A comparison of human factors evaluation approaches for nuclear power plant control room assessment and their relation to levels of design decision specificity

This document has been downloaded from Chalmers Publication Library (CPL). It is the author’s version of a work that was accepted for publication in:

Citation for the published paper:

Downloaded from: http://publications.lib.chalmers.se/publication/251365

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.
A comparison of human factors evaluation approaches for nuclear power plant control room assessment and their relation to levels of design decision specificity

Simonsen, Eva
Chalmers University of Technology, Department of Industrial and Materials Science, Division of Design & Human factors
eva.simonsen@chalmers.se

Many design decisions must be made and repeatedly evaluated during the development process to form a nuclear power plant control room system that supports safe operation. The purpose of this paper is to compare utilised approaches to evaluate nuclear power plant control room systems and explore how they relate to design decisions at different levels of specificity. The method used was a review of academic literature. The result showed that evaluation of more specific design decisions is largely addressed. However, there is a need to further develop methodologies and methods for formative evaluation of more general design decisions to support assessment earlier in the development process.

Keywords: Nuclear power, control room, human factors, evaluation method, development process, design decision, early evaluation

1. Introduction
Nuclear power is a high-risk industry where safe operation is crucial. The purpose of a nuclear power plant is to produce electricity, but this process involves risks. The International Atomic Energy Agency (2007) defines nuclear safety as “the achievement of proper operating conditions, prevention of accidents or mitigation of accident consequences, resulting in protection of workers, the public and the environment from undue radiation hazards” (ibid, pp.133). Widening the definition to include hazards other than radiation, safe operation of a nuclear power plant is thus the production of electricity without exposing workers, the public or the environment to undue hazards.

A nuclear power plant is operated from a central control room: the plant’s core functional entity with associated physical structure, where operators are stationed to carry out centralised control, monitoring and administrative responsibilities (International Standard Organisation, 2000). The physical structure of the control room, together with the operators and organisational structures such as work routines, can be viewed as a socio-technical system (for example using the definition by Hendrick and Kleiner, 2001), a control room system.

According to the IAEA International Nuclear Safety Advisory Group (1999), human factors is one of the underlying principles of nuclear safety. The discipline of human factors as a profession is concerned with optimising human well-being and overall system performance through the application of theory, principles, data and methods to design (International Ergonomics Association, 2016). Applying a human factors perspective when designing a control room system is thus a necessary approach to achieve safe operation.

Nuclear power plants do not remain unchanged during their lifetime, components must be upgraded when spare parts are no longer available and new technology allows
modernisations that enhance efficiency. This is true for equipment in the control room as well. In addition, since the control room must maintain the ability to control the process, changes in the plant’s process systems will consequently require changes in the control room. Due to its central function in the operation of the plant, all changes to the control room system have the potential to impact safety (Norros and Nuutinen, 2005).

According to Ullman (1997), design is the successive development and application of constraints to reduce the number of potential solutions to a problem, until only one unique product remains. Developing and applying constraints is another way of saying that the possible values of design variables are constrained, in other words it is the making of design decisions (Bligård et al., 2016). When successively narrowing the solution space through the application of constraints, the design variables considered are more and more specific. Constricting the value of a design variable will in turn limit the possible underlying and dependent design variables that can be considered. A design decision is when a design variable is given a specific value. For example, the design variable “colour” can be given the value “red” – making “red” a design decision. A further specification of the variable would be to decide on a specific colour code. The natural order in a development process is to gradually move from more general to more specific design decisions, and phases in development processes are often differentiated based on the specificity of design decisions considered in that phase.

The need to address safety issues during the design process is supported by reviews of accident and incident data, which suggests that 20-50% of accidents and incidents have their root causes in design (Kinnersley and Roelen, 2007; Taylor, 2007). A multitude of design variables must be decided to form a control room system, and identifying and specifying them to create a control room system design that supports safe operation is not a trivial task.

Evaluation is an activity that can be used to navigate among this multitude of design decisions in the development process. A report from a Nuclear Energy Agency committee (OECD/NEA Committee on Safety of Nuclear Installations, 2005), stated that the process for modification of nuclear power plants should include actions to verify the fulfilment of requirements and validate the appropriateness of the modification. Hale et al. (2007), when comparing development processes for complex technical systems involving major accident hazards, noted that one similarity between the processes was “the idea of waypoints at which the safety of the design is checked, before moving on to the next phase” (ibid, pp. 312). According to Hale et al. (2007), this iteration of safety checks ensures that safety issues are kept in focus as the design process progresses.

In a development process, the ideal is to evaluate constraints on design variables as soon as they are set, in order to avoid having to reconsider the constraints on underlying design variables. Because of this, during the development process constant evaluation is not only important for safety issues (as stated by Hale et al., 2007), but for all constraints on design variables that form the solution, to ensure that the design achieves the desired effect. For this to be possible the evaluation approach used must be able to assess design decisions at different levels of specificity.

Design decisions must be represented in some way for evaluation to be possible. This representation can take many forms (see for instance Nielsen, 1993; Broberg et al., 2011), for example a description in text form or a wooden mock-up with paper printouts representing the operator interface. Because of the gradual specification of design variables in the development process, it is possible to use a representation of more specific design decisions to evaluate the more general design decisions preceding them, but not the other way round. The
representation of design decisions used in an evaluation is thus a determining factor for the level of design decisions the evaluation activity is able to assess.

The purpose of this paper is to compare utilised approaches to evaluate control room systems in the nuclear power industry and to explore how they relate to design decisions at different levels of specificity. The assumption behind this purpose is that identified gaps shown by this comparison and mapping of evaluation approaches to design decision levels should indicate needs for further development of evaluation approaches.

2. Evaluation, Verification, and Validation
The dictionary definition of “evaluate” is to “determine the value or condition of usually by careful study” (Britannica Online, 2017). Nielsen (1993) differentiates between two kinds of evaluation activities depending on their purpose: formative and summative evaluation. Formative evaluations are done to improve the design as part of an iterative development process. Summative evaluations are done to assess the overall quality of the design. Both formative and summative evaluations may be performed during the course of the development process, and an evaluation activity may fulfil both formative and summative purposes at the same time. There is little use, however, in performing a formative evaluation at the end of the development process because of the limitations in available time and resources to actually improve the design.

The terms ‘verification and validation’ (V&V) are commonly used in literature dealing with control room evaluation. In an ergonomics standard for the ergonomic design of control centres (International Standard Organisation, 2006, p.1) the evaluation process is defined as the “combined effort of all verification and validation (V&V) activities in a project using selected methods and the recording of the results”. Evaluation here is an overall concept, and verification and validation are specific activity types. Verification is defined as “confirmation, through the provision of objective evidence, that specified requirements have been fulfilled” (International Standard Organisation, 2006, p. 2) and validation as “confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled” (ibid). These definitions propose that verification focuses on details whereas validation is a more holistic assessment. The purpose of verification and validation activities is to assess and provide proof of the quality of the design, thus making them summative according to the definition given by Nielsen (1993).

3. Method
The method used in this paper was a literature review of academic literature on the subject. More specifically, the comparison of utilised approaches to evaluate control room systems in the nuclear power industry and the exploration of how they relate to design decisions at different levels of specificity was executed in two steps.

In the first step, approaches were utilised to evaluate control room systems in the nuclear power industry sought through a search in a scientific database. The Scopus database was searched using the search string “("nuclear power plant" AND "control room") AND ("evaluation" OR "validation" OR "verification" OR "assessment")”, limiting the search to texts from the last ten years (2007-2017). From 241 papers in the list of search results, 35 papers concerning evaluations proposed or performed in industry were identified based on the contents of titles and abstracts. Papers describing more research-oriented control room design and evaluation method studies were omitted to focus the review on approaches better adapted to conditions in industry projects. Papers not in English or unrelated to the topic of the study were omitted.
The review of the 35 papers focused on determining if the proposed or performed evaluation activities were formative or summative, and on comparing the methodology used, especially the system representation used for the assessment.

The second step of the methodology explored how the identified evaluation approaches related to design decisions at different levels of specificity. A common reference for different design decision levels was needed to allow comparison between different evaluation approaches. The design levels of the ACD³-framework (a product development mapping tool visualising design decisions) and development process phases derived from them was chosen to serve as this reference (Bligård et al., 2016). In the present paper, the design levels are viewed as “falling”, with design decisions being more general at the top (higher levels) and more specific at the bottom (lower levels). The ACD³-framework defines the different design levels in the following way:

- **Effect (Needfinding phase):** The effect that the machine is intended to achieve in the context (the term ‘machine’ is defined as the artefact the end users will be interacting with, i.e. the product being developed)
- **Usage (Design of use phase):** The use of the machine by humans
- **Architecture (Overall design phase):** The technical architecture of the machine
- **Interaction (Detailed design phase):** The interaction between human/context and the machine in detail
- **Element (Structural design phase):** The technical elements of the machine

Each of the 35 papers in the review was mapped to the design levels the assessed design was in when it was evaluated. In many of the papers the descriptions of the specificity of the evaluated designs were not very elaborate. For these cases the specificity of the evaluated design was assumed to be the same as the representation used for the assessment. For example, if a mock-up containing detailed operator interfaces was used, the design decisions to be evaluated were assumed to be on the interaction level. In the review, a distinction was made between clear mappings (where the level of specificity of the evaluated design was clearly specified in the reviewed paper) and estimated mappings (where the level of specificity of the evaluated design was difficult to determine and had to be estimated).

### 4. Results

This section describes similarities and differences between the stated purpose of the reviewed evaluation approaches and other aspects of the methodology used (such as the system representation used). The mapping of the design levels the assessed design was in when it was evaluated for the different papers is also presented.

#### 4.1. Evaluation of Nuclear Power Plant Control Rooms in Industry

The review revealed that the papers by De La Garza et al. (2012) and Labarthe and De La Garza (2011) were based on the same project; so too were the papers by Hwang et al. (2009) and Yang et al. (2009); as were the papers by Rivere (2015) and Rivere et al. (2015); as also Gunnarsson and Eliasson (2010) and Hill et al. (2009); and Jia et al. (2014) and Jia and Zhang (2014).

Of the reviewed papers, some focused on verification and validation, and described these as summative activities (Ha et al., 2007; Lin et al., 2009; Lee and Chung, 2012; Suh et al., 2013; Jia et al., 2014; Jia and Zhang, 2014; Sun et al., 2016). Others described evaluation activities, sometimes called V&V, that fulfilled both summative and formative purposes (Alonso et al., 2008; Dobos et al., 2010; Zhang et al., 2010; Labarthe and De La Garza, 2011; De
La Garza et al., 2012; Liu et al., 2012; Eisner et al., 2015; Rivere, 2015; Rivere et al., 2015). Some of the papers described approaches where formative and summative evaluations were separate activities (Carvalho et al., 2008; Chuang and Chou, 2008; Manrique and Valdivia, 2008; Gray and Basu, 2009; Hill et al., 2009; Gunnarsson and Eliasson, 2010; Rejas, 2010; Roth et al., 2010; Song and Zhang, 2010; Pfledderer, 2012; Boring, 2014). Five papers described approaches where the main purpose was formative, to provide input to design (Huang et al., 2007; Hwang et al., 2009; Yang et al., 2009; Liang and Chen, 2010; Hanes et al., 2015). In the remaining papers the purpose of the evaluations was not clearly stated (Anokhin and Marshall, 2007; Jones et al., 2007; Anokhin and Marshall, 2009).

In the reviewed papers simulators of some sort, often full-scale, were used when performing validation activities. The accounts of summative activities (especially validation) were also more detailed in terms of describing the methodology used than the accounts of formative evaluations. Overall, the reviewed papers focused more on the overall structure of the evaluation process and the system representation used, than on the detailed methodology used in the different evaluation activities.

4.2 Relation to levels of design decisions

The specificity of the design assessed in evaluation activities in the reviewed papers is presented in Table 1. To denote the difference in certainty in the mappings, clear mappings are indicated with dark grey and estimated mappings with light grey.

Table 1: The specificity of the design assessed in evaluation activities in the reviewed papers (dark grey indicates a clear mapping, light grey indicates an estimated mapping). Each evaluation approach is mapped against the design levels of Bligård et al. (2016).

<table>
<thead>
<tr>
<th>Reviewed papers</th>
<th>Effect (Needfinding)</th>
<th>Usage (Design of use)</th>
<th>Architecture (Overall design)</th>
<th>Interaction (Detailed design)</th>
<th>Element (Structural design)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manrique and Valdivia (2008); Song and Zhang (2010)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Chuang and Chou (2008), Dobos et al. (2010), Liu et al. (2012), Pfledderer (2012), Hanes et al. (2015)</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Hill et al. (2009), Gunnarsson and Eliasson (2010), Zhang et al. (2010), Boring (2014), Eisner et al. (2015)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Ha et al. (2007)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Sum of evaluation approaches mapped to the level</strong></td>
<td><strong>5</strong></td>
<td><strong>5</strong></td>
<td><strong>14</strong></td>
<td><strong>30</strong></td>
<td><strong>30</strong></td>
</tr>
</tbody>
</table>
Counting papers based on the same project as one, the papers reviewed showed that five evaluation approaches could be mapped to the Effect and Usage levels, 14 to the Architecture level, and 30 to the Interaction and Structure levels respectively. Also, the mappings for the higher levels were not as clear as for the lower levels. For example none of the papers clearly described an approach that assessed a design at the Effect and Usage levels.

5. Discussion
The purpose of this paper was to compare utilised approaches to evaluate control room systems in the nuclear power industry and to explore how they relate to design decisions at different levels of specificity. However, not all control room evaluations executed at nuclear power plants are reported in academic literature. Consequently, this search does not provide a comprehensive account of all control room evaluations in practice, but it does shed some light on the experiences that are shared for others to learn from and build upon.

The 35 reviewed evaluation approaches could to a greater degree be mapped to lower levels of design decision specificity. Mapping to higher levels was also more uncertain due to less detail in the descriptions. For many of the reviewed control room evaluations in practice, the descriptions focused on the need for evaluation and the system representations used, and did not reveal much detail on the exact method used. This was especially true for evaluation of design decisions at higher levels.

Even though lower level design decisions may be considered before the higher level decisions, the higher level design decisions must be finalised before the lower level decisions. This explains the connection between the design levels and the process phases of the ACD3-framework (Bligård et al., 2016). The possibility to make changes in a design decreases with its finalisation, which makes formative evaluation more worthwhile earlier in a development process. There is thus a connection between higher level design decisions, formative evaluation, and earlier process phases. Consequently, the lack of details on methodology for assessing higher level design decisions is a lack of detail in methodology for early evaluations, especially formative ones. The most detailed evaluation approach descriptions were given for the activity Integrated System Validation, a final summative assessment of the control room. The same observation was made in one of the reviewed papers, where Eisner et al. (2015) note that there is less guidance from codes and standards to be found for verification and validation activities that are not Integrated System Validation.

Some researchers have focused specifically on early evaluation of control room systems. Two such examples are Boring and Lau (2017) and Boring (2017), who argued for evaluation approaches that relieve the Integrated System Validation activity of some of its burden in providing evidence of acceptability, promoting complementary evaluation activities earlier in the development process. In the paper by Boring (2017) some suggestions for suitable methods are given, but the focus is on the overall evaluation process structure and general advice for this rather than detailed description of methods and methodology.

Another group of researchers who have addressed the issue of early evaluation are Laarni et al. (2011), who conducted a series of small-scale usability tests preceding the final validation. More specifically, the methods used were usability test, expert evaluation, cognitive walkthrough, focus group, and usability questionnaires. Simulator testing was seen as a central task, but human-system interface-oriented walkthroughs using screen/paper mock-ups were also mentioned. The approach presented in Laarni et al. (2011) was set in the context of the verification and validation process in Laarni et al. (2014). Here, the authors
described a stepwise validation approach where sub-systems were validated successively before the final validation. The paper focused on how this stepwise approach builds evidence for the final assessment of design acceptability, but also acknowledged the evaluation activities’ contribution to improving the design.

Simulator testing was seen as a central task in the evaluation approach proposed by Laarni et al. (2011). Validation activities in the reviewed control room evaluations were also reliant on simulators of some sort, often full-scale. This is a trait shared by other research on nuclear power plant control room evaluation which describes methods or methodology that require the use of a physical simulator of some sort, often a full-scale simulator (for example Le Blanc et al., 2010; Jang et al., 2011; Braarud et al., 2015; Gibson, 2015). Other approaches rely on 3D models and virtual simulators (for example Tran et al., 2007; Luquetti dos Santos et al., 2009; Gatto et al., 2013; Chen et al., 2014; Yan et al., 2014).

A simulator for a proposed design is typically not developed or upgraded until later in the development project, making these types of approaches unsuitable for evaluating design decisions at higher levels of specificity. 3D models or virtual simulators, since they are cheaper to develop and change, can be created and used earlier than physical simulators. Evaluation approaches using this kind of representation can thus be used for design decisions at higher levels. Still, if the virtual simulator includes operator interfaces it cannot be used until the design is at the Interaction level, thus limiting the possibility to assess design decisions at higher levels when they are taken (often earlier in the development process).

Human reliability analysis (HRA) methods are a group of methods typically used to identify sources of errors in use (typically called human errors) and quantify how likely it is for such errors to occur. In the nuclear industry, it is typically a part of the probabilistic safety assessment and not a part of the development process (Boring and Bye, 2008). Some have however argued that HRA should be done as part of the design process (see for example United States Nuclear Regulatory Commission (2012) and (Boring and Bye, 2008)). Boring and Bye (2008) propose that “a thorough HRA at the design phase can help ensure that an unbuilt system is safe” and that “such an analysis may also allow effective comparisons between competing designs or prioritization of design issues” (ibid, pp 736). One of the evaluation approaches reviewed in this paper, by Manrique and Valdivia (2008), utilised HRA in a manner similar to this, to evaluate the risk of use errors early on. Another reviewed approach, by De La Garza et al. (2012), proposed better integration of HRA and other human factors activities. Other reviewed approaches mentioned HRA as a source for input to V&V (Song and Zhang, 2010; Zhang et al., 2010; Lee and Chung, 2012; Li et al., 2012; Rivere, 2015), or merely that it is a part of the human factors engineering process (Rejas, 2010; Roth et al., 2010; Pfledderer, 2012; Rivere et al., 2015). The rest of the reviewed articles do not mention HRA at all.

Many decisions must be made when planning an evaluation activity, and they must be made consciously in order not to risk skewing the evaluation result. More detailed information on how to implement methods for evaluating design decisions at higher levels in control room designs would benefit evaluation planning and execution. HRA methods are one topic that is interesting to explore further, together with the reviewed evaluation approaches mapped to higher design decision levels found in this paper. Further work should also endeavour to develop or identify methods that do not require the use of system representations developed later in the development process (such as simulators), to allow formative evaluations earlier and thus reduce the risk of late and expensive changes.
6. Conclusions
The literature review presented in this paper showed that formative evaluation approaches for design decisions of higher levels of specificity are less common and not described in as much detail as summative evaluations for lower level design decisions. This gap has to some extent been addressed by academia, but guidance can be further detailed and improved, for example by further investigating evaluation approaches utilising system representations available in earlier project phases (when more general design decisions are normally made).

Much can be gained from assessing control room system design decisions at higher levels, since this means design concepts can be evaluated earlier in the development process, making changes easier and cheaper to implement. There is a need to further develop methodologies and methods suitable for formative evaluation of design decisions at higher levels, and to assess their applicability for control room system evaluation.

Acknowledgements
The author would like to thank Dr Lars-Ola Bligård and Professor Anna-Lisa Osvalder for valuable discussions and feedback during the development of this paper.

References


