from Placeless Documents. Computer Supported Cooperative Work (CSCW), Vol. 12, No. 4, pp. 465–490. ISSN 0925-9724. doi: 10.1023/A:1026149119426. http://dx.doi.org/10.1023/A%3A1026149119426.

Dourish, P. 2001. Where the Action Is – The Foundations of Embodied Interaction. Cambridge: MIT Press.

Dul, J., Bruder, R., Buckle, P., Carayon, P., Falzon, P., Marras, W.S., Wilson, J.R. & van der Doelen, B. 2012. A strategy for human factors/ergonomics: developing the discipline and profession. Ergonomics, Vol. 55, No. 4, pp. 377–395.

Eckert, C. & Clarkson, J. (2005). The reality of design. In J. Clarkson, & C. Eckert (Eds.), Design process improvement - A review of current practice (pp. xvi). London: Springer-Verlag.

Edmondson, W.H. & Beale, R. 2008. Projected Cognition – extending Distributed Cognition for the study of human interaction with computers. Interacting with Computers, Vol. 20, No. 1, pp. 128–140. ISSN 0953-5438. doi: 10.1016/j.intcom.2007.10.005. http://www.sciencedirect.com/science/article/pii/S0953543807000781.

Endsley, M.R. & Kiris, E.O. 1995. The out of the loop performance problem and level of control in automation. Human factors, Vol. 37, No. 2, pp. 381–394.

Engeström, Y., Virkkunen, J., Helle, M., Pihlaja, J. & Poikela, R. 1996. The change laboratory as a tool for transforming work. Lifelong learning in Europe, Vol. 1, No. 2, pp. 10–17.

Falzon, P. 2008. Enabling safety: issues in design and continuous design. Cognition, Technology & Work, Vol. 10, No. 1, pp. 7–14. ISSN 1435-5558. doi: 10.1007/s10111-007-0072-1. http://dx.doi.org/10.1007/s10111-007-0072-1.

Farrington-Darby, T. & Wilson, J. 2009. Understanding social interactions in complex work: a video ethnography. Cognition, Technology & Work, Vol. 11, No. 1, pp. 1–15. ISSN 1435-5558. doi: 10.1007/s10111-008-0118-z. http://dx.doi.org/10.1007/s10111-008-0118-z.

Filippi, G. 2006. Studying computerized emergency operating procedures to evaluate the impact of strong procedure guidance on operators' work practices. Proceedings of the American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Controls, and Human-Machine Interface Technologies. Albuquerque, NM.

Flach, J., Mulder, M. & van Paassen, M.M. 2004. The Concept of the Situation in Psychology. A Cognitive Approach to Situation Awareness: Theory and Application. Burlington. Pp. 42–60.

Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N. & Patey, R. 2004. Rating non-technical skills: developing a behavioural marker system for use in anaesthesia. Cognition, Technology & Work, Vol. 6, No. 3, pp. 165–171. ISSN 1435–5558. doi: 10.1007/s10111-004-0158-y. http://dx.doi.org/10.1007/s10111-004-0158-y.

Fuld, R. 2000. The fiction of function allocation, revisited. International Journal of Human-Computer Studies, Vol. 52, No. 2, pp. 217–233. ISSN 1071-5819. doi: 10.1006/ijhc.1999.0286. http://www.sciencedirect.com/science/article/pii/S1071581999902860.

Furniss, D., Back, J., Blandford, A., Hildebrandt, M. & Broberg, H. 2011. A resilience markers framework for small teams. Reliability Engineering & System Safety, 1, Vol. 96, No. 1, pp. 2–10. ISSN 0951-8320. doi: 10.1016/j.ress.2010.06.025.

Furuta, K., Sasou, K., Kubota, R., Ujita, H., Shuto, Y. & Yagi, E. 2000. Human Factor Analysis of JCO Criticality Accident. Cognition, Technology & Work, Vol. 2, No. 4, pp. 182–203. ISSN 1435-5558. doi: 10.1007/PL00011501. http://dx.doi.org/10.1007/PL00011501.

Gatto, L.B.S., Mól, A.C.A., Santos, I.J.A.L.d., Jorge, C.A.F. & Legey, A.P. 2013. Virtual simulation of a nuclear power plant's control room as a tool for ergonomic evaluation. Progress in Nuclear Energy, Vol. 64, pp. 8–15. ISSN 0149-1970. doi: 10.1016/j.pnucene.2012.11.006". http://www.sciencedirect.com/science/article/pii/S0149197012001412.

Gay, G. & Hembrooke, H. 2004. Activity-Centered Design: An Ecological Approach to Designing Smart Tools and Usable Systems. Cambridge, MA: MIT Press.

Gibbons, M., Limoges, C., Novotny, H., Schwartzman, S., Scott, P. & Trow, M. 1994. The new production of knowledge – The dynamics of science and research in contemporary societies. London: Sage Publications Ltd.

Gore, J., Banks, A., Millward, L. & Kyriakidou, O. 2006. Naturalistic Decision Making and Organizations: Reviewing Pragmatic Science. Organization Studies, Vol. 27, No. 7, pp. 925–942. doi: 10.1177/0170840606065701. http://oss.sagepub.com/cgi/content/abstract/27/7/925.

Grote, G., Weichbrodt, J., Günter, H., Zala-Mezö, E. & Künzle, B. 2009. Coordination in high-risk organizations: the need for flexible routines. Cognition, Technology & Work, Vol. 11, No. 1, pp. 17–27. ISSN 1435-5558. doi: 10.1007/s10111-008-0119-y. http://dx.doi.org/10.1007/s10111-008-0119-y.

Grudin, J. 2008. A Moving Target: The Evolution of HCI. In: Anonymous The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications. Second edition ed. Boca Rayon: CRC Press. Pp. 1–24.

Ha, J.S., Seong, P., Lee, M.S. & Hong, J.H. 2007. Development of Human Performance Measures for Human Factors Validation in the Advanced MCR of APR-1400. IEEE Transactions on Nuclear Science, Vol. 54, No. 6, pp. 2687–2700. ISSN 0018-9499. doi: 10.1109/TNS.2007.907549.

Hahn, S.E. & Murphy, L.R. 2008. A short scale for measuring safety climate. Safety Science, 8, Vol. 46, No. 7, pp. 1047–1066. ISSN 0925-7535. doi: 10.1016/j.ssci.2007.06.002.

Hamilton, W.I., Kazem, M.L.N., He, X. & Dumolo, D. 2013. Practical human factors integration in the nuclear industry. Cognition, Technology & Work, Vol. 15, No. 1, pp. 5–12. ISSN 1435-5558. doi: 10.1007/s10111-012-0213-z. http://dx.doi.org/10.1007/s10111-012-0213-z.

Han, S.H., Yang, H. & Im, D. 2007. Designing a human–computer interface for a process control room: A case study of a steel manufacturing company. International Journal of Industrial Ergonomics, Vol. 37, No. 5, pp. 383–393. ISSN 0169-8141. doi: 10.1016/j.ergon.2006.12.006. http://www.sciencedirect.com/science/article/pii/S0169814106002678.

Harrison, S., Sengers, P. & Tatar, D. 2011. Making epistemological trouble: Third-paradigm HCI as successor science. Interacting with Computers, 9, Vol. 23, No. 5, pp. 385–392. ISSN 0953-5438. doi: 10.1016/j.intcom.2011.03.005.

Hart, S.G. & Staveland, L.E. 1988. Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In: Anonymous Human Mental Workload. Amsterdam: Elsevier. Pp. 139–183.

Hassenzahl, M. & Tractinsky, N. 2006. User experience – a research agenda. Behaviour & Information Technology, Mar, Vol. 25, No. 2, pp. 91–97. ISSN 0144929X. doi: 10.1080/01449290500330331. http://search.ebscohost.com/login.aspx?direct=true&db=asr&AN=19277459&site=ehost-live.

Hayes, J. 2011. Use of safety barriers in operational safety decision making. Safety Science, No. 0ISSN 0925-7535. doi: 10.1016/j.ssci.2011.10.002.

Healey, A., Cacciabue, C. & Berman, J. 2013. Human factors in nuclear safety: present and future. Cognition, Technology & Work, No. 15, pp. 1–3.

Heath, C. & Luff, P. 1992. Collaboration and control: Crisis Management and Multimedia Technology in London Underground Line Control Rooms. Journal of Computer Supported Cooperative Work, Vol. 1, No. 1, pp. 22–48.

Hessels, L. & van Lente, H. 2008. Re-thinking new knowledge production: A literature review and a research agenda. Research Policy, Vol. 37, pp. 740–760.

Hevner, A., March, S.T., Park, J. & Ram, S. 2004. Design Science in Information Systems Research. MIS Quaterly, Vol. 28, No. 1, pp. 75–105.

Hoc, J. 1996. Supervision et contrôle de processus: la cognition en situation dynamique. France: Presses Universitaires de Grenoble, Sciences et Technologies de la connaissance.

Hoc, J. 2001a. Towards a cognitive approach to human-machine cooperation in dynamic situations. International Journal of Human-Computer Studies, No. 54, pp. 509–540.

Hoc, J. 2001b. Towards ecological validity of research in cognitive ergonomics. Theoretical Issues in Ergonomics Science, Vol. 2, No. 3, pp. 278–288.

Hoc, J. 1989. Strategies in controlling a continuous process with long response latencies: needs for computer support to diagnosis. International Journal of Man-Machine Studies, Vol. 30, pp. 47–67.

Hoc, J. & Carlier, X. 2002. Role of a Common Frame of Reference in Cognitive Cooperation: Sharing Tasks between Agents in Air Traffic Control. Cognition, Technology & Work, Vol. 4, pp. 37–47.

Hoffman, R.R., Crandall, B. & Shadbolt, N. 1998. Use of the critical decision method to elicit expert knowledge: a case study in the methodology of cognitive task analysis. Human factors, Vol. 40, No. 2, pp. 254–276.

Hollnagel, E. 2013. A Tale of Two Safeties. Nuclear Safety and Simulation, Vol. 4, No. 1.

Hollnagel, E. 2008. Risk + barriers = safety? Safety Science, 2, Vol. 46, No. 2, pp. 221–229. ISSN 0925-7535. doi: 10.1016/j.ssci.2007.06.028.

Hollnagel, E. (ed.). 2003. Handbook of Cognitive Task Design. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hollnagel, E., Paries, J. & Woods, D. (eds.). 2011. Resilience Engineering in Practice: A Guidebook. Burlington: Ashgate. Safety Culture and Behavioral

Change at the Workplace: A selection of papers from the 23rd NeTWork Workshop, 2004. Pp. 637–652. ISSN 0925-7535. http://www.sciencedirect.com/science/article/pii/S0925753507000197.

Hollnagel, E. & Woods, D.D. 2005. Joint Cognitive Systems: Foundations of Cognitive Systems Engineering. Boca Raton, FL: CRC Press.

Hollnagel, E. & Woods, D.D. 1983. Cognitive Systems Engineering: New wine in new bottles. International Journal of Man-Machine Studies, Vol. 18, pp. 583–600.

Hollnagel, E., Woods, D. & Leveson, N. 2006. Resilience Engineering: Concepts and Precepts. Padstow, GB: Ashgate.

Huang, F. & Hwang, S. 2009. Experimental studies of computerized procedures and team size in nuclear power plant operations. Nuclear Engineering and Design, No. 239, pp. 373–380.

Huang, F., Hwang, S., Yenn, T., Yu, Y., Hsu, C. & Huang, H. 2006. Evaluation and comparison of alarm reset modes in advanced control room of nuclear power plants. Safety Science, 12, Vol. 44, No. 10, pp. 935–946. ISSN 0925-7535. doi: 10.1016/j.ssci.2006.06.002.

Huang, F., Lee, Y., Hwang, S., Yenn, T., Yu, Y., Hsu, C. & Huang, H. 2007. Experimental evaluation of human–system interaction on alarm design. Nuclear Engineering and Design, Vol. 237, No. 3, pp. 308–315. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2006.04.027. http://www.sciencedirect.com/science/article/pii/S002954930600313X.

Hutchins, E. 1995. Cognition in the Wild. Cambridge MA: MIT Press.

Hwang, S., Liang, S.M., Liu, T.Y., Yang, Y., Chen, P. & Chuang, C. 2009a. Evaluation of human factors in interface design in main control rooms. Nuclear Engineering and Design, Vol. 239, No. 12, pp. 3069–3075. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2009.09.006. http://www.sciencedirect.com/science/article/pii/S0029549309004336.

Hwang, S., Liang, S.M., Liu, T.Y., Yang, Y., Chen, P. & Chuang, C. 2009b. Evaluation of human factors in interface design in main control rooms. Nuclear Engineering & Design, Vol. 239, No. 12, pp. 3069–3075. ISSN 0029-5493. http://search.ebscohost.com/login.aspx?direct=true&db=asr&AN=45422816&site=ehost-live.

Hwang, S., Lin, J., Liang, G., Yau, Y., Yenn, T. & Hsu, C. 2008a. Application control chart concepts of designing a pre-alarm system in the nuclear power plant control room. Nuclear Engineering and Design, Vol. 238, No. 12, pp. 3522–3527. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2008.07.011. http://www.sciencedirect.com/science/article/pii/S0029549308003087.

Hwang, S., Yau, Y., Lin, Y., Chen, J., Huang, T., Yenn, T. & Hsu, C. 2008b. Predicting work performance in nuclear power plants. Safety Science, 8, Vol. 46, No. 7, pp. 1115–1124. ISSN 0925-7535. doi: 10.1016/j.ssci.2007.06.005.

Hyysalo, S. 2002. Transforming the object in product design. Outlines, Critical Social Studies, Vol. 4, No. 1, pp. 59–83.

IEA 2010. Definition of Ergonomics. Available: http://www.iea.cc/01_what/What%20is%20Ergonomics.html [2013, September/17].

ISO 9241-210 2010. Human-centred design for interactive systems. European Standard. Geneve: International Standardization Organization.

ISO 9241-11 1998. ISO 9241 Ergonomic Requirements for Office Work with Visual Display Terminals, part 11: Guidance on Usability. Geneve: International Standardization Organization.

J. Robert, T. 2007. Statistics of design error in the process industries. Safety Science, 2, Vol. 45, No. 1–2, pp. 61–73. ISSN 0925-7535. doi: 10.1016/j.ssci.2006.08.013.

Jamieson, G.A. 2007. Ecological Interface Design for Petrochemical Process Control: An Empirical Assessment. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, Vol. 37, No. 6, pp. 906–920. ISSN 1083-4427. doi: 10.1109/TSMCA.2007.897583.

Jamieson, G.A., Miller, C.A., Ho, W.H. & Vicente, K.J. 2007. Integrating Task- and Work Domain-Based Work Analyses in Ecological Interface Design: A Process Control Case Study. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, Vol. 37, No. 6, pp. 887–905. ISSN 1083-4427. doi: 10.1109/TSMCA.2007.904736.

Jang, G., Suh, S., Kim, S., Suh, Y. & Park, J. 2013a. A proactive alarm reduction method and its human factors validation test for a main control room for \SMART\. Annals of Nuclear Energy, Vol. 51, pp. 125. ISSN 0306-4549. doi: 10.1016/j.anucene.2012.07.035. http://www.sciencedirect.com/science/article/pii/S0306454912003210.

Jang, I., Kim, A.R., Harbi, M.A.S.A., Lee, S.J., Kang, H.G. & Seong, P.H. 2013b. An empirical study on the basic human error probabilities for \NPP\ advanced main control room operation using soft control. Nuclear Engineering and Design, Vol. 257, pp. 79–87. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2013.01.003. http://www.sciencedirect.com/science/article/pii/S0029549313000186.

Jang, I., Park, J. & Seong, P. 2012. An empirical study on the relationships between functional performance measure and task performance measure in \NPP\ \MCR\. Annals of Nuclear Energy, Vol. 42, pp. 96. ISSN 0306-4549. doi: 10.1016/j.anucene.2011.10.004. http://www.sciencedirect.com/science/article/pii/S0306454911004026.

Jou, Y., Lin, C.J., Yenn, T., Yang, C., Yang, L. & Tsai, R. 2009a. The implementation of a human factors engineering checklist for human–system interfaces upgrade in nuclear power plants. Safety Science, 8, Vol. 47, No. 7, pp. 1016–1025. ISSN 0925-7535. doi: 10.1016/j.ssci.2008.11.004.

Jou, Y., Yenn, T., Lin, C.J., Yang, C. & Chiang, C. 2009b. Evaluation of operators' mental workload of human–system interface automation in the advanced nuclear power plants. Nuclear Engineering and Design, Vol. 239, No. 11, pp. 2537–2542. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2009.06.023. http://www.sciencedirect.com/science/article/pii/S0029549309003136.

Julien, E. 2008. Knowledge transfer in organisational reliability analysis: From post-accident studies to normal operations studies. Safety Science, 12, Vol. 46, No. 10, pp. 1420–1434. ISSN 0925-7535. doi: 10.1016/j.ssci.2007.10.002.

Kaptelinin, V. & Nardi, B. 2012. Activity theory in HCI – Fundamentals and Reflections. Morgan & Claypool. Synthesis Lectures on Human-Centered Informatics.

Kaptelinin, V. & Nardi, B.A. 2006. Acting with Technology: Activity Theory and Interaction Design. Cambridge, MA: MIT Press.

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, 7, Vol. 23, No. 4, pp. 289–298. ISSN 0953-5438. doi: 10.1016/j.intcom.2011.04.008.

Karvonen, H., Koskinen, H. & Haggrén, J. 2012. Enhancing the user experience of the crane operator: comparing work demands in two operational settings. Proceedings of the 30th European Conference on Cognitive Ergonomics.

Edinburgh, United Kingdom, New York, NY, USA: ACM. ECCE '12. Pp. 37–44. ISBN 978-1-4503-1786-3. http://doi.acm.org/10.1145/2448136.2448144.

Kecklund, L.J., Edland, A., Wedin, P. & Svenson, O. 1996. Safety barrier function analysis in a process industry: A nuclear power application. International Journal of Industrial Ergonomics, Vol. 17, No. 3, pp. 275–284. ISSN 0169-8141. doi: 10.1016/0169-8141(94)00112-X. http://www.sciencedirect.com/science/article/pii/016981419400112X.

Kestenbaum, V. 1977. The Phenomenological Sense of John Dewey: Habit and Meaning. Atlantic Highlands, NJ: Humanities Press.

Kim, S.K., Byun, S.N. & Lee, D.H. 2011. Development and validation of a heuristic model for evaluation of the team performance of operators in nuclear power plants. Annals of Nuclear Energy, Vol. 38, No. 9, pp. 2005–2016. ISSN 0306-4549. doi: 10.1016/j.anucene.2011.04.016". http://www.sciencedirect.com/science/article/pii/S0306454911001411.

Kim, S.K., Suh, S.M., Jang, G.S., Hong, S.K. & Park, J.C. 2012. Empirical research on an ecological interface design for improving situation awareness of operators in an advanced control room. Nuclear Engineering and Design, Vol. 253, pp. 226–237. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2012.08.026. http://www.sciencedirect.com/science/article/pii/S0029549312004578.

Kinnersley, S. & Roelen, A. 2007. The contribution of design to accidents. Safety Science, 2, Vol. 45, No. 1–2, pp. 31–60. ISSN 0925-7535. doi: 10.1016/j.ssci.2006.08.010.

Kirlik, A. 2012. Relevance versus generalization in cognitive engineering. Cognition, Technology & Work, Vol. 14, No. 3, pp. 213–220. ISSN 1435-5558. doi: 10.1007/s10111-011-0204-5. http://dx.doi.org/10.1007/s10111-011-0204-5.

Kirwan, B. 2007. Safety informing design. Safety Science, 2, Vol. 45, No. 1–2, pp. 155–197. ISSN 0925-7535. doi: 10.1016/j.ssci.2006.08.011.

Klemola, U. & Norros, L. 2002. Activity-based analysis of information characteristics of monitors and their use in anaesthetic practice. Paper presented at the 11th European Congress on Cognitive Ergonomics. Catania, Italy.

Klemola, U. & Norros, L. 2001. Practice-based criteria for assessing anaesthetists' habits of action: outline for a reflexive turn in practice. Medical education, Vol. 35, No. 5, pp. 455–464. ISSN 03080110. doi: 10.1046/j.1365-

2923.2001.00894.x. http://search.ebscohost.com/login.aspx?direct=true&db=asr&AN=4537871&site=ehost-live.

Knudsen, F. 2009. Paperwork at the service of safety? Workers' reluctance against written procedures exemplified by the concept of 'seamanship'. Safety Science, 2, Vol. 47, No. 2, pp. 295–303. ISSN 0925-7535. doi: 10.1016/j.ssci.2008.04.004.

Kontogiannis, T. & Kossiavelou, Z. 1999. Stress and team performance: principles and challenges for intelligent decision aids. Safety Science, 12, Vol. 33, No. 3, pp. 103–128. ISSN 0925-7535. doi: 10.1016/S0925-7535(99)00027-2.

Körvers, P.M.W. & Sonnemans, P.J.M. 2008. Accidents: A discrepancy between indicators and facts! Safety Science, 8, Vol. 46, No. 7, pp. 1067–1077. ISSN 0925-7535. doi: 10.1016/j.ssci.2007.06.004.

Koski-Jännes, A. 1999. From addiction to self governance. In: Anonymous Perspectives on activity theory. Cambridge, UK: Cambridge University Press.

Kuutti, K. 2010. Where are the Ionians of user experience research? Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. Reykjavik, Iceland, New York, NY, USA: ACM. NordiCHI '10. Pp. 715–718. ISBN 978-1-60558-934-3. http://doi.acm.org/10.1145/1868914.1869012.

Kuutti, K. 2007. Design Research, Disciplines, and the New Production of Knowledge. Proceedings of the International Association of Societies for Design Research (IASDR) Conference "Emerging Trends in Design Research". Hong Kong, Pp. paper #39.

Kuutti, K. 1991. The concept of activity as a basic unit of analysis for CSCW research. Proceedings of the second conference on European Conference on Computer-Supported Cooperative Work. Amsterdam, The Netherlands, Norwell, MA, USA: Kluwer Academic Publishers. Pp. 249–264. ISBN 0-7923-1439-5. http://portal.acm.org/citation.cfm?id=1241910.1241929.

Laar, D.V. & Deshe, O. 2002. Evaluation of a visual layering methodology for colour coding control room displays. Applied Ergonomics, Vol. 33, No. 4, pp. 371–377. ISSN 0003-6870. doi: 10.1016/S0003-6870(01)00048-5. http://www.sciencedirect.com/science/article/pii/S0003687001000485.

Laarni, J., Savioja, P., Karvonen, H. & Norros, L. 2011. Pre-validation of nuclear power plant control room design. Proceedings of 9th International Conference on Engineering Psychology and Cognitive Ergonomics, EPCE 2011, held as part of CHI International. Orlando, FL, Vol. 6781. Pp. 404–413.

Laitio, P. 2013. User interface solution for supporting operators' automation awareness in nuclear power plant control rooms. Espoo: Aalto University.

Laitio, P., Savioja, P. & Lappalainen, J. 2013. Exploring Automation Awareness in Nuclear Power Plant Control Rooms. Proceedings of EHPG 2013. Storefjell, Norway.

Lau, N., Jamieson, G.A., Skraaning, G. & Burns, C.M. 2008. Ecological Interface Design in the Nuclear Domain: An Empirical Evaluation of Ecological Displays for the Secondary Subsystems of a Boiling Water Reactor Plant Simulator. IEEE Transactions on Nuclear Science, Vol. 55, No. 6, pp. 3597–3610. ISSN 0018-9499. doi: 10.1109/TNS.2008.2005725.

Lau, N., Jamieson, G.A. & Skraaning, G. 2012. Situation Awareness in Process Control: A Fresh Look. Proceedings of NPIC&HMIT 2012. San Diego.

Law, E.L. & Sun, X. 2012. Evaluating user experience of adaptive digital educational games with Activity Theory. International Journal of Human-Computer Studies, Vol. 70, No. 7, pp. 478–497. ISSN 1071-5819. doi: http://dx.doi.org/10.1016/j.ijhcs.2012.01.007. http://www.sciencedirect.com/science/article/pii/S1071581912000304.

Le Bot, P. 2004. Human reliability data, human error and accident models – illustration through the Three Mile Island accident analysis. Reliability Engineering & System Safety, Vol. 83, No. 2, pp. 153–167. ISSN 0951-8320. doi: http://dx.doi.org/10.1016/j.ress.2003.09.007. http://www.sciencedirect.com/science/article/pii/S0951832003002035.

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, Vol. 46, No. 1, pp. 50–80. doi: 10.1518/hfes.46.1.50_30392. http://hfs.sagepub.com/content/46/1/50.abstract.

Lee, S.W., Kim, A.R., Ha, J.S. & Seong, P.H. 2011. Development of a qualitative evaluation framework for performance shaping factors (PSFs) in advanced \MCR\ \HRA\. Annals of Nuclear Energy, Vol. 38, No. 8, pp. 1751–1759. ISSN 0306-4549. doi: 10.1016/j.anucene.2011.04.006. http://www.sciencedirect.com/science/article/pii/S0306454911001319.

Leontjev, A.N. 1978. Activity, Consciousness, and Personality. Upper Saddle River, NJ: Prentice-Hall.

Leplat, J. & Hoc, J. 1981. Subsequent verbalization in the study of cognitive processes. Ergonomics, Vol. 24, No. 10, pp. 743–755.

Leveson, N. 2011a. Engineering a Safer World: Systems thinking applied to safety. Cambridge, MA: MIT Press.

Leveson, N.G. 2011b. Applying systems thinking to analyze and learn from events. Safety Science, 1, Vol. 49, No. 1, pp. 55–64. ISSN 0925-7535. doi: 10.1016/j.ssci.2009.12.021.

Li, X., Powell, M.S. & Horberry, T. 2012. Human Factors in Control Room Operations in Mineral Processing: Elevating Control From Reactive to Proactive. Journal of Cognitive Engineering and Decision Making, Vol. 6, No. 1, pp. 88–111. doi: 10.1177/1555343411432340. http://edm.sagepub.com/content/6/1/88.abstract.

Li, X., Sanderson, P., Memisevic, R., Wong, W. & Choudhury, S. 2006. Evaluating functional displays for hydropower system: model-based guidance of scenario design. Cognition, Technology & Work, Vol. 8, No. 4, pp. 269–282. ISSN 1435-5558. doi: 10.1007/s10111-006-0044-x. http://dx.doi.org/10.1007/s10111-006-0044-x.

Lin, C.J., Yenn, T. & Yang, C. 2010a. Automation design in advanced control rooms of the modernized nuclear power plants. Safety Science, 1, Vol. 48, No. 1, pp. 63–71. ISSN 0925-7535. doi: 10.1016/j.ssci.2009.05.005.

Lin, C.J., Yenn, T. & Yang, C. 2010b. Optimizing human–system interface automation design based on a skill-rule-knowledge framework. Nuclear Engineering and Design, Vol. 240, No. 7, pp. 1897–1905. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2010.03.026. http://www.sciencedirect.com/science/article/pii/S0029549310002049.

Lind, M. 2003. Making sense of the abstraction hierarchy in the power plant domain. Cognition, Technology & Work, Vol. 5, No. 2, pp. 67–81. ISSN 1435-5558. doi: 10.1007/s10111-002-0109-4. http://dx.doi.org/10.1007/s10111-002-0109-4.

Lind, M. 1999. Plant Modeling for Human Supervisory Control. Institute of Measurement and Control. Transactions, Vol. 21, No. 4–5, pp. 171–180. doi: 10.1177/014233129902100405.

Luff, P., Heath, C. & Jirotka, M. 2000. Surveying the scene: technologies for everyday awareness and monitoring in control rooms. Interacting with Computers, Vol. 13, No. 2, pp. 193–228. ISSN 0953-5438. doi: 10.1016/S0953-5438(00)00038-2. http://www.sciencedirect.com/science/article/pii/S0953543800000382.

Luff, P. & Heath, C. 2000. The collaborative production of computer commands in command and control. International Journal of Human-Computer Studies, Vol. 52, No. 4, pp. 669–699. ISSN 1071-5819. doi: 10.1006/ijhc.1999.0354. http://www.sciencedirect.com/science/article/pii/S1071581999903543.

Luojus, S. 2010. From a Momentary Experience to a Lasting One: The Concept of and Research on Expanded User Experience of Mobile Devices. Oulu: University of Oulu.

Määttänen, P. 2009. Toiminta ja kokemus (Action and expereince). Helsinki: Gaudeamus.

MacIntyre, A. 1984. After Virtue: Study in Moral Theory. Notre Dame, IN: University of Notre Dame Press.

MacKay, R.B. & McKiernan, P. 2004. The role of hindsight in foresight: refining strategic reasoning. Futures, Vol. 36, No. 2, pp. 161–179. ISSN 0016-3287. doi: 10.1016/S0016-3287(03)00147-2". http://www.sciencedirect.com/science/article/pii/S0016328703001472.

Maddox, M. & Brickey, M. 1982. The Integration of Human Factors Methodology into Nuclear Power Plant Control Room Reviews. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Vol. 28.

March, S.T. & Smith, G.F. 1995. Design and natural science research on information technology. Decision Support Systems, Vol. 15, No. 4, pp. 251–266. ISSN 0167-9236. doi: http://dx.doi.org/10.1016/0167-9236(94)00041-2. http://www.sciencedirect.com/science/article/pii/0167923694000412.

Martínez-Córcoles, M., Gracia, F., Tomás, I. & Peiró, J.M. 2011. Leadership and employees' perceived safety behaviours in a nuclear power plant: A structural equation model. Safety Science, 10, Vol. 49, No. 8–9, pp. 1118–1129. ISSN 0925-7535. doi: 10.1016/j.ssci.2011.03.002.

McCarthy, J. 2000. The paradox of understanding work for design. International Journal of Human-Computer Studies, Vol. 53, No. 1, pp. 197. ISSN

1071–5819. doi: 10.1006/ijhc.2000.0371. http://www.sciencedirect.com/science/article/pii/S1071581900903719.

McCarthy, J. & Wright, P. 2004. Technology as Experience. Cambridge: MIT Press.

Megill, A. 1997. Four Senses of Objectivity. In: Anonymous Rethinking Objectivism. Durham: Duke University Press. Pp. 1–20.

Meister, D. 1995. Cognitive behavior of nuclear reactor operators. International Journal of Industrial Ergonomics, Vol. 16, No. 2, pp. 109–122. ISSN 0169-8141. doi: 10.1016/0169-8141(94)00089-L. http://www.sciencedirect.com/science/article/pii/016981419400089L.

Merleau-Ponty, M. 1986. The Phenomenology of Perception. London: Routledge & Kegan Paul.

Morel, G., Amalberti, R. & Chauvin, C. 2009. How good micro/macro ergonomics may improve resilience, but not necessarily safety. Safety Science, Vol. 47, No. 2, pp. 285–294. ISSN 0925-7535. doi: 10.1016/j.ssci.2008.03.002.

Morineau, T., Beuzet, E., Rachinel, A. & Tobin, L. 2009. Experimental evaluation of a tide prediction display based on the ecological interface design framework. Cognition, Technology & Work, Vol. 11, No. 2, pp. 119–127. ISSN 1435-5558. doi: 10.1007/s10111-007-0103-y. http://dx.doi.org/10.1007/s10111-007-0103-y.

Mumaw, R.J., Roth, E.M., Vicente, K., J. & Burns, C.M. 2000. There Is More to Monitoring a Nuclear Power Plant than Meets the Eye. Human factors, Vol. 42, No. 1, pp. 36. ISSN 00187208. http://search.ebscohost.com/login.aspx?direct=true&db=asr&AN=3361051&site=ehost-live.

Naikar, N. 2006. Beyond interface design: Further applications of cognitive work analysis. International Journal of Industrial Ergonomics, Vol. 36, No. 5, pp. 423–438. ISSN 0169-8141. doi: 10.1016/j.ergon.2006.01.006. http://www.sciencedirect.com/science/article/pii/S0169814106000205.

Naito, N., Itoh, J., Monta, K. & Makino, M. 1995. An intelligent human-machine system based on an ecological interface design concept. Nuclear Engineering and Design, Vol. 154, No. 2, pp. 97–108. ISSN 0029-5493. doi: 10.1016/0029-5493(94)00903-C. http://www.sciencedirect.com/science/article/pii/002954939400903C.

Nardi, B.A. 1996. Activity Theory and Human-Computer Interaction. In: Anonymous Context and Consciousness. Cambridge, MA: MIT Press. Pp. 7–16.

NEA 5334. 2004. Nuclear Regulatory Challenges Related to Human Performance. Paris: OECD NEA. 21 p.

Neale, D.C., Carroll, J.M. & Rosson, M.B. 2004. Evaluating computer-supported cooperative work: models and frameworks. Proceedings of the 2004 ACM conference on Computer supported cooperative work. Chicago, Illinois, USA, New York, NY, USA: ACM. CSCW '04. Pp. 112–121. ISBN 1-58113-810-5. http://doi.acm.org/10.1145/1031607.1031626.

Nelson, J., Buisine, S. & Aoussat, A. 2013. Anticipating the use of future things: Towards a framework for prospective use analysis in innovation design projects. Applied Ergonomics. ISSN 0003-6870. doi: 10.1016/j.apergo. 2013.01.002". http://www.sciencedirect.com/science/article/pii/S0003687 013000033.

Nemeth, C., Wears, R., Patel, S., Rosen, G. & Cook, R. 2011. Resilience is not control: healthcare, crisis management, and ICT. Cognition, Technology & Work, Vol. 13, No. 3, pp. 189–202. ISSN 1435-5558. doi: 10.1007/ s10111-011-0174-7. http://dx.doi.org/10.1007/s10111-011-0174-7.

Nemeth, C. 2012. Adapting to change and uncertainty. Cognition, Technology & Work, Vol. 14, No. 3, pp. 183–186. ISSN 1435-5558. doi: 10.1007/ s10111-011-0200-9. http://dx.doi.org/10.1007/s10111-011-0200-9.

Nielsen, J. 1993. Usability Engineering. Boston MA: Academic Press.

Norman, D.A. 2006. Logic versus usage: the case for activity-centered design. Interactions, November, Vol. 13, No. 6, pp. 45-ff. ISSN 1072-5520. doi: http://doi.acm.org/10.1145/1167948.1167978. http://doi.acm.org/10.1145/1167948.1167978.

Norros, L. 2013. Developing human factors/ergonomics as a design discipline. Applied Ergonomics. ISSN 0003-6870. doi: http://dx.doi.org/10.1016/ j.apergo.2013.04.024. http://www.sciencedirect.com/science/article/pii/ S0003687013000975.

Norros, L. 2012. Analysis of work practices from the resilience engineering perspective. International electric journal of nuclear safety and simulation, Vol. 3, No. 4.

Norros, L. 2005. The concept of habit in the analysis of situated actions. Theoretical Issues in Ergonomics Science, Vol. 6, No. 5, pp. 385–407. doi: 10.1080/14639220500076520. http://www.tandfonline.com/doi/abs/10.1080/14639220500076520.

Norros, L. 2004. Acting under Uncertainty. The Core-Task Analysis in Ecological Study of Work. VTT Publications 546. Espoo: VTT. http://www.vtt.fi/inf/pdf/publications/2004/P546.pdf.

Norros, L. & Klemola, U. 1999. Methodological considerations in analysing anaesthetists' habits of action in clinical situations. Ergonomics, Vol. 42, No. 11, pp. 1521–1530.

Norros, L., Liinasuo, M. & Hutton, R. 2011. Evaluating the potential of new technological tools for safety critical work. Interacting with Computers, 7, Vol. 23, No. 4, pp. 299–307. ISSN 0953-5438. doi: 10.1016/j.intcom.2011.05.003.

Norros, L., Liinasuo, M. & Savioja, P. 2013. Operators' conceptions of Procedure Guidance in NPP Process Control. Naturalistic Decision Making, NDM 2013.

Norros, L., Norros, I., Liinasuo, M. & Seppänen, K. 2012. Impact of human operators on communication network dependability. Cognition, Technology & Work, pp. 1–10. ISSN 1435-5558. doi: 10.1007/s10111-012-0225-8. http://dx.doi.org/10.1007/s10111-012-0225-8.

Norros, L. & Nuutinen, M. 2005. Performance-based usability evaluation of a safety information and alarm system. Int.J.Hum.Comput.Stud., September, Vol. 63, No. 3, pp. 328–361. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2005.03.004. http://dl.acm.org/citation.cfm?id=1096337.1096340.

Norros, L. & Salo, L. 2009. Design of joint systems: a theoretical challenge for cognitive systems engineering. Cognition, Technology & Work, Vol. 11, No. 1, pp. 43–56.

Norros, L., Salo, L., Savioja, P. & Laarni, J. 2009. The Ecological Interface Design Experiment (2005) – Qualitative Analysis of Operator Practices in Three Interface Conditions. In: Tilannetietoisuutta tukevat näytöt prosessiteollisuuden valvomoissa. VTT Tiedotteita 2495. Espoo: VTT. Pp. 154–222. ISBN 978-951-38-7312-7.

Norros, L. & Savioja, P. 2004. Usability evaluation of complex systems – a Literature Review. Helsinki: Radiation and Nuclear Safety Authority (STUK).

Novotny, H., Scott, P. & Gibbons, M. 2001. Re-Thinking Science – Knowledge and the Public in an Age of Uncertainty. Cambridge, UK: Polity Press.

Nuutinen, M. 2006. Expert identity in development of core-task oriented working practices for mastering demanding situations. Six empirical studies on operator work. Doctoral dissertation, University of Helsinki. VTT Publications 604. Espoo: VTT. http://www.vtt.fi/inf/pdf/publications/2006/P604.pdf.

Nuutinen, M. 2005a. Contextual assessment of working practices in changing work. International Journal of Industrial Ergonomics, Vol. 35, No. 10, pp. 905–930. ISSN 0169-8141. doi: 10.1016/j.ergon.2005.04.001. http://www.sciencedirect.com/science/article/pii/S0169814105000661.

Nuutinen, M. 2005b. Expert Identity construct in analysing prerequisites for expertise development: a case study of nuclear power plant operators' on-the-job training. Cognition, Technology & Work, Vol. 7, No. 4, pp. 288–305. ISSN 1435-5558. doi: 10.1007/s10111-005-0013-9. http://dx.doi.org/10.1007/s10111-005-0013-9.

Nuutinen, M. & Norros, L. 2009. Core task analysis in accident investigation: analysis of maritime accidents in piloting situations. Cognition, Technology & Work, Vol. 11, No. 2, pp. 129–150. ISSN 1435-5558. doi: 10.1007/s10111-007-0104-x. http://dx.doi.org/10.1007/s10111-007-0104-x.

Nuutinen, M. & Norros, L. 2001. Co-operation on Bridge in Piloting Situations: Analysis of 13 Accidents on Finnish Fairways. CSAPC 01 8th Conference on Cognitive Science Approaches to Process Control. Munich,

O'Connor, P., O'Dea, A., Flin, R. & Belton, S. 2008. Identifying the team skills required by nuclear power plant operations personnel. International Journal of Industrial Ergonomics, Vol. 38, No. 11–12, pp. 1028–1037. ISSN 0169-8141. doi: 10.1016/j.ergon.2008.01.014. http://www.sciencedirect.com/science/article/pii/S0169814108000346.

Obrist, M., Reitberger, W., Wurhofer, D., Förster, F. & Tscheligi, M. 2011. User Experience Research in the Semiconductor Factory: A Contradiction? In: Anonymous Human-Computer Interaction – INTERACT 2011. Vol. 6949. Lecture Notes in Computer Science. Berlin/Heidelberg: Springer. Pp. 144–151. ISBN 978-3-642-23767-6. http://dx.doi.org/10.1007/978-3-642-23768-3_12.

O'Hara, J. 1999. A quasi-experimental model of complex human-machine system validation. Cognition, Technology & Work, Vol. 1, No. 1, pp. 37–46.

O'Hara, J., Brown, W., Lewis, P. & Persensky, J. 2002. The Effects of Interface Management Tasks On Crew Performance and Safety in Complex, Computer-Based Systems. Upton, NY: Brookhaven National Laboratory.

O'Hara, J., Higgins, J. & Brown, W. 2008. Human Factors Considerations with Respect to Emerging Technology in Nuclear Power Plants. Upton, NY: Brookhaven National Laboratory.

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

Paay, J., Sterling, L., Vetere, F., Howard, S. & Boettcher, A. 2009. Engineering the social: The role of shared artifacts. International Journal of Human-Computer Studies, Vol. 67, No. 5, pp. 437–454. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2008.12.002. http://www.sciencedirect.com/science/article/pii/S1071581908001754.

Palmqvist, H., Bergström, J. & Henriqson, E. 2012. How to assess team performance in terms of control: a protocol based on cognitive systems engineering. Cognition, Technology & Work, Vol. 14, No. 4, pp. 337–353. ISSN 1435-5558. doi: 10.1007/s10111-011-0183-6. http://dx.doi.org/10.1007/s10111-011-0183-6.

Palviainen, J. & Väänänen-Vainio-Mattila, K. 2009. User Experience in Machinery Automation: From Concepts and Context to Design Implications. In: Anonymous Human Centered Design. Vol. 5619. Berlin/Heidelberg: Springer. Lecture Notes in Computer Science. Pp. 1042–1051. ISBN 978-3-642-02805-2. http://dx.doi.org/10.1007/978-3-642-02806-9_119.

Papin, B. 2011. Balancing human and technical reliability in the design of advanced nuclear reactors. Nuclear Engineering and Design, Vol. 241, No. 12, pp. 5238–5244. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2011.08.075. http://www.sciencedirect.com/science/article/pii/S0029549311007497.

Papin, B. & Quellien, P. 2006. The operational complexity index: A new method for the global assessment of the human factor impact on the safety of advanced reactors concepts. Nuclear Engineering and Design, Vol. 236, No. 10, pp. 1113–1121. ISSN 0029-5493. doi: 10.1016/j.nucengdes.

2005.10.015. http://www.sciencedirect.com/science/article/pii/S0029549 305003870.

Park, J. 2009. The Complexity of Proceduralized Tasks. London: Springer-Verlag. ISBN 978-1-84882-790-5.

Park, J. & Cho, S. 2010. Investigating the effect of task complexities on the response time of human operators to perform the emergency tasks of nuclear power plants. Annals of Nuclear Energy, Vol. 37, No. 9, pp. 1160–1171. ISSN 0306-4549. doi: 10.1016/j.anucene.2010.05.001. http://www.science direct.com/science/article/pii/S030645491000157X.

Pasquini, A., Pozzi, S. & McAuley, G. 2008. Eliciting information for safety assessment. Safety Science, 12, Vol. 46, No. 10, pp. 1469–1482. ISSN 0925-7535. doi: 10.1016/j.ssci.2007.11.014.

Paternò, F. & Santoro, C. 2006. Analysing context-dependent deviations in interacting with safety-critical systems. Reliability Engineering & System Safety, Vol. 91, No. 12, pp. 1546–1553. ISSN 0951-8320. doi: 10.1016/ j.ress.2006.01.015. http://www.sciencedirect.com/science/article/pii/S09 51832006000391.

Patrick, J., James, N. & Ahmed, A. 2006. Human processes of control: tracing the goals and strategies of control room teams. Ergonomics, 10/10, Vol. 49, No. 12, pp. 1395–1414. ISSN 0014-0139. http://ejournals.ebsco.com/ direct.asp?ArticleID=429A88EC463883F3677E.

Patterson, E., Woods, D., Cook, R. & Render, M. 2007. Collaborative cross-checking to enhance resilience. Cognition, Technology & Work, Vol. 9, No. 3, pp. 155-162. ISSN 1435-5558. doi: 10.1007/s10111-006-0054-8. http://dx.doi.org/10.1007/s10111-006-0054-8.

Paunonen, H. 1997. Roles of informating process control systems. PhD ed. Tampere University of Technology.

Paunonen, H. & Oksanen, J. 2011. Usability and user experience of process control systems and the feeling of control. Proceedings of 10th International Technical Conference on Pulp, Paper and allied Industry. Pragati Maidan, New Delhi, India. Pp. 299–305.

Peffers, K., Tuunanen, T., Rothenberger, M. & Chatterjee, S. 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems, Vol. 24, No. 3, pp. 45–77. ISSN

07421222. http://search.ebscohost.com/login.aspx?direct=true&db=asr&AN=28843849&site=ehost-live.

Peirce, C.S. 1998a. The Harvard Lectures on Pragmatism. In: Anonymous The Essential Peirce. Selected Philosophical Writings. Vol. 2. Bloomington & Indianapolis: Indiana University Press. Pp. 133–241.

Peirce, C.S. 1998b. The Peirce Edition Project. Introduction. In: Anonymous The Essential Peirce. Selected Philosophical Writings. Bloomington: Indiana University Press.

Peirce, C.S. 1958. Letters to Lade Welby. In: Anonymous Selected writings of C.S. Peirce. New York: Dover Publications. Pp. 380–432.

Pentland, B.T. & Feldman, M.S. 2008. Designing routines: On the folly of designing artifacts, while hoping for patterns of action. Information and Organization, Vol. 18, pp. 235–250.

Petersen, J. 2004. Control situations in supervisory control. Cognition, Technology & Work, Vol. 6, No. 4, pp. 266–274. ISSN 1435-5558. http://dx.doi.org/10.1007/s10111-004-0164-0.

Petersen, J. & May, M. 2006. Scale transformations and information presentation in supervisory control. International Journal of Human-Computer Studies, Vol. 64, No. 5, pp. 405–419. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2005.09.001. http://www.sciencedirect.com/science/article/pii/S1071581905001692.

Pirus, D. 2004. Functional HSI for Computerised Operation. Proceedings of NPIC&HMIT 2004. Columbus, OH.

Randell, R., Wilson, S. & Fitzpatrick, G. 2010. Editorial – Evaluating New Interactions in Health Care: Challenges and Approaches. International Journal of Human-Computer Interaction, 05, Vol. 26, No. 5, pp. 407–413. ISSN 1044-7318. doi: 10.1080/10447311003719847. http://search.ebscohost.com/login.aspx?direct=true&db=asr&AN=49261413&site=ehost-live.

Räsänen, M. & Nyce, J.M. 2006. A new role for anthropology?: rewriting context and analysis in HCI research. Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles. Oslo, Norway, New York, NY, USA: ACM. NordiCHI '06. Pp. 175–184. ISBN 1-59593-325-5. http://doi.acm.org/10.1145/1182475.1182494.

Rasmussen, J. 1986. Information processing and human-machine interaction. An approach to cognitive engineering. Amsterdam: Elsevier.

Rasmussen, J. 1985. The role of hierarchical knowledge representation in decision making and system management. IEEE Transactions on Systems, Man and Cybernetics, No. 15, pp. 234–243.

Rasmussen, J. 1983. Skill, rules, and knowledge: signals, signs and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man and Cybernetics, Vol. 15, No. 3, pp. 257–266.

Reason, J. 1990. Human Error. New York, NY: Cambridge University Press.

Reiman, T. & Oedewald, P. 2007/8. Assessment of complex sociotechnical systems – Theoretical issues concerning the use of organizational culture and organizational core task concepts. Safety Science, Vol. 45, No. 7, pp. 745–768.

Rempe, J., Farmer, M., Corradini, M., Ott, L., Gauntt, R. & Powers, D. 2012. Revisiting Insights from Three Mile Island Unit 2 Postaccident Examinations and Evaluations in View of the Fukushima Daiichi Accident. Nuclear Science and Engineering, Vol. 172, No. 3, pp. 223–248.

Riera, B. 2001. Specifications, Design and Evaluation of an Advanced Human-Adapted Supervisory System. Cognition, Technology & Work, Vol. 3, No. 1, pp. 53–65. ISSN 1435-5558. doi: 10.1007/PL00011522. http://dx.doi.org/10.1007/PL00011522.

Riihiaho, S. 2000. Experiences with usability evaluation methods. Licenciate ed. Helsinki University of Technology.

Rollenhagen, C. 2011. Event investigations at nuclear power plants in Sweden: Reflections about a method and some associated practices. Safety Science, Vol. 49, No. 1, pp. 21–26. ISSN 0925-7535. doi: 10.1016/j.ssci.2009.12.012.

Ross, S., Ramage, M. & Rogers, Y. 1995. PETRA: participatory evaluation through redesign and analysis. Interacting with Computers, 12, Vol. 7, No. 4, pp. 335–360. ISSN 0953-5438. doi: DOI: 10.1016/0953-5438(96)87697-1.

Roth, E.M., Easter, J., Hall, R.E., Kabana, L., Mashio, K., Hanada, S., Clouser, T. & Remley, G.W. 2010. Person-in-the-Loop Testing of a Digital Power

Plant Control Room. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 54, No. 4, pp. 289–293. doi: 10.1177/154193121005400405. http://pro.sagepub.com/content/54/4/289.abstract.

Roth, E., Mumaw, R.J. & Lewis, P. 1994. Empirical investigation of operator performance in cognitively demanding simulated emergencies. Washington DC: Nuclear Regulatory Commission.

Roth, E. & O'Hara, J. 2002. Integrating Digital and Conventional Human-System Interfaces: Lessons Learned from a Control Room Modernization Program.

Roto, V., Law, E., Vermereen, A. & Hoonhout, J. 2011. User Experience White Paper – Bringing Clarity to the Concept of User Experience. http://www.allaboutux.org/files/UX-WhitePaper.pdf.

Rückriem, G. 2009. Digital Technology and Mediation: A Challenge to Activity Theory. In: Anonymous Learning and Expanding with Activity Theory. New York, NY: Cambridge University Press. Pp. 88–111.

Rückriem, G. 2003. Tool or Medium? The Meaning of Information and Telecommunication Technology to Human Practice. A Quest for Systemic Understanding of Activity Theory. Paper presented in FISCAR '03.

Salmon, P., Stanton, N., Walker, G. & Green, D. 2006. Situation awareness measurement: A review of applicability for \C4i\ environments. Applied Ergonomics, Vol. 37, No. 2, pp. 225–238. ISSN 0003-6870. doi: 10.1016/j.apergo.2005.02.001. http://www.sciencedirect.com/science/article/pii/S0003687005000554.

Salo, L. & Savioja, P. 2006. Practises of process control in digital control room: possibilities and threats. ECCE 13 – Trust and Control in Complex Socio-Technical Systems. Zurich. Proceedings of ECCE 2006.

Saurin, T.A. & Gonzalez, S.S. 2013. Assessing the compatibility of the management of standardized procedures with the complexity of a sociotechnical system: Case study of a control room in an oil refinery. Applied Ergonomics, Vol. 44, No 5, pp. 811–823. ISSN 0003-6870. doi: 10.1016/j.apergo.2013.02.003. http://www.sciencedirect.com/science/article/pii/S0003687013000264.

Schön, D. 1983. The Reflective Practitioner. New York: Basic Books.

Schraagen, J., Militello, L., Omerod, T. & Lipshitz, R. 2008. Naturalistic Decision Making and Macrocognition. Aldershot: Ashgate.

Sears, A. & Jacko, J.A. 2008. The Human-Computer Interaction Handbook. Fundamentals, Evolving Technologies and Emerging Applications. Second edition. New York, NY: Lawrence Erlbaum Associates.

Shattuck, L.G. & Miller, N.L. 2006. Extending Naturalistic Decision Making to Complex Organizations: A Dynamic Model of Situated Cognition. Organization Studies, Vol. 27, No. 7, pp. 989–1009. doi: 10.1177/0170840606065706. http://oss.sagepub.com/cgi/content/abstract/27/7/989.

Sheridan, T.B. 1992. Supervisory control. In: Anonymous Handbook of human factors and ergonomics. 2nd ed. Hoboken, New Jersey: John Wiley & Sons. Pp. 1295–1327.

Silva Junior, L.C., Borges, M.R.d.S. & Carvalho, P.V.R. 2012. Human factors evaluation in nuclear power plant control rooms using a mobile system to support collaborative observation. Progress in Nuclear Energy, Vol. 55, pp. 93–101. ISSN 0149-1970. doi: 10.1016/j.pnucene.2011.11.007". http://www.sciencedirect.com/science/article/pii/S0149197011002411.

Skjerve, A.B., Kaarstad, M., Størseth, F., Wærø, I. & Grøtan, T.O. 2011. Planning for resilient collaboration at a new petroleum installation – A case study of a coaching approach. Safety Science, Vol. 50, No. 10, pp. 1952–1959. ISSN 0925-7535. doi: 10.1016/j.ssci.2011.11.006.

Skjerve, A.B.M. & Skraaning, Gyrd Jr. 2004. The quality of human-automation cooperation in human-system interface for nuclear power plants. International Journal of Human-Computer Studies, Vol. 61, No. 5, pp. 649. ISSN 1071–5819. doi: 10.1016/j.ijhcs.2004.06.001. http://www.sciencedirect.com/science/article/pii/S1071581904000485.

Smith, K. & Hancock, P.A. 1995. Situation awareness is adaptive, externally directed consciousness. Human factors, Vol. 37, No. 1, pp. 137–148.

Sørensen, C. & Al-Taitoon, A. 2008. Organisational usability of mobile computing – Volatility and control in mobile foreign exchange trading. International Journal of Human-Computer Studies, Vol. 66, No. 12, pp. 916–929. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2008.07.002. http://www.sciencedirect.com/science/article/pii/S1071581908000785.

Stanton, N.A. & Ashleigh, M.J. 2000. A field study of team working in a new human supervisory control system. Ergonomics, 08/01, Vol. 43, No. 8, pp. 1190–1209. ISSN 0014-0139. http://ejournals.ebsco.com/direct.asp?ArticleID=A3JR4VEHFGDKJ1TKAK9N.

Stanton, N.A., Chambers, P.R.G. & Piggot, J. 2001. Situational awareness and safety. Safety Science, Vol. 39, pp. 189–204.

Stanton, N.A., Salmon, P.M., Walker, G.H. & Jenkins, D. 2009. Genotype and phenotype schemata as models of situation awareness in dynamic command and control teams. International Journal of Industrial Ergonomics, Vol. 39, No. 3, pp. 480–489. ISSN 0169-8141. doi: 10.1016/j.ergon.2008.10.003. http://www.sciencedirect.com/science/article/pii/S0169814108001595.

Stanton, N., Salmon, P., Jenkins, D. & Walker, G. 2010. Human Factors in the Design and Evaluation of Central Control Room Operations. Boca Raton, FL: CRC Press.

Suchman, L. 1987. Plans and Situated Actions: The Problem of Human Computer Interaction. Cambridge: Cambridge University Press.

Suomen Automaatioseura ry 2010. Valvomo – Suunnittelun periaatteet ja käytännöt (in Finnish).

Theureau, J., Filippi, G., Saliou, G. & Vermersch, P. 2001. A methodology for analyzing the dynamic collective organization of nuclear power plant operators in simulated accidental situations. The Cognitive Work Process: Automation and Interaction. Neubiberg, Germany. 307 p.

Theureau, J., Filippi, G., Saliou, G. & Vermersch, P. 2002. Cultural issues of nuclear power plant collective control in accidental situations and their impact upon design issues. Eleventh European Conference on Cognitive Ergonomics ECCE-11. Catania, Italy.

Upton, C., Doherty, G., Gleeson, F. & Sheridan, C. 2010. Designing decision support in an evolving sociotechnical enterprise. Cognition, Technology & Work, Vol. 12, No. 1, pp. 13–30. ISSN 1435-5558. doi: 10.1007/s10111-008-0124-1.

Vaitkunas-Kalita, S., Landry, S.J. & Yoo, H. 2011. Coincidence Between the Scientific and Folk Uses of the Term “Situation(al) Awareness” in Aviation Incident Reports. Journal of Cognitive Engineering and Decision Making, Vol. 5, No. 4, pp. 378-400. doi: 10.1177/1555343411424694. http://edm.sagepub.com/content/5/4/378.abstract.

van Aken, J.E. 2004. Management Research Based on the Paradigm of the Design Sciences: The Quest for Field-Tested and Grounded Technological Rules. Journal of Management Studies, Vol. 41, No. 2, pp. 219–246. ISSN 1467-6486. doi: 10.1111/j.1467-6486.2004.00430.x. http://dx.doi.org/10.1111/j.1467-6486.2004.00430.x.

Van Aken, J.E. 2005. Management Research as a Design Science: Articulating the Research Products of Mode 2 Knowledge Production in Management. British Journal of Management, Vol. 16, No. 1, pp. 19–36. ISSN 1467-8551. doi: 10.1111/j.1467-8551.2005.00437.x. http://dx.doi.org/10.1111/j.1467-8551.2005.00437.x.

Vicente, K.J. 1999. Cognitive Work Analysis. Toward a Safe, Productive, and Healthy Computer-Based Work. Mahwah, NJ: Lawrence Erlbaum Publishers.

Vicente, K.J., Roth, E. & Mumaw, R.J. 2001. How do operators monitor a complex, dynamic work domain? The impact of control room technology. International Journal of Human-Computer Studies, Vol. 54, pp. 831–356.

Vicente, K.J. 2004. The Human Factor. New York, NY: Routledge.

Vink, P. & van Eijk, D.J. 2007. The effect of a participative product design process on user performance. Safety Science, 6, Vol. 45, No. 5, pp. 567–577. ISSN 0925-7535. doi: 10.1016/j.ssci.2007.03.001.

von Wright, G.H. 1998. In the Shadow of Descarte. Essays in the Philosophy of Mind. Dordrecht: Klower Academic Publisher.

Vrazalic, L. 2004. Towards holistic human-computer interaction evaluation research and practice: development and validation of the distributed usability evaluation method. Doctor of philosophy ed. Wollongong University.

Vygotsky, L.S. 1978. Mind in Society. The Development of Higher Psychological Processes. Cambridge, MA: Harvard University Press.

Wahlström, B., Norros, L. & Reiman, L. 1992. Human Factors Research in the Nuclear Power Field in Finland. Proceedings of the IEEE Fifth Conference on Human Factors and Power Plants. Monterey, CA, USA, 7–11 June 1992. Pp. 537–542.

Walker, G.H., Stanton, N.A., Salmon, P.M. & Jenkins, D.P. 2009. An evolutionary approach to network enabled capability. International Journal of Industrial Ergonomics, Vol. 39, No. 2, pp. 303–312. ISSN 0169-8141. doi: 10.1016/j.ergon.2008.02.016. http://www.sciencedirect.com/science/article/pii/S0169814108000644.

Wickens, C. & Hollands, J. 2000. Engineering Psychology and Human Performance. Upper Saddle River, NJ: Prentice-Hall Inc.

Wilson, J.R. 1995. Solution ownership in participative work redesign: The case of a crane control room. International Journal of Industrial Ergonomics, Vol. 15, No. 5, pp. 329–344. ISSN 0169-8141. doi: 10.1016/0169-8141(94)00080-M. http://www.sciencedirect.com/science/article/pii/016981419400080M.

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

Wright, P., Dearden, A. & Fields, B. 2000. Function allocation: a perspective from studies of work practice. International Journal of Human-Computer Studies, Vol. 52, No. 2, pp. 335–355. ISSN 1071-5819. doi: 10.1006/ijhc.1999.0292. http://www.sciencedirect.com/science/article/pii/S1071581999902926.

Wright, P. & McCarthy, J. 2010. Experience-Centered Design: Designers, Users, and Communities in Dialogue. Morgan & Claypool. ISBN 9781608450442.

Yang, C., Hsieh, T., Lin, S., Lin, C.J., Teng, H. & Chiu, Y. 2011. Operators' signal-detection performance in video display unit monitoring tasks of the main control room. Safety Science, 12, Vol. 49, No. 10, pp. 1309–1313. ISSN 0925-7535. doi: 10.1016/j.ssci.2011.04.010.

Yang, C., Yang, L., Cheng, T., Jou, Y. & Chiou, S. 2012. Assessing mental workload and situation awareness in the evaluation of computerized procedures in the main control room. Nuclear Engineering and Design, Vol. 250, pp. 713–719. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2012.05.038. http://www.sciencedirect.com/science/article/pii/S0029549312003147.

Ylirisku, S., Halttunen, V., Nuojua, J. & Juustila, A. 2009. Framing design in the third paradigm. Proceedings of the 27th international conference on Human factors in computing systems. Boston, MA, USA, New York, NY, USA: ACM. CHI '09. Pp. 1131–1140. ISBN 978-1-60558-246-7. http://doi.acm.org/10.1145/1518701.1518874.

Zerger, B., Maqua, M. & Wattrelos, D. 2013. Modifications to nuclear power plants: A benefit which is at risk. An extensive analysis of the operating experience. Nuclear Engineering and Design, Vol. 256, pp. 141–147. ISSN 0029-5493. doi: 10.1016/j.nucengdes.2012.12.011. http://www.sciencedirect.com/science/article/pii/S002954931200605X.

PAPER I

Systems Usability Framework for Evaluating Tools in Safety-Critical Work

In: Cognition, Technology and Work 15,
pp. 255–275.

Paper I of this publication is not included in the PDF version.

PAPER II

Systems Usability Concerns in Hybrid Control Rooms

In: Proceedings of the 8th International Topical Meeting on Nuclear Plant Instrumentation and Control and Human-Machine Interface Technologies. San Diego, USA.

SYSTEMS USABILITY CONCERNS IN HYBRID CONTROL ROOMS

Paula Savioja, Iina Aaltonen, Hannu Karvonen, Hanna Koskinen, Jari Laarni, Marja Liinasuo, Leena Norros
VTT Technical Research Centre of Finland
P.O.Box 1000
FI-02044 VTT, Finland
paula.savioja@vtt.fi; iina.aaltonen@vtt.fi; hannu.karvonen@vtt.fi; hanna.koskinen@vtt.fi; jari.laarni@vtt.fi, marja.liinasuo@vtt.fi, leena.norros@vtt.fi

Leena Salo
Fortum, Power Division
P.O.Box 100
FI-00048 FORTUM, Finland
leena.salo@fortum.com

ABSTRACT

This paper presents findings from two large scale simulator test series which were carried out in the Finnish NPPs. Altogether 18 professional operating crews took part in the experiments and each of them conducted three scenarios at the training simulator of the respective plant. The studies provide evidence on how well-trained operating crews actually manage accident and incident situations in a hybrid control room. The tests were organized in order to have a solid reference in the possible future validation efforts regarding the human factors in the modernization of automation and control room systems. The systems usability (SU) evaluation framework has been developed to understand the role of tools in complex activity. In this paper, we describe the test results of these reference tests for the two plants: both, separately and by comparing the findings of the two cases using the SU framework.

Key Words: hybrid control room, evaluation, human factors, systems usability

1 INTRODUCTION

Digital user interface technology made its way to the main control rooms (MCR) of nuclear power plants (NPP) several decades ago. For instance, the process monitoring systems which analyze and present process data to be used by control room operators have been based on digital technology for a long time. Lately, control room upgrades have increased the significance of digital technology by introducing applications that exploit digital tools also in the interaction between the operator and the process. At the same time conventional solutions have also partly been maintained in the control rooms. Thus, the MCRs of both Finnish NPPs constitute now *a hybrid control room concept*.

A hybrid control room is a control room which simultaneously contains analogical and digital technologies. The hybridity can take many different forms; it is up to the modernization philosophy of the respective plants which systems or control room functions have been digitalized and which have remained in the original form. The primary motivation for the studies reported in this paper was originally the fact that not so much comprehensive systematically gathered data existed about the operating activity in a hybrid control room. On one hand there was evidence that operators are able, in their daily work, to cope with the hybridity of the control. This consideration is based simply on the fact that no major incidents had been reported to have been caused by the hybridity in the control room. On the other hand it can be hypothesized that the more hybridity is introduced in the form of new systems exploiting different generations of technology, the more confusing the overall situation will become from the operating point of view. Some anecdotal evidence from operators exists to suggest that operating so many different kinds of systems takes up too large portion of the cognitive resources. Also it had been thought that implementing evolutionary upgrades in the control room would only have an effect on the secondary tasks of the operators and thus the change from an operating point of view could be interpreted as minor. Secondary tasks comprise of user interface manipulations such as navigation. The scientific motivation for the studies presented here was to gain better understanding how the hybridity of the tools (the control room as an operating tool) has an effect on the operating practices. The practical motivation was to gather baseline data concerning operator activity in order to be used as a reference in later validation tests. Yet, another aim was a development of a comprehensive control room evaluation method which would highlight the role of tools in operating activity. In this paper we present two baseline evaluation studies which were conducted in the Finnish NPPs for the above mentioned purposes.

2 BACKGROUND

Acquiring reliable and valid information on the usability of control rooms in general and on hybrid ones in particular, is challenging due to at least two reasons. Firstly, operators tend to have divergent opinions on the usability of specific interface solutions [1]. For instance, some operators consider the monitoring and controlling of the processes to be easier in a digitalized control room as there the operator is able to perform all operations at the same display. However, others think that the fact of not being able to find directly some specific information from some specific display, opposite to the way information is found in the conventional control room, complicates operator’s work [1]. Both opinions are well grounded, based on experience and feeling of control. Consequently, it can be demanding to find an appropriate solution as during designing, it is advisable to take into account conceptions of operators, even if they can be opposite to each other.

Secondly, the very way information is expressed in the user interface is unique in each control room. Each control room is tailored and for safety reasons, not very detailed information is even allowed to publish. Furthermore, even if the conventional hard-wired control room may force the existence of some general features shared in practically all control rooms, digitalization brings along the possibility of designing very different solutions, resulting in even more unique control rooms. Thus, it is hard to acquire or develop information that would be valid for control rooms in general. This is also reflected in the scientific discussion. For instance, it can be maintained both that there is easily too much data available in a digitalized control room which makes it difficult to find the relevant information [2, 3] but also that digitalized user interfaces diminish the mental load resulting from searching for information[4] - depending on the user-interface solutions in the control room.

Some general findings have been, however, identified. Examples of information or some kind of principles or facts on a general level seem to be that digitalized solutions offer more possibilities to edit the ways information is expressed [5]; that changes in user interfaces affect work practices of the operators [2]; that efficient [6-9] or perhaps even standardized [10, 11] communication improves performance; and that operators move less in a digitalized control room [5, 12, 13]. The variability in control rooms, or simulator control rooms, presumably slightly diminishes when knowledge of the most usual drawbacks has reached all stakeholders and lessons from poor design solutions are learned by designers , and when the most urgent or easily-realizable improvements on the new solutions are made and the working practices are settled. To reach the most positive end result, i.e., efficient and usable control rooms in the nuclear domain, information should be delivered in such a general level that it is valid to as large number of control rooms as possible. This paper aims at delivering this kind of general knowledge about operating activity in a hybrid control room.

3 THE THEORETICAL APPROACH AND THE METHODS IN THE STUDIES

This paper does not provide full theoretical justification for the methods applied in the studies. The aim is only to briefly describe the SU evaluation framework which provides comprehensiveness in the assessment and the kinds of methods used in the studies.[1]

3.1 Systems Usability Evaluation Framework

The SU approach [14] assumes that a tool in an activity has three separate but intertwined functions: instrumental, psychological, and communicative. *The instrumental function* refers to the tool's ability to have an effect on the environment for which it has been designed e.g. a hammer's ability to make the nails penetrate into a surface or a control systems user interface's ability to allow the user to manipulate process components. *The psychological function* refers to the effect that the tool has on the user. In order to be able to use the tool the user creates mental models and schemas about the tool as such. Even further, in the psychological function the tool shapes the user's understanding about the controlled process and the physical phenomena taking place in the process. Thus, the tool affects the user's conceptualization of both the tool itself and the controlled environment. *The communicative function* refers to the effect that the tool has on the community of the users. It means that by using the joint tool the community adopts and shares vocabulary, knowledge and even values concerning the work. Most importantly, the communicative function stresses how meanings are conveyed by the tool in the community.

In evaluating SU (i.e. considering the fulfillment of the three above mentioned functions) the usage activity must be approached from multiple perspectives. The first perspective is concerned with the outcome of the performance. With outcome we mean concrete measurable results of activity which can be observed by an external observer. This perspective answers to the question of what happened during the course of activity. In a safety-critical context the outcome alone is not sufficient to describe the activity, thus the second perspective is concerned with the way of acting. The way of acting answers to the question of how, concerning the mechanisms producing the outcome of activity. The reason why way of acting is so important is the fact that since the CR operators are experts in their field, all the participating crews tend to reach the acceptable level of performance outcome at least in operational situations that are well trained. This means that more subtle means of analyzing the activity are needed to understand the internal quality of activity. Way of acting analyses the activity from the point of view of practice and its orientation to the core demands of the domain. The third perspective to activity is the user experience. This perspective is concerned with how the operating crews themselves view the technology in testing. We are interested in the qualities of experience and awareness that is accumulated in action, because it reveals inherent features of action that cannot be reached by observation from outside.

3.2 Data Collection Methods Utilized in the Studies

Similar data collection methods were utilized in both studies to be described in the following sections. Basically both studies were usage experiment simulator studies conducted within the training program of the operating crews of the respective plants.

3.2.1 Orientation interview

In the beginning of each experiment session the whole participating crew was interviewed in an orientation interview. These interviews were conducted individually for each operator. The orientation interview is a fairly short interview concerning the operator's personal epistemic attitude towards work and the controlled process [see more in 15]. Orientation interviews lasted from five to approximately 25 minutes and contained six defining questions concerning NPP process operator work. All the interviews were audio recorded, transcribed and analyzed qualitatively. The scale of orientations was: interpretive,

[1] A more detailed theoretical elaboration of the method is under preparation [18].

confirmative, and reactive [see more in 16]. An interpretive orientation emphasizes interpretation of the general rules of the domain always in the light of the particular situation at hand. Confirmative orientation emphasizes rules and norms. Reactive orientation reflects passivity towards work. The orientation interview results were utilized in analyzing the communicative function of the tool.

3.2.2 Observed process control activity in simulated accident and incident scenarios

The operating activity during simulated scenarios was video recorded with overview cameras and head mounted cameras. This data was later scored and analyzed both qualitatively and quantitatively. The observation data was utilized in order to construct a dynamic account of a crew's operating activity over time. The observation data was used in understanding all three functions of a tool.

3.2.3 Expert evaluation of crew performance

One of the simulator trainers acted as an expert performance evaluator in both plants. The expert had a pre-defined judgment scale which was defined separately for each scenario. For all the scenarios the scale contained the same themes: Detection, Diagnosis, Utilizing procedures, Stabilizing, and Co-operation. Each theme had several scenario relevant sub-measures. The expert observed the process control activity and simultaneously fulfilled the rating scale online. All the data were treated quantitatively and utilized in the analysis of instrumental function of the tool.

3.2.4 Task load assessment

Each operator's task load was measured utilizing the NASA TLX [17] measurement procedure after each separate simulator run. Task load data was utilized in the analysis of psychological function of the tool.

3.2.5 Process tracing interview

In order to gather the operating crew's conception about the simulated process control, a process tracing interview was conducted after each separate simulator run. In plant A the process tracing was conducted as a group interview for the whole operating crew and in plant B the procedure was conducted individually for each operator. This change in research procedure was a result from lessons learnt in plant A where the situation was such that typically one member of the crew took a leading role in answering the questions. In the process tracing interview the whole simulator run was reconstructed by asking the operators event by event what had happened in the run, what was the significance of each event, what control room features they had utilized in taking care of the event, and how they considered the adequacy of that specific control room feature. All the process tracing interviews were both audio and video recorded. The audio files were transcribed and the data treated qualitatively. The data was utilized in the analysis of psychological and communicative functions of the tool.

3.2.6 Systems usability questionnaire

In the end of each experiment session the operators individually filled in a systems usability questionnaire. The questionnaire contains 49 positive statements regarding the control room. The statements have been constructed with the aid of systems usability approach i.e. to consider all three functions of the tool. This means that there were separate statement sets for each instrumental, psychological, and communicative function. All the statements were formulated so that they posed a positive assessment of the control room and the operators were instructed to score each statement on a four point scale ranging from completely agree to completely disagree. The basic assumption behind the statements was that generally quite positive result should be obtained as the evaluation was conducted on a control room which is in daily use at both plants. All the data was treated statistically and utilized in the analysis of all three tool functions.

4 THE STUDY IN PLANT A

4.1 The Specifics of Plant A

Plant A is a two-unit VVER plant originating from the late 1970's and producing currently close to 500MW electrical power in each unit. The concept of operation is such that in a main control room (MCR) (one per unit) a standard operating crew consisting of three operators (turbine, reactor, and supervisor) operates the plant in normal conditions. In case of abnormal conditions a safety engineer joins the normal crew. The MCR is a hybrid composition of digital and analogue user interfaces. Process monitoring system constitutes perhaps the most important view to the process and is based on digital information presentation. All operations, except maneuvering of control rods, are conducted via original hard wired panels and operating desks. The emergency operating procedures (EOPs) are utilized in paper format and are in flow chart format. The EOPs had recently been renewed from event based at the time of the study. The control room and more widely the control concept is a product of extensive in-house development effort within the company. During the study the modernization (digitalization) of the MCR had been started.

The specific research questions concerning plant A was to formulate a performance based reference which could be utilized in the validation of new control room design. The plant had extensive plans on modernizing the control room in the near future. An additional research question was related to the usage of the recently renewed EOPs.

4.2 The Scenarios

Three different scenarios were developed in the training organization of plant A. The scenarios at plant A were: 1. Loss of coolant accident, 2. Primary secondary leak, 3. Electrical bus bar failure. Scenarios 1 and 2 are so-called design basis accidents, which means that specific emergency operating procedures cover the necessary process interventions which operators are required to carry out. An additional plant protection signal failure was added in the scenario 1 to increase complexity. In scenario 3 it is not evident which operating procedure would best suit the situation and the event is not part of the normal training program of the crews.

The scenarios were analyzed carefully by the researchers and the simulator trainer who acted as an expert evaluator of performance in order to formulate the evaluation bases for performance evaluation and the evaluation of operating practices.

4.3 The Participants

All twelve operating crews of the plant participated in the study which means that altogether 46 operators acted as users in the experiment. In general an operating crew consists of three licensed operators, but in many of the crews there were trainees who also took part in the simulator exercises. The operating experience of the operators in plant A varied from 1 to 32 years of experience. There were 18 participants in the experience group 1 - 9 years, 13 participants in the experience group 10 - 19 years, and 13 participants in the experience group over 19 years.

4.4 The Results Concerning Plant A

The data analysis was conducted by the researchers and reported back to the operating unit as part of operators' yearly class room training. In this paper we concentrate mainly on those results showing difficulties in the operating practices and which seem to be caused by the hybridity of the control room solution.

The performance of the crews (expert rating) was overall good (mean values between 3 and 4 out of 5), and the performance differences between crews were not statistically significant. Based on the results of the expert performance rating it can be said that the performance outcome in each scenario was on a satisfactory level. In this conclusion the measure of satisfactory level performance is that of not endangering safety e.g. completing the requirements expressed in the emergency operating procedures.

In the two accident situations (scenarios 1 and 2) the crews' task loads were overall lower than in the electric bus bar system failure (Fig. 1). For the accidents emergency operating procedures exists but for the electric bus bar failure there is no one specific procedure. The electric bus bar failure scenario as such was not as severe as the accident situations but the task loads were still higher. The scenario's effect on each task load factor except frustration was statistically significant ($p<0,05$) or very significant ($p<0,001$).

In the analysis of operating practices there were some differences between the crews. Differences existed in the amounts of communication and movement around the control room, especially for the shift supervisor. Also there were differences in the extent of using and communicating of process information for decision making. One remarkable difference between the crews in operating practice was found. In scenario 1 (loss of coolant, LOCA) there was a small additional failure. One of the plant protection signals (containment isolation) did not function correctly. The checking of the protection signals is not in the beginning of the LOCA procedure, but nevertheless, the operators, if they are aware of the endangered safety functions, should be aware of the status of the containment isolation. Only one crew managed to notice this failure and take the measure to correct the situation before it was mentioned in procedure. This was considered to be variation in the operating practice.

There were qualitative differences in the orientations of the operators. When all the answers of all the operators were coded 29% reflected an interpretive orientation, 48% a confirmative orientation, and 23% a reactive orientation. The result is not poor as such, but among professional operators with quite long experience in process control work we might expect the interpretive orientation to be more common. As the confirmative orientation is by far the dominant it can be concluded that operators tend to maintain an attitude according to which process control work is confirmatory activity in which it is enough to follow and obey the rules and procedures.

According to the responses of the systems usability questionnaire, the operators experienced the usability and functionality to be more positive than negative. 31% of the positive statements were completely agreed with, 56% somewhat agreed with (Fig. 2). The most negative experiences concerned complexity caused to the operators' work by the user interface, error possibilities and especially recovery from errors in the user interface, difficulty to learn to use the systems, support for finding right operative solution in an unclear situation, help of the procedure in understanding process situation, support for personal styles, and support for adaptive activity.

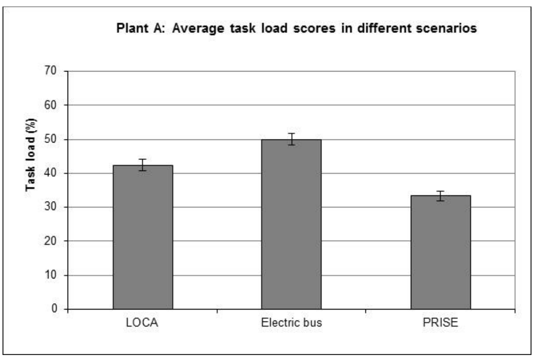

Figure 1. Task load scores in different scenarios. LOCA refers to loss of coolant accident and PRISE to primary-secondary leak, both of which are accidents. Electrical bus bar failure scenario resulted in higher task load values than the other two ($p<0.05$).

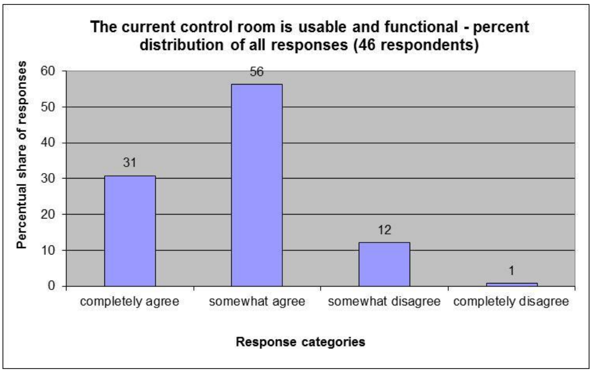

Figure 2. Summary of all answers in the systems usability questionnaire (Plant A).

5 THE STUDY IN PLANT B

5.1 The Specifics of Plant B

Plant B is a two-unit BWR plant originating from the turn of 1980s producing close to 1700MW electrical power. The concept of operation is similar to the plant A as there are three main operators in a normal crew (turbine operator, reactor operator, and shift supervisor). One significant difference is that in addition to the three MCR operators one extra person is almost always present in the control room: a supervisor for the field operators. This supervisor often even has a license of a turbine operator and depending on the level of expertise s/he might take part in the turbine operations or the overall problem solving conducted by the crew.

The MCRs of plant B are truly hybrid in nature. The turbine side operating interface has been digitalized some years ago and thus all turbine operations (excluding some manual back-up) are conducted via soft control methods. At the same time all primary side operations are conducted in traditional user interface. In plant B there is also a process monitoring system based on digital information presentation which is used by both turbine and reactor operators. In the turbine side modernized user interface there are embedded many kinds of additional information for the operator e.g. procedures and information concerning automation (logic diagrams).

The specific questions of plant B were to gather operating experience of the current control room solution as there had not been comprehensive performance based studies conducted since the non-safety related turbine side user interface modernization. At the same time a reference data might be needed some time in the future if further control room modernizations would at some point be needed.

5.2 The Scenarios

In plant B the scenarios were developed with the help of the simulator training organization and the specific questions mentioned above in mind. The scenarios were decided to be smaller-scale incidents which, if not responded adequately by the operating crew, would lead to automatic scram and of course to production losses. The scenarios were: 1. Failure in decay heat removal system; 2. An ejector failure; and 3. An automation failure in a pre-heater line. Scenarios 2 and 3 were combined so that they were represented as one continuous flow of events. In comparison to plant A the plant B was not interested in the utilization of EOPs and thus specific accident scenarios were not chosen.

5.3 The Participants

In plant B six operating crews participated in the study. Altogether there were 24 participants and the operating experience varied from 0 to 31 years. 12 operators had operating experience of 0 – 9 years, 3 operators had operating experience of 10 – 19 years, and 9 operators had operating experience exceeding 19 years.

5.4 The Results Concerning Plant B

In plant B the overall performance ratings of the expert varied between 2 and 5, but the average ratings were quite good varying between 3.5 and 3.9 (out of five). The lowest ratings were given for detection and use of procedures.

In plant B the scenario 3 was by far the most difficult for the operators to handle. It involved a failure in the automation system which was very difficult for the operators to discriminate from a process failure. The new digital interface would have given the operators an opportunity to detect that it was indeed an automation failure. However, only one crew was able to make this detection and thus avoid an unnecessary reactor scram. In the further analysis of the operating practice of this one crew it was detected that the amount of communication in this crew was higher than in the other crews. The same difference was reflected in the time that the crew spent physically together, communicating and interpreting information provided in the user interfaces. This finding suggests that this one crew had developed superior collaboration practices.

The effect of the user interface on operating practice was evident in the amount of movements of the operators. The turbine operators who use the digitalized user interface moved less than other operators. Another notable result of the analysis of movements was that of total time of movement of the shift supervisor. The shift supervisor was away from his own station more than either of the two other operators in scenarios 2 and 3 which both involved more the turbine side (digital user interface) of the process.

There were quite large differences in the task loads when compared between the different operator roles (Fig. 3). The turbine operators reported the highest loads in all subscales. The effect of the operator role was statistically significant for subscale frustration ($p<0,01$): The turbine operators were significantly more frustrated than reactor operators and shift supervisors. In addition the effect of the operator role was statistically nearly significant concerning physical demand ($0,1 < p < 0,05$): The turbine operators considered the operating conditions physically more demanding than reactor operators. This is an interesting finding considering that the turbine operators were the ones utilising the modernised user interface.

In the orientation of the operators towards the process there were also differences. When all the answers were coded and the results calculated, 23% of answers reflected an interpretive orientation, 58% a confirmative orientation, and 19% reactive orientation. The confirmative orientation is by far the dominant which reflects the prevailing attitude of understanding process control work as confirmatory activity in which it is enough to follow and obey the rules and procedures.

In answering the usability questionnaire the operators of plant B perceived the current control room quite usable (Fig 4). Concerning the usability and functionality 35% totally agreed with the positive statements, 52% somewhat agreed, 12% somewhat disagreed and 1% totally disagreed. Thus majority of the conceptions lay on the positive side. Also the answers of the turbine operators, who have the modernized user interface in use contained more completely agree answers than those of reactor operators and shift supervisors (40% vs. 27% and 30%).

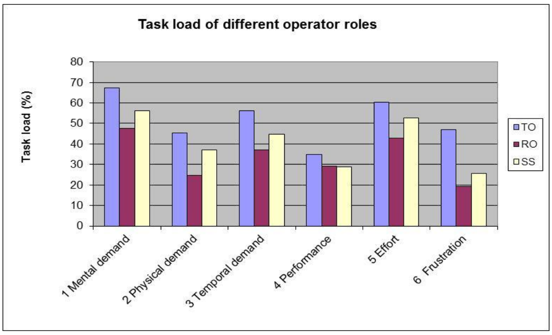

Figure 3. Task load of different operator roles, all scenarios combined. TO = turbine operator, RO = reactor operator, SS = Shift supervisor. The turbine operators have the highest task loads.

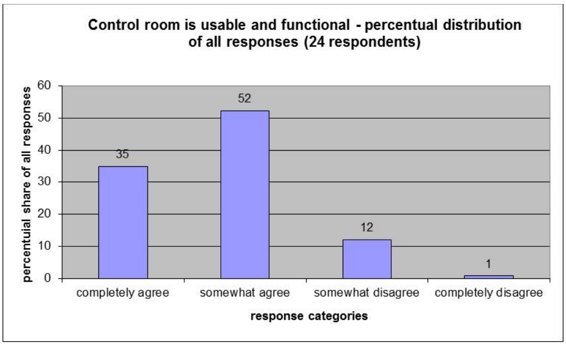

Figure 4. Summary of questionnaire answers in plant B.

In the open comments and process tracing interviews some critique on the user interface was nevertheless presented. The main problem was that the needed information is spread all around the control room. Also, the information monitor might be far from the operating interface. Also alarm system presents itself as problematic to the operators. Alarms are too many, and the information value is not high enough especially in the situations when it would be needed.

6 THE RESULTS COMBINED: SYSTEMS USABILITY

6.1 The Instrumental Function

In the instrumental function both of the evaluated control rooms worked quite well. This means that the operating crews were able to effectively carry out the control operations considered necessary in the simulated accident and incident situations. This issue is featured in relatively good operator performance results obtained in both plants from expert judgment. As in plant B there were some low performance scores obtained (detection, use of procedures) this may reflect the hybridity as different operators utilized different type tools. Generally the results from systems usability questionnaire were also on quite good level. When the statements concerning instrumental function were analyzed separately for both plants it was found out that the percentage distribution of answers was such that in plant A 35% of the statements were responded with completely agree and in plant B 30%. Some problematic issues concerning instrumental function were however identified. The most problematic issues concerning instrumental function were problems with feedback in a digital touch screen solution (in plant A). The operators stated that it is not always clear whether a command has been implemented. Another problematic issue was the possibility of errors in using especially the digital interface. The recovery from errors was in addition experienced problematic. Also the operators claimed that the systems in the control room do not help in focusing on the relevant phenomena when support in that is needed. These issues were common in both plants, and as both constitute a hybrid solution it may be one of the causes of the problems.

6.2 The Psychological Function

Although the psychological function was also supported relatively well it was the most problematic of the three general tool functions. In the questionnaire data only 25% in plant A and 28% in plant B completely agreed with the positive statements concerning psychological function. The learnability of the control room systems was considered in both plants not to be on an optimal level. Even though the control room is a complex system and thus cannot be effortlessly learnt to use, the operators made statements that in fact it is confusing to have different information presentation principles and notations within the systems of the same control room. The learnability is affected by hybridity as it induces more learning challenges in the form of multiple different systems utilizing different interfaces and operating logics.

In plant B where there were more digital user interfaces in use, the understanding of the automation information provided in the interface was not completely in an optimal level. Only one crew was able to solve an automation related problem in scenario 3 and was thus able to avoid production loss related to an automatic scram. In the analysis it was found out that the problem was related to the psychological function because the other crews' abilities to utilize the system were not on the same level with the system capabilities. This is related to hybridity as it might polarize the skill levels between the different operating crews. In plant A there were differences on micro-level in the crews' operating practices in proceduralised scenarios. We considered this also as a problem in the psychological function of the tool: Not all crews had been able to create higher levels of way of acting which would be manifested in the crews mitigating actions concerning the core safety functions of the plant.

A positive issue concerning psychological function was the new state-based EOPs in plant A. They truly seemed to lessen the task load of the users.

6.3 The Communicative Function

Overall, the communicative function lies somewhere between the psychological and the instrumental in terms of successfulness. In plant A 29% of the questionnaire answers were completely agreed upon and in plant B 44%. Most problematic communicative features according to the questionnaire data concerned the interpretation of the utilized sounds in the control room, operator support in unclear situations, and EOPs ability to find new solutions. Also the fact that operator orientations lay mostly in the confirmatory category was considered to be a problem of the communicative features of the control room. The tool should support more firmly an operating attitude which sees process control work as a higher level functioning the aim of which is to ensure nuclear safety. The differences in the collaboration practices (movements, communication, time spent together) show weakness in the communicative function of the tool: Not all crews are able to collaborate effectively and efficiently in the control room.

7 CONCLUSIONS CONCERNING HYBRIDITY

Although both studied MCRs constitute a unique case, some general conclusions can be drawn concerning hybrid control rooms in general. After a careful analysis we considered hybridity not to be problematic from the operating point of view per se (operators seem to cope well), but it is also evident that some of the problems in the user interface and the human performance may be caused by the hybridity of the control room.

The problems in instrumental level might be caused by hybridity. The detection problems in plant B and problems in use of procedures are areas in operator work which are heavily affected by the tool that is used. Also in plant B it was witnessed that turbine operators utilizing the modernized interface were more satisfied with the tool. This is a hybridity related concern as it might cause polarization in operating practices.

It was evident in both plants that the psychological function of the control room was not on an optimal level. This is an indication that hybridity causes confusion in the operating work. This means that more careful attention should be paid both to the consistency throughout the different systems (in overall operation logic) and to the training processes. We claim that in a well-functioning psychological tool there would not be inconsistencies in the operating practices between the crews because the functionalities would be such that they could be exploited by all the crews to the same extent. In the current solutions there were tendencies in the operation practices to break into two opposing groups. This is a weakness on the overall level because in the long run it means variation in the performance outcome also.

The communicative function is reflected in the ways operating crews collaborate in the control room and also in the ways the control room systems manage to convey meaningful information for operations. Based on the findings in the reported usage experiments the communicative functions of both studied control rooms could be improved. The improvement needs were evident in the results of the orientation interview results of both plants. The dominant orientation was confirmative which reflects an attitude towards work according to which process control work is confirmatory in nature. This means that the benefit of hybridity, the variety of information presentation formats and operating interfaces has not been exploited by the operators. Bu rather, it works against the development of the interpretative orientations.

In this paper we have presented two simulator studies which have been conducted in Finnish NPPs. The results show that hybridity as such is not extremely problematic but it poses new demands on processes of training and learning. These problems may not become evident in the performance outcome of operating activity but rather are visible on the level of operating practices. Also consistency concerning different features of the control room becomes easily compromised within a hybrid control room. As a solution for these problems we propose more continuous monitoring (evaluations) of the prevailing operating practices and the practice point of view to be exploited also in the design of new control room features.

8 ACKNOWLEDGMENTS

This paper was prepared under the National Nuclear Power Plant Safety Research 2007–2010 (SAFIR) funded mainly by the Finnish nuclear companies. The authors would like to express their gratitude to every participant of the studies.

9 REFERENCES

1. F. Huang and S.-L. Hwang, "Experimental studies of computerized procedures and team size in nuclear power plant operations," *Nuclear Engineering and Design*, pp. 373-380, (2009).
2. M. Kaarstad, E. Nystad and S. Strand, "Work practices and cooperation in a near future and far future operational environment," in Simulator-Based Human Factors Studies Across 25 Years, A. B. Skjerve and A. Bye, Eds. London: Springer, (2011).
3. J. O'Hara and J. Persensky, "Human performance and plant safety performance. establishing a technical basis and framework for evaluating new human-system interfaces," in Simulator-Based Human Factors Studies Across 25 Years, A. B. Skjerve and A. Bye, Eds. London: Springer, (2011).
4. C. W. Yang, T. -. Yenn and C. J. Lin, "Assessing team workload under automation based on a subjective performance measure," *Safety Science.* **vol. 48**, pp. 914-920, (2010).

5. K. Vicente, E. Roth and R. Mumaw, "How do operators monitor a complex, dynamic work domain? The impact of control room technology," *Int. J. Human-Computer S.*, **vol. 54**, pp. 831-356, (2001).
6. E. Roth and J. O'Hara, "Integrating digital and conventional human-system interfaces: Lessons learned from a control room modernization program," Tech. Rep. NUREG/CR-6749, (2002).
7. D. S. Morriseau, "Hybrid control rooms: The effects of introducing new technology into existing control rooms," Halden Reactor Project, Tech. Rep. HWR, (2001).
8. D. Min, Y. H. Chung and W. C. Yoon, "Comparative analysis of communication at main control rooms of nuclear power plants," in *Proceedings of IFAC/IFIP/IFORS/IEA Symposium*. (2004)
9. P. O. Braarud, H. Broberg and S. Massaiu, "Performance shaping factors and masking experiment 2006: Project status and planned analysis," in *Proceedings of the EHPG*, Storefjell Norway, (2007)
10. M. C. Kim, J. Park, W. Jung, H. Kim and Y. J. Kim, "Development of a standard communication protocol for an emergency situation management in nuclear power plants," *Annals of Nuclear Energy*, **vol. 37**, pp. 888-893, (2010).
11. S. Kim, J. Park and Y. J. Kim, "Some insights about the characteristics of communications observed from the off-normal conditions of nuclear power plants," *Human Factors and Ergonomics in Manufacturing & Service Industries*, **vol. 21**, pp. 361-378, (2011).
12. M. Kaarstad, S. Strand and C. Nihlwing, "Work practices in computerbased control rooms – insights from workshops with operators," Halden Reactor Project, Tech. Rep. HWR, (2008).
13. A. Sebok, "Team performance in process control: influences of interface design and staffing levels," *Ergonomics*, **vol. 43**, pp. 1210-1236, (2000).
14. P. Savioja and L. Norros, "Systems usability - promoting core-task oriented work practices," in *Maturing Usability. Quality in Software, Interaction and Value*, E. Law, E. Hvannberg and G. Cockton, Eds. London: Springer, (2008).
15. L. Norros, *Acting Under Uncertainty.* Espoo: VTT, (2004).
16. L. Norros, M. Liinasuo and P. Savioja, "Operators' conceptions of procedure guidance in NPP process control," in *Proceedings of Enlarged Halden Project Group Meeting*, (2011).
17. S. G. Hart and L. E. Staveland, "Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. Amsterdam: Elsevier, (1988).
18. P. Savioja and L. Norros, "Systems Usability Framework for Evaluating Tools in Safety-Critical Work" (accepted for publication in Cognition Technology & Work).

PAPER III

Systems Usability

Promoting Core-Task Oriented Work Practices in Complex Work

In: Maturing Usability: Quality in Software, Interaction, and Value. Pp. 123–143.

Paper III of this publication is not included in the PDF version.

PAPER IV

Functional Situation Models in Analyses of Operating Practices in Complex Work

In: Proceedings of ECCE 2012
Re-Thinking Cognition. Pp. 58–64.

Functional Situation Models in Analyses of Operating Practices in Complex Work

Paula Savioja
VTT Technical Research Centre of Finland
P.O.Box 1000
FI-02044 VTT
Finland
Paula.Savioja@vtt.fi

Leena Norros
VTT Technical Research Centre of Finland
P.O.Box 1000
FI-02044 VTT
Finland
Leena.Norros@vtt.fi

Leena Salo
Fortum,
Power Division
P.O.Box 100
FI-00048 FORTUM
Finland
Leena.Salo@fortum.com

ABSTRACT

Motivation – In safety critical work it is common to evaluate human activity based on the concrete outcomes it produces. But, in order to understand more thoroughly the possible implications for safety, also profound perspectives concerning the mechanisms producing the outcome are needed. In this paper we introduce a model of control situations that connects human actions with the purposes rising from the domain. This model, labelled functional situation model (FSM) enables analysis of operating activity from the perspective of way of acting i.e. work practice. Analysis of work practices complements the analysis of outcome of activity (e.g. task completeness, errors, time). The aim is to promote adoption of resilient work practices by analysing which ways of acting in a given situation are aiming for the general objective of safety.

Research approach – Research approach is constructive: a formative modelling technique has been created which draws from theoretical roots of functional domain modelling. The exploitation of the models in analyses of operating activity draws from the pragmatist conception of habit.

Design – A FSM denotes a control situation from the point of view of critical functions which are endangered in a situation. The human actions are also depicted in the model, and connected to the critical functions which are aimed to maintain.

Implications – The practical implication of an FSM is that it enables analyses (and evaluation) of operating practices and characterisation of them according to how they take the critical functions and the general objectives of the domain into account.

Take away message –Resilience in operating practice assumes that actors are able to make the connection between situational goals of actions and the general objectives of the domain. FSM makes this connection explicit and thus enables analyses of resilience features in practices.

Keywords

Functional situation model, work practice, resilience

INTRODUCTION

This paper addresses the issue of situation models in design and evaluation of complex systems in safety critical work. When a new tool is tested (e.g. a control room system in an industrial plant) it is typically put into an evaluation in some particular situation. For example in the safety critical nuclear industry, complex scenarios are run in high fidelity simulators in order to evaluate whether the integrated system, the operating crew and the technological systems function adequately. In this type of test it is usually found important to carry out an analysis of the test situation (the simulated scenario) prior to the simulator run in order to understand what is the desired behaviour of the whole system. The aim of the analysis is to understand the characteristics of good operating performance in the given situation. The model of the operating performance is the result of this analysis and it depicts the expected activity in the situation.

The analysis of operating performance demands can be carried out referring to the results of task analysis, for example, a hierarchical task analysis (HTA). HTA distinguishes task hierarchies by recognising which sub-tasks belong together and which kinds of higher level tasks comprise of the lower level tasks. The problem of hierarchical task models is that they do not explicitly connect the human activity to the phenomena in the environment which is a driver for the activity. Thus, a hierarchical task model is not able to portray the contextual significance of the actions modelled.

We maintain that a model of operating performance is absolutely necessary. But, this model should be such that in it the role of environmental features to the organisation of performance is acknowledged. Approaches that consider human acting as being situated (Suchman, 1987), embodied (Dourish, 2001), distributed (Hutchins, 1995), and socially constructed and driven by meanings that a technology is able to make the users experience (McCarthy & Wright, 2004), are examples of approaches that fulfil this requirement. In addition, we agree with hermeneutic approaches which acknowledge that researchers as external observers always bring a pre-understanding of the situation with them. The modelling aims at making this pre-condition explicit. By making it explicit it also becomes possible to realise, during the actual analyses, the incompleteness of the initial model. Utilisation of the model in the analyses of operating practices is a process of interpretation in which the model of the situation and the researchers' understanding of the realised practice co-evolve dialogically.

In this paper we present a situation model labelled functional situation model (FSM) which draws from different theoretical roots. FSM has been developed in order to analyse and understand situational work practices in safety-critical industries.

BACKGROUND

The main inspiration for FSM approach is the Rasmussen's (Rasmussen, 1986) control domain modelling, the abstraction hierarchy (AH). AH represents the controlled system and its environment at multiple levels of *means-end* abstractions. This is fundamentally different from HTA in which the relationship of subsequent levels is *part-whole*. In AH the different levels constitute a qualitatively different abstraction of the system. In AH the specific view point is *constraints* governing the functioning of the system.

AH is a general model of the domain. For the purpose of understanding demands of activity in *particular situations*, specific models of situations are needed. Flach et al. (2004) present a theory- based depiction how Rasmussen's levels of abstraction could be used to describe resources within situations. Petersen (Petersen, 2004) has also developed modelling in order to understand control situations for the purpose of design of human-machine systems. Later Bjørkli et al (2007) have added a level of strategy to Petersen's model based on empirical work in maritime environment.

The development of the FSM modelling approach has been conducted in different research projects in NPP and other safety critical industries. In evaluating quality of human technology interaction it is evident that some kind of model of human activity is needed in order to understand the demands concerning it. The first studies in which we utilized functional modelling of situations are reported in an overview by Norros (Norros, 2004; see alsoNorros & Nuutinen, 2005). These were studies of human error, process information presentation, and validation of an alarm handling system conducted in the 1980's and 90's. The modelling technique utilised in those studies, aimed at explicating the information concerning the environment (the controlled process) and optional operations which was required to adequately handle the process situation. Later similar situational information requirement modelling was utilised in the development of operator training (Norros, 2004). Already early on, the situation modelling method aimed at describing the characteristics of the environment which afford activity from the operators' part. In the analyses of operating *practices* this view point is extremely important. What differentiates practice from action is the connection to the *meaning*. The meaning and purpose the action strives for constitute its contextual value.

In the next sections we present the functional situation modelling technique as it has been recently applied (Savioja & Norros, accepted for publication) in the analysis of operating practices.

FUNCTIONAL SITUATION MODELS

When new technology is introduced as a tool in safety critical work, it is necessary to understand the implications on the work from multiple perspectives and levels of abstraction. It has been outlined that the operating activity shall be examined from three perspectives: performance, way of acting, and experience (Savioja & Norros, accepted for publication). Performance evaluation refers to outcome oriented perspective to human activity and a reference for it can be for example a HTA model of required operator actions. But, in order to understand the way of acting (work practice) a different reference is required which connects the actions and their meaning in the specific situational context. For this purpose we present the FSM approach. A general schema of a FSM is presented below (Figure 1).

A FSM combines two viewpoints to the modelled situation; chronological and functional. These two dimensions define a two-dimensional space in which the most important operator actions and process phenomena are mapped for the purpose of understanding both operating performance and practice.

Chronological View to the Situation

The chronological view is the most obvious way to analyse the operating activity. In the chronological view the scenario is divided roughly into different phases in which the operating actions have a specific goal. In the FSM the chronological phases are labelled according to the goals: Detection, mitigation of effect, diagnosis, and stabilising the process state.

Although the distinction is made between the phases it must be acknowledged that in real activity all the goals are somehow present simultaneously. But for the purpose of making sense of the process situation the distinction is nevertheless made in FSM modelling. Also, in a different circumstance the phases might have different goals, thus different labels for the phases could be utilised.

Detection

In the model, detection phase denotes that the crew identifies some process events requiring operator actions. The process information presented to the operators is typically alarm information and notifications. Simultaneously, monitoring of all process information is conducted by the crew.

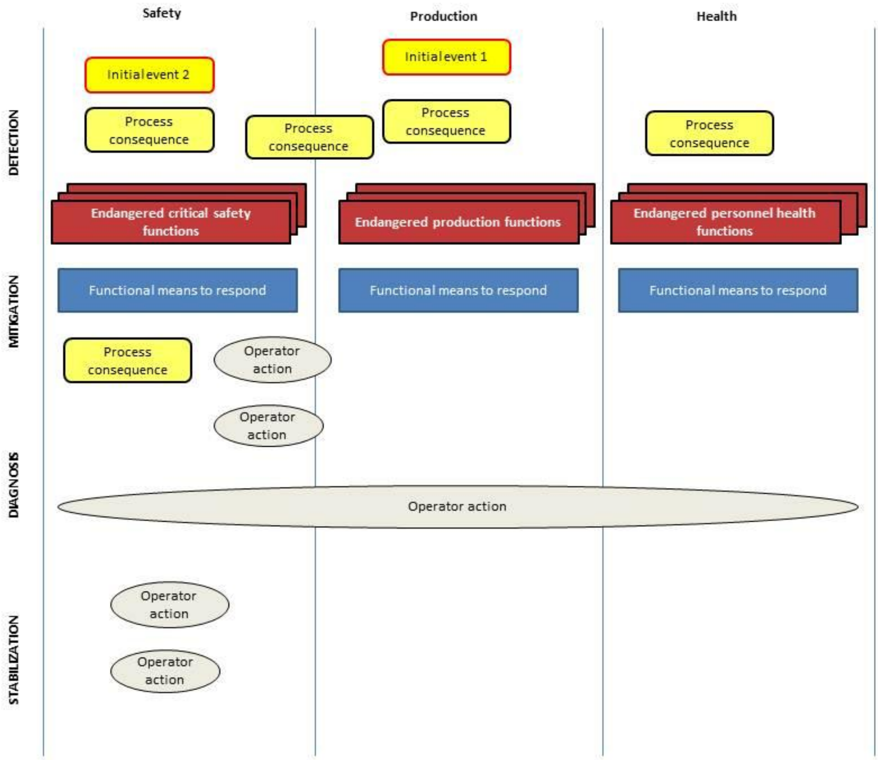

Figure 1. The general schema of a FSM. In modelling a scenario with an FSM, the boxes describing process events and information (yellow), critical functions (red), and operating crew actions (grey) are depicted connected to the critical functions (read: safety, production, health) and the operative means to mitigate the functions (blue). The description is organised according to the chronological phases (detection, mitigation, diagnosing, and stabilization) according to the scenario in question.

In the FSM the detection phase is concerned with the information that is available for the operators to understand the deviations in the process. It is also important to identify and explicate the initial event(s). The most important alarms informing the operators are included in the model. The implications of the initial event on the process state are depicted on parameter and overall process level.

Mitigation of the effect

The line between mitigation and detection is not always distinct as the operating activity actually constitutes a continuous cycle of monitoring and acting. Thus, the exact point in time when the mitigation phase "starts" is not so important, but we want to make a distinction that some operating actions are tuned more towards mitigating the situation than perceiving information about the situation. It is also typical that automatic functions handle some of the actions in this phase. In this case important operator actions are to confirm the fulfilment of the automatic functions. In the mitigation phase of an FSM the operating actions that mitigate the process situation are mapped under the specific process information and initial events of the detection phase that they are connected to.

Diagnosis

As the ultimate operating goal in an accident situation is to bring the process into a safe stabile state, diagnosis related actions are required from the operating crew. It is important to realise what the process situations is, in order to identify the required actions. In the diagnosis phase these actions are depicted in the model under the specific parameters of detection phase that they are related to.

Stabilisation

Stabilisation phase refers to the operating activities which aim at bringing the process into a safe and stable state. These actions are connected to the relevant initial events also.

Functional View to the Situation

A functioning sociotechnical system has the objectives of production, safety, and health (Vicente, 1999). These objectives form the basis of the functional view of the FSM (the three “lanes” in Figure 1).

The most important items in the functional view are the critical functions (see e.g. Norros, 2004 p.111) of the process which are endangered in the specific situation. These functions can be, depending on the initial events, related to safety, production, or health. Typically, in a complex situation there are critical functions related to all of the above objectives which are endangered.

Operating activity, on a high level, is oriented towards maintaining the critical functions of the process. Thus, the operating actions required in the situation can be collated under general means to respond to the endangerment of the critical functions. This is depicted in the FSM by presenting a specific functional level in the model “functional means to respond”. This level connects the individual operating actions to the critical functions. The relationship with critical function level upward and with individual actions downward is of type *means-ends*.

Connecting the Chronological to the Functional

By connecting the individual level actions of the crew and the functional view to the situation it is made explicit what is the *meaning* of each action in the wider context of the scenario. This is the important aspect of an FSM which enables the analysis of operating activity on the level of *practice*. We maintain that making this connection in action, is actually what makes operating practice *resilient* in a situation.

If and when operating crew is tuned towards the critical functions of the domain they always take into account them in their actions and follow how the critical functions behave and are affected during the course of complex process control activity. This is a characteristic of resilient practice. Resilience means that activity is able to respond to external events that may not have been rehearsed or pre-planned. We maintain that in order for the operating crew to be able to handle all possible situations they must make the connection always, even in a situation which is well rehearsed or planned.

In order to identify the critical functions in a situation it is important, first, to analyse the domain on a general level. We therefore consider the domain from the point of view of three general characteristics. These are the dynamicity, complexity and uncertainty of the system (DCU characteristics). The core task of a work accomplished in the particular work domain composes of taking these characteristics into account in all situations. Coping with the DCU characteristics requires that the actors are capable and willing to mobilise resources related to skills, knowledge and collaboration. (Norros 2004). The central reference of appropriate practice is that the DCU features are taken into account. *Resilience* emerges from the ability of the actors to tackle the DCU features of a *particular situation*, and thus to fulfil the core task. Hence, in Functional Situation Modelling the DCU features need to be considered, and the practices of tackling the situation are assessed with reference to managing the core-task demands.

UTILISING FSMS IN ANALYSES OF OPERATING PRACTICE

As stated above in this paper, earlier versions of functional models of operating situations have been utilised in analysis of human errors, alarm system designs, control strategies etc. This particular form of FSM which is introduced in this paper has been utilised in the evaluation of control room adequacy from the point of view of *systems usability*, concept that we have created to define the quality of complex technologies from a holistic point of view (Savioja and Norros, accepted for publication).

In the study, a series of reference tests was carried out in a nuclear power plant full scope training simulator prior to main control room modification. The aim was to gather evidence of the current control room functioning in a way that comparisons could be made with the new design solutions. One view point in these evaluations was the control room’s ability to support and promote core-task oriented work practice. Altogether 12 operating crews took part in the usage experiment and conducted each three simulated scenarios producing over hundreds of hours of video material.

The FSMs had multiple other roles in the test series in addition to enabling evaluation of operating practices. The FSM, first of all, provided the researchers an excellent opportunity to understand what is going on in the complex process system, and how the operating crew is expected to manage the situations. Second of all, as the FSM depicts the most important process events it guides focusing of the detail level analysis of operating activity.

Construction of an FSM

In the experiment series the most detailed scenario model was constructed for a loss of coolant accident (LOCA) which is a design bases accident typically well-rehearsed by the professional operating crews. The particular FSM is depicted in below (Figure 2).

As the experiment was conducted under the yearly training program of the particular plant the scenario design was mainly conducted by the simulator trainers of the plant. In the experiment, there were specific research questions concerning the usage of emergency operating procedures and thus the LOCA situation was chosen to be one of the scenarios to be conducted.

The following is an explanation of the FSM depicted in Figure 2. In the particular LOCA the initial event is a leak in one of the primary loops. The leak is of size 50kg/s. The process consequence visible to the operating crew is that the pressurizer level drops dramatically. Also some alarms go off but they are not depicted in the model. Another initial event was also included in the scenario. This was a failure in a plant protection signal which governs the automatic protection systems of the plant. The failure was such

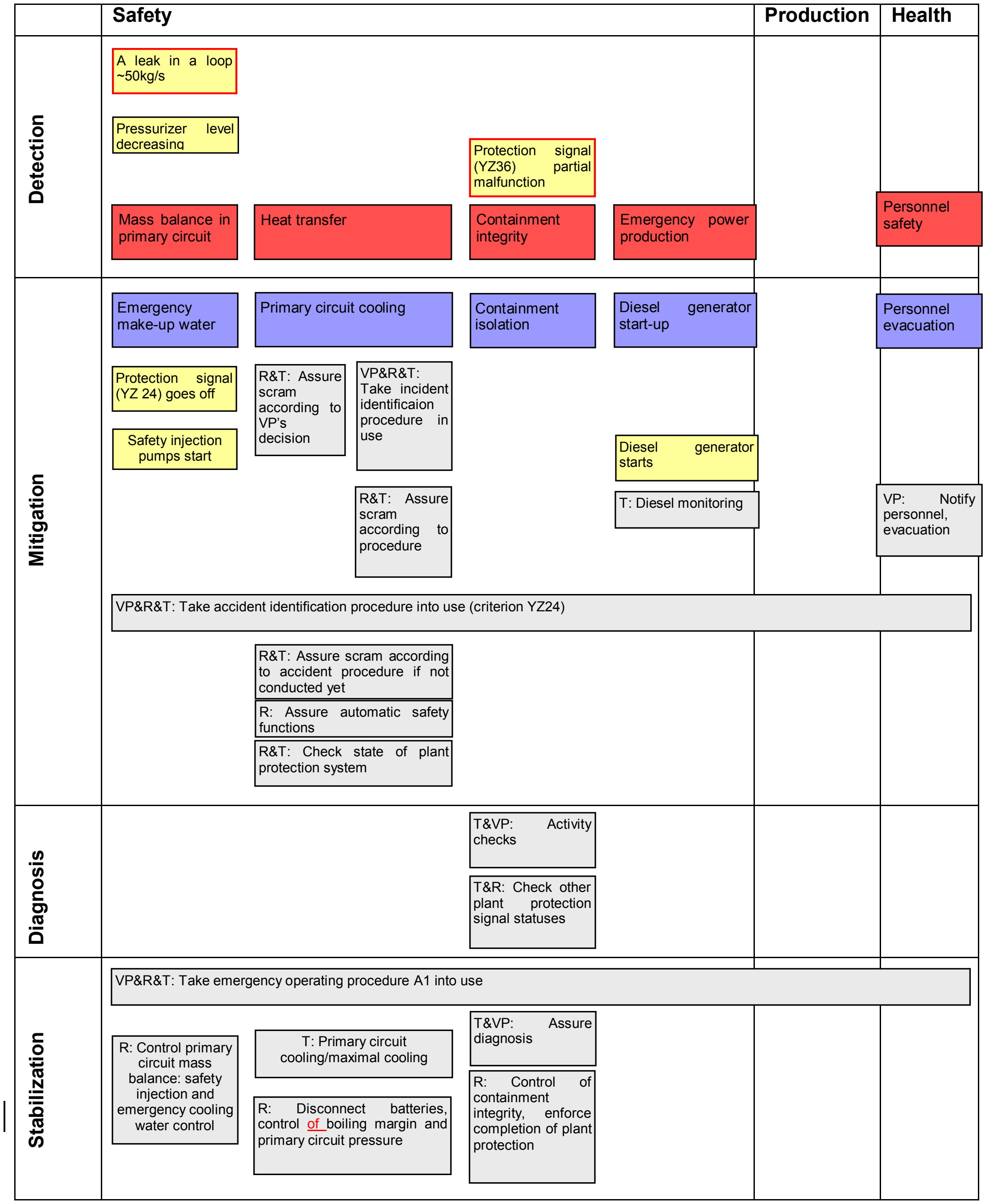

Figure 2. A LOCA conducted in the experiment depicted as an FSM. T= turbine operator, R=reactor operator, VP= shift supervisor. Yellow refers to process events, grey to operator actions, blue and red to functional means ends relationships concerning the process situation.

that the signal was initiated but it failed to launch the implications.

A critical function which is primarily endangered in the LOCA is that of primary circuit mass balance. This function is supposed to maintain that there is always enough of mass (cooling water) in the primary circuit in order to cool down the reactor core. In addition heat transfer is endangered as the coolant is leaking in the containment building and is not capable of transferring the heat to the secondary circuit. Although there is no failure in the systems related to heat transfer.

The failure in the plant protection signal (modelled as secondary initial event) endangers the critical function of containment integrity. When the automatic plant protection system is not functioning correctly the containment is not isolated properly.

The critical function of emergency power was also considered to be endangered because power systems are crucial in any accident situation in order to enable functioning of safety systems which require electricity.

Production related critical functions were not considered in the LOCA scenario as it is an accident and operating activities are mainly directed towards ensuring safety.

A personnel well-being related function was identified as in a LOCA situation the whole plant is evacuated for safety reasons.

Several means to mitigate the endangered safety functions were identified. In order to maintain mass balance in the primary circuit it is necessary to take the auxiliary feed water systems into use. This happens via automatic plant protection functions and the operating crews' responsibility is to monitor and detect that safety injection pumps are starting. Heat transfer is mitigated by concentrating on primary circuit cooling. In the beginning the important operating task is to assure that the automatic scram has functioned correctly. This can be done either by shift supervisor's decision or based on an emergency operating procedure. Containment integration is mitigated by the means of isolation and emergency power by the means of manually assuring the starting up of the diesels. Personnel safety is mitigated by evacuation means.

In the diagnosis and stabilisations phases the relevant operating crew actions were picked from the relevant emergency operating procedure and depicted under the critical function which they mostly refer to.

In the LOCA FSM depicted in Figure 2 process events are depicted as yellow boxes. The initial events have red line. The endangered critical functions are marked as red boxes and the relating operative means are depicted below as blue boxes. The operating crews' actions are marked as light grey boxes. In the text it is always marked whether the action is the duty of the turbine operator (T), reactor operator (R) or the shift supervisor (VP).

The Procedure of Utilising a FSM in the Analyses of Operating Practices

Data collection methods

During the simulator run meticulous recording of each crew's activity was conducted. Each crew member carried a head mounted camera which enabled post hoc analyses on a detailed level for example, direction of gaze. In addition the whole activity of the crew was recorded with two overview cameras from which the crew members' discussions and movements in the control room could be discovered.

During the simulator run a process expert (simulator trainer of the plant) acted as an expert judge of crew's process control activity. The process expert judged whether the crew conducted all the expected actions. This analysis constitutes part of the assessment of the performance of the crew.

Later, after the simulator run had been completed a process tracing interview was conducted in which the crew members gave their own account about what had happened in the scenario.

Data analysis methods

The analysis of operating practices was conducted selectively. First one crew's conduct was transcribed on a detailed level including nearly all operations, communications, directions of gazes and movements. Analysis for this one particular crew was conducted initially from the point of view of utilising operating procedures (Salo, Norros, & Savioja, 2009).

Based on both the FSM and this one crews activity, several episodes in the process control activity were selected for more thorough analyses concerning every crew participating in the experiment.

In the analyses of operating practices the semiotic model of habit (Peirce, 1991) was exploited (Figure 3). In this model *interpretant* refers to how people make sense of the signs of the environment. This is perceived in the realised behaviour of actors (operations, communications etc.) The *sign* identifiable in the environment always refers to an *object*. The way people use (interpret) signs reveals the objects they strive for. Thus; sign, object, and interpretant pattern (habit) emerges as a result of a continuous action-perception-interpretation cycle which connects the actor and the environment (Figure 3) (Määttänen, 2009).

In the analysis of operating practices we utilised the semiotic model as proposed by Norros (2004) and looked at crew members' behaviour in connection to the signs exploited in activity. Based on this information we deduced which was the object in the activity. And this object is a characteristic of the practice. The object can be either such that it promotes resilience in practice or such that it is more tuned towards realising lower level immediate goals of actions.

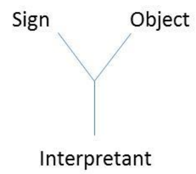

Figure 3. The semiotic model of habit (Norros, 2004; Peirce, 1991) **utilised in the micro level analysis of operating practices.**

For example, in the situation described above in Figure 2 one of the essential actions in mitigation phase is to take into use the correct emergency operating procedure (EOP). In the experiment most of the crews took the correct EOP into use (interpretant), but there were differences between the crews in the information (signs) based on which they took the EOP into use. While some crews utilised information that was related to alarm texts pointing directly to the particular EOP others utilised in addition process information relating to the operational means of responding to the endangered critical functions(object).

The method of making an assessment of operating practice

Based on the analysis described above it was possible to evaluate the operating practices. In this form we differentiate *practice* from *action* by reference to the purpose. Action can be judged by understanding whether it was completed or not, but in judging practice the reference is the objective it strives for.

In the example of taking EOPs into use, the crews which were striving for the objective of understanding whether the initial situation assessment is valid i.e. interpreting process parameters thoroughly had a more core-task oriented operating practice than those crews which simply took the EOP into used based on alarm information solely. The process information reflection of which was considered significant in the analysis was that depicted in the FSM and connected to the specific action of taking the EOP into use.

Thus, inspired by the semiotic model of habit and the interconnection of sign, object, and interpretation we compared the signs that the crews exploited in their activity to the FSM. If the crew utilised information that was related either to the functional means to respond to critical functions or the critical function as such the operating practice was considered more adequate than in the case when such information was not utilised. The analyses were completed with the data gathered via process tracing interviews in which the operators had explicated their activity right after the scenario.

We maintain that good operating practice is such that it is oriented to the core-demands of the domain. These rather permanent characteristics of the domain can be identified by analytical work i.e. core-task modelling (see also Karvonen et al., 2011; Norros, 2004). For example, in an accident situation we maintain that good operating practice is orienting to maintaining the endangered critical safety functions. This is a contextual interpretation of the concept of practice. By applying FSMs in analyses of practices it was possible to evaluate the contextual relevance of each crew's operating practice

DISCUSSION

The way of analysing human activity in a situation presented in this paper, aims at understanding the demands and possibilities that the environment sets on the operating activity in a specific scenario (The FSM). The models are situational instantiations of the more general domain models such as abstraction hierarchy of Rasmussen and the DCU description of the domain.

In the field of human computer interaction several modelling methods to depict tasks and also the wider organizational context have been developed (Beyer & Holtzblatt, 1998). These models however, do not show the connection of the task to the meaning and objective. And additionally, they are models of specific instantiations of actions in a situation whereas the FSM aims at depicting the generalized demands on actions in the form of critical functions of the process.

The modern understanding of human behaviour is that it is contextual. The power of FSMs is that they portray human activity in connection to the environmental phenomena toward which the activity is oriented. The environment is the context in which the activity takes place. In practice, in the example described above, the environment is the process system which is controlled by the automation and human operators. By describing the functioning of this joint system the FSM connects the description of human activity and a technological system. This makes the model compatible with the principles of ecological psychology which strives for understanding the joint functioning of human and the environment.

The FSM also addresses the collaborative and distributed character of human activity. In the model activity is portrayed on the level of an operating crew but the actions of individual persons can also be depicted. Similarly, connections that exist between the actions of different individuals can be modelled. In complex work, cognition can be interpreted to be distributed among human actors, automation systems and for example operating procedures. In an FSM the role of each of these actors can be depicted in connection to the critical functions of the domain. In fact, cognition is in our semiotic analyses not comprehended as internal processing of information but rather from the point of view of how meaning related to the appropriate control of the process is created as a function of a collaborative activity of the team, and how meaning of the situation is distributed within the team and, with further participants of the activity.

In the end, the most important feature of an FSM is that it makes explicit what is the *meaning* of each specific action carried out by the human actors and the process automation systems in the *situation*. The meaning is the objective to which the action is connected, and acting is required to make sense of the situation, perception alone is not enough. This approach liberates the analyst from utilising and specifying the concept of situation awareness. Evaluation of awareness is not necessary because adequate sense making in the situation is explicit in the actors' behaviour. We maintain, that the orientation of practice towards the critical functions of the domain makes the practice resilient.

Creating FSMs about operating situations, it lays ground for different kinds of data collection and analyses methods which are oriented towards understanding meaning making processes in an activity. Previously the FSMs have been utilised in selecting the most important process events for micro level analyses operating activity (Salo et al., 2009). We believe that this kind of explications of the meaningfulness of actions could be made use of also in training and teaching situations. Some evidence of FSM applicability was gained later in another reference study in a different nuclear power plant, where the simulator trainers created the FSMs with only short introduction by the researchers to the modelling technique.

CONCLUSION

In this paper we have introduced a model that is needed for describing human activity in a situation. This model is labelled functional situation model (FSM). FSM depicts the activity in connection to the environmental phenomena which is aimed to be controlled by the activity. By enabling connecting actions to their purpose FSM supports analysis of operating activity on the level of practice.

ACKNOWLEDGMENTS

This research has been conducted within the Finnish Nuclear Safety Research Program SAFIR. The authors wish to express to their gratitude to all project members and personnel of the participating organisations.

REFERENCES

Beyer, H., & Holtzblatt, K. (1998). Contextual design: Defining customer-centered systems. San Francisco: Morgan Kaufman.

Bjørkli, C., Øvergård, K., Røed, B., & Hoff, T. (2007). Control situations in high-speed craft operation. Cognition, Technology & Work, 9(2), 67-80. doi:10.1007/s10111-006-0042-z

Dourish, P. (2001). Where the action is - the foundations of embodied interaction. Cambridge: MIT Press.

Flach, J., Mulder, M., & van Paassen, M. M. (2004). The concept of the situation in psychology. In S. Banbury, & S. Tremblay (Eds.), A cognitive approach to situation awareness: Theory and application (pp. 42-60). Burlington:

Hutchins, E. (1995). Cognition in the wild. Cambridge MA: MIT Press.

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. doi:10.1016/j.intcom.2011.04.008

McCarthy, J., & Wright, P. (2004). Technology as experience. Cambridge: MIT Press.

Määttänen, P. (2009). Toiminta ja kokemus (Action and experience). Helsinki: Gaudeamus The Helsinki University Press.

Norros, L. (2004). Acting under uncertainty. the core-task analysis in ecological study of work. Espoo: VTT.

Norros, L., & Nuutinen, M. (2005). Performance-based usability evaluation of a safety information and alarm system. Int.J.Hum.-Comput.Stud., 63(3), 328-361. doi:10.1016/j.ijhcs.2005.03.004

Peirce, C. S. (1991). Peirce on signs. In J. Hooper (Ed.), Writings on semiotic by charles sanders peirce. Chapel Hill: The University of North Carolina Press.

Petersen, J. (2004). Control situations in supervisory control. Cognition, Technology & Work, 6(4), 266-274. Retrieved from http://dx.doi.org/10.1007/s10111-004-0164-0

Rasmussen, J. (1986). Information processing and human-machine interaction. an approach to cognitive engineering. Amsterdam: Elsevier.

Salo, L., Norros, L., & Savioja, P. (2009). Using operating procedures in NPP process control. Paper presented at the European Conference on Cognitive Ergonomics: Designing Beyond the Product --- Understanding Activity and User Experience in Ubiquitous Environments, Helsinki, Finland. pp. 24:1-24:9. Retrieved from http://portal.acm.org/citation.cfm?id=1690508.1690541

Savioja, P., & Norros, L. Systems usability framework for evaluating tools in safety-critical work. Cognition, Technology & Work 2012. DOI: 10.1007/s10111-012-0224-9

Suchman, L. (1987). Plans and situated actions: The problem of human computer interaction Cambridge University Press.

Vicente, K., J. (1999). Cognitive work analysis. Toward a safe, productive, and healthy computer-based work. Mahwah, NJ: Lawrence Erlbaum Publishers.

PAPER V

Identifying Resilience in Proceduralised Accident Management Activity of NPP Operating Crews

In: Safety Science 68,
pp. 258–274.

Safety Science 68 (2014) 258–274

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/ssci

Identifying resilience in proceduralised accident management activity of NPP operating crews

Paula Savioja *, Leena Norros, Leena Salo [1], Iina Aaltonen

VTT Technical Research Centre of Finland, Vuorimihentie 3, Espoo, P.O. Box 1000, FI-02044 VTT, Finland

ARTICLE INFO

Article history:
Received 22 December 2013
Received in revised form 6 April 2014
Accepted 14 April 2014

Keywords:
Emergency operating procedures
Nuclear power plant operation
Habit of action
Practice approach
Resilience

ABSTRACT

This paper reports a study in which the usage of emergency operating procedures (EOP) in a simulated accident scenario in a nuclear power plant was carefully analysed. The background assumption was that in an accident situation, even within the closely defined envelope of the EOPs, crews may act differently. The empirical findings demonstrate the variance in the operating performance of twelve professional crews all complying with the EOP. The variance was identified by analysing how crews acted in situations and what particular meanings they associated to the situations. Variance was identified in six tasks: information usage, situation identification, dealing with automation, decision making, communication, and leadership. Corresponding habits, the operationalizations of practices, were identified in the analysis. The discovered operating practices in EOP usage were analysed from the point of view of increasing system level resilience. For example, a strong inclination towards connecting operations with process responses will eventually increase the capability of the whole system to adapt to external disturbances. The conclusion is that on a micro-level the activities of crews are different even within EOP following and also, that some practices seem to have more capability produce resilience in the system. A practical implication is that in the training of proceduralized work, attention should be paid to integrating the EOP usage to the operating work, not only to adherence to the procedure. The contribution of the paper is the insight into the role of proceduralized tasks in the construction of resilient emergency operating activity.

1. Introduction

Emergency operating procedures (EOP) constitute one of the most important defences for coping with extreme situations in safety critical systems. The mechanism of producing safety is the pre-planning of critical courses of action for different demanding situations. As the actions are represented in EOPs, and learned in the operator training, the operating crews will be capable of maintaining process safety in the forthcoming situations. Hence, the purpose of EOPs is to help operators deal with demanding situations and, by ensuring adequate human performance, assure the safety in the system. The underlying assumption is, that having the EOP in use in a situation reduces operator workload and stress and thus lessens deviations from the pre-planned and accepted way of handling emergencies, i.e., *variation* in operator behaviour becomes minimised.

This line of reasoning assumes that variation in operator behaviour will result in negative consequences. In other words, variation constitutes a human error. Following this line of reasoning variation should be reduced to minimum because it poses a threat to safety.

Starting from a different premise concerning variation, it may also be argued that this very variation in operator behaviour may actually have positive effects on the overall safety. Differences in operating behaviour can also result from adapting and fine-tuning activity to more accurately suit the particular situational demands. The operating crew might deviate from the pre-planned course of action because it is what the situation demands. This "positive variance" in operating behaviour may be a considered an indication of adaptation which is characteristic for resilient acting. Therefore it may increase the ability of the system to survive also unexpected situations.

Leading on from the possibility of variance having both positive and negative consequences for safety, an obvious question regarding development of human activity within safety critical systems is that how the positive variance could be promoted while simultaneously reducing the variance leading to negative consequences.

* Corresponding author. Tel.: +358 40 5000 847.
E-mail addresses: paula.savioja@vtt.fi (P. Savioja), leena.norros@vtt.fi (L. Norros), leena.salo@fortum.com (L. Salo), iina.aaltonen@vtt.fi (I. Aaltonen).
[1] Present address: Fortum, Power Division, Finland.

http://dx.doi.org/10.1016/j.ssci.2014.04.008
0925-7535/

To answer this question the intricacies of the positive variance must be thoroughly examined. In this paper, we present a study in which the usage of EOPs in a simulated accident situation by 12 professional operating crews were analysed as *practices of following EOPs*. From a practice theory point of view rule-following never equals the written rule because "the understanding of how to follow a rule is always against the background of what is taken-for-granted", following EOP is embodied rather than reflective (Nicolini, 2013). As it is inherent to the concept of practice that these regular behaviours express meaning of acting and support sense – making of the situation (Norros, 2004, pp. 79–84), it is possible to analyse how good EOP-following practices are contributing to resilience.

The present study focused on first identifying and secondly characterising variation in the operating practices of nuclear power plant (NPP) operating crews in conducting proceduralized accident management activity. In the secondary analyses the viewpoint was the ability of the identified operating practice to increase resilience in the system. The contribution of the paper is (1) in the contextual description of the differences in operating practices of crews in following EOPs i.e. showing that differences do exist, (2) in analysis of the different practices from resilience point of view, and (3) in generalised characterisation of the resilience producing mechanisms in operating practices.

The paper is divided into four main parts. Section 2 provides a short review of the relevant literature concerning procedure utilisation, procedure design and operating practices in safety–critical industries. The empirical study is presented Sections 3 and 4 of this paper. Section 3 concerns the research process and Section 4 the findings of the study. The paper finishes by discussing how the findings of the empirical study may reflect on the development of safety–critical socio-technical systems on a more general level.

2. Background

For NPP operations EOPs constitute a *tool* with the aid of which the goals of process control in a demanding situation are met (Carvalho et al., 2009). According to Park (2009), the function of EOPs in an NPP is to help the qualified operators to bring the plant to an established operating boundary by providing practical actions to cope with the emergency event. Studies concerning design and usage of EOPs (de Brito, 2002; De Carvalho, 2006; Dekker, 2003; Dien, 1998; Filippi and Gody, 2010; Le Blanc et al., 2012; Norros et al., 2011; Salo et al., 2009; Wright and McCarthy, 2003) in the nuclear domain and other safety–critical industries have shown that conducting proceduralized tasks is demanding and requires particular expertise from the operating crews and they thus demonstrate the importance of the topic.

2.1. Complexity of using of operating procedures

A proceduralized task refers to an operating task for which a procedure has been designed. Thus, it also implies that the particular situation has been expected in design, and technical and organizational defences have been developed for the situation. For example, in nuclear domain, a range of potential accidents is always considered in the design of the plant, the automation systems, the control rooms, the procedures, and the operator training programs. These accident situations are chosen based on deterministic and probabilistic reliability analyses and they constitute the within design base collection of operational conditions. From the process operators' point of view these situations are the ones for which EOPs exists and for this reason operating activity is referred to as being proceduralized. These situations are also regularly exercised in full-scope simulators to develop routines for the control of the conditions. In NPPs EOPs are also developed state-based, for instance, to identify situations.

International Atomic Energy Association's (IAEA) safety requirements (IAEA, 2000b) and Safety Guide (IAEA, 2000a) concerning operating procedures state that the availability and correct use of operating procedures is an important factor contributing to plant safety, and therefore operating procedures applying comprehensively for normal, abnormal, and emergency conditions must be developed. Although IAEA clearly states that procedures must be followed, acting according to a procedure in a process control task is not simple (Park, 2009) and requires special expertise. Learning to act according to procedures is an extensive process taking place during both the normal daily operations and the simulated emergency handling exercises.

It is known that procedures do not cover the entirety of operating work. The gap between procedures and practice has been presented, for instance, by Suchman (1987) in the context of maintaining complex machines, and also later by e.g. Dekker (2003) and Pentland and Feldman in several articles (Feldman and Pentland, 2003; Feldman, 2003; Pentland and Feldman, 2008), who refer to *ostensive* and *performative* parts of routines and thus make explicit that what people do (performative part) is not equal per se to the artefact (ostensive part). In research, routines are typically dealt with from the ostensive point of view (Becker, 2004) and therefore, although the performative part is acknowledged to exist, research concerning its nature is not very common. We interpret, that the emphasis on performative aspect of routines means analysing rule-following from a practice point of view, which we do in this work.

Wright and McCarthy (2003) have explored the inherent paradoxes in the proceduralization of activity in the domain of aviation. In their study, procedures were explored from the perspectives of different stakeholders: designers of procedures, operating companies, trainers, and end users. Their finding is that the authors of procedures express contradictory conceptions regarding procedures: on the one hand procedures are considered as prescriptions of activity which are completed in real usage (procedures as such are incomplete), and on the other hand that they contain all the information that is needed in the situation (procedures as such are complete). This result shows how the gap between procedures and practice may be interpreted differently by relevant stakeholders.

The gap between procedures and practice reflects even a more general issue concerning proceduralization of activity, the spirit of which resonates in many elaborations concerning safety of complex systems (Hollnagel et al., 2006; Norros, 2004; Vicente, 1999): Obviously, it is good to pre-plan tasks in order to help the system survive the known disturbances and make the handling of those situations efficient, but simultaneously, in order to maintain safety also in the unexpected non-planned situations, pre-planning cannot be the only defence. As the complex socio-technical systems are open to disturbances from the environment, it is impossible to pre-plan all the possible future operating conditions. The problem is, that if the enactment of the pre-planned tasks is understood in the organisation and, e.g., in training, as if the effective behaviour that follows were the same as what is written in the procedure, and as if enactment were invariant across the actors and situations, the procedure usage is reduced to the issue of either following or non-following. This leaves many of the reasons for deviations of procedures uncovered similarly with the possible variance within following procedures.

2.2. Identifying resilience in operating practices

When procedure-following is taken as a practice which is not identical with the written procedure, attention is given to the qualitative differences in practices with regard to the sense of acting and to the weight put by the actors to various situationally relevant constraints. This perspective allows analyses and conclusions concerning, for instance, fluency and efficiency of using the procedures and the sensitiveness of procedure-following to realising if the conditions are changing to the extent in which the procedure does not apply anymore. This means that it becomes possible to analyse the contribution of actual practices to the system resilience.

At least two lines of research in which the performative aspect of routines in complex work have been tackled can be identified. One such approach exploits the concept of *non-technical skills* (NTS). These skills refer to personal, social and cognitive skills that have been found to play a critical role in maintaining safety, especially for individuals working in teams in high risk domains (Fletcher et al., 2004). The concept is oriented towards individual cognition but also so-called macro-cognitive concepts (Schraagen et al., 2008) have been incorporated to their definition. As a result, categories of non-technical skills such as task management, team working, situation awareness, decision-making, are used to define behavioural markers for these skills (Flin and Maran, 2004; Flin, 2007). In a recent article, Saurin et al. (2013) make an attempt to specify the concept of NTS that they consider too individualistic and not sufficiently considering the context in which the activity takes place. Drawing especially on Dekker (2011) the authors claim that NTS neglect the context because they are assumed to be exploited as automatic behaviours, and that NTS are not generalizable in unexpected situations. The Resilience Engineering approach (Hollnagel et al., 2006) emphasising the need for adaptation in coping with complex systems, is proposed as a means to specify NTS more adequately. From this perspective and drawing on ideas of distributed cognition (Hutchins, 1995) "NTS are the individual and team skills which are necessary to adjust performance in order to make sense of the working situation not covered by the formal design of the socio-technical system, such as formal procedures and formal training." (Saurin et al., 2013). This new definition of NTS is very close to the performative aspect of routines contributing to system safety in which we assume to find differences when studying proceduralized accident management of professional crews.

The NTSs are relevant from procedure and resilience point of view because they are connected to aspects of behaviour which are not described in procedures. We conclude that notwithstanding the attempt to elaborate the performative aspect of routines, activity is in the NTS-studies analysed as a sequence of actions. The authors do not recognize the need to accept that activity is not fully describable in the terms used. The current NTS approach does not take advantage of the practice theory approach as an alternative to the cognitive approach of describing behaviour.

A second line of research that could be interpreted as taking the performative perspective to analysis of routines in action concerns identification of *resilience markers* in behavioural data by Back et al. (2008) and Furniss et al.(2011). The aim in the two papers is to find evidence for resilience in the observable behaviour. The concept of *practice* has been adopted as the key unit of analysis in the studies even though it is not very profoundly explained what constitutes practice in the behavioural data. Special interest in analysis of practices is in understanding the capabilities of tools and instruments to support resilient behaviour. Even if not very explicit, the authors appear to consider the theoretical advantage of the concept of practice to identify generic regularities in actual behaviour that could provide means to cope with the un-expected.

We see that both the analysis of NTS and the resilience markers approach are very relevant for our attempt to understand the mechanisms of possible positive variation in operating practices. We agree with the authors about the need to define actual behavioural expressions of the potential of an organisation to act safely, i.e., manifest resilience. In agreement with the recommendations of both Saurin et al. (2013), Saurin and Gonzalez (2013) and Furniss et al. (2011) we consider the role of instruments in practices – in our case procedures. We see, however that neither the NTS approach nor the resilience markers approach fully follow the idea of Becker (2004) and Grote (2012) that routines are an intrinsic part in the structure of behaviour and that ostensive and performative parts of routines serve basically the same generic functions. Even though utilising the concept of practice, the authors do not make a methodological conclusions of turning to this concept, and they do not make operationally explicit how to identify practices in observational data. We also see that – probably due to at least implicitly contrasting procedures/routines (as predefined) and skills/practices (as adaptive) – these studies do not provide characterisations of practices that would support resilience.

2.3. Intelligent procedure use

In their comprehensive studies on NPP operators' procedure usage Dien (1998) and Filippi et al. (2006) came very close to the practice theory oriented approach we would consider useful. The conclusion Dien draws on his studies is that operators are not able to, and that it is not reasonable to require them to, follow procedures literally. Instead, operators are required in their procedure application to "make up their oversights", i.e., to complete procedures to fit the demands of the situation. Dien coined the term "intelligent application of procedures" to mean the operating activity which is successful in fulfilling the contradictory objectives in procedure use. The intelligent application of procedures was a research issue also when implementing computerised procedures to the N4 NPPs. According to Perin (2005) even the company's (Electricité de France, EdF) own top managers expressed concerns of a too tight of guidance of operators by the procedures (Perin, 2005). Filippi (2006) synthesises findings of N4 procedure studies by defining the psychological characteristics of procedure usage in incident and accident situations. Our interpretation of these characteristics, which is also influenced by other authors' work (De Carvalho, 2006; Nuutinen, 2006; O'Hara et al., 2000; Roth et al., 1994; Throneburg and Jones, 2006) is that intelligent application of procedures requires from operating crews.

- *connecting the situation with the procedure* which requires the operator to create a context for interpretation, i.e. to make the procedures meaningful in the current situation and a relevant framework for action, which requires cognitive effort,
- maintaining *coherence in acting* despite of interruptions that require changes in attention and maintaining unfinished tasks active for retrieval; the interruptions are often induced by other team members which requires anticipation and control of other persons' acting,
- organising efficient *collaboration and coordinating* with other team members according to the prescriptions of the procedure which typically require parallel independent acting with occasional joint updates between the team at prescribed phases,
- developing *trust in automation and procedures* is important for the development self-confidence and professional identity which have effect on operators' the ability to act in uncertain situations.

We see, moreover, that the above demands can act as bases for developing indicators for judging the capability of rule-following practices to support resilience of the system. Thus, in the detailed study of operating crews utilising EOPs in accident management, the background assumption was that the aforementioned requirements of intelligent procedure usage may be fulfilled to different extent by the different operating crews. If so, these differences would be evident in the operating crews' practices of utilising EOPs to handle emergency situations.

2.4. Research problem: Practices of conducting proceduralized tasks

The studies concerning procedure usage have shown that utilising procedures in an activity is not simple or straight forward indeed: Special expertise is required from the operating crews in completing procedures in action and conducting the proceduralized tasks intelligently.

Studies on performative aspects of routines have shown that there is more to human activity than is assumed in treating proceduralized activity merely as procedure following. But few background assumptions implicit in their conductance can be questioned. Many of the studies seem to contradict rule-following with not obeying or even non-usage of procedures. The importance of rule-following cannot be argued against, it is self-evident, but we see that the important issue is to understand *how* rules should be followed. All NPP process control is governed by procedures but these procedures may be followed in different ways: In different situations different crews may interpret the rules differently. Whether these different interpretations in action are significant from the point of view of safety cannot be discovered if procedure utilisation is studied merely from the point of view of following/not following procedures. Therefore, the research approach must be such that it can capture the variance within the procedure following.

Another issue concerns the concepts of resilience and practice. In simple terms, resilience is the ability of the system to flourish even in unexpected situations. It is clear that resilience is a critical characteristic of socio-technical systems but the current literature does not make very explicit links between the actual practices of work and resilience in the system. It is necessary to understand more thoroughly how the behaviour of the individual crews is connected to the safety maintaining functions of the whole system. And this leads to the concept of practice. None of the reviewed studies explicitly define what they mean by practice even though the concept is used in exploring and identifying safety ensuring mechanisms in observed behaviours.

Drawing on the above reasoning, our assumption is that (1) there may be differences between operating crews in how procedures are utilised and (2) it is possible that some of the ways of using procedures are more beneficial for safety of the system than some others. In the study presented in this paper, we investigate how exactly professional NPP operating crews conduct proceduralized emergency operating tasks. The research problem concerns the operating practices in conducting proceduralized accident management tasks. It is formulated as two operative research questions as follows:

1. What kind of variation can be identified in the accident management practices between the different operating crews (who all fulfil the demands prescribed in the operating procedures)?
2. What kind of practices can be identified that can be argued to contribute to system level resilience?

3. Research process

In order to better understand *how* emergency operating procedures are used in complex work, a study at a NPP training simulator was conducted. This chapter describes the research process utilised in the study.

3.1. Data collection

3.1.1. The plant, the participants, and the EOPs

The particular nuclear power plant is of type pressurised water reactor consisting of two separate units. The plant originates from the late 1970s and produces currently close to 500 MW electrical power in each unit. A normal operating crew (excluding field operators) consists of three operators: shift supervisor (SS), reactor operator (RO), and a turbine operator (TO). The responsibilities of the operators are divided so that RO takes care of the primary circuit: heat generation and cooling. TO's responsibility is the turbine operation and electricity generation. SS has a leading role in making crucial operative decisions and ensuring the duties of both RO and TO.

All twelve operating crews of the plant participated in the study which means that altogether 44 operators acted as users in the experiment (in addition to the normal crew the trainees in crews also took part in the exercises). Thus, the operating experience of the participants varied from 1 to 32 years of experience. There were 18 participants in the experience group 1–9 years, 13 participants in the experience group 10–19 years, and 13 participants in the experience group over 19 years.

The EOPs of the plant have been designed so that there are two different identification procedures: Incident identification (I0) and Accident identification (A0). Depending on particular automatic signals and alarms the operators choose either of the above. In choosing which identification procedure to take into use, operators receive support from automation system; when the respective plant protection signal is launched a support display appears on the operating screen which commands to take either I0 or A0 into use. Also, the operators tend to know by heart the criteria (the automatic plant protection signals) which indicate which EOP should be taken into use. Each operator role has a designated flowchart type procedure which has been designed for the specific operator tasks. The identification EOPs of TO and RO prescribe actions related to respective sections of the power plant process. The identification EOP of the SS prescribes actions that further ensure the actions of TO and RO and bring thus redundancy to the activity of the crew.

3.1.2. The simulated accident scenario

The scenario which was analysed from EOP usage point of view is an emergency situation in which there is a leak in the primary circuit which means that the circulation of the coolant is decreased and thus the cooling of reactor is endangered. This particular accident scenario is a typical design basis accident and has been labelled loss of coolant accident (LOCA). The specific LOCA utilised in the simulator study was midsize, which means that reactor and turbine scrams were automatically released, containment isolation was completed, and safety injection water systems were initiated by the automation system. Also, diesel generators were started up in order to assure energy supply for safety systems. The operators' tasks in this type of scenario consist mainly of double-checking and assuring that all the automatic safety systems are functioning as required and of further identification of the situation e.g. locale of the leak. In this scenario there was one additional simulated failure in the safety systems: A particular plant protection signal did not function correctly and thus containment isolation was not completed automatically. Scenario is presented in more detail below.

In the beginning the plant is running on full power. Some minutes after the start of the scenario there is a leak of 50 kg/s in one of the six primary loops. The pressurizer level will drop quickly. The leak endangers the mass balance on the primary circuit side of the plant. Signal YZ36 will go off but fails to release a plant protection chain that is related to the automatic isolation of the containment. Very *soon plant protection signal YZ24 goes off and starts safety injection pumps to compensate the loss of water in the primary circuit. The operators should detect the abnormalities in the process state and perform appropriate actions to verify the functioning of safety systems. YZ36 signal is a criterion for taking the incident identification procedure I0 into use and YZ24 for taking the accident identification procedure A0. After detecting that the process is in disturbance and that shutting down the plant is required the operators can perform scram (and verify automatic shut down) on the basis of either the shift supervisor's decision or the procedure I0, or, if they have not released scram before noticing the YZ24 signal, on the basis of the procedure A0. Identification of the accident situation is performed with A0, which then leads the operators to the A1 primary circuit leak procedure. With A1 the crew runs the plant towards a safe state by cooling the primary circuit and by manually releasing the YZ36 signal that finalises the isolation of the containment.*

As a starting point, the simulated accident scenario was carefully analysed from the perspective of critical functions which are endangered in the situation and the required respective operator actions. This analysis produced a *functional situation model (FSM)* (Savioja et al., 2012) of the accident situation. This model depicts the generic critical functions of nuclear power production in the light of this particular emergency situation. The model also has a temporal dimension and it depicts the main operations that the operators are supposed to conduct in the situation. In the model the operations are connected to the functions. Thus the model describes the *meaning* of each operation: The model makes explicit both *what* actions operators take in order to gain control of the process and for which operational *purpose*. The model is the reference for analysing *how* operating crews in their activity take into account important process information.

3.1.3. Observation of the operating crew activity

The operating activity of each crew in the simulated accident scenario was observed both online and via recordings. The recordings were both in audio and in video format. Each operator carried a head mounted camera which enabled analysis of direction of gaze. In addition there were overview video cameras and audio recorders registering operating activity. The process events and all operations were recorded in simulator logs.

Later, selected parts of the activity were transcribed into spread sheets in which the courses of action for each crew were depicted on a detailed level. This description included process operations, verbal communications, movements (person's position in the control room), and directions of gaze (when distinguishable in the data) for each crew.

3.1.4. Selection of relevant episodes in the data

The analyses were focused by selecting relevant episodes from the scenario. Identification of the relevant episodes commenced with the FSM. First, it was considered which parts of the chosen scenario were relevant from the point of view utilising procedures. Next, the videoed process control activity of one crew was carefully transcribed and analysed by two researchers (see Salo et al. (2009)). The findings concerning the operating practices of the one crew suggested that although the performance was impeccable in the sense that all proceduralized actions were carried out, it seemed that towards the later episodes the crew adopted an operating practice which was not very sensitive to the process information and was tuned more towards the procedure than actually the process.

In the detailed analysis all the distinguishable decision making points were marked in a spread sheet containing the transcribed activity. The points were compared with the relevant parts of the FSM and as a result four episodes were identified as being crucial for successful accident management in the given scenario: (1) Initial detections and scram, (2) Coping with the plant protection failure, (3) Detection of new indications concerning the nature of situation and (4) Diagnosis and choosing the event based procedure. Later the episodes 1 and 2 were combined and the analysis of the activity of all crews was conducted with 3 episodes.

3.2. Data analysis

The data analyses consist of two phases. First, the differences in the behaviour of the crews were identified with regard to critical process control tasks found relevant in the episodes, and secondly the differences were classified and graded.

3.2.1. Identification of the crews' operating practices (in process control tasks)

In identifying whether and how the crews' behaviours in conducting the proceduralized tasks differed from each other, the semiotic concept of habit (Peirce, 1958) was utilised as an operationalization of the concept of practice (Norros, 2004). Habit is the notion that has been proposed to express the way an organism is organised to meet the changing and unexpected features of the environment. The philosophers Charles Sanders Peirce and John Dewey saw habit as a fundamental principle of human thinking and action (Dewey, 2002; Peirce, 1998): Human actors connect themselves to the possibilities of the environment by continuous action–perception cycles, during which the outcomes of action are observed. Relatively stable anticipatory schemes, habits, emerge when an appropriate response to and interpretation of the environment is reached. According to Peirce's theory, habit has a triadic structure containing three components: sign, interpretant, and object. The three components are interconnected as depicted below (Fig. 1). Interpretant is the element of habit which is directly observable in the behaviour of the actor. Acting of a person is driven by interpreting the signs in the environment. The sign refers to any perceivable element or feature in the environment. It can be an action (e.g. a communication, a movement) of another person or it can be a technology mediated sound or a visual feature. The idea or object to which the sign can be connected via the interpreting behaviour reveals the objective which is sought for. In other words: the way the crew acts based on particular signs reveals what they consider the situation to demand from them.

Identification of habits was conducted for each episode and for each crew separately. The treatment was carried out first for all crews in episode one, then for all crews in the next episode.

In the identification of habits we analysed the perception–action cycles in the particular episode and focused first at what people did (interpretant). In this step the key actions (verbalisations, operations, movements, etc.) of each crew member in the episode were written out in textual format. The next step was to look at the very moments prior to the deed, and identify the signs based on which the particular actions were carried out. The signs were previous actions, process information, information in the procedure, sounds, communications, etc. It is not always obvious

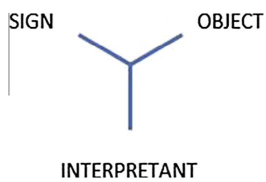

Fig. 1. The semiotic structure of habit (Peirce, 1958).

what sign has triggered the interpretation process because human acting is continuous, but in most cases we were able to come to a conclusion what constitutes the sign on a suitable level of abstraction. In the last step, we deduced the objective in the habit by analytically comparing the interpretant, the sign, and the demands of the process situation described in the FSM.

As this analysis of perception–action cycles was carried out for the entirety of the data we identified that there was notable variance in the crews' behaviour with respect to the generally known critical process control tasks that we found relevant in each episode, i.e. in information usage, situation identification, dealing with automation, decision making, communication, and leadership. All crews conducted these process control tasks complying with the procedure, but the micro-level of the behaviour, the ways of conducting the tasks, i.e. the habits of decision making, habit of situation identification, habit of dealing with automation, habits of decision making, habits of communication, and habits of leadership, were internally different. Sometimes crews acted based on different signs, sometimes the same signs were interpreted differently, and so on.

3.2.2. Classifying and grading habits

The next task in the analysis was to grade the variance within the habits. We started by analysing similarities and differences qualitatively. In nearly all habits (habit of decision making, identification of situation, etc.) we identified classes which fit the classes of interpretative, confirmative, and reactive previously defined on the basis of the Peircean theory for differences in the creative strength of habits (Norros, 2004): An interpretative habit is such that behaviours can be identified which point in direction of expressing interest on the *present* situation and urging to own *interpretation* of situational demands, *questioning* the observed phenomena, *building expectations* of future events. Nearly an opposite type of habit is a reactive habit. Reactive habit is such that it reflects *passivity* and *lack of expectations* concerning the situation. No indications of own interpretations and reflections are identified in the data. Weakness of reactive habit is that it is able only to react to situations; there are no anticipatory aspects. A confirmative habit is such that neither reactive nor interpretative characteristics can be identified; therefore it constitutes a third type. Confirmative habit can be described as taking the situation *for granted*, acting in a *pre-defined* way, and over emphasising *rules* and *procedures.* In a confirmative habit repetition is dominant whereas in the interpretative habit also the adaptive potential of a habit is represented. We utilised these pre-defined labels of habit types as the analysis frame, but the qualitative characteristics of each class were grounded in the observational data.

The background assumption in using this specific internal typology of habits is that interpretativeness in operating habits may be a mechanism of producing resilience in the system. For example an alarm as a sign might be interpreted by seeking other process indications confirming the particular alarm, an interpretative habit, which is a resilience-producing habit of action because it is oriented to the process information and making thorough sense of the situation. Whereas, the same alarm could also be interpreted by conducting exactly and only the action the alarm indicates. This kind of habit of reacting to alarms can be argued to have less capacity to produce system level resilience because it takes the provided alarm information for granted and does not aim for contextual interpretation of the situation.

As an end-result of all the analyses it came possible to formulate results such as: Crew A portrays certain types of habits in the process control tasks required in the studied episodes and give a contextual description of these habits. We were also able to assess whether the crews' practices would support resilience of the system.

4. Findings

The findings of the research are presented in two sections. The first section (4.1) presents the habits of the crews episode by episode, and crew by crew. It shows how different habits are distributed among the different crews of the plant. In the second section (4.2) the differences in habits are contemplated for their capability of producing system level resilience in the activity.

4.1. Different habits in each episode

4.1.1. Episode 1: Initial detections and scram

The first episode covers a time frame from the detection of the first signs of disturbance to the point of time when SS makes a decision of first actions and the reactor scram is conducted. During this time, the operators will detect that the process is in disturbance, that operator actions are required, and that specific procedure should be taken into use. Therefore, the name of the first episode is "Initial detections and scram".

Using the triadic structure of habit as an instrument the detailed analyses of crews' process control activity revealed differences in crews' behaviour in four process control tasks relevant in this episode: information usage, situation identification, dealing with plant protection, and justifications for scram. In the following section each of the emerging habit kinds are presented and the internal differences within them are demonstrated by grading them on the scale of interpretative, confirmative and reactive and by contextually defining what, based on the observational data, constitutes e.g. interpretative information usage. The summary of the graded habits in episode 1 is presented (Table 3).

4.1.1.1. Habits of information usage. In the first episode there were differences between the crews in the extent to which process information was gathered before determining what to do, and before conducting any process interventions. There was variation from utilising only the alarm information to extensive and diverse process parameter observations and dialogue (examples in Table 1). In analysing the data several behavioural indications were identified based on which the habit could be classified being one of the three types: interpretative, confirmative or reactive. These indications were utilised in the grading of the habits as depicted below (Fig. 2).

Some crews gathered redundant and diverse information before conducting any process interventions. This was considered interpretative because it reflects an objective of validating the initial observations by gathering process information profoundly. Also these crews typically gathered different *types* and *abstraction levels* of information, e.g. alarms, display support system, process parameter values, trends, automation information, procedures. An important characteristic was also utilisation of both redundant and diverse information sources. Most importantly it was typical to the crews to jointly reflect on the acquired information i.e. construct own understanding of the situation and create assumptions concerning it. This reflection process is a knowledge creation process which increases the opportunity to spot any initially false basic assumptions. It also shows that the crew is tuned towards understanding the dynamic process phenomena as it is taking place. If the crew conducted information gathering this way, the habit of information usage was considered interpretative.

Confirmative habits in information usage were also identifiable in the data. Some crews seemed to conducts double-checking more as rule. While this habit is adequate as such, it does not indicate a deeper strive to understand the situation and being one with the real-time process. The problem with the confirmative habit in this case is that it may lead the operators to only consider such new information that validates the existing one, which is a well-known

Table 1
Examples of crew activities in information utilisation.

Crew I in episode 1	Crew B in episode 1	Crew A in episode 1
First, all three operators are looking at the process monitoring system **SS**: “The doors of the ice condenser” [reads aloud the alarm text] “Do we have a small leak in the primary circuit, maybe a big one?”	First, all three operators are looking at the process monitoring system **TO:** “Ice condenser doors [reading from the alarm list].” **RO:** “The pressuriser level and circuit pressure are dropping”	All operators are sitting at their respective desks **SS:** “Now came moisture alarm [reading from the alarm list].” **TO:** “Ice condenser doors [reading from the alarm list].”
SS looks at trends on process monitoring system and plant protection signals. RO looks at display support system which displays a command to take I0 into use. RO looks at plant protection signals **RO**: “YZ 36 [name of the protection signal], I0” **TO**: “The pressurizer level crashed” **SS**: “Primary circuit pressure is 120 bars”	RO looks to the plant protection signals. **SS:** “Yes, we have some kind of a leak.” **TO:** “Seems like a leak in a primary circuit” **SS:** “Conduct the scrams”	RO looks at display support system displaying command to take I0 into use. RO is silent. **TO:** “YZ36 [name of the protection signal].” **SS:** “Drop in pressurizer level”.
TO turns to face SS, who listens actively. **TO**: “Now there is a big leak going on” **SS:** [in reply to TO]: “Ice condenser doors are open there. Shall we scram the reactor? Yes, do it.” **RO:** “Yes”		SS looks to the protection signal panel **SS:** “YZ36” **TO:** “Scram?” **SS:** “Yes”.
Analysis: crew utilises different kinds and types of information and both RO and SS check the plant protection signal status. Information usage is also dialogical → interpretative habit	Analysis: Crew utilises both alarm and parameter information but draws the conclusion to conduct the scram without considering the plant protection signal status. → reflects a confirmative habit	Analysis: Crew makes hasty observations based on alarm information mainly. The information provided by different systems is only read aloud. No one seems to be listening what the other person is saying. → reactive habit

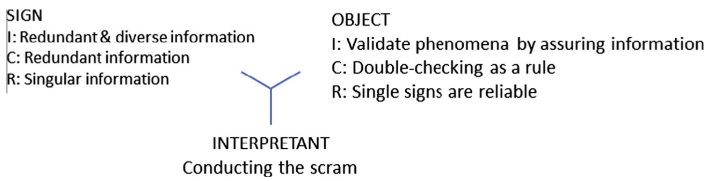

Fig. 2. Habits of information usage: differences identified in the data. I = interpretative, C = confirmative, R = reactive.

source of erroneous situation interpretation. The confirmative habit also utilises multiple sources of information but the observations are not as profound as with the interpretative habit because diversity of sources is not utilised to the same extent. Most importantly, there are no signs of joint knowledge creation process in the behaviour of the crew. As is characteristic for a confirmative habit, the rule-based approach generally produces a good result because in most cases the rule applies.

In analysing the data, an indicator pointing to reactive habit was utilising alarm information only. This was considered reactive as it reflects considering singular information reliable enough to base decisions on. The reactive habit in information usage means that information in alarms and the support systems is taken for granted and no additional double checking concerning accuracy of observation is made. Characteristic for the crews acting this way was that observations are paced by alarms and thus the crew is always a little bit lagging behind the process events.

4.1.1.2. Habits of identification of process situation. There were differences in crews' habits in situation identification based on which the same classification of interpretative, confirmative, and reactive could be made. The analysis was carried out based on verbalisations concerning the process situation. The reference in determining the adequacy of the habit was the FSM in which the characteristics of the scenario had been thoroughly analysed.

In the analysis of verbalisations and communications there were differences between crews from reading aloud alarm texts to contemplating the general characteristics and also consequences of the situation (examples in Table 2). In determining whether the habit of situation identification was interpretative, confirmative, or reactive, several behavioural indications emerged. They were utilised in the grading of the habits as depicted below (Fig. 3).

If the operating crews referred to safety functions in their verbalisations, the situation identification was considered to be a threat to mass balance. This indicates a thorough and holistic functional understanding of the process situation and thus the habit was analysed to be interpretative. The same judgement was made if the size and the location of the leakage were somehow referred to. The interpretative habit in identification of the process situation means that the information concerning process, e.g. parameter values are not only considered for face value but also their functional meaning is e.g. the meaning of a certain parameter value from the point of view of the overall safety is taken into consideration explicitly. This provides resilience in the system because it enables treatment of situations which do not fit the pre-existing typology of possible events.

Confirmative habits were such that situation was identified to be some kind of leakage as this word was utilised in verbalisations but no general linking to the safety functions was made. The confirmative habit in this case reflects a strive to fit the on-going situation to an existing typology of possible events and thus it does not build capability in the system to survive unexpected events. The confirmative habit in process situation identification means that process events are identified, but not really functionally. A habit was seen to be confirmative if there was no evidence of the crew contemplating e.g. size of the leak.

If, in the other end of the spectrum, the verbalisations concerned only alarm information which is directly readable in the alarm system, the interpretation was considered to be that there is a process disturbance, and the habit was considered to be reactive. The reactive habit in process situation interpretation meant

Table 2
Examples of crew activities in interpreting the process situation.

Crew C in episode 1	Crew G in episode 1	Crew H in episode 1
A few seconds silence after the first alarm. **SS**: "Containment isolation signal is on."	Some seconds of silence after the first alarm. **TO:** "Ice condenser doors are open, TL22."	SS looks at the alarm list. **SS:** "How did you guess it was going to be I0?" **RO**: "Ice condenser doors [reads directly from the list]. All batteries are on. Isolation signal" **SS:** "Did I0 come already?" **RO:** "Yes."
SS turns to plant protection signal panel and looks also to the side where safety injection pumps are located. RO is looking at trend information. RO looks at display support system where to command to take I0 into use is displayed. **RO:** "TZ20 and 30 are flowing in."	Everybody looks at the alarm list. TO looks at trends, SS reads alarm messages carefully. **RO:** "I0, criterion YZ containment isolation[reads directly from the display]."	
SS looks again to the side panel. **RO:** "Is there more waters?"	RO looks at the plant protection panel behind him. **TO:** "Primary circuit leak."	TO and SS turn to the plant protection signal panel. RO looks also in that direction. **SS**: "Let's take the I0 now."
SS looks at trend information concerning safety injection pumps. TO points to the display containing I0 commandment. **TO:** "Take I0 into use, it says there." **SS:** "How is the pressurizer level behaving?" **RO:** "Level is dropping." **TO:** "It is dropping fast." **SS:** "Going down fast, you say?" **RO:** "There's I0." **SS:** "Presumably there's leak in the primary circuit."	SS follows what RO is doing and looks also to the back panels. **SS:** "So, it is YZ36, let's take the I0s," **SS:** "So, it seems like a leak."	
Analysis: Crew is carefully considering what the process information means before taking action. This is evident in the way the SS is considering the pressurizer level behaviour and safety injection pump status. → interpretative	Analysis: Crew makes a conclusion that there is a leak but does not really go further towards functional interpretation of the situation. → confirmative	Analysis: The main observation of the crew is that the situation is I0 which is the name of the procedure to use in this particular incident. → reactive

that the crew did not explicitly consider the nature nor severity or functional meaning of the incident.

4.1.1.3. Habit of dealing with automation. The third habit in episode one concerned the plant protection signals. All the crews observed that the plant protection signals had gone off. What in this episode was thoroughly analysed was the crews' behaviour during and immediately after realising that plant protection had gone off.

The variation in behaviour of the crews was as follows: Some crews seemed to take a notice of the on-going plant protection chain which in this case was total containment isolation, and immediately check if the automatic sequence was functioning adequately. This behaviour indicates that the crew understands the function of plant protection and takes an active situation specific questioning stance in their relation to automatic safety system. This is evidence of agency which indicative of interpretative habit.

On the other end of the spectrum some other crews interpreted the protection signal as a direct signal to perform the scram. By no means is this a wrong interpretation but it reflects total reliance on automation and thus this habit was considered reactive. In the middle class the containment isolation was mentioned but no active role of ensuring that it was functioning was taken by the crew. This reflects a stance in which the human and the automation remain in their separate pre-defined roles, and neither bothers the other by questioning the adequacy of its functioning. Results of all crews are again found in Table 1.

In determining whether the habit in dealing with automation was interpretative, confirmative, or reactive, several behavioural indications emerged. They were utilised in the grading of the habits as depicted below (Fig. 4).

4.1.1.4. Habit of decision making. The fourth function in which crews' activities contained variation concerned way of deciding to conduct scram. Although all the crews conducted the manual scram to ensure the automatic one, the different crews took different measures of making the decision. The main difference in behaviour of the crews was that some crews took the initiative to conduct the scram by SSs' judgement of the situational demands. The other way to conduct the scram was to do it instructed by procedure. This difference means that some crews conducted the scram prior to taking the abnormal operating procedures into use. It was analysed that SS's discretion concerning the need to conduct the scram is a sign of interpretative habit. It reflects understanding of the situational needs, anticipation, and prioritizing safety relevant tasks and most importantly, human agency in controlling the automated process. Conducting the scram after the instruction from the procedure was considered confirmative habit because it reflects a rule-based attitude to decision making: The conductance and e.g. timing of actions is controlled by the procedure. In the habit of decision making we did not discover any reactive habit, but as depicted below, a reactive habit might be hesitance in conducting the scram which would mean that there is insufficient system level control.

In determining whether the habit of decision making was interpretative, confirmative, or reactive, several behavioural indications emerged. They were utilised in the grading of the habits as depicted below (Fig. 5).

The summarised crew level results in episode 1 are presented below (Table 3). In the case of information usage four crews portrayed an interpretative habit (C, D, F, H, I), five crews a confirmative habit (B, E, G, K, L), and two crews a reactive habit (A, J). In situation interpretation there were two interpretative crews (C and I), six confirmative crews (B, D, E, F, G, K), and four reactive crews (A, H, J, L). In interpretation of the meaning of the plant protection signal there was only on interpretative crew (H), four confirmative crews (C, E, F, G) and seven reactive crews (A, B, D, I, J, K, L). In conducting scram six of the crews carried an interpretative habit (A, B, C, E, I, L) and six a confirmative habit (D, F, G, H, J, K).

4.1.2. Episode 2: Detection of new indications concerning the nature of situation

Episode 2 is about realising that situation is deteriorating. The leak is so large that the safety injection pumps start to compensate

Table 3
Summarized results concerning habits of the crews in episode 1.The grading of habits is depicted as colouring of the table cell. White = reactive, grey = confirmative, black = interpretative.

Crew	Habit of information usage	Habit of situation interpretation	Habit of dealing with automation	Habit of decision making
A	plant protection, pressurizer level	Disturbance	Cue to perform scram	SS
B	emergency cooling, pressurizer level, primary circuit pressure, plant protection	Leakage	Cue to perform scram	SS
C	containment isolation, plant protection coolant flows, pressurizer level, pressurizer level gradient	Mass balance	Realise isolation	SS
D	plant protection, pressurizer level, primary circuit pressure, several trends in PMS	Leakage	Cue to perform scram	procedure
E	alarms, plant protection, pressurizer level	Leakage	Realise isolation	SS
F	pressurizer level, plant protection, emergency cooling, pressurizer level	Leakage	Realise isolation	procedure
G	emergency cooling, plant protection	Leakage	Realise isolation	procedure
H	emergency cooling, primary circuit pressure, containment isolation, plant protection	Disturbance	Ensure isolation actively	procedure
I	emergency cooling, pressurizer level, primary circuit pressure	Mass balance	Cue to perform scram	SS
J	emergency cooling, plant protection	Disturbance	Cue to perform scram	procedure
K	plant protection, alarm info	Leakage	Cue to perform scram	procedure
L	alarms, plant protection signals, primary circuit pressure	Disturbance	Cue to perform scram	SS

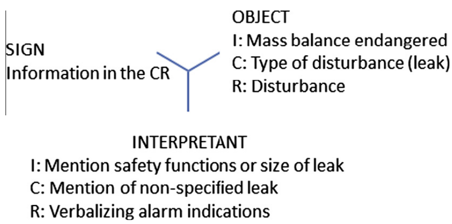

Fig. 3. Habits of process situation identification: differences identified in the data.

the loss of water. This defines the situation an accident as opposed to an incident which it had been up to this point. This new information yields a procedure shift from I0 (incident identification) to A0 (accident identification). This procedure change is a crucial point in the scenario as an accident procedure has been designed to tackle more severe situations than an incident procedure.

In close analysis of crews' behaviours in episode 2, it was possible to see differences concerning two issues: Information usage and Communication.

4.1.2.1. Habit of information usage. Information usage in episode 2 refers to the information based on which the crew decides to take A0 into use. There are several possibilities: (1) The procedure shift is announced in the display support system of the RO. In the display a text "take A0 into use" is displayed and connected to an audio alarm. (2) The same signal which makes the procedure shift instruction appear on RO's display also starts the safety injection pumps. It is possible that the crew notices that these pumps are running and from this process information deduces that the situation must have deteriorated and thus a procedure shift is needed. (3) The crew may also observe the signal (YZ 24) which the initiator for the display system and the safety injection pumps.

The analysis of which information was utilised was based on observing individual persons' moving and whereabouts in the control room and directions of gazes. Similar differences in crews' information usage were identified in the data concerning episode 2 as previously in episode 1 (Fig. 2): Some crews carried out all the above mentioned three ways of using information and consequently the habit was graded interpretative. It was analysed that for interpretative habit it was necessary that the crew made notice of safety injection pumps which is indicative of following actual process response and also checked the plant protection signals which is indicative of double-checking that automation is functioning correctly. If the crew changed procedures only based on display support system the habit was judged to be reactive as there is no indication of strive for understanding the functional meaning of the information displayed. If the crew only observed the safety injection pumps the habit was graded as confirmative. Results of all crews are displayed below (Table 5).

4.1.2.2. Habit of communication. Communication habits of the crews were observed throughout the scenario but as specific indicator they were utilised in the analysis of episode 2. This episode had a special demand for communication as it is about shifting the procedure. In a heavily proceduralised task, in which almost all the operations are described in the procedure, the selection of correct procedure becomes the critical point in which the common understanding of the crew could be used as a way of testing the decision. But this cannot happen if the crew does not communicate adequately. In analysis of communication it was considered important that the crew would use the resource of collaboration in ensuring that they are moving into a right direction when shifting to use the procedure A0. Therefore the same information sources were utilised as in analysis of information usage in episode 2 but in the analysis of communication the question was whether the individual crew member who became aware of the particular information made it available to the other crew members also. In other words the question was to what extent the information was communicated within the crew. Examples of different communication habits are depicted below (Table 4). In analysing the habits in communication several differences between crews were

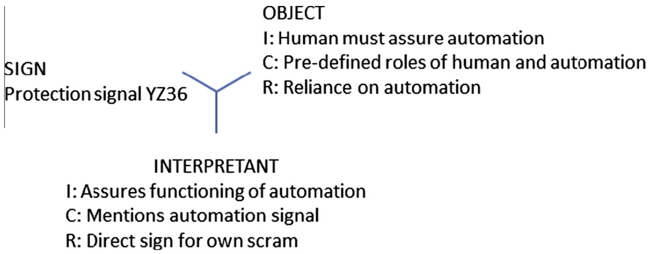

Fig. 4. Habits of dealing with automation: differences identified in the data.

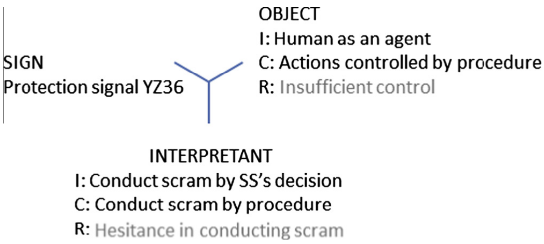

Fig. 5. Habits of decision making: differences identified in the data.

identified. Based on the differences the classes of interpretative, confirmative, or reactive, were formulated which were utilised in the grading of the habits as depicted below (Fig. 6).

The interpretative habit of communicating all the relevant information reflects an objective of creating joint awareness of the situation. Interpretative habit in communication was about depth of issues which were talked about and dialogue. Dialogue in communication enables creation of new knowledge and is thus considered interpretative. Creation of new knowledge is especially important resilience characteristic because resilience assumes that system can survive even totally unprecedented situations about which no previous knowledge exists.

The confirmative habit in communication was type of conversation which did not include real dialogue concerning process status or e.g. projected upcoming process behaviour or required crew activity. In the conversation lacking dialogue the contents are merely repetitions of own interpretations and it seems that the objective in the communication is to confirm what is already known. In the analyses of communication also the content was taken into consideration. For example if there was evidence (based on direction of gaze) that the crew was making observations concerning functional status of the process but did not communicate this level information, the communication was seen to be confirmative.

The reactive habit of saying aloud display support system information reflects an objective of information transfer. The reactive habit in communication was that the crew made the decision to switch procedures without contemplating together whether the

Table 4
Examples of crew habits in communication in episode 2.

Crew I in episode 2	Crew B in episode 2	Crew G in episode 2
TO: "Shall we start cooling already [asking from SS]?" **SS:** "Let's take I0 first. And ensure scram." **TO:** "Yes, ensure." **SS:** "Clear, the scrams have been ensured, let's take I0 next".	The crew conducts the scram ensuring in total silence. **TO:** "Turbine trip ensured." **SS:** "Good." **RO:** "Reactor scram ensured." **SS:** "Oh well."	RO is looking at the display support system. **RO:** "Now there is the A0 criterion on. Shall we switch to A0?" **SS:** "Should we then discard this I0? Doesn't A0 always overrule I0?" **RO:** "Yes"
SS turns to fetch the I0 procedures, TO and RO are waiting by SS's desk. SS gives the I0 procedures to TO and RO. **RO:** "Take A0 into use [reading from the display support system]"	TO talks to the field operator via radio. **RO:** "Take A0 into use [reads from the display] is the command." **SS:** "A0?"	SS turns to fetch the new procedure. RO and TO stand around SS's desk waiting for the new procedures
SS turns to look at the wall panels where the safety injection pump indications are located. **SS:** "Did we have YZ24?" **RO:** "We should check, but yes..." **SS:** "Let's take A0"	RO points to the display and repeats**:** "A0 into use". SS looks at the status of safety injection pumps. **SS:** "Oh, we have safety injection pumps running."	
TO and RO return the I0 procedures to the SS who gives them A0 procedures. **SS:** "There are your procedures. So, we have conducted the scarms. At what time?"	SS fetches the A0 procedures. TO is not aware of the situation change because he has been dealing with field operations all the time	
SS checks the time from process monitoring system. **TO:** "Make the announcement [to SS]. **SS:** "Eight something [checking time], yes I will announce [to TO]. Did it order the ensurance yet? **TO:** "Yes."		
SS starts to make the general announcement concerning accident situation		
Analysis: crew is working closely together, talking about both process and automation related issues, making questions to each other, and listening and answering actively. A specific characteristic of this crew was that both operators where in active dialogue with the SS. → interpretative	Analysis: crew members state aloud their observations but do not really talk to each other or build on previous statements. → confirmative	Analysis: No real discussion, only mention procedure hierarchies → reactive

Table 5
Summarized results concerning habits of the crews in episode 2. The grading of habits is depicted as colouring of the table cell. White = reactive, grey = confirmative, black = interpretative.

Crew	Habit of information usage	Habit of communication
A	Safety injection pumps	Display system only
B	Safety injection pumps	Safety injection pumps
C	Safety injection pumps	Safety injection pumps
D	Safety injection pumps	Display system only
E	Display system, Safety injection, Plant protection	Display system, Safety injection, Plant protection
F	Display system, Safety injection, Plant protection	Display system, Safety injection, Plant protection
G	Display system only	Display system only
H	Safety injection pumps	Safety injection pumps
I	Display system, Safety injection, Plant protection	Display system, Safety injection, Plant protection
J	Display system only	Display system only
K	Display system, Safety injection, Plant protection	Display system, Safety injection, Plant protection
L	Safety injection pumps	Display system only

process situation really requires it. They did not discuss together the process parameters. Only the display support system information was made remarks about, thus it seems that the style of communication enforces the interpretation that the process situation is what the procedure has named it.

Concerning communication it is obvious that if the crew has not observed particular information (reactive in information usage) they cannot communicate about it either (see Table 2). But it is interesting that for some crews it seemed to be typical (based on directions of gaze) that individual members may make observations which are not communicated to other members (crews A, D, L). It is also a fact that the crews which were interpretative in utilising information were of that also in communication.

4.1.3. Episode 3: Diagnosis and selecting the event based procedure

The third episode is about approaching the end of procedure A0 accident identification procedure. The operators are supposed to identify the situation based on process parameters and choose the respective EOP into use. This is an important phase in accident management because based on the result of A0 the crew (SS) shall choose the next EOP into use. Making the right decision is crucial for successful accident management. In this case, it should be the procedure A1 as the situation is LOCA. How the SS utilises his crew as a resource in his decision making was the centre of analyses in episode 3.

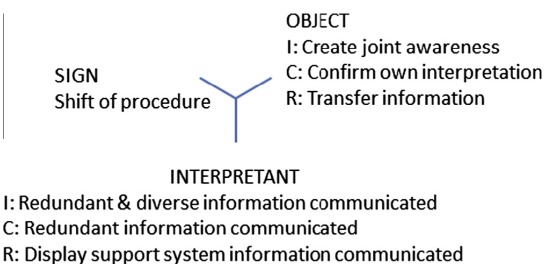

Fig. 6. Habits of communication: differences identified in the data.

In the analysis of the behaviour of the crews in episode 3 there were great differences in the ways the SSs behaved in the situation. The main differences were in coordination and collaboration which in this connection are combined and labelled the function of crew leadership.

The crews' work in this part of the scenario is quite strictly dictated by the A0 procedure. The relevant part of the procedure is depicted below (Fig. 7). The procedure is used by following lines and making choices on which arrow to follow after each statement box. Each box contains a statement and if the statement is true the line downward is followed, if it is false the line to the side is followed. This way, at the end point of this procedure one of the A1, I4, or A2 procedure will be taken into use next.

The SS uses this procedure. The correct path is to procedure A1 through all the boxes i.e. parameter *Y* is not below 20 (false), time is greater than 5 min (true), protection signals *XX–YY* are not on (false), protections signals *ZZ–AA* are not on (false), and parameter *X* is less than 110 bar (true). The way the SSs conducted this task involving different kinds of process and automation checks varied. Examples of crew activities are depicted below (Table 6).

In analysing the habits leadership there were again differences between the crews based on which the classification to interpretative, confirmative, and reactive was applied (Fig. 8).

One way of utilising the procedure by the SS was following the lines with finger and pen, saying aloud each statement, and asking some parameter values from the operators, and after arriving to the conclusion (which is the next procedures) seeking confirmation from the operators also for example by saying: "I get A1, do you agree?". This habit indicates apt use of procedure as a tool to control own behaviour, dialogue both with self and other crew members to avoid misinterpretations and using the operators' collaboration as a resource in ensuring the diagnosis. Thus this habit was analysed to be interpretative. The interpretative habit entails that the SS has a leading role in the decision making but that the whole crew is involved. This brings resilience into the system because it is acknowledged that procedure shift is a critical point in the activity and the diagnosis must be assured utilising each crew members' point of view.

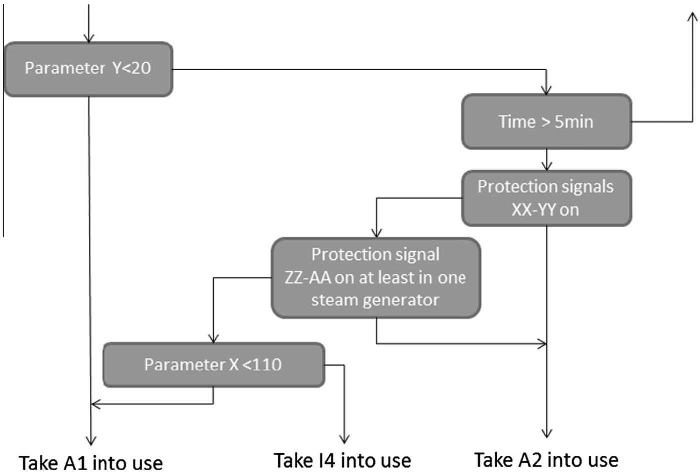

Fig. 7. Part of procedure A0 which is used in episode 3. The procedure is used so that each box is statement which is either true or false. If it is true the line to the side from that box shall be followed, and if the statement is false the line downwards is followed. This way one of the procedures A1, I4, or A2 will be taken into use after A0.

Table 6
Examples of crew activities in leadership.

Crew A in episode 3	Crew B in episode 3	Crew D in episode 3
SS is looking at his own monitors and procedures. The operators are in their own positions looking at process monitoring system displays. TO closes the lids of the scram buttons and returns to his seat. All operators monitor their own displays. **SS:** "Primary-secondary [leak] identification complete, and primary-secondary leak has been identified, John for your information, no primary-secondary leak." **RO [John]:** "OK."	**SS [to TO]**: "Have you conducted primary-secondary identification?" **TO:** "Yeah, no [primary-secondary leak]." **SS:** "And did not identify?" **TO:** "Did not identify." **SS:** "Primary-secondary identification conducted, not identified. RO, did you hear?" **RO:** "Yes, I heard." **SS:** "OK."	SS is walking towards his own desk. He looks at the procedure simultaneously. He speaks in a very low voice, not audible to the others. **SS:** "boiling margin, time."
All three are reading their own procedure. RO is looking at alarm information in the desk alarm system. **SS:** "Boiling margin is, I assume, above XX [real parameter value removed]?" **RO:** "Yes, it is above. And the time is more than five minutes." **SS:** "Yes, it is above five minutes. Let's continue."	All the operators are at their own desks. SS is looking at procedure and monitors. **TO**: "I'm getting A1 from this." **SS**: "Yes."	SS checks the time of the scram from his own display. **SS:** "Yes, continue."
RO walks to the plant protection signal panel, TO is standing next to SS's desk following the discussion alertly. **RO:** "77 [referring to plant protection signal number] no. **SS:** "77, no 91 through 96 have not launched, any of them?" **RO:** "No." **SS:** "Pressure?" **RO:** "It is under XX[real parameter value removed]. Take A1 into use."		RO looks at the protection signal panel and walks up to it. **TO [to SS]**: "What is the situation?" **SS:** "Let's continue." **SS**: "Below 110. Let's take A1 into use.
RO walks back to his desk. **SS:** Yes, let's take A1 into use.		The shift supervisor fetches the new procedures to the operators.
SS turns to fetch the A1 procedures from the shelf behind him		
Analysis: Crew goes through each decision making box in the procedure together. Even though the TO is not saying anything in this case, he is standing with the RO and SS, listening the conversation actively, and following the progress in his own procedure. → interpretative	Analysis: the crew is collaborating as all the operators are involved but they are merely stating their own individual observations. → confirmative	Analysis: SS is talking almost like to himself and is not getting responses from RO and TO. → reactive

The confirmative habit in this scenario was something that which did not have clear inclinations towards either interpretative or reactive. For example confirmative habit was about communicating some of the values to be checked with the crew members but not really dialogically or reflectively i.e. some transparency in diagnosis was made available to the crew. The confirmative habit in leadership meant that in this particular situation the end result was discussed with the whole crew but not the decision making points which lead to it. This is confirmative leadership because the aim is more to acquire confirmation for own diagnosis than to construct the diagnosis together.

Some SSs conducted the whole task silently on their own and only announced the end result: "Take A1 into use". Typical for this behaviour was also that neither pen nor finger was utilised in reading the procedures and following the lines. It may even be suspected that the SS did not truly follow the whole chain but leant

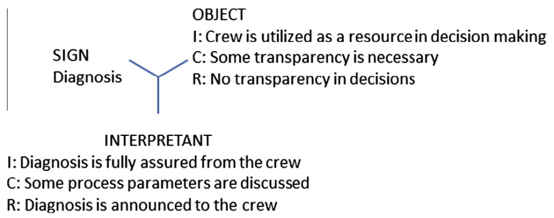

Fig. 8. The shift supervisors' habits of leadership: the differences identified in the data.

on a previously made diagnosis of the situation which determines the next procedure. Also, this way of using the procedure leaves room for possibility for a mistake as no opportunity for dialogue is created in the habit. Thus this habit was analysed to be reactive due to lack of transparency in decision making. The reactive habit in leadership was such which was not really collaborative work. In some crews the SS only announced the next procedure and did not involve the other operators in the decision making process at all.

The summarised results of episode 3 are depicted below (Table 7). Five crews had a interpretative habit (A, E, I, K, and L), four crews a confirmative habit (B, F, G, and H) and three crews a reactive habit (C, D, and J). In the end all the crews arrived in the correct diagnosis and took the correct procedure into use next but as this point in the scenario is very crucial it is very important that the crew and especially the SS adopt habits which are less prone to mistakes and misinterpretations of the process situation and which take full advantage of the available resources.

4.1.4. Summarised crew level results throughout the episodes

The summary of the different crews' habits in the different episodes with regard to the emerged 7 functions is presented below (Table 8). Most of the crews had constructed operating habits that vary on the scale from reactive to interpretative during the scenario. Almost all the crews had habit characteristics from all three classes. Nevertheless, some inclinations towards different habit profiles can be identified. In the following a conclusion has been made that if the crew has 6 or more indicators of certain habit there is a strong inclination towards the respective habit. Secondly, if the majority of indicators is in certain category this is the dominant habit characteristic of the crew. Thirdly, if different habits are represented equally often the habit profile is diffuse. According to this analysis (Table 8) crew J is strongly reactive. Crews A and D are dominantly reactive. Crew B is strongly confirmative as is also crew G. Dominantly confirmative are F and H. Crews C, K, and L are quite diffuse between confirmative and interpretative and crews E and I are dominantly and strongly interpretative respectively (see Table 8).

The point of this analysis is not to make comparisons between the operating crews or to be able to say which crew has constructed the best habits but rather to have an idea of the variation and dominant characteristics of the habits of action of the operating crews. On the top level altogether 7 separate functions were analysed for each of the 12 crews covering 84 separate habit characteristics. Out of those 27 interpretative which yields to 32.1%, 34 confirmative which yields to 40.5%, and 23 were reactive which yields to 27.4%. This result shows that confirmative way of acting is the prevailing habit of the operating crews (Fig. 9).

4.2. Summary: Resilience characteristics in habits of action

In this previous section the differences in operating crews' habits of action in proceduralized accident management were presented. The classification to interpretative, confirmative, and reactive was made contextually based on the profoundness of the connection with the dynamic process evident in the observational data. The present section presents the argumentation based on which the interpretative habits identified in the behavioural data to add to the system level resilience. The summary of the habit characteristics is presented in the end of this section (Table 9).

Altogether in the analyses of the crews' habits of action six different tasks were identified for which the habits were different and could be graded on the scale from interpretative through confirmative to reactive. They are: Information usage, Process situation identification, Dealing with automation, Decision making, Communication, and Leadership. These functions are not, of course, unheard of, as they resemble for example the macro-cognitive functions described in naturalistic decision making (Schraagen et al., 2008) and individual level non-technical skills described by in previous literature (see e.g. Fletcher et al., 2004). In this study, the functions emerged in the micro-level examination of proceduralized activity of operating crews is evidence for their importance

Table 7
Summarized results concerning habits of the crews in episode 3. The grading of habits is depicted as colouring of the table cell. White = reactive, grey = confirmative, black = interpretative.

Crew	Habit of crew leadership
A	All indicators in the loop are communicated, procedure shift is assured from operators
B	Process parameters are discussed but not to full extent
C	SS "announces" the procedure change
D	SS "announces" the procedure change
E	All indicators in the loop are communicated, procedure shift is assured from operators
F	Process parameters are discussed but not to full extent
G	Process parameters are discussed but not to full extent
H	Process parameters are discussed but not to full extent
I	All indicators in the loop are communicated, procedure shift is assured from operators
J	SS "announces" the procedure change
K	All indicators in the loop are communicated, procedure shift is assured from operators
L	All indicators in the loop are communicated, procedure shift is assured from operators

Table 8
Summarised crew level results throughout the episodes. reactive 23, confirmative 34, interpretative 27.

Crew	EPISODE 1				EPISODE 2		EPISODE 3	Habit profile
	Info use	Situa tionin trp.	Deal. w. autom.	Decis. making	Info use	Comm.	Leadership	
A								dominantly reactive
B								dominantly confirmative
C								diffuse btw. confirmative and interpretative
D								diffuse btw. confirmative and reactive
E								dominantly interpretative
F								dominantly confirmative
G								dominantly confirmative
H								dominantly confirmative
I								strong interpretative
J								strong reactive
K								diffuse btw. confirmative and interpretative
L								diffuse between all classes

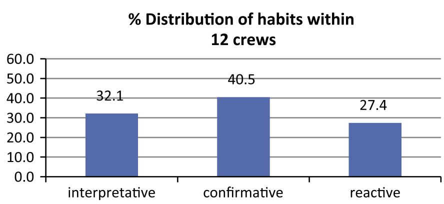

Fig. 9. Percentual distribution of habits of action of the 12 operating crews in a simulated LOCA scenario.

and for the claim that even within proceduralized acting there are differences in how the demands of the procedures are fulfilled.

Characteristic for the interpretative habits is that the crew activity is profoundly connected to the process situation. On a general level this characteristic enables resilience because it becomes possible to observe the possible deviation from the expected process phenomena which is a fundamental pre-requisite for dealing with the unexpected situations. This means that in order to create resilience in the system, process information produced by the automation system shall be widely used: both redundancy and diversity in the types of information to be used are important. Cross-checking is utilised for gathering re-assurance. The interpretations made concerning the process status aim for deep functional understanding of the situation and the possible consequences to the other parts of the process. Dealing with automation reflects joint intelligence in the collaboration. Human agency is reflected in the decision making. The group work can be characterised as active and dialogical. All of these features of habits are such that they may help the operating crew to survive unexpected situations which have not been considered during the design of the system.

The confirmative habits are such that they also result in good outcomes of activity in most cases, but in comparison to the interpretative habits the confirmative habits are less tuned to understanding the process phenomena on a profound functional level and are more concerned with the events that are taking place. Thus, controlling the events with the pre-defined measures (procedures and practices) becomes the focus and objective of activity.

The reactive habits may also produce a relatively good end result in the situations in which the process behaves in a way that has been fully anticipated in the design of the system. In the reactive habit there is little potential for creating new ways of handling the work if it happened for some reason to be required at one point in time.

5. Discussion

In this study the utilisation of pre-defined emergency operating procedures by professional NPP operating crews in simulated

Table 9
Characterization of habits in process control.

	Information usage (episodes 1 and 2)	Interpretation of process situation	Dealing with automation	Decision making	Communication	Leadership
Interpretative	Variety of sources, redundancy and diversity in information sources, dialogue in interpretation of information	Interpretation by considering functional meaning of process events	Human assures the automatic functions. Shared responsibility of human and automation	SS makes decision to scram the process. Human as an active, present agent in decision making	Dialogue concerning process status in the situation. Diverse and redundant information communicated. Reflects creation of joint awareness	Active engagement of each operator in all the decision points. Transparency in contemplation enables to spot false conceptions
Confirmative	Multiple source but taken for granted	Identify the process events based on an existing typology of possible events e.g. a leak	Automation functioning is observed but not taken action on. Reliance on the pre-defined roles of human and automation	Scram is conducted paced by the procedure. Actions are controlled by the procedure	Statements made aloud concerning process parameters. Reflects confirmation of own interpretations	The end result of the decision making process is stated and confirmed by all the operators
Reactive	Variation in information sources not sufficient	Identify that something is going on but now strive to understand or label the situation	Automation information is taken for granted, reflects total reliance on automation	Not identified in the data	Process state is not explicitly mentioned. Transfer of support system information	No real collaboration. SS announces the next steps

accident situations has been scrutinised. The main finding in this study was that within the group of 12 operating crews there were differences in the habits of conducting proceduralized tasks, and specifically within certain identified process control tasks. This means that even though proceduralising drives operating crews towards similar behaviours, individual and crew level variations remain. In the following the results are discussed in the light of previous research, practical value, and methodical reflections.

5.1. Interpretativeness as a source of resilience

In the findings section of this paper we have characterised the differences in the crews' habits by exploiting human–environment relationship characterisation developed originally by Peirce (1998). This characterisation arranges habits on a continuum between reactive and interpretative according to how the environment is considered in the habit. Our main claim is that an interpretative habit of the crew is a mechanism which enables system level resilience in the activity (see also Norros, 2012). Confirmative, and even reactive habits in some cases, may produce adequate end results in the within design bases situations but if a situation is novel, interpretative habits are more appropriate and provide the required capabilities to survive the out-of-design bases situations.

In the preface of a recent guidebook to resilience engineering (Hollnagel et al., 2011) Hollnagel defines resilience 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" (Hollnagel et al., 2011). This definition is a synthesis of several earlier contributions to defining the concept (Hollnagel et al., 2006). The essential abilities for adjustment are identified as follows: the ability to *address the actual* situation, the ability to *address* (and identify) *the critical* factors for safety, the ability to (anticipate) and *address the potential* for safety, and the ability to *address the factual* experience and learn from it (Hollnagel et al., 2011). The interpretativeness characteristics in the habits of action of the crews identified in this study fit well with Hollnagel's more general categories. Addressing the actual situation can be interpreted to be interpretative information usage and decision making: active engagement and presence in the situation. Addressing the critical deals with at least the functions of information usage and situation interpretation identified in this study: The interpretative habits concerning those functions are truly characterised by a capability to assess which information and events are critical and in which ways. To address the potential is reflected at least in the functions of dealing with automation and decision making. By taking an active engaged stance in both the aforementioned functions the crew actually seeks to take into account the possibility for failures in the technical systems. To address the factual means to learn from experience and this is in fact related to interpretativeness in general. Interpretativeness builds on profound understanding of the factual, and reflecting that with the profound perceptions of the current situation in order to create new knowledge and continually develop habits. This way the understanding of what is going on, and

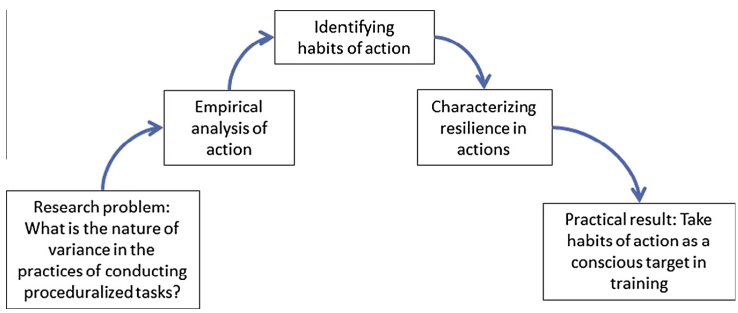

Fig. 10. Research process utilised in the study: From a practical problem, to theoretical considerations on habits of action and finally to practical recommendations.

what is required from the human operators in the situation, builds on experience but is firmly attached to the current situation.

Resilience in the observable behaviour has been previously addressed by Back et al. (2008) and Furniss et al. (2011). The framework to identify and trace resilience composes of three-level conceptual hierarchy: markers, strategy and observations of resilience. Applying the Furniss's resilience framework the characterisations of interpretative habits of action in our findings are on the level of observations of resilience. The difference of the approaches is in the way the observations of behaviour are generalised. Furniss et al. aim for a generalised set of resilience markers which are not connected to any specific context. Our aim in this study was to generalise resilience producing habits specifically in this context and the characterisations of the habits of action are in line with this purpose. Nevertheless, the strategy category level of the framework as it was utilised in the case study in Furniss et al. (2011) closely resembles the grading of habits of action presented in this paper.

An important finding of this study is the identification of the tasks in which the crews' behaviours manifested different habits of action. The tasks (information usage, process situation interpretation, dealing with automation, decision making, communication, and leadership) resemble the non-technical skills described in other domains (Fletcher et al., 2004; Flin and Maran, 2004; Flin, 2007; Saurin et al., 2013) but in essence we agree with Saurin et al. (2013) who make a claim that it is not always possible to differentiate between technical and non-technical skills. Our results are in line with this notion: the value of interpretative habits in these general functions is created by actually combining the "technical" with the "non-technical". For example, in our analysis communication is claimed interpretative only if the content is relevant and meaningful from the point of view of the on-going process situation. Dialogue and joint reflection do not suffice if the content does not have contextual value.

5.2. Methodical reflections

The analyses of the data were time consuming and required quite profound knowledge concerning the nuclear power production process, the simulated accident situation, and the EOP philosophy utilised in the plant, as well as the particular EOP in use in different phases of the scenario. Each video concerning the crew activity was watched repeatedly in order to spot the behaviours of all the three main crew members. Although directions of gazes were also recorded we must of course acknowledge the fact that not all activity is observable for an outsider. As a result of this it must also be said that it is possible that the crews which we analysed to have reactive habits in may have other merits which remained un-identified by us. Also, as the research situation was part of the operating crews' yearly training it may have affected the crew behaviours somehow. For example, it might have been intimidating to have external observers in place.

The analyses of operating practices were conducted using the semiotic concept of habit. Habit as an analytical tool is a powerful concept: it connects the phenomena in the environment (signs) and actions of people (interpretant) and the meaning of both (objective) with the same concept and thus fundamentals of human–environment relationship can be revealed. In characterising different ways to describe nuclear power plant operators' work Carvalho et al. (2009) state that the situated models – based on the characteristics of the environment rather than on the preprogrammed action sequences – have been considered the best way to model the actual control room operators' work. The semiotic concept of habit as it makes explicit the connection between the environment and human behaviour is an important new conceptual tool to develop and further exploit the situated models of human activity.

The research described in this paper travelled on three different conceptual levels (Fig. 10). Firstly, on the first (lowest) level there is the practical research problem: If there is variance in the ways crews conduct proceduralized tasks, and is it significant for safety? Then, on the second level, the empirical data concerning accident management practices is analysed utilising the concept of action. On this level the observable behaviour of the crews was written out as such, with as little prejudice as possible about what is good performance and what is possibly not. On the third level, the level of habits of action, by comparing the observable behaviour of the different crews, the demands of the situation and the signs the crews utilised, different habits of action were identified in the data. The next phase was to come back to the level of actions and answer to a question concerning resilience. Is there something in the particular habit of action which adds to system level resilience? After this the dissecting of the actions and habits of the operating crews it becomes possible to return to the practical level: How should the crews be trained in conducting proceduralized tasks?

5.3. Practical consequences of the findings

In the rhetoric of the safety critical system operations it is common to put proceduralization of activities and professional skills into contradiction (Dekker, 2003). Particularly safety culture training has emphasised the procedure following as an attitude required from the personnel in all situations. But typically there are two sides to this debate: One is to promote proceduralisation by arguing that in order to increase safety, natural human variability shall be decreased by proceduralisation. The whole human reliability analysis (HRA) tradition is based on this idea of strict execution of rules and procedures (De Carvalho, 2006). The opposing argument is that too strict proceduralization lessens the operators' autonomy and freedom to act in a way they themselves consider adequate in a situation and thus decreases the ability to handle unexpected situations.

In the light of the results of this study it is possible to see that perhaps procedures and professional skills are not the opposing ends of the line depicting the locus of control and intelligence in conducting the demanding work. Truly, there were different habits in acceptably conducting the proceduralized tasks. The habits adopted by the operating crews solve the trade-off between proceduralization and autonomy (Papin, 2010): a crew has developed an interpretative habit to conduct the proceduralized task in order to fulfill the requirements of resilience and at the same time adhere to the procedure.

The problem is, in the operative sense, that only about a third of the identified habits were interpretative in nature. The confirmative habits are the most prevailing (40%). It is a challenge to the training of operating crews and also to the design of procedures and user interfaces how these numbers could be turned around and interpretativeness become the prevailing habit of action in process control. This process may start by identifying and characterising the interpretativeness characteristics that should be aimed at. This research contributes to this effort by contextual descriptions of interpretative habits for six different tasks. We see that the level of habits could and should be a conscious target of development in the training of operating crews. The procedure is the backbone on which the crew must be able to rely on, but the existence of procedures does not make the work less demanding. Quite contrary; the skills of intelligent procedure usage must be trained, learned, and maintained.

Similarly, the support for interpretative habits should be taken into account in procedure design. Different procedure design paradigms should be explored concerning their ability to support interpretativeness in acting. For example, the prevailing flow chart based EOP design may be such that it promotes confirmatory habits as it implies that all questions concerning process parameters (decision making points) have a clear "yes or no answer". In most cases this is true but the problem of e.g. oscillating parameter values is left to the operator to solve in the situation.

6. Conclusion

This paper has presented a study in which detailed analysis of NPP operating crew behaviour in simulated accident situations was carried out. Several differences were identified in the habits of carrying out proceduralized tasks. The differences are connected to six different tasks which are a reflection of the general demands in operating work of controlling a complex safety–critical process.

The differences in practices could be identified by exploiting a theory-based grading of operators' interactions with the process, and with each other, that portrays different levels of interpretative power. Based on the theory and drawing on the empirical results it could be claimed that interpretative practices create resilience into the system. The proposed method enables a more concrete and systematic way of identifying practices and their effects on safety compared to other methods available.

The findings show that contrasting procedure guidance and creative operator competence as alternative means of maintaining is superficial. Instead, both are elements of practice and function jointly towards stabilising the environmental variability and better anticipation of the results of action when applied and developed in accordance with the situational demands.

Acknowledgements

This research has been conducted within the Finnish national nuclear safety research program (SAFIR) in projects O'Practice and Hacas. We express our gratitude to project manager Jari Laarni and our colleagues who participated in the data collection of the study: Marja Liinasuo, Hannu Karvonen, and Hanna Koskinen. We would like to acknowledge the crucial role of all the personnel who helped us in organizing the study at the plant: the operators, the simulator trainers and the control room engineers.

References

Back, J., Furniss, D., Hildebrandt, M., Blandford, A., 2008. Resilience markers for safer systems and organisations. SAFECOMP 2008, LNCS 5219, pp. 99–112.
Becker, M.C., 2004. Organizational routines: a review of the literature. Ind. Corp. Change 13 (4), 643–678. http://dx.doi.org/10.1093/icc/dth026.
Carvalho, P.V.R., de Oliveira, M.V., Luquetti dos Santos, I.J., 2009. A computerized tool to evaluate the cognitive compatibility of the emergency operational procedures task flow. Prog. Nucl. Energy 51 (3), 409–419. http://dx.doi.org/10.1016/j.pnucene.2008.09.007.
de Brito, G., 2002. Towards a model for the study of written procedure following in dynamic environments. Reliab. Eng. Syst. Saf. 75 (2), 233–244. http://dx.doi.org/10.1016/S0951-8320(01)00097-7.
De Carvalho, P.V.R., 2006. Ergonomic field studies in a nuclear power plant control room. Prog. Nucl. Energy 48 (1), 51–69. http://dx.doi.org/10.1016/j.pnucene.2005.04.001.
Dekker, S., 2003. Failure to adapt or adaptations that fail: contrasting models on procedures and safety. Appl. Ergon. 34 (3), 233–238. http://dx.doi.org/10.1016/S0003-6870(03)00031-0.
Dekker, S., 2011. Drift into Failure – From Hunting Broken Components to Understanding Complex Systems. Ashgate, Burlington, VT.
Dewey, J., 2002. Human Nature and Conduct. Dover Publications Inc., Toronto.
Dien, Y., 1998. Safety and application of procedures, or how do "they" have to use operating procedures in nuclear power plants. Saf. Sci. 29 (3), 179–187.
Feldman, M.S., 2003. A performative perspective on stability and change in organizational routines. Ind. Corp. Change 12 (4), 727–752, <http://www.ingentaconnect.com/content/oup/indcor/2003/00000012/00000004/art00727>.
Feldman, M.S., Pentland, B.T., 2003. Reconceptualizing organizational routines as a source of flexibility and change. Adm. Sci. Q. 48 (1), 94–118, <http://www.jstor.org/stable/3556620>.
Filippi, G., 2006. Studying computerized emergency operating procedures to evaluate the impact of strong procedure guidance on operators' work practices. In: Proceedings of the American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Controls, and Human–Machine Interface Technologies, Albuquerque, NM.
Filippi, G., Gody, A., 2010. Teamwork and Collaborative Settings. Procedures as a Support for Collaboration of Control Room Crews. Enlarged Halden Program Group Meeting.
Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N., Patey, R., 2004. Rating non-technical skills: developing a behavioural marker system for use in anaesthesia. Cogn. Technol. Work 6 (3), 165–171. http://dx.doi.org/10.1007/s10111-004-0158-y.
Flin, R., 2007. Teaching surgeons about non-technical skills. Surgeon 1 (April), 86–89.
Flin, R., Maran, N., 2004. Identifying and training non-technical skills for teams in acute medicine. Qual. Saf. Health Care 13 (1), i80–884.
Furniss, D., Back, J., Blandford, A., Hildebrandt, M., Broberg, H., 2011. A resilience markers framework for small teams. Reliab. Eng. Syst. Saf. 96 (1), 2–10. http://dx.doi.org/10.1016/j.ress.2010.06.025.
Grote, G., 2012. Safety management in different high-risk domains – all the same? Saf. Sci. 46 (3), 450–460. http://dx.doi.org/10.1016/j.ssci.2011.07.017.
Hollnagel, E., Woods, D., Leveson, N., 2006. Resilience Engineering: Concepts and Precepts. Ashgate, Padstow, GB.
Hollnagel, E., Paries, J., Woods, D., Wreathall, J. (Eds.), 2011. Resilience Engineering in Practice: A Guidebook. Ashgate, Burlington.
Hutchins, E., 1995. Cognition in the Wild. MIT Press, Cambridge, MA.
IAEA, 2000a. Operational Limits and Conditions and Operating Procedures for Nuclear Power Plants. International Atomic Energy Agency, Vienna, Austria.
IAEA, 2000b. Safety of Nuclear Power Plants: Operation. IAEA Safety Standards Series No. Requirements No. NS-R-2. International Atomic Energy Agency, Vienna, Austria.
Le Blanc, K., Oxstrand, J., Waicosky, T., 2012. Model of procedure usage – results from a qualitative study to inform design of computer-based procedures. In: Proceedings of the 8th International Topical Meetings on Nuclear Plant Instrumentation and Control and Human–Machine Interface Technologies.
Nicolini, D., 2013. Practice Theory, Work, & Organization. Oxford University Press, Oxford, UK.
Norros, L., 2004. Acting Under Uncertainty. The Core-task Analysis in Ecological Study of Work. VTT, Espoo.
Norros, L., 2012. Analysis of work practices from the resilience engineering perspective. Int. Electric J. Nucl. Saf. Simul. 3 (4).
Norros, L., Liinasuo, M., Savioja, P., 2011. Operators' conceptions of procedure guidance in NPP process control. In: Proceedings of Enlarged Halden Project Group Meeting, 1, pp. C.
Nuutinen, M., 2006. Expert Identity in Development of Core-task Oriented Working Practices for Mastering Demanding Situations. Six Empirical Studies on Operator Work (Doctoral, University of Helsinki).
O'Hara, J., Higgins, J.C., Stubler, W.F., Kramer, J., 2000. Computer-based Procedure Systems: Technical Basis and Human Factors Review Guidance (Technical Report No. NUREG/CR-6634). Nuclear Regulatory Commission, Washington, DC.
Papin, B., 2010. Operator Guidance in Plant Operation: Characterization and Orientations for Better Team Performance and Socio-technical Systems Resilience. Enlarged Halden Program Group Meeting.
Park, J., 2009. The Complexity of Proceduralized Tasks. Springer-Verlag, London.
Peirce, C.S., 1958. Letters to lade welby. In: Wiener, P. (Ed.), Selected Writings of C.S. Peirce. Dover Publications, New York, pp. 380–432.
Peirce, C.S., 1998. The Harvard lectures on pragmatism. In: Project, T.P.E. (Ed.), The Essential Peirce. Selected Philosophical Writings. Indiana University Press, Bloomington & Indianapolis, pp. 133–241.
Pentland, B.T., Feldman, M.S., 2008. Designing routines: on the folly of designing artifacts, while hoping for patterns of action. Inf. Organ. 18, 235–250.
Perin, C., 2005. Shouldering Risks. The Culture of Control in the Nuclear Power Industry. Princeton University Press, Princeton, NJ.
Roth, E., Mumaw, R.J., Lewis, P., 1994. Empirical Investigation of Operator Performance in Cognitively Demanding Simulated Emergencies (Technical Report No. NUREG/CR-6208). Nuclear Regulatory Commission, Washington, DC.
Salo, L., Norros, L., Savioja, P., 2009. Using operating procedures in NPP process control. Paper presented at the European Conference on Cognitive Ergonomics: Designing Beyond the Product—Understanding Activity and User Experience in Ubiquitous Environments, Helsinki, Finland, pp. 24:1–24:9. <http://portal.acm.org/citation.cfm?id=1690508.1690541>.
Saurin, T.A., Gonzales, 2013. Assessing the compatibility of the management of standardized procedures with the complexity of a sociotechnical system: Case study of a control room in an oil refinery. Appl. Ergon. 44 (5), 811–823. http://dx.doi.org/10.1016/j.apergo.2013.02.003.
Saurin, T.A., Wachs, P., Henriqson, É., 2013. Identification of non-technical skills from the resilience engineering perspective: a case study of an electricity distributor. Saf. Sci. 51 (1), 37. http://dx.doi.org/10.1016/j.ssci.2012.06.010.
Savioja, P., Norros, L., Salo, L., 2012. Functional Situation Models in Analyses of Operating Practices in Complex Work. ECCE 2012 Re-Thinking Cognition.
Schraagen, J.-., Militello, L., Omerod, T., Lipshitz, R., 2008. Naturalistic Decision Making and Macrocognition. Ashgate, Aldershot.
Suchman, L., 1987. Plans and Situated Actions: The Problem of Human Computer Interaction. Cambridge University Press.
Throneburg, E.B., Jones, J.M., 2006. Situation awareness and trust in computer-based procedures in nuclear power plant operations. In: Proceedings of the American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Controls, and Human–Machine Interface Technologies, Albuquerque.
Vicente, K.J., 1999. Cognitive Work Analysis. Toward A Safe, Productive, and Healthy Computer-based Work. Lawrence Erlbaum Publishers, Mahwah, NJ.
Wright, P., McCarthy, J., 2003. Analysis of procedure following as concerned work. In: Hollnagel, E. (Ed.), Handbook of Cognitive Task Design. Lawrence Erlbaum Associates, Mahwah, New Jersey, p. 679.

PAPER VI

User experience: does it matter in complex systems?

Accepted for publication 2013 in
Cognition, Technology and Work.

Paper VI of this publication is not included in the PDF version.

Series title and number

VTT Science 57

Title	**Evaluating systems usability in complex work** **Development of a systemic usability concept to benefit control room design**
Author(s)	Paula Savioja
Abstract	The design of industrial control rooms assumes fulfilling the goals of production, safety, and human well-being. Control rooms and the user interfaces within them should enable the effective and efficient conduct of work in all the operating conditions which can be foreseen during specification and design. At the same time, the user interfaces should enable the control of the process in unprecedented and totally unexpected situations while simultaneously maintaining the safety of the process. During the design or modification of the control room, the potentiality of the emerging solution to fulfil the objectives is assessed by conducting empirical evaluations. This dissertation presents the development of an evaluation methodology which enables developing the control room towards meeting these objectives. Industrial process control constitutes a socio-technical system in which people and technologies have multiple and sometimes overlapping roles. In order to meet the demands of maintaining safety in all situations, the socio-technical system should have built-in capability of dealing with the unexpected. The control room, and the user interfaces within it, are an integral part of the socio-technical system. Thus, they have a role in construing and maintaining the safety of the system. The concept of systems usability (SU) is introduced in the dissertation to evaluate the systemic effects of control room solutions. SU is a human-centred quality attribute of user interfaces and control rooms attributed to technology, but the value in the use of the technology is evidenced in the success of the activity in which the technology is used. Thus, the research makes sense of the significance of the individual technological solutions in, and for, the entirety of an activity system. Systems usability means that a tool in an activity serves the functions of 1) an instrument, 2) a psychological tool, and 3) a communicative tool. The meaning of each function in the specific domain is contextually defined. Furthermore, the quality of the tool can be assessed utilising different perspectives on the usage activity: performance, way of acting, and user experience. By combining the functions of the tool and perspectives on activity, a systemic framework for developing contextual indicators for a good control room is construed. Utilising the concept of SU in the control room requires a model-based evaluation approach. This means that the general contextual work demands are considered in defining the reference of a successful process control activity. In addition, the scenarios utilised in the evaluation are modelled also taking the general work demands into account. A specific scenario modelling method, functional situation modelling (FSM), is presented in this dissertation. FSM combines a functional and a chronological view of the activity of an operating crew in a particular situation. By making explicit the connection between required operating actions and critical domain functions, the model lays the ground for analysing operating activity from the point of view of maintaining the critical functions. On aspect of control room evaluation is identifying whether the operating practices are attuned to maintaining safety in all situations. For this purpose, a data analysis method is presented in this dissertation. The analysis is based on identifying how operators identify and interpret signs depicted in the control room. Signs are, for example, information technological representations of process information, but they may also be the activities of other crew members. If the interpretation of signs has an identifiable global safety-related aspect, it may be concluded that the operating practice, even though situated, is also attuned to the general functions of the work which must always be maintained. A particular viewpoint in evaluating SU is provided by analysing user experiences (UX) which emerge in the complex work. For this purpose, a UX questionnaire was developed within the SU evaluation framework. The questionnaire is based on UX indicators which reflect experiences of appropriateness concerning the three tool functions (instrument, psychological, communicative). The importance of UX as a measure of germinating, not yet existing tool appropriateness, is presented in the dissertation. The contribution this dissertation makes is in the intersection of the research fields of human factors and ergonomics (HF/E) and usability engineering. The theoretical foundations of the research are in activity theory and cognitive systems engineering. The empirical work has been conducted by following the control room modernization efforts of Finnish nuclear power plants, during which evaluations have been carried out.
ISBN, ISSN	ISBN 978-951-38-8145-0 (Soft back ed.) ISBN 978-951-38-8146-7 (URL: http://www.vtt.fi/publications/index.jsp) ISSN-L 2242-119X ISSN 2242-119X (Print) ISSN 2242-1203 (Online)
Date	June 2014
Language	English, Finnish abstract
Pages	168 p. + app. 116 p.
Keywords	Control room evaluation, human factors, systems usability, control room design, nuclear power plant
Publisher	VTT Technical Research Centre of Finland P.O. Box 1000, FI-02044 VTT, Finland, Tel. +358 20 722 111

Julkaisun sarja ja numero

VTT Science 57

Nimeke	**Systeemikäytettävyyden arviointi monimutkaisessa työssä** **Systeemisen käytettävyyskäsitteen kehittäminen valvomosuunnittelun tueksi**
Tekijä(t)	Paula Savioja
Tiivistelmä	Teollisuudessa käytettävien valvomoiden suunnittelu tähtää tuotannon, turvallisuuden ja ihmisten hyvinvoinnin tavoitteiden saavuttamiseen. Valvomoiden ja niissä käytettävien käyttöliittymien tulisi olla sellaisia, että ne mahdollistavat tuloksellisen ja tehokkaan työskentelyn kaikissa suunnitteluvaiheessa ennakoitavissa olevissa tuotannollisissa tilanteissa. Samalla käyttöliittymien tulisi mahdollistaa prosessinhallinta myös ennalta odottamattomissa tilanteissa turvallisuustavoitteista tinkimättä. Kehittyvää valvomoratkaisua arvioidaan valvomon suunnittelu- ja muutosprojekteissa tekemällä empiirisiä kokeita. Tämä väitöstyö käsittelee sellaisen valvomoarviointimenettelyn kehittämistä, joka mahdollistaa valvomon kehittämisen aikaisen arvioinnin edellä mainittujen yleisten tavoitteiden suhteen. Teollinen prosessinhallinta muodostaa sosioteknisen järjestelmän, jossa ihmisillä ja teknologioilla on useita, usein päällekkäisiä rooleja. Jotta sosiotekninen järjestelmä voisi saavuttaa turvallisuustavoitteet kaikissa tilanteissa, siinä tulisi olla sisäänrakennettua kyvykkyyttä käsitellä epävarmuuksia. Valvomo ja sen sisältämät käyttöliittymät ovat keskeinen osa sosioteknistä järjestelmää. Siten niiden rooli turvallisuuden rakentumisessa ja ylläpidossa on olennainen. Tässä väitöstyössä esitellään systeemikäytettävyyden käsite, jonka avulla voidaan eritellä valvomoratkaisujen systeemisiä vaikutuksia suunnittelun aikaisissa valvomoarvioinneissa. Systeemikäytettävyys on valvomoiden käyttöliittymien ihmiskeskeinen laatutekijä ja sinällään teknologian ominaisuus, mutta teknologian käyttöarvo liittyy aina toimintaan, jossa sitä käytetään. Systeemikäytettävyyden käsitteen kautta tämä tutkimus käsittelee siis teknologioiden merkitystä osana toimintajärjestelmää ja sen tavoitteiden saavuttamista. Systeemikäytettävyys erittelee työvälineen funktioita toimintajärjestelmässä toiminnan teorian pohjalta: Työväline on 1) instrumentti, 2) psykologinen väline ja 3) kommunikatiivinen väline. Jokaisen funktion merkitys tietyssä toiminnassa määritellään asiayhteydessään. Työvälineen laatua voidaan tarkastella kolmesta eri näkökulmasta toimintaan: suoritus, toimintatapa ja kokemus. Yhdistämällä työvälineen funktiot ja tarkastelunäkökulmat saadaan kehikko kontekstuaalisten hyvän valvomon indikaattoreiden määrittämiseen. Systeemikäytettävyyden käsitteen käyttäminen valvomoarvioinnissa vaatii mallipohjaista arviointia, jossa työn yleiset vaatimukset ymmärretään ja otetaan huomioon arviointiskenaarioiden suunnittelussa. Tässä väitöstyössä on kehitetty erityinen skenaariomallinnustekniikka: funktionaalinen tilannemallinnus (FSM, Functional situation model). FSM yhdistää funktionaalisen ja kronologisen näkökulman operaattorivuoron toimintaan tietyssä tilanteessa. Tekemällä eksplisiittiseksi yhteydet vaadittujen operaatioiden ja prosessin kriittisten funktioiden välillä malli luo perustan operaattoritoiminnan analyysille prosessin kriittisten funktioiden näkökulmasta. Arvioinnissa on kiinnitettävä huomiota kehittyviin operaattorityön käytäntöihin ja siihen, että ne yleisesti tähtäävät turvallisuuden varmistamiseen kaikissa tilanteissa. Tätä tarkoitusta varten kehitetty operaattoritoiminnan analyysimenetelmä esitellään tässä väitöstyössä. Analyysi perustuu siihen, miten operaattorivuorot havaitsevat ja tulkitsevat valvomossa esiintyviä prosessista kertovia merkkejä. Merkit voivat olla informaatioteknologisia, mutta ne voivat olla myös esimerkiksi vuoron muiden jäsenten tekoja. Jos merkkien tulkinnassa pystytään tunnistamaan globaaleja yleisiä piirteitä, voidaan päätellä, että toiminta suuntautuu yleisten tavoitteiden, esimerkiksi turvallisuuden, ylläpitämiseen. Samalla tunnistetaan myös välittömien prosessinhallintaan liittyvien tekojen toteutuminen tai toteutumatta jääminen. Valvomoarvioinneissa on myös syytä tarkastella monimutkaisessa työssä ilmeneviä käyttäjäkokemuksia. Ydinvoimalaitoksen operaattoreiden käyttäjäkokemuksia tutkittiin kolmessa eri teknologiaolosuhteessa. Käyttäjäkokemuksia tutkittiin kyselymenetelmän avulla, joka on kehitetty systeemikäytettävyyskehikon pohjalta. Työssä esitellään käyttäjäkokemusindikaattorit kaikille kolmelle välinefunktiolle (instrumentti, psykologinen väline, kommunikatiivinen väline) ja eritellään kolmen eri tutkimuksen tuloksia. Työssä tarkastellaan myös käyttäjäkokemuksen merkitystä kehittyvän, ei vielä valmiin työvälineen arvioinnissa. Väitöstyön kontribuutio on inhimillisten tekijöiden ja ergonomian (human factors and ergonomics) ja käytettävyyssuunnittelun (usability engineering) välialueella. Työn teoreettinen perusta on toiminnan teoriassa ja kognitiivisessa järjestelmäsuunnittelussa (cognitive systems engineering). Empiirinen osuus on tehty seuraamalla suomalaisten ydinvoimalaitosten valvomokehityshankkeita, joiden yhteydessä erilaisia systeemikäytettävyysarviointeja on suoritettu.
ISBN, ISSN	ISBN 978-951-38-8145-0 (nid.) ISBN 978-951-38-8146-7 (URL: http://www.vtt.fi/publications/index.jsp) ISSN-L 2242-119X ISSN 2242-119X (painettu) ISSN 2242-1203 (verkkojulkaisu)
Julkaisuaika	Kesäkuu 2014
Kieli	Englanti, suomenkielinen tiivistelmä
Sivumäärä	168 s. + liitt. 116 s.
Avainsanat	Control room evaluation, human factors, systems usability, control room design, nuclear power plant
Julkaisija	VTT PL 1000, 02044 VTT, puh. 020 722 111

Evaluating systems usability in complex work

Development of a systemic usability concept to benefit control room design

Today, digital technologies are used widely in industrial settings and they are making their way also to the control rooms in nuclear power plants (NPP). In developing technologies for NPP control rooms it is crucial to address safety issues as well as the goals of production and human well-being. This dissertation presents an evaluation approach that can be used to 1) assess safety implications of new technologies and 2) improve design solutions based on feedback from the usage situations.

Industrial process control constitutes a socio-technical system in which people and technologies have multiple and sometimes overlapping roles. In order to meet the demands of maintaining safety in all situations, the socio-technical system must have built-in capability of dealing with the unexpected. The control room, and the user interfaces within it, are an integral part of the socio-technical system. Thus, they have a role in construing and maintaining the safety of the system. The dissertation introduces the concept of systems usability (SU) through which the systemic effects of control room solutions can be assessed. SU is a human-centred quality attribute of user interfaces and control rooms attributed to technology, but the value-in-use of the technology is evidenced in the success of the activity system in meeting its objectives.

The contribution this dissertation makes is in the intersection of the research fields of human factors and ergonomics (HF/E) and usability engineering. The theoretical foundations of the research are in activity theory and cognitive systems engineering. The empirical work has been conducted by following the control room modernization efforts of Finnish nuclear power plants, during which evaluations have been carried out.

ISBN 978-951-38-8145-0 (Soft back ed.)
ISBN 978-951-38-8146-7 (URL: http://www.vtt.fi/publications/index.jsp)
ISSN-L 2242-119X
ISSN 2242-119X (Print)
ISSN 2242-1203 (Online)