Safety Science 134 (2021) 105030

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

Systems usability case in stepwise control room validation

Hanna Koskinen[a,*], Jari Laarni[a], Leena Norros[b], Marja Liinasuo[a], Paula Savioja[c]

[a] *VTT Technical Research Centre of Finland Ltd, Tekniikantie 21, Espoo, P.O.Box 1000, 02044 VTT, Finland*
[b] *University of Helsinki, PL 3, Fabianinkatu 33, 00014 Helsingin yliopisto, Finland*
[c] *Radiation and Nuclear Safety Authority (STUK), Laippatie 4, Helsinki, P.O.Box 14, 00881 Helsinki, Finland*

ARTICLE INFO

Keywords:
Human factors
Systems Usability Case
Stepwise validation
Control room systems

ABSTRACT

We present a Systems Usability Case approach enabling a requirement-based human factors evaluation of complex socio-technical systems. The approach is especially suitable for stepwise verification and validation activities of nuclear power plant control-room systems. Systems Usability Case is based on the Safety Case approach and exploits the Systems Usability construct. Systems Usability Case is a conceptual procedure within which a reasoning process takes place that enables moving from abstract understanding of human performance and usability of complex tools, to concrete proof of a specific system, and finally to interpreting, in elaborated abstract terms, what is learned about the overall acceptability and level of Systems Usability of the targeted system. The paper introduces the Systems Usability Case approach and demonstrates an application of it to real data from a stepwise control room validation. The results suggest that Systems Usability Case has some advantages: it helps to conduct the validation activities in a more systematic fashion; it makes the reasoning process more explicit and transparent; we are able to build a longitudinal view of the progress of the design process; and constructing the Systems Usability Case enables monitoring of the fulfilment of the requirements from the human factors (i.e., Systems Usability) perspective.

1. Introduction

When people are talking about complex and safety-critical socio-technical systems, a nuclear power plant (NPP) typically presents an ideal example. Because of the complexity, its construction is a considerable effort that takes time, and the design process is difficult to plan. Moreover, since the life cycle of a NPP is expected to be long (i.e., several decades), upgrades of different degrees are needed throughout the life cycle, for example, as a response to technological advancements or changes in operational and safety demands. These characteristics lay down criteria for the design and construction process: It must be systematic, comprehensive, and divided into manageable chunks; it must proceed in gradual stages from one chunk to another, and it must be easy to monitor and well documented.

During the NPP life cycle, it is expected that changes and upgrades are required to the control room human-system interface (CR HSI) systems. The extent of these changes may vary from just upgrading the existing displays to new ones to fully digitalizing the CR HSI systems in connection with a larger automation modernization project. Regardless of the project scope, similar requirements to those for the overall NPP design process can also be laid down for the CR HSI systems design, that is, it must be systematic, comprehensive, and iterative; be easy to monitor throughout the system life-cycle, be well-documented and, in the end, reach a conclusion about the system's suitability for use (Grote, 2012; Law et al., 2008).

Human factors engineering (HFE) is an essential part of the integrated design process in which the diverse design streams, tasks and disciplines are coordinated in order to achieve an optimal design solution. Due to the human operators' significant and yet evolving role in ensuring the ultimate safety of highly automated plant operations, HFE is important when design solutions are developed for a NPP CR (Moray, 1997). Control room verification and validation (V&V) is a critical HFE activity in ensuring the proper functioning of the CR HSI systems (Norros et al., 2015; USNRC, 2012). Thus, systematic efforts are needed to gather and document validation and other human factors data at different phases of the plant's operational life, and during the design and implementation of different CR and HSI upgrades. In this paper, we concentrate on CR HSI validation activity that may enable us as independent human factors experts to make a statement about the acceptability and usability of the target CR system. Validation of the system

* Corresponding author.
E-mail address: hanna.koskinen@vtt.fi (H. Koskinen).

https://doi.org/10.1016/j.ssci.2020.105030
Received 20 December 2019; Received in revised form 24 September 2020; Accepted 1 October 2020
Available online 5 November 2020

against performance-based criteria is a challenging task both theoretically and practically and has thus motived us to develop new systematic methods for conducting human factors validations.

1.1. Goals and functions of CR validation

From a practical point of view, validation of a CR has four main functions: it provides (1) an independent and comprehensive evaluation of the CR HSIs, (2) support and continuous feedback for an iterative CR design, (3) input to human reliability analysis (HRA), and (4) a reference for human performance monitoring. Even though the validation approach presented in this paper may generally support HRA, we concentrate here on the first two functions which set equally important goals for validation, that is, CR validation should not only provide justifications for the proposed design solutions, but also help the designers to improve the maturing system (e.g., Woods and Sarter, 1993). A critical question is how to achieve these two goals in the most practical and economically feasible way. With regard to the first goal, the aim is to achieve a reasonable confidence of the system safety; with regard to the second goal, the aim is to promote a continuous maturation of the design solution with allocated resources and within the planned time frame.

Even though activities concerning these two goals, that is, arriving to a safety justification and furthering the design, are tightly interwoven, the two are typically clearly separated processes. In the NPP context, independence of design and formal evaluation is considered desirable. That is, according to the traditional model of systems engineering, the design process consists of a series of steps that follow each other sequentially and concludes with an integrated system validation (ISV) of the resulted design solution (i.e., system) before commissioning.

If we look at the history of CR validation in the nuclear domain (O’Hara, 1999), the focus in the licensing process has long been on the V&V, and specifically on the demonstration of the validity of the integrated system at the end of the design process. Guides and standards have provided explicit guidance on how to accomplish V&V in the licensing process, and there has been quite much research on the topic, for example, under the OECD NEA Halden Reactor Project (Braarud and Strand, 2011; OECD/NEA, 2017; Simonsen and Osvalder, 2018). One of the most advanced description of the approach are given in NUREG/CR-6393 by John O’Hara (USNRC, 1997; Wise and Wise, 1993), who presented a conceptual model for final validation based on quasi-experimental research methodology. In the past few years, these documents have been reconsidered due to the challenging demands that the question of V&V continues to set out, for example, how to aggregate and systematize large amounts of data produced during the validation process (OECD/NEA, 2017; 2019).

The ultimate aim of all these efforts has been to refine validation methodology in such a way that we can achieve reasonable confidence of the acceptability of the new design. Conventionally ISV has served the licensing process and it has been considered as a completely separate enterprise from the testing conducted during the design, which is mainly accomplished at the sub-system level.

However, there have been claims that these two enterprises, that is, design and evaluation for acceptability, are not that different from each other, as might seem at the first glance. Therefore, we should try to develop a common framework for both of them, which would be based on a step-by-step testing of the quality of a system – in order to hit two birds with the same stone and potentially also to cut the costs of the design work. Many researchers have advocated this kind of multi-stage approach for system validation and developed methodology for it. The first systematic effort for the introduction of a multi-stage validation approach in the nuclear domain has been made under the OECD/NEA Working Group of Human and Organizational Factors (WGHOF) Task Group by an international group of human factors experts in 2016–18 (OECD, NEA, 2019). One of the main challenges for multi-stage validation identified by this task group is the aggregation of validation evidence over the validation activities throughout the design process, which is the topic of this paper.

1.2. Approaches to CR validation

Regarding the first function of CR validation, how are we able to achieve a reasonable confidence of the safety of a complex system? Wise and Wise (Wise and Wise, 1993, see also Everdij et al., 2009) presented four general approaches to validation, which they associated and thus named based on four Western philosophers, Locke, Leibnitz, Kant and Hegel according to the epistemic orientation they portray. Here we focus on the “Lockean” and “Kantian” approaches, which present two opposite ways of thinking about validation. A “Lockean” approach is based on the traditional scientific method, and according to it, the validation process should produce one definite answer to the critical question of whether the system is safe or not. If the system is relatively simple and well-developed, the “Lockean” approach would provide straightforwardly an answer to the validation question. There are several examples of the application of this approach in various domains. The construction of a decision model within the U.S. Marine Corps Operational Test and Evaluation Activity (MCOTEA) (United States Marine Corps, 2007) is an example of this type of approach, which is based on detailed, quantitative data that can be expressed in the form of mathematical formulas. Similar example in nuclear domain can be found, Ha et al. (Ha et al., 2007) introduced a calculation of a discrepancy score for the assessment of plant performance, which was based on the discrepancy between acceptable values and the observed values of specific process parameters. Models in the “Lockean” approach are descriptive in nature in that they provide a quantitative summary of the experimental evidence.

According to the “Kantian” approach to validation, there is no single answer to the question of system validity, but instead many answers that consider the system and its safety from different points of views. Safety Case provides a good example of a Kantian approach (Bishop and Bloomfield, 1998). We propose that because a NPP CR with its HSIs is a highly complex system whose safety and efficiency depend on the interplay of both technical and human and organizational features, a more heterogeneous Kantian approach for its validation is needed. The approach should be based on a diverse set of validation questions - coming from various sources and expressed in a form of requirements - and of answers to these questions. For each requirement, a set of criteria for their acceptance is derived, and the system is tested against these criteria. A reasonable confidence of the system safety is not achieved until all the criteria are met.

1.3. Multi-stage validation in CR design process

As already discussed above, carrying out a comprehensive evaluation of CR HSI systems involve several methodological challenges. First, each CR with its respective HSIs is a unique entity, and in practice it may be difficult to find a suitable reference system that could be directly used as a baseline in the assessment of the system’s usability. The design process and the completion of a CR design project may be very long and time-consuming processes and result in an abundance of information and data to be handled and managed. Furthermore, it is often not operatively possible or economically feasible to accomplish all the products of design at once; instead, the design solutions need to be evaluated in a focused way and the required evaluations are implemented in several stages. Consequently, the human factors efforts and the validation activities need to be planned and organized in a stepwise manner. We have coined the term “sub-system validation” (SSV) to describe a validation activity targeting to evaluate a subset of design solutions, and believe that typically a series of SSVs is needed (Fig. 1) (Laarni et al., 2014).

Conducting SSV as part of the validation process has the following four characteristics: Firstly, the validation process is divided into several steps, which focus on the different parts of the CR HSI system, including the concept of operation for the new CR (Laarni et al., 2013). Secondly, the validation activities aim at identifying and localizing design flaws

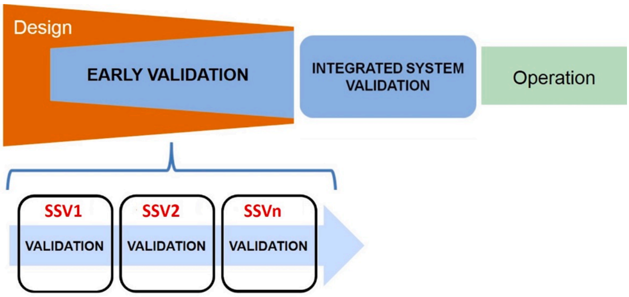


Fig. 1. Stepwise validation in CR design (see Laarni et al., 2013). The arrow indicates the timeline of the validation activities. SSV means Sub-System Validation.

and places for improvements as early as possible, that is, at the conceptual phase (Simonsen et al., 2020), when the possible problems may still be cost-effectively acted upon and solved. Third, a graded approach is followed in focusing of validation activities, according to which, design solutions that are most safety–critical and have the highest novelty value will be evaluated most comprehensively. Fourthly, requirements are systematically used as a reference in the assessment of acceptability of CR solutions.

Many viewpoints have to be considered when performing SSV as a part of a stepwise CR HSI systems validation process. For example, the different stages of the design process may set different demands on the validation process (e.g., validation of design with mockups vs. full scale simulator), and different amounts of time are needed in the design of different parts of the system. Furthermore, there are different human factors focus areas, such as user interface and procedure design, simulator facilities and training of operations personnel that are greatly affected by the design project. All these areas should also be considered, one way or the other, when planning and implementing validation activities for CR HSI systems.

Our approach to CR HSI system validation will provide an effective means for the systemization of the validation evidence in order to reach reasonable confidence of the system safety. Furthermore, it will give support for incremental, stepwise, comprehensive and iterative design of NPP CR HSI systems. Therefore, it may be a particularly suitable methodology for a multi-stage validation of CR HSI systems. In this paper, we will focus on our approach from the perspective of how to effectively and systematically gather, aggregate and manage human factors data and validation evidence to draw conclusions about the usability of the system. We first present the conceptual background of the Systems Usability Case (SUC) approach to CR HSI systems validation, then we describe the logic of the reasoning process and explain in more detail how the SUC approach may be implemented in CR HSI validation. In this connection, we also discuss how the SUC supports the progress of the validation process through a number of SSVs to the ISV. After that, we provide a demonstration of SUC's practical application in a real life CR modernization project. In the final section, we summarize the benefits that the SUC may hold and discuss its methodological limitations and developmental needs that guide our future research.

2. Systems usability case methodology

This section will define the key concepts of the SUC methodology. Primarily, the SUC approach promotes transparency and rigor in CR HSI system validations, and encourages documenting the validation result in the form of typical of safety case demonstrations. In Fig. 2, theories and research approaches have been illustrated that provide inspiration for the SUC. The SUC is a result of decades of human factors work that we have advanced both theoretically and in practice through several research and validation projects. We first introduce the concept of "safety case", which has been our main starting point for the development of SUC.

2.1. Safety case

A safety case is a definitive requirement in many safety standards, and therefore safety cases are required to be produced in many safety–critical domains such as rail transport, military, oil and gas drilling and nuclear industries. According to one definition, a safety case is *"a documented body of evidence that provides a convincing and valid argument that a system is adequately safe for a given application in a given environment"* (Bishop and Bloomfield, 1998). The idea of a safety case is to gather safety-related information into one document that is usable during the system's life-cycle to demonstrate the safety of the system. The safety case provides a frame within which the safety related information may be documented and organized in a structured way. Furthermore, in the safety case the abstraction level and the interconnections between pieces of information may be taken into account and made explicit. Safety cases are built and structured based on three main elements (Bishop and Bloomfield, 1998): claims, evidence and arguments. *A claim* is an entity that expresses a property of the system. *Evidence* is data about the system's ability to support safety, and it is used as the basis of making a safety argument. An *argument* is a description that connects the pieces of evidence to a claim.

Both Safety Case approach (Bishop and Bloomfield, 1998) and the Generic Methodology for V&V (Roza et al., 2013) apply argumentation and hierarchical 'tree-like' presentation of results (see Fig. 2), both of which have also been applied in the SUC.

2.2. Systems usability concept

Systems Usability (SU) characterizes the appropriateness of technology from the human factors point of view (Savioja, 2014; Savioja and Norros, 2008). Systems Usability denotes the capability of a technology to meet the core-task requirements (Norros, 2004) of a work activity in any situation. In essence, in nuclear energy production the core task is the safe and efficient functioning of the plant. In addition to supporting the fulfillment of the core-task demands, SU is defined by the generic functions of the technology as a tool in human activity (Norros, 2017). To put it more accurately, Systems Usability expresses *"the capability of a technology to support fulfilment of the work demands so that the objectives of the activity are met, and the technology has the capability to support the three theory-based general tool functions, instrumental, psychological, and*

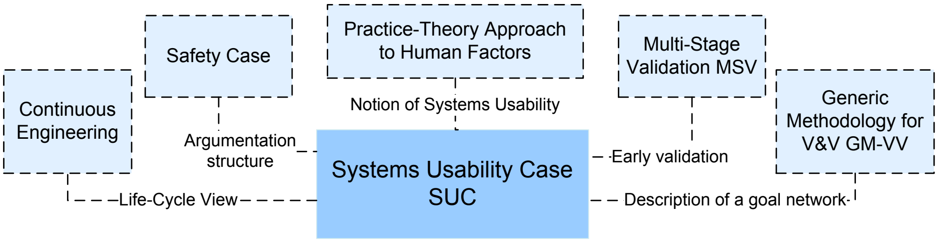

Safety Case
Practice-Theory Approach to Human Factors
Multi-Stage Validation MSV
Continuous Engineering
Generic Methodology for V&V GM-VV
Notion of Systems Usability
Argumentation structure
Early validation
Systems Usability Case SUC
Life-Cycle View
Description of a goal network

Fig. 2. A visual overview of theories and streams of research that has influenced the development of SUC (Bishop and Bloomfield, 1998; Laarni et al., 2014; Norros et al., 2015; Roza et al., 2013; Shamie, 2014).

communicative functions"(Savioja, 2014, p.87).

From the perspective of system validation, the SU provides a contextually defined human factors requirement towards which the design solution should be steered, and which the validated end-product should fulfill. The SU construct consists of nine generic indicators of SU that are derived by cross-tabulating the above mentioned three generic tool functions, that is, instrumental, psychological and communicative function, and three perspectives on activity, that is, performance (outcome), way of acting, and user experience (UX), expressing the use of a tool in the fulfillment of the core task. The nine generic SU indicators of good quality of a tool are shown in Fig. 3. Furthermore, a set of more specific indicators may be defined that are relevant for the work and technology under consideration (e.g., NPP process control work), as shown in Fig. 3. CR HSI systems of good SU could potentially, through promoting an appropriate way of the personnel acting, increase the socio-technical system's ability to meet situational demands (Savioja et al., 2014). As indicated, for example, by Hollnagel et al. (Hollnagel et al., 2011) - who define resilience "*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*" - such ability of a system speaks of its resilience that in turn strengthens its overall safety.

2.3. Systems usability case

A Systems Usability Case delivers the basic inference structure of the safety case methodology, that is, we make explicit a set of claims, produce evidence as a basis for making arguments and apply it to building a Systems Usability Case. Like a safety case, a SUC draws on positive evidence that, in our case, supports the fulfilment of SU. We label this positive evidence as a human engineering conformity (HEC) which indicates that the design solution achieves its potentials and supports resilient and safe performance. However, we have modified and extended the idea of safety case to better meet the demands of informing design: We take into account design deficiencies identified in the validation process, and take them as negative evidence of the fulfilment of SU. This kind of negative piece of evidence is called a human engineering discrepancy (HED) as commonly used in human factors engineering. HEDs may be deviations from optimal operator performance, requirements or design conventions. This means that we do not only generate a final statement or a description of the performance of the tested system during its operation as is done in safety case, but SUC also aims at producing information relevant to the design process.

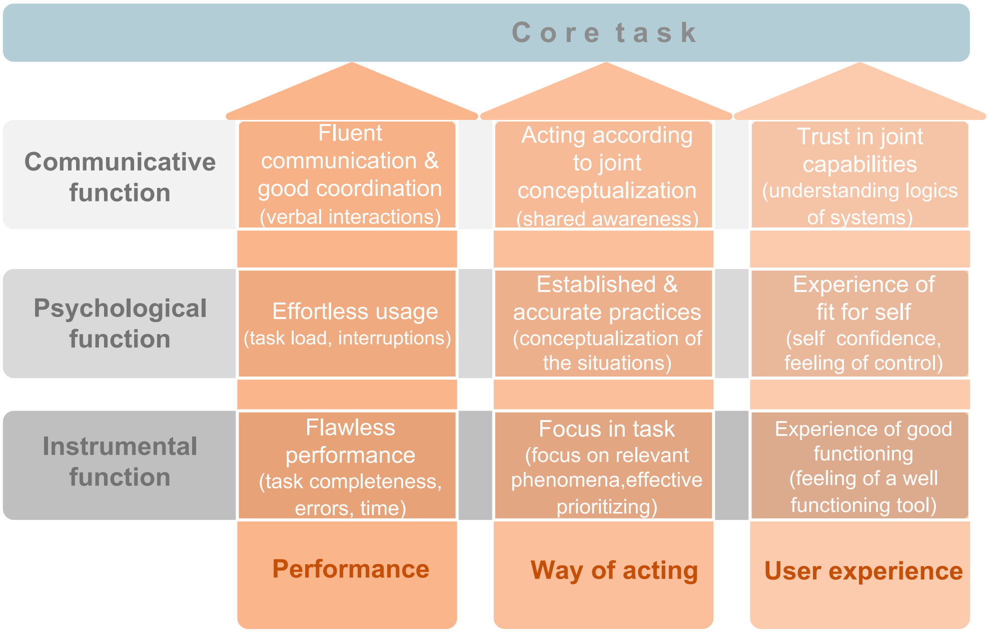
Based on the above-mentioned elaboration, SUC is defined as aiming at *"creating an accumulated documented body of evidence throughout the design that provides a convincing and valid argument of the degree of usability of a system for a given application in a given environment"* (Liinasuo and Norros, 2007). SUC provides a foundation for systematic and comprehensive human factors data management by enabling the consolidation of various types of information and by building up an overall picture of the progress of the design work. Furthermore, the approach provides continuous support for ongoing iterative design of CR HSI systems by producing regular feedback to design. All this increases transparency and supports better monitoring of the proceeding of the design project.

3. Constructing a Systems Usability Case

Constructing a SUC within one validation study (e.g., one SSV in a stepwise validation process) is divided into two stages of reasoning and an intervening practical testing of the system (see Fig. 3): the first stage of reasoning is the formulation of the *goal structure*, established before the actual testing of the system. Then the system is tested in practical experiments and trials, for example, in a series of validation tests (Laarni et al., 2015; Roza et al., 2013) for gaining evidence. Then follows the second stage of reasoning aiming at the establishment of the *claim structure.* This makes explicit the arguments that show how the achieved evidence supports/rejects the fulfilment of the claims. The goal and claim structures of the SUC and the intervening test activities are depicted in Fig. 4 and described in more detail below.

3.1. Goal structure

The goal structure part of the SUC (see upper part of the Fig. 4) provides the reference base for the validation, and it is an essential element of the independent CR HSI validation process. Creating and working through the goal structure helps human factors experts to plan the test activities and to identify the focus areas important to be included



Core task

	Performance	Way of acting	User experience
Communicative function	Fluent communication & good coordination (verbal interactions)	Acting according to joint conceptualization (shared awareness)	Trust in joint capabilities (understanding logics of systems)
Psychological function	Effortless usage (task load, interruptions)	Established & accurate practices (conceptualization of the situations)	Experience of fit for self (self confidence, feeling of control)
Instrumental function	Flawless performance (task completeness, errors, time)	Focus in task (focus on relevant phenomena,effective prioritizing)	Experience of good functioning (feeling of a well functioning tool)

Fig. 3. Nine general indicators of Systems Usability and some NPP specific performance measures (Norros et al., 2015 modified from Savioja, 2014) The case-specific indicators are presented in parentheses as an example.

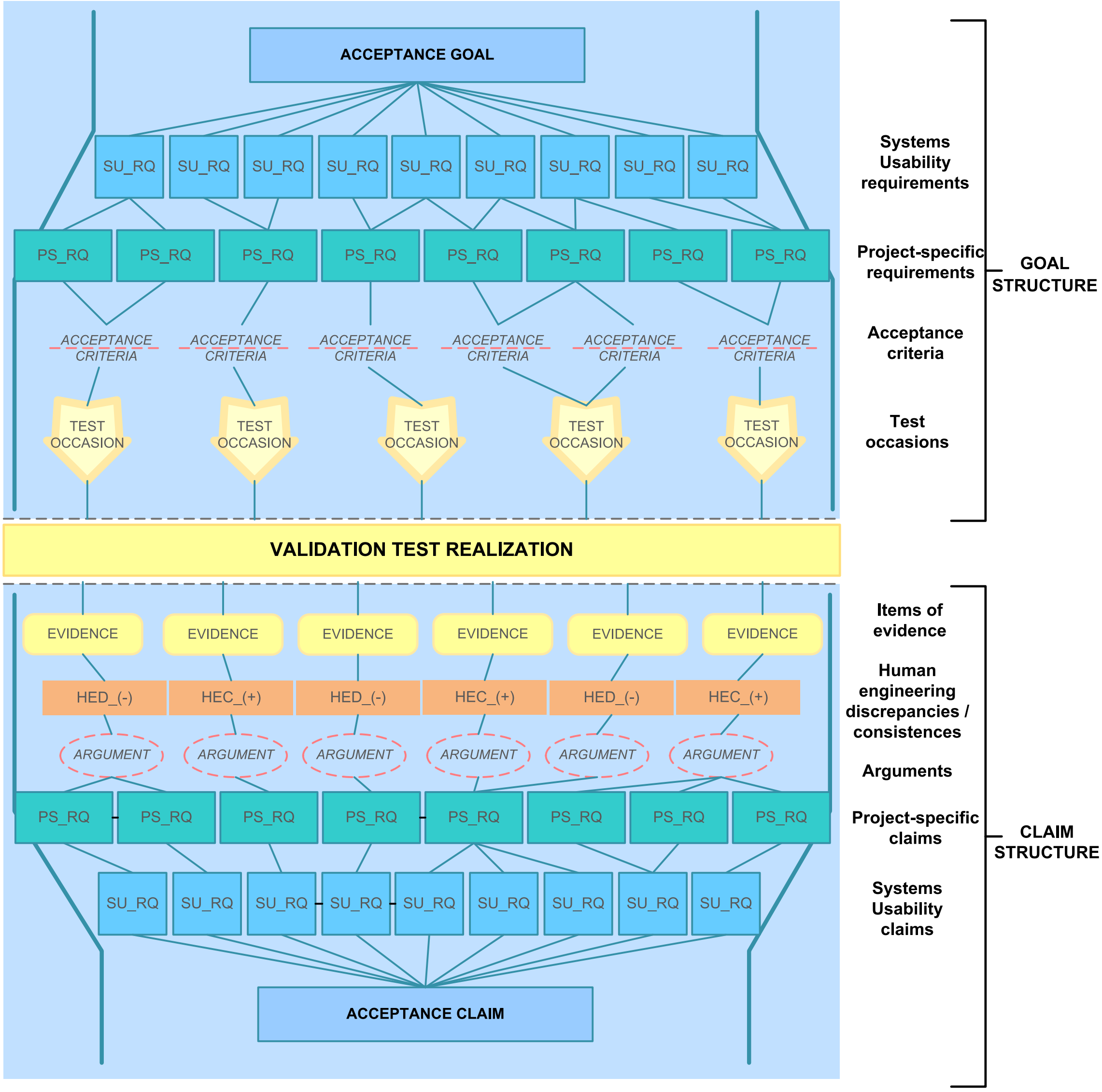


Fig. 4. General SUC framework including goal structure and claim structure parts.

into the validation test. The high-level acceptance goal (topmost in Fig. 4) describes the ultimate objective for validation, which, in our case, means, the promotion of system safety through human performance. The general idea is that the high-level acceptance goal is divided into a hierarchy of several sub-level goals, that is, into SU and project-specific requirements. SU requirements consist of the nine generic indicators of SU (see Fig. 3) which should be expressed in a contextualized form reflecting the specifics of the work domain. The specific, conceptualized statements link these abstract indicators to the concrete reality of CR settings and operator work (as an example SU indicator of effortless usage: the new HSI should not induce higher levels of workload and stress, measured by NASA-TLX, among operators than the old HSI).

The project-specific requirements are drawn from the design documentation developed by the design organization, typically on the basis of operating experience, standards and guidelines. In constructing the goal structure, the human factors experts thematically organize the project-specific requirements under the SU requirements in such a way that each selected project-specific requirement is linked to at least one related SU requirement.

For each project-specific requirement, a specific acceptance criterion, or a set of criteria, are then derived. Acceptance criteria make explicit statements about acceptable human performance in the new CR, and they make it possible to determine whether and to what degree the design conforms to the requirements. Acceptance criteria are derived from the SU and project-specific requirements. The fulfilment of the acceptance criteria can be tested by a defined set of performance measures (e.g., performance time or number of errors in performance). A validation test occasion is determined for each of these criteria which provides a test condition where a particular criterion can be assessed. Proper handling of nuclear accident events, such as loss of coolant accidents (LOCA) and steam generator tube rupture (SGTR), consists of a set of operator activities through which the acceptance criteria can be tested.

The actual validation test is then conducted, for example, in a simulator environment with licensed CR operators. The implementation of the validation test is illustrated with a horizontal yellow text box in the middle of Fig. 4. Both performance measures and the availability of test occasions set constraints for the specification of acceptance criteria (e.g., for their accuracy): Acceptance criteria have to be defined in such a way that makes it possible to measure them by existing performance measures and also takes into account the details of the selected operational conditions. To take a simple example of one of our recent validation studies: it was required that the shift supervisor is able to complete his/her duties and responsibilities in the beginning of a simulated accident quickly enough so that the other operators need not wait for their completion before proceeding to the next page of the procedure.

In the upper part of the Fig. 4 (i.e., goal structure part), the colored shapes illustrate the different phases of constructing the goal structure of SUC. The human factors experts need to devote considerable effort to the analysis of the design documentation in developing the goal structure and defining the acceptance criteria and relevant test occasions (e.g., scenario runs). Only after these tasks have been accomplished, can the actual validation tests be conducted.

3.2. Claim structure

The claim structure (see the lower part of the Fig. 4), is a kind of mirror image of the goal structure part and enables a case-based structuring of the validation test results. The claim structure comprises a set of evidence, arguments and claims. In this stage of reasoning, the requirements serve a new role as claims. Constructing the claim structure aims to create a documented body of evidence providing a valid argument of the degree of SU of the system under consideration.

An item of evidence is a description of operator or system performance in the context of a particular operational test occasion (see middle part of Fig. 4). Examples of items of evidence are task completion times, errors in performance and subjective evaluations of, for example, mental workload, situation awareness and teamwork. For each test occasion, at least one item of evidence is associated. Items of evidence may be either positive or negative from the point of view of the design depending on whether the corresponding acceptance criterion is met or not. Positive evidence, that is, a HEC, includes signs of positive surprises and clear evidence of future potentials, whereas a HED is any indication of problems or deviations regarding the user interface, procedures or user performance. However, it is expected that completely new HEDs also emerge in the test that may not be connected to any predefined acceptance criteria. Consequently, new requirements need to be defined in the goal structure.

In SUC, a claim is a generic property of the design solution, or its use, which has been defined prior to design (cf. requirements in the goal structure). Claims regarding what is considered a “good” system provides a reference against which the acquired evidence is assessed. An argument is the human factors experts’ explanation of how evidence makes explicit the fulfillment of a claim (i.e., whether a HEC or HED is identified). In other words, it defines a generic mechanism in the use or design of the system that may cause the specific HEC/HED. Arguments are needed to bridge the gap between a specific piece of evidence and a more generic claim.

The arguments emerge from the data as each HED/HEC is individually analyzed. In our validation tests, we categorized the arguments into three classes: 1) a conflict/coherence with design principles or good design practices, 2) a conflict/coherence with or threat/no threat to the generic human performance conditions, and 3) an error or a near miss that hindered/no hindrance of the functioning of the system. For each argument type, a couple of more specific arguments were introduced. For example, to name some negative examples, under the first category (i.e., conflict with design principles or good design practices) the more specific arguments were: “violates style guide”, “violates established usability principles”, and “discrepancy between procedure (paper/electronic) and operational display”. Examples of the specific arguments in the second category were: “increases mental load”, “hampers the fluency of performance”, “hinders the coordination of teamwork”, and “hinders control of own performance”. Finally, in the third argument category, following arguments emerged: “possibility for/actual misinterpretation”, “possibility for/actual false execution of action”, “possibility for/actual confusion with regard to automatic operations”, ”decrease in the reliability of actions”.

To summarize, the claim structure part of SUC aims to organize the information acquired in the validation tests, and the expert knowledge about human factors into a systematic and meaningful general view and to make the reasoning behind the validation results explicit.

3.3. Systems usability case in multi-stage validation process

As noted, a case-based approach seems particularly helpful, when the system to be evaluated is unique and comparison to other systems is difficult. One of the main aims of establishing a SUC is to bring to the front the arguments and evidence for safety in such a way that the argumentation is transparent and logically valid, and the fulfilment of regulatory requirements can be evaluated. Another aim of the SUC is to make decisions about safety traceable throughout the life-cycle of a product.

Fig. 5 below demonstrates how the validation evidence is aggregated and design solutions are maturated over a variety of validation test activities. It also demonstrates how the individual SSV test activities are interrelated. There is a progression of fulfilment of human factors requirements through the time of validation. Some of the HEDs are resolved in one stage (i.e., a particular SSV test), but the revised design solutions have to be retested in the next stage (i.e., following SSV test). In addition, project-specific requirements may change or become more precise, and new project-specific requirements may emerge during the validation process that need to be included in the SUC. It is quite typical that operators raise concerns about impacts of a particular design solution on aspects of operator work that have not been adequately addressed in specification of requirements and thus in establishment of the SUC goal structure. These concerns have to be put into requirements in the next iteration of the SUC. However, the idea is that at in the end of the validation process all the identified HEDs are resolved and new critical requirements do not emerge any more. This kind of systematic and incremental process provides support both for the continuous evaluation and maturation of the design solutions and for the independent validation of the CR HSI systems (see the black arrow at the bottom of the Fig. 5).

4. Demonstration

Next, we will present a simplified real-life example to illustrate the application of the SUC approach in CR HSI systems V&V. The presented example is related to validation activity targeted to new safety HSIs and emergency operating procedures (EOPs) introduced in connection to one particular NPP modernization project. Fig. 6 presents a simplified goal structure of a SUC regarding the example SSV test activity.

The main acceptance goal and the SU requirements are common to all SSVs (see the two upper rows in the Fig. 6), but the tabulated project-specific requirements (PS_RQ) are specific to the example SSV test case (i.e., requirements concerning the subsystem to be validated). There are dozens of relevant project-specific requirements, however, only three of them are displayed in Fig. 6. For the specific operational concept and safety HSI and EOPs, the following requirements were defined:

- **PS_RQ1** If the main CR cannot be used the unit shall be able to be operated to safe state by procedure guidance,
- **PS_RQ2** A dedicated safe shutdown procedure shall be designed to each operator qualification (SS, RO and TO) separately and
- **PS_RQ3** Entrance criteria shall be defined for safe shutdown procedure (see Fig. 5).

The acceptance criteria (C) are described as precisely as it is reasonable and possible to do. In Fig. 6, three acceptance criteria are depicted:

- **C1** The main CR is left without any delays and endangering own health,
- **C2** The unit can be operated to safe state (outlet temperature of core less than 140 °C) and

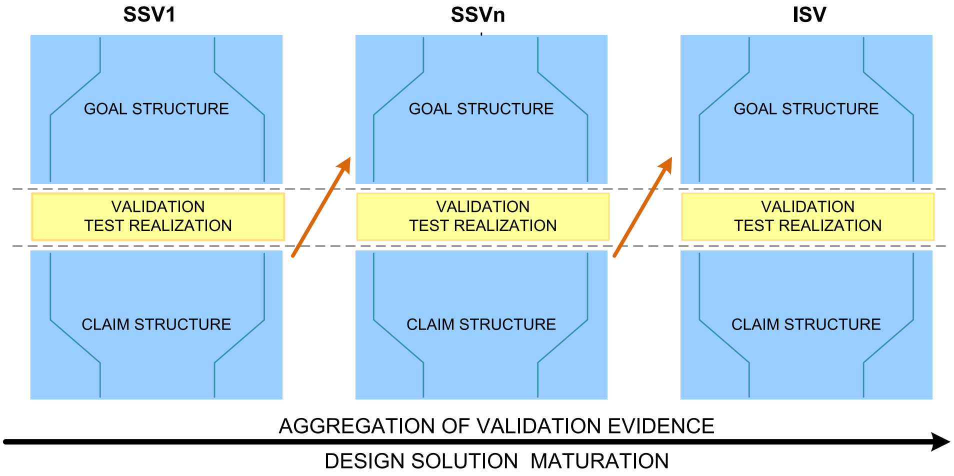


Fig. 5. SUC development process illustrating the aggregation of validation evidence and maturation of the design solution over individual SSV tests towards the ISV.

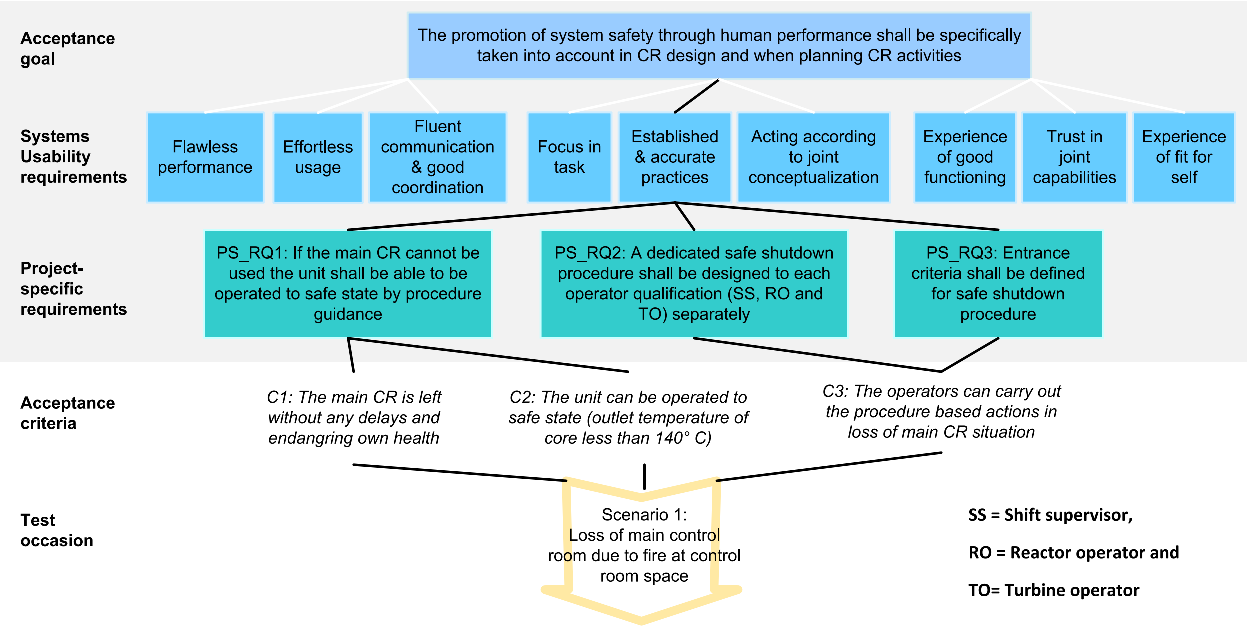

Acceptance goal
The promotion of system safety through human performance shall be specifically taken into account in CR design and when planning CR activities
Systems Usability requirements
Flawless performance
Effortless usage
Fluent communication & good coordination
Focus in task
Established & accurate practices
Acting according to joint conceptualization
Experience of good functioning
Trust in joint capabilities
Experience of fit for self
Project-specific requirements
PS_RQ1: If the main CR cannot be used the unit shall be able to be operated to safe state by procedure guidance
PS_RQ2: A dedicated safe shutdown procedure shall be designed to each operator qualification (SS, RO and TO) separately
PS_RQ3: Entrance criteria shall be defined for safe shutdown procedure
Acceptance criteria
C1: The main CR is left without any delays and endangring own health
C2: The unit can be operated to safe state (outlet temperature of core less than 140° C)
C3: The operators can carry out the procedure based actions in loss of main CR situation
Test occasion
Scenario 1: Loss of main control room due to fire at control room space
SS = Shift supervisor,
RO = Reactor operator and
TO= Turbine operator

Fig. 6. Example of a goal structure part of a SUC in a CR modernization project.

- **C3** The operators can carry out the procedure based actions in loss of main CR situation.

Several acceptance criteria can be created for one project-specific requirement in order to improve the reliability and validity of the assessment. For example, in Fig. 6, two acceptance criteria (C1 and C2) are linked to PS_RQ1, but only one (C3) to the other two requirements (i.e., PS_RQ2 and PS_RQ3). It can also be seen that the acceptance criteria C3 is common to PS_RQ2 and PS_RQ3. The procedure by which acceptance criteria are created and linked to the requirements is also important for weighing the feasibility and functionality of the project-specific requirements. For example, if the requirements are of too general level, it is impossible to verify them by a reasonable set of criteria. On the other hand, the requirements can also be too specific so that it is difficult to generate relevant testable acceptance criteria to them. In Fig. 6, a test occasion is identified and depicted under the operational conditions in which the three acceptance criteria can be assessed. Typically, test occasions are designed in such a way that a multitude of acceptance criteria can be tested in a single test occasion: in our example, the three acceptance criteria (C1, C2 and C3) are all tested in the same scenario run (Scenario1).

With regard to a specific SSV test, it is possible that only a subset of the requirements becomes validated, fulfilment of the rest of them is not supported by evidence. It is the task of the design organization to decide what should be done to solve the HEDs so that the design will fulfil the requirements. Even though all the HEDs are included in the SUC, their criticality and importance should be determined. The classification of HEDs according to their safety-criticality and importance is based on an expert elicitation process, that is, negotiations between the validation team and representatives of the design organization. Arguments play a mediating role between HEDs and claims (i.e., SU and project-specific requirements). Sometimes a particular HED does not cause any further actions. However, in most cases, the HED cannot be ignored, but instead it must be settled through a redesign of the HSI, operator-training program or procedures.

4.1. First ssv test realization

The evidence from the first SSV test realization showed that operators had some performance difficulties in the main CR before they moved to the emergency control post. Consequently, three HEDs (HED1, 2 and 3 in Fig. 7) were identified. The current operating procedure was shown to be too lengthy, some steps were in a wrong order, and generally, there were too many actions to take in considering the urgency of the situation (loss of the main CR, e.g., hasty fire). For these HEDs, a new requirement emerged (see the project-specific claims row in Fig. 8) regarding the procedure guidance, and it was formulated in the following way (NEW PS_RQ4): Procedure guidance shall be as clear and concise as possible and applicable to all kinds of loss of the main CR conditions. In order to maintain requirements' traceability it is important that the existing requirement is left as it is, and the new requirement is added as a subordinate one to the parent requirement. The new requirement has to be considered in procedure design, and possibly it has to be evaluated at a later phase of the V&V process (i.e., in the following SSV tests or in the final ISV test realized at the end of the modernization project).

With regard to the SU claims, the critical question is whether they were threatened in that particular condition, and how much positive evidence (i.e., HECs) is piled up to support them and how much negative evidence (i.e., HEDs) against them. In order to make the SUC framework complete, the main acceptance claim is depicted at the bottom of the framework (Fig. 8), and, as can be seen, it is the same as the acceptance goal depicted on the topmost row in the goal structure (Fig. 6).

4.2. Second ssv test realization

Because the HEDs identified in the first SSV test were considered relatively safety–critical, further development of EOPs was required. Consequently, it was also necessary to re-evaluate the EOPs and the safety HSIs in the second SSV test. A new, more specific acceptance criterion (C4) making a reference to the operator performance in the first SSV test was formulated for the new requirement (NEW PS_RQ4, see Fig. 9). The comparative form is justified as the design work was still ongoing and it was important to see whether the design is maturing and converging to an acceptable solution. In the second SSV test, more focus should be placed on those operational conditions and issues that were recognized as the most problematic in the first SSV test and, thus, get an in-depth understanding of the effects of the design changes. However, later on in the validation process it is likely that the acceptance criterion takes a more objective form.

According to the results of the second SSV test (see + HEC1 in Fig. 10), the new version of the EOP could be considered acceptable (Fig. 10). In particular, human factors experts found that no major problems/issues emerged in coordinating the crew activities when preparing to leave the main CR. However, there was still some room for improvement: only the Shift Supervisor (SS) had a dedicated procedure in use at the beginning of the simulator run, when the crew still stayed in the main CR and performed some immediate actions. The SS is expected to instruct the other operators and give commands to them. Operators thought that even though they were able to do the required operations quickly and fluently, it would have helped them to form a precise mental model of the situation and prepare for the actions ahead, if they had had their own version of the procedure (- HED4). Even though the new HED was not considered to be high in importance or safety-criticality, it should be carefully reconsidered in the later validation tests.

5. Discussion

It appeared through our experiences on carrying out human factors validations of complex safety–critical systems, such as validation of NPP CR HSI systems, that more systematic and holistic methods are needed to achieve comprehensive, transparent and well-documented assessment of a system's usability and safety. We have also recognized that in a validation process a lot can be learned about the design solution that may help to further develop it and this information is too valuable to be lost in the process, and thus should be fed back to the design. We have deployed the SUC approach for the aggregation of validation results and information to improve the design, and systematize the validation process. In this section, we discuss conducting human factors validations of complex systems by means of the SUC approach, and pinpoint the practical implications and challenges associated with its use.

It is impossible to completely validate a system and demonstrate that it is infallible and perfect for its purpose of use. But it is possible to look at the system from different perspectives, put it in challenging situations, ask different validation questions and try to answer to them as best we can. The more complex the system is, the more there will be gaps that may not be fully covered in V&V. As said, Wise and Wise (Wise and Wise, 1993) suggested a "Kantian" approach to validation of systems where the system is highly complex and where novel technologies are used. According to this approach, there are several validation questions, and each validation question has necessarily several answers, because it is impossible to know beforehand the correct questions and the correct answers to these questions. However, if we take a "Kantian" approach, there is the danger of chaos of heterogeneous and incompatible evidence that is difficult to master and to draw any final conclusions. What we are advocating here is that the SUC may save us from despair, because it is especially suitable for structuring and managing heterogeneous validation evidence from the validation of complex systems (or systems of systems) such as NPP CR HSI systems throughout their life-cycle. Within SUC the collection and organization of rich and diverse data is enabled.

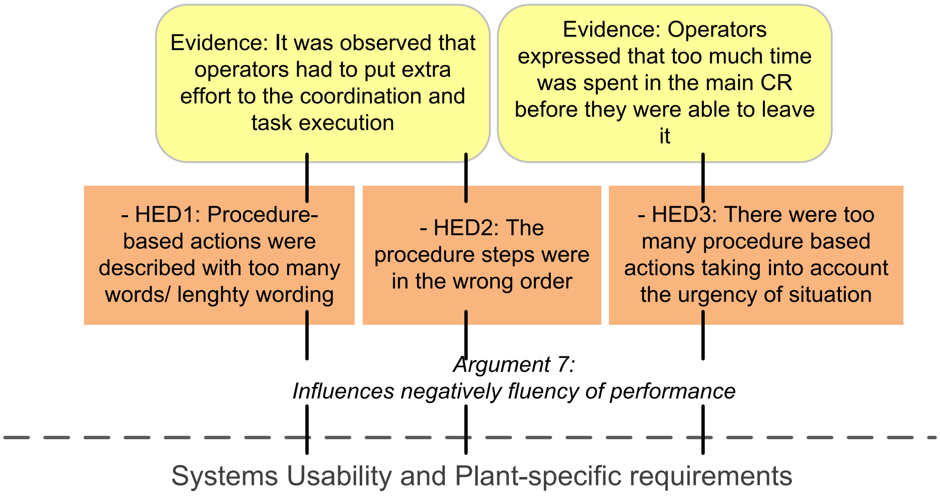

Evidence: It was observed that operators had to put extra effort to the coordination and task execution
Evidence: Operators expressed that too much time was spent in the main CR before they were able to leave it
- HED1: Procedure-based actions were described with too many words/ lenghty wording
- HED2: The procedure steps were in the wrong order
- HED3: There were too many procedure based actions taking into account the urgency of situation
Argument 7:
Influences negatively fluency of performance
Systems Usability and Plant-specific requirements

Fig. 7. Demonstration of Evidence-HED-Argument linkages to claims concerning the HSI and EOP solution in our example case.

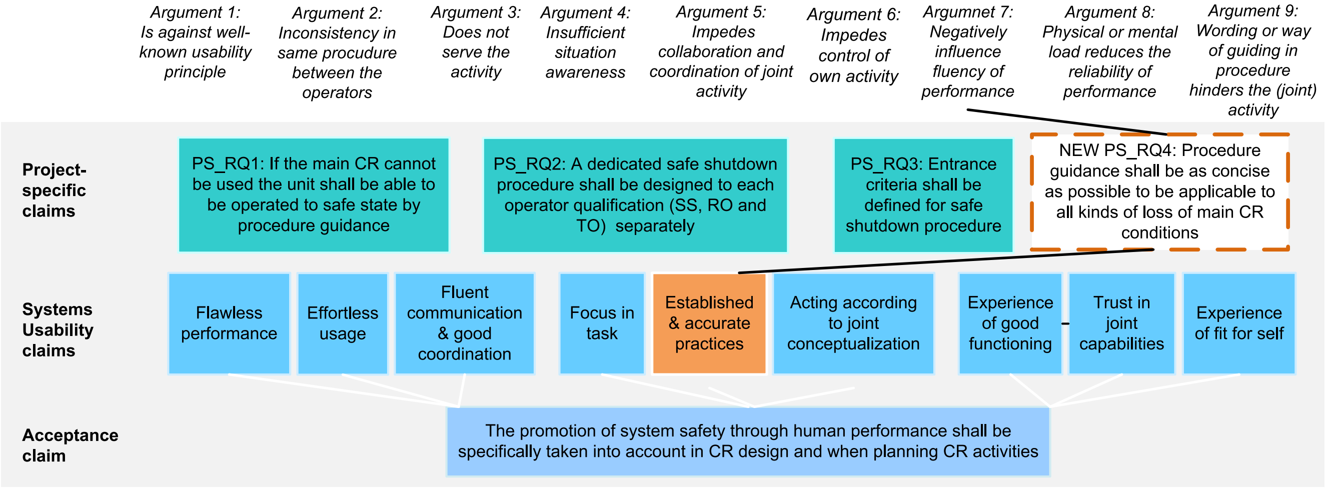

Argument 1: Is against well-known usability principle
Argument 2: Inconsistency in same procudure between the operators
Argument 3: Does not serve the activity
Argument 4: Insufficient situation awareness
Argument 5: Impedes collaboration and coordination of joint activity
Argument 6: Impedes control of own activity
Argumnet 7: Negatively influence fluency of performance
Argument 8: Physical or mental load reduces the reliability of performance
Argument 9: Wording or way of guiding in procedure hinders the (joint) activity
Project-specific claims
PS_RQ1: If the main CR cannot be used the unit shall be able to be operated to safe state by procedure guidance
PS_RQ2: A dedicated safe shutdown procedure shall be designed to each operator qualification (SS, RO and TO) separately
PS_RQ3: Entrance criteria shall be defined for safe shutdown procedure
NEW PS_RQ4: Procedure guidance shall be as concise as possible to be applicable to all kinds of loss of main CR conditions
Systems Usability claims
Flawless performance
Effortless usage
Fluent communication & good coordination
Focus in task
Established & accurate practices
Acting according to joint conceptualization
Experience of good functioning
Trust in joint capabilities
Experience of fit for self
Acceptance claim
The promotion of system safety through human performance shall be specifically taken into account in CR design and when planning CR activities

Fig. 8. Claim structure part of the SUC. Right most is presented the new requirement emerging from the first SSV test.

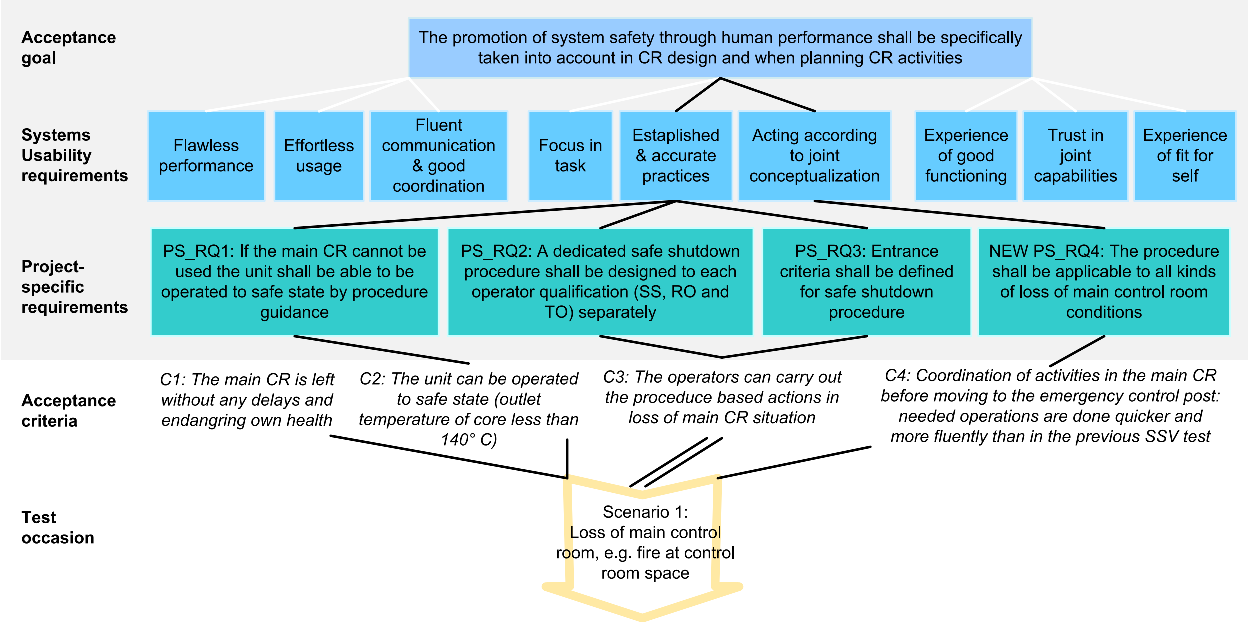

Acceptance goal
The promotion of system safety through human performance shall be specifically taken into account in CR design and when planning CR activities
Systems Usability requirements
Flawless performance
Effortless usage
Fluent communication & good coordination
Focus in task
Established & accurate practices
Acting according to joint conceptualization
Experience of good functioning
Trust in joint capabilities
Experience of fit for self
Project-specific requirements
PS_RQ1: If the main CR cannot be used the unit shall be able to be operated to safe state by procedure guidance
PS_RQ2: A dedicated safe shutdown procedure shall be designed to each operator qualification (SS, RO and TO) separately
PS_RQ3: Entrance criteria shall be defined for safe shutdown procedure
NEW PS_RQ4: The procedure shall be applicable to all kinds of loss of main control room conditions
Acceptance criteria
C1: The main CR is left without any delays and endangring own health
C2: The unit can be operated to safe state (outlet temperature of core less than 140° C)
C3: The operators can carry out the proceduce based actions in loss of main CR situation
C4: Coordination of activities in the main CR before moving to the emergency control post: needed operations are done quicker and more fluently than in the previous SSV test
Test occasion
Scenario 1: Loss of main control room, e.g. fire at control room space

Fig. 9. A goal structure part of a SUC created for the second SSV test including the new requirement and additional and refined acceptance criteria.

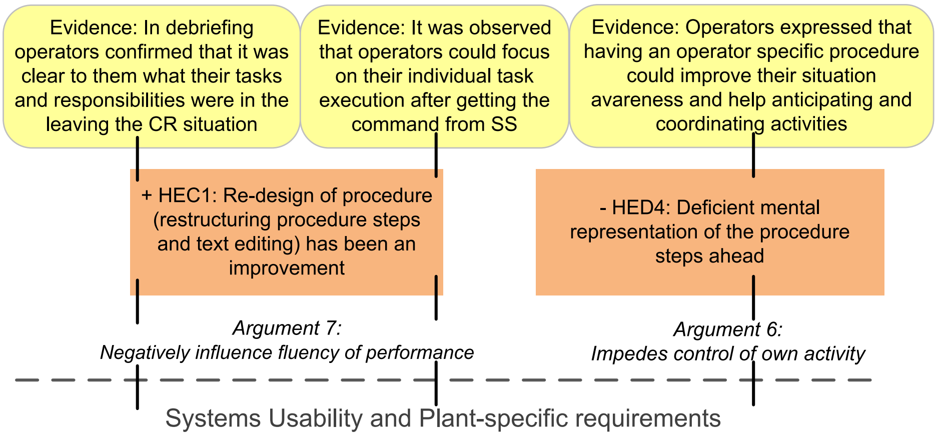

Evidence: In debriefing operators confirmed that it was clear to them what their tasks and responsibilities were in the leaving the CR situation
Evidence: It was observed that operators could focus on their individual task execution after getting the command from SS
Evidence: Operators expressed that having an operator specific procedure could improve their situation avareness and help anticipating and coordinating activities
+ HEC1: Re-design of procedure (restructuring procedure steps and text editing) has been an improvement
- HED4: Deficient mental representation of the procedure steps ahead
Argument 7: Negatively influence fluency of performance
Argument 6: Impedes control of own activity
Systems Usability and Plant-specific requirements

Fig. 10. Demonstration of Evidence-HED-Argument linkages to claims concerning the HSI and EOP solution in second SSV test.

Furthermore, in SUC the chain of reasoning based on the evidence is made explicit and transparent, and it helps to draw attention to the main problems and important questions regarding the validated design solution.

5.1. Quality of requirements

SUC is a requirement-based approach, and SUC is as good as the requirement base on which it is built. Thus, quality of SUC depends on the quality of the requirements set for the design at hand (i.e., theory-driven and project -specific requirements). First, the theory-driven SU requirements have to be formulated in such a way that with the help of them it is possible to evaluate the coverage of the design requirements (i.e., project-specific requirements), to monitor the fulfilment of these requirements in a longitudinal manner, and to assess the scope of project-specific requirements in relation to the theory-driven SU requirements. Second, the project-specific requirements are often written in a way that it is impossible to evaluate them by performance-based validation tests. Therefore, a lot of effort and exercising is often required before a sufficient and satisfactory requirement and acceptance criterion base has been achieved. The requirement selection and categorization exercise is an essential part of constructing the goal structure, and hence the formation of a basic understanding of capabilities and limitations of the target system. In addition, the more accurately the requirements specify the expected behavior of the crew, the more accurately the acceptance criteria can be determined. According to our practical experiences, it is often a real challenge for design organizations to establish a requirement base that would effectively guide the design of CR HSI systems. Thus, the requirement elicitation and management may be areas in which development work is still needed.

5.2. Improved understanding and base for making assessment

It is apparent that constructing a SUC helps in the systematic execution of the validation process and in comprehension of validation results. With the help of SUC, we are able to monitor the progress of the design process, and to achieve a gradually deepening understanding of the strengths and possible weaknesses of the system. At the beginning of the validation process, the independent human factors evaluators' understanding of the system to be validated and its level of usability is insufficient. Thus, the validation questions they are asking may be too general, lacking concreteness and precision. This would be the case regardless of human factors experts being involved in the validation process early on, for example, having accomplished the SSV type of validations or being involved only in the end of the design process in the ISV. For that reason, one apparent benefit of SUC is starting the validation effort by creating the goal structure by working through the project-specific requirements and thematically organizing them according to the theory-driven SU requirement. As a result, the human factors experts become better acquainted with the validated system at hand. Thus, SUC urges the evaluator to think about the critical validation questions beforehand and in a more informed way. We would even tend to suggest that through construction of SUC and especially of its goal structure, the validation questions may be set more accurately at the first place. Whereas after the actual validation test, building the claim structure makes explicit the reasoning process, in which the concrete validation evidence is connected to more abstract/generic understanding of human activity and the usability of a system. At the same time, human factors experts' understanding of the strengths and possible weaknesses of the system increase and they are able to set more accurate validation questions so that more targeted and comprehensive testing of the system is possible.

With the aid of such an explicit reasoning process that SUC entails, the conceptual understanding of the system is increased so that it can be debated and agreed upon in more detail with the progress of the validation process. In addition, the reasoning process becomes better documented, and thus can be traced back at a later point in time. This evolving conceptualization of the system's validity increases the stakeholders' ability to understand the strengths and weaknesses of the system.

Validation activities, the results of which are documented as a SUC, provide also a continuous feedback to the design work. Through SUC, we are thus able to track the progressive improvement of the system's usability throughout the design process. In the SUC, the design solution also becomes captured in different levels of abstraction, while the project-specific requirements may describe the solution in more generic terms, the evidence from the validation tests often tell about the concrete HSI issues.

5.3. Suc in the tradition of safety demonstration

We are not the first to argue that systematic methods are needed for the accumulation and systematization of validation evidence and drawing conclusions from the evidence: similar ideas have been suggested elsewhere. According to a group of safety authority experts (Bel et al., 2018), the assessment of nuclear instrumentation and control (I&C) software cannot focus only on the evaluation of the end product, but also the quality of the design process and methods used in it have to be covered. The term 'safety demonstration' was coined to describe an artefact in which a set of arguments and evidence items support a set of claims on the system safety of the operation of safety–critical systems used in a NPP (Bel et al., 2018). According to this definition, safety demonstration is a set of information items stored in a database or a specific document (Valkonen et al., 2016). We propose that the SUC approach we are advocating is an example of this kind of safety demonstration document.

There are also some concrete applications of the safety case approach to the structuring and management of validation evidence. For example, the Generic Methodology for V&V (GM-VV) developed under the Simulation Interoperability Standards Organization (SISO) (Roza et al., 2013) is based on argumentation theories, and its structure of argumentation is similar to the SUC with its goal and claim structure. Based on our reading of Roza et al., (Roza et al., 2013) the GM-VV does not differentiate different forms of requirements, but deals with acceptability criteria at a general level. Neither does it address different forms of claims, but talks generally about acceptability claims. However, since the argumentation process is similar in both cases, their differences are less important than similarities (Table 1). It is encouraging that even though these two approaches for V&V which concern very different central questions, that is, the usability of systems and the correctness of models and simulations, have led to similar ways of organizing and structuring empirical data for use in V&V, and specifically when seeking answer to the acceptability of the system.

Table 1
A comparison of the Systems Usability Case (SUC) and the Generic Methodology for V&V (GM-VV) (Roza et al., 2013).

Compared aspect	SUC	GM-VV
Theoretical background	Practice theory, Safety Case	Systems engineering, Argumentation theories
Aim/ Central question	Usability (design feedback) and appropriateness of CR solution for use	Correctness of the model and simulation products
Scope	V&V (operational activities)	V&V (technical, project and enterprise level activities)
Structure	Hierarchical network	Hierarchical network
Life-cycle phase	Can be applied to validation needs throughout the life-cycle	Can be applied to validation needs throughout the life-cycle

5.4. Open issues

The SUC approach is still under development. We have applied it and demonstrated its use in this paper with two real-life validation projects concerning the modernization of a nuclear power plant CR. By following the SUC structure, we were able to take the design requirements as a base for the validation and form a set of criteria and test scenarios that truly test the system from the usability point of view. The customer found the SUC approach to be systematic and informative, supporting the collaboration in dealing with the HEDs, as well. Yet, it is necessary to gain more use experiences of SUC in order to better understand its suitability for different types of validations (e.g., SSVs in different phases of the design process and the final ISV) and how it works over the system life cycle.

There are some open issues that need to be considered and resolved: First, since the project-specific requirements can be of several types (e. g., general CR level requirements or HSI specific requirements such as safety HSI), it would be useful to classify them according to their type in order to facilitate the construction of the goal structure. However, this kind of classification has not been done in the depiction shown in Fig. 5. Second, neither have we prioritized the requirements nor the acceptance criteria derived from these requirements, even though it would be important to grade them, for example, according to their safety importance. The classification of the requirements and acceptance criteria would also be helpful in guiding the prioritization and weighing of HEDs and HECs identified during the validation process. Third, one important question we have not yet properly touched upon is how the positive evidence (i.e., HECs) are considered and taken into account in deriving the final conclusion of the validity of the design solution: If the positive and negative evidence are pitted against each other, how to do that?

Fourth, it still remains somewhat open what is meant by “final conclusion” that is drawn based on the ISV. From the more traditional point of view the final conclusion of an ISV is an estimation of whether all HEDs have been resolved, and there is no signs of new undetected design deficiencies. However, in reality, there are always some HEDs that cannot be completely resolved, and which we have to learn to live with. In addition, it is possible that some emerging HECs manifest the potentials of the new design. In the latter case, that also the SUC represents, the final conclusion could, in a positive case, be that, despite some remaining deficiencies, the system has a lot of potential that provides a reasonable confidence of its suitability for use. This kind of conclusion provides grounded reasons for considering the system’s validity, and we see that it offers a reasonable solution to express a system’s validity in case of continuous engineering of complex systems. In fact, the conclusion is well in accordance with a life-cycle approach and the idea of continuous improvement, according to which the design has the potential to mature and evolve through its lifecycle, and specific phases can be identified in this process. Human performance monitoring is one central HFE activity that in a way continues from what was learned about the operation of the system in an ISV. It provides an opportunity to carefully follow the development of the issues identified in the ISV but also a way to register new emergent HF issues. Thus, ISV may not be even expected to provide what is called a “final conclusion”.

Finally, we have not yet fully addressed how and in which form most efficiently to store and document the SUC. Several software programs are available for the management and documentation of a SUC. We especially intend to find a documenting format for SUC allowing us better recording all the phases of constructing SUC (i.e., goal structure and claim structure). Ideally the software program should also aid in making illustrative summaries from the data base that would help comprehending the results and making the assessments. In validations that we have conducted so far, a software tool developed by Bishop and Bloomfield (1998) was considered suitable starting point for constructing SUC.

6. Conclusions

Because a huge amount of human factors related data regarding the CR HSI systems is collected throughout the life-cycle of a NPP, systematic methods are needed for the accumulation and systematization of validation evidence and drawing conclusions from that evidence. During the development of the SSV approach, it was reasoned that a case-based approach (i.e., SUC) is particularly suitable when the system to be evaluated is unique and comparison to other systems is difficult. SUC also supports formative evaluation, in which the interest is to steer the development of the system by successive evaluations, and nevertheless remain independent of the design process itself. We have established a SUC for real life CR and HSI design cases, however, more experience needs to be gained and the approach tested throughout the design process (i.e., in number of SSVs and a final ISV). Based on the lessons learnt so far, we believe SUC promotes continuous engineering and maturation of a design solution to improve system safety.

Acknowledgments

We would like to acknowledge the SAFIR programme (The Finnish Research Programme on Nuclear Power Plant Safety) that has funded the research projects within which the SUC methodology has been developed. We also esteem the work of Fortum Power and Heath Ltd, designers, training personnel and operators who have participated in the evaluations.

References

Bel, V., BfE, CNSC, CSN, ISTec, KAERI, KINS, ONR, SSM, STUK, NSC, 2018. Licensing of safety critical software for nuclear reactors - Common position of seven European nuclear regulators and authorized technical support organisations.

Bishop, P., Bloomfield, R.A., 1998. Methodology for Safety Case Development, in: Safety-Critical Systems Symposium. Birmingham, UK.

Braarud, P.O., Strand, S., 2011. Human Factors Integrated System Validations - Lessons Learned NPP modernisation Projects HWR-986.

Everdij, M.H.C., Blom, H.A.P., Scholte, J.J., Nollet, J.W., Kraan, B., 2009. Developing a framework for safety validation of multi-stakeholder changes in air transport operations. Saf. Sci. 47, 405–420. https://doi.org/10.1016/j.ssci.2008.07.021.

Grote, G., 2012. Safety management in different high-risk domains - All the same? Saf. Sci. 46, 450–460.

Ha, J.S., Seong, P.H., Lee, M.S., Hong, J.H., 2007. Development of human performance measures for human factors validation in the advanced MCR of APR-1400. IEEE Trans. Nucl. Sci. 54, 2687–2700.

Hollnagel, E., Pariès, J., Woods, D.D., Wreathall, J., 2011. Resilience Engineering Perspectives Volume 3: Resilience Engineering in Practice. Ashgate, Farnham, UK.

Laarni, J., Karvonen, H., Koskinen, H., Liinasuo, M., Norros, L., Savioja, P., Salo, L., Laakso, A., Lehtonen, M., 2013. A stepwise validation process for the main control room of Fortum Loviisa nuclear power plant. in. Enlarged Halden Project Group (EHPG) Meeting.

Laarni, J., Norros, L., Salo, L., 2015. Multi-stage Approach to Control Room Validation, in: OECD/NEA WGHOF Task Group.

Laarni, J., Savioja, P., Norros, L., Liinasuo, M., Wahlström, M., Salo, L., 2014. Conducting multistage HFE validations - constructing Systems Usability Case. In: ISOFIC/ISSNP 2014. Jeju, Korea, pp. 1–10.

Law, E., Hvannberg, E., Cockton, G., 2008. Maturing Usability: Quality in software, interaction and value. Springer-Verlag, London.

Liinasuo, M., Norros, L., 2007. Usability case – integrating usability evaluations in design. In: Law, E., Larusdottir, K., Norgaard, M. (Eds.), COST294-MAUSE Workshop on Downstream Utility. Institute of Research in Informatics of Toulouse IRIT, Toulouse France.

Moray, N., 1997. Human Factors in Process Control. In: Salvendy, G. (Ed.), Handbook of Human Factors and Ergonomics. Wiley, New York, pp. 1944–1971.

Norros, L., 2017. Understanding Acting in Complex Environments: Building a Synergy of Cultural-Historical Activity Theory. Peirce, and Ecofunctionalism. Mind Cult. Act. 1–18.

Norros, L., 2004. Acting under Uncertainty: The Core-Task Analysis in Ecological Study of Work. VTT publications 546.

Norros, L., Savioja, P., Koskinen, H., 2015. Core-Task Design: A Practice-Theory Approach to Human Factors. Synth. Lect. Human-Centered Informatics 8, 1–141. https://doi.org/10.2200/S00631ED1V01Y201502HCI027.

O’Hara, J., 1999. A quasi experimental model of complex human-machine system validation. Cogn. Technol. Work 1, 37–46.

OECD/NEA, 2019. Multi-Stage Validation of Nuclear Power Plant Control Room Designs and Modifications. NEA No. 7466.

OECD/NEA, 2017. Human Factors Validation of Nuclear Power Plant Control Room Designs and Modifications, in: Proceedings of the Expert Workshop, NEA/CSNI/R

(2016)17/ADD1. Nuclear Energy Agency NEA, Charlotte, United States 19-21 February 2015.

Roza, M., Voogd, J., Sebalj, D., 2013. The Generic Methodology for Verification and Validation to support acceptance of models, simulations and data. J. Def. Model. Simul. 10, 347–365. https://doi.org/10.1177/1548512912459688.

Savioja P., Evaluating Systems Usability in Complex Work – Development of a systemic usability concept to benefit control room design VTT Science 2014, Aalto university PhD thesis, VTT.

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

Savioja, P., Norros, L., Salo, L., Aaltonen, I., 2014. Identifying resilience in proceduralised accident management actitivity of NPP operating crews. Saf. Sci. 68, 258–274.

Shamie, C., 2014. Continuous Engineering For Dummies. IBM Limited Edition. Wiley, Hoboken, NJ.

Simonsen, E., Bligård, L.-O., Osvalder, A.-L., 2020. Feasibility of methods for early formative control room system evaluation. Int. J. Ind. Ergon. 75, 102890 https://doi.org/10.1016/j.ergon.2019.102890.

Simonsen, E., Osvalder, A.L., 2018. Categories of measures to guide choice of human factors methods for nuclear power plant control room evaluation. Saf. Sci. 102, 101–109. https://doi.org/10.1016/j.ssci.2017.10.006.

United States Marine Corps, 2007. UNITED STATES MARINE CORPS, MCOTEA, Operational Test & Evaluation Manual. Quantico.

USNRC, (United States Nuclear Regulatory Commission), 2012. Human factors engineering program review model. USNRC, Washington, D.C., NUREG-0711, Rev. 3.

USNRC, (United States Nuclear Regulatory Commission), 1997. Integrated System Validation: Methodology and Review Criteria. USNRC, Washington, D.C., NUREG/CR-6393.

Valkonen, J., Tommila, T., Linnosmaa, J., 2016. Safety Demonstration of Nuclear I&C - An Introduction. VTT-R-00167-16.

Wise, J.A., Wise, M.A., 1993. Basic considerations in verification and validation. In: Wise, J.A., Hopkin, V.D., Stager, P. (Eds.), Verification and Validation of Complex Systems: Human Factors Issues. Springer, Berlin.

Woods, D., Sarter, N.B., 1993. Evaluating the impact of new technology on human-machine cooperation. In: Wise, J.A., Hopkin, V.D.H., Stager, P.S. (Eds.), Verification and Validation of Complex Systems: Human Factors Issues. Springer, Berlin.