Advances in Intelligent Systems and Computing 495

Sacit M. Cetiner
Paul Fechtelkotter
Michael Legatt *Editors*

Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries

Proceedings of the AHFE 2016 International Conference on Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries, July 27–31, 2016, Walt Disney World®, Florida, USA

Springer

Advances in Intelligent Systems and Computing

Volume 495

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

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

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

Sacit M. Cetiner · Paul Fechtelkotter
Michael Legatt
Editors

Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries

Proceedings of the AHFE 2016 International Conference on Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries, July 27–31, 2016, Walt Disney World®, Florida, USA

Springer

Editors
Sacit M. Cetiner
Oak Ridge National Laboratory
Oak Ridge, TN
USA

Michael Legatt
Electric Reliability Council of Texas, Inc.
Taylor, TX
USA

Paul Fechtelkotter
IBM Corporation
Medfield, MA
USA

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-41949-7 ISBN 978-3-319-41950-3 (eBook)
DOI 10.1007/978-3-319-41950-3

Library of Congress Control Number: 2016944353

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Advances in Human Factors and Ergonomics 2016

AHFE 2016 Series Editors

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

7th International Conference on Applied Human Factors and Ergonomics

Proceedings of the AHFE 2016 International Conference on Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries, July 27–31, 2016, Walt Disney World®, Florida, USA

Advances in Cross-Cultural Decision Making	*Sae Schatz and Mark Hoffman*
Advances in Applied Digital Human Modeling and Simulation	*Vincent G. Duffy*
Advances in Human Factors and Ergonomics in Healthcare	*Vincent G. Duffy and Nancy Lightner*
Advances in Affective and Pleasurable Design	*WonJoon Chung and Cliff(Sungsoo) Shin*
Advances in Human Aspects of Transportation	*Neville A. Stanton, Steven Landry, Giuseppe Di Bucchianico and Andrea Vallicelli*
Advances in Ergonomics In Design	*Francisco Rebelo and Marcelo Soares*
Advances in Ergonomics Modeling, Usability & Special Populations	*Marcelo Soares, Christianne Falcão and Tareq Z. Ahram*
Advances in Social & Occupational Ergonomics	*Richard Goossens*
Advances in Neuroergonomics and Cognitive Engineering	*Kelly S. Hale and Kay M. Stanney*
Advances in Physical Ergonomics and Human Factors	*Ravindra Goonetilleke and Waldemar Karwowski*
Advances in The Ergonomics in Manufacturing: Managing the Enterprise of the Future	*Christopher Schlick and Stefan Trzcielinski*
Advances in Safety Management and Human Factors	*Pedro Arezes*
Advances in Human Factors, Software, and Systems Engineering	*Ben Amaba*
Advances in Human Factors and Sustainable Infrastructure	*Jerzy Charytonowicz*
Advances in The Human Side of Service Engineering	*Tareq Z. Ahram and Waldemar Karwowski*

(continued)

(continued)

Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries	*Sacit Cetiner, Paul Fechtelkotter and Michael Legatt*
Advances in Human Factors in Sports and Outdoor Recreation	*Paul Salmon and Anne-Claire Macquet*
Advances in Human Factors and System Interactions	*Isabel L. Nunes*
Advances in Human Factors, Business Management, Training and Education	*Jussi Kantola, Tibor Barath, Salman Nazir and Terence Andre*
Advances in Human Factors in Robots and Unmanned Systems	*Pamela Savage-Knepshield and Jessie Chen*
Advances in Design for Inclusion	*Giuseppe Di Bucchianico and Pete Kercher*
Advances in Human Factors in Cybersecurity	*Denise Nicholson, Janae Lockett-Reynolds and Katherine Muse*

Preface

Human factors in energy that focus on the oil, gas, nuclear, and electric power industries aims to address the critical application of human factors knowledge to the design, construction, and operation of oil and gas assets, to ensure that systems are designed in a way that optimizes human performance and minimizes risks to health, personal or process safety, or environmental performance. The conference focuses on delivering significant value to the design and operation of both onshore and offshore facilities. Energy companies study the role of human behavior for safety and accident prevention; however, third party providers and different operators have different standards and different expectations. While oil and gas exploration and production activities are carried out in hazardous environments in many parts of the world, offshore engineers are increasingly taking human factors into account when designing oil and gas equipment. Human factors such as machinery design, facility and accommodation layout, and the organization of work activities have been systematically considered over the past twenty years on a limited number of offshore facility design projects to minimize the occupational risks to personnel, support operations and maintenance tasks, and improve personnel well-being. Better understanding for human factors issues also support the nuclear industry's move from analog to digital control rooms. Human considerations like lighting, temperature, even ergonomics, play important parts in the design. This book will be of special value to a large variety of professionals, researchers, and students in the broad field of energy modeling and human performance. The book is organized into four sections.

Section 1: Reducing Human Error Through Situation Awareness, Training, and Simulations
Section 2: Applying Human Factors: Building Better Processes, Procedures, and Organizations in Energy
Section 3: Human Factors in Energy
Section 4: Simulation and Interface Design for Safety Focused Research

This book will be of special value to a large variety of professionals, researchers, and students in the broad field of energy research, error prevention, and human performance who are interested in situation awareness, training, and simulations. We hope this book is informative, but even more—that it is thought-provoking. We hope it inspires, leading the reader to contemplate other questions, applications, and potential energy solutions in creating good designs for all.

We would like to thank the editorial board members for their contributions.

Oak Ridge, TN, USA — Sacit M. Cetiner
Medfield, MA, USA — Paul Fechtelkotter
Taylor, TX, USA — Michael Legatt
July 2016

Contents

Part I
Reducing Human Error Through Situation Awareness, Training, and Simulations

Toward a Descriptive Measure of Situation Awareness in Petrochemical Refining

Tristan Grigoleit, Hector Silva, Mary Ann Burress and Dan Chiappe

Abstract The petrochemical field is seeking to increase efficiency, improve safety, and lessen environmental impacts. One way to improve the performance of operators is to investigate their situation awareness (SA). Research has shown that SA is a predictor of performance. However, there is little consensus on how to measure SA. This study investigated two prominent techniques for measuring SA: the Situation Present Assessment Method (SPAM) and the Situation Awareness Global Assessment Technique (SAGAT). These techniques were examined for their psychometric properties in assessing SA among operators. Results of this investigation showed both SAGAT and SPAM could predict certain performance variables exclusively of each other. It was also found that SPAM and SAGAT were not sensitive to changes in SA resulting from differences in task workload. However, neither measure was significantly intrusive on primary task performance, suggesting that these metrics can be used in future experiments in petrochemical refining with further refinement.

Keywords Human performance · Situation awareness · Measurement · Petrochemical refining · SPAM · SAGAT · Simulation

1 Introduction

In hazardous operating environments, awareness of the current situation is often the only thing keeping workers and those they serve from becoming the victims of tragedy. Particularly grave are the potential results of a loss of situation awareness in the process control environment of petrochemical refining. In the case of the 2005 BP Texas City refinery explosion, the deficient situation awareness of

T. Grigoleit (✉) · H. Silva · M.A. Burress
Chevron Products Company, El Segundo, CA 90245, USA
e-mail: Tristan.Grigoleit@chevron.com

T. Grigoleit · H. Silva · D. Chiappe
California State University Long Beach, 1250 Bellflower Blvd, Long Beach, CA 90840, USA

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_1

operators was identified as a contributing factor leading to the deaths of 15 and injury of 180 people [1]. Situation Awareness is considered a major factor affecting performance in a number of complex systems. As a construct, it is of increasing interest to professionals attempting to define and measure it, as well as derive design recommendations from its implications. With increased government regulations coming to the petrochemical refining field [2] and the lack of Human Factors research conducted in a refining context, Human Factors professionals will be looking to previous research in other domains for guidance on how to address potential situation awareness pitfalls. This work covers the results of a study assessing the applicability of SA measures when used in conjunction with a simulated petrochemical process control task.

SA is a major factor influencing the likelihood of safe operation for console operators in petrochemical refining. SA is a factor that impacts the outcome of human performance without guaranteeing a successful outcome [3]. Operator performance in petrochemical refining is especially important due to the operator's role in ensuring safe operation and efficiency. Human error is a natural component of human performance. The inherent likelihood of human error and the profit-driven need to increase efficiency has pushed industries to employ the use of automation in conjunction with human operation. This increased reliance on automation can create a perilous environment for operators of petrochemical plants, as processes once requiring active engagement now more heavily rely on the monitoring of plant state changes [4]. If automation is employed in an improper manner and operator task load and awareness is not taken into account, it can lead to out of the loop syndrome, where operators have difficulty taking over a task when automation fails [5]. Insight into the nature of petrochemical operator SA can be applied toward the development of safety systems and training that combats the challenges of operating a process console in today's environment.

A review of SA theory presents two promising measurement techniques for measuring the SA of individual operators: the Situation Awareness Global Assessment Technique [6] and the Situation Present Assessment Method (SPAM; [7]). While SAGAT is the more popular measure of SA, research suggests that SPAM measures aspects of SA that are dependent on operator interaction with the environment [8]. Nuclear process control, which serves as a close analogy to petrochemical process control, has its own measure of SA in the Situation Awareness Control Room Inventory (SACRI; [9]). Since SACRI is strikingly similar to SAGAT in that both measures of SA are administered via freeze-probe techniques (i.e., pausing of the simulation when administering SA probes), they can be viewed as having similar strengths and limitations. These "offline" techniques differ significantly from "online" measures like SPAM that use real-time assessment (i.e., SA probes are administered while the operator performs task).

To determine the applicability of SAGAT and SPAM, assessments of these measures' intrusiveness, validity, sensitivity, and relationship to other constructs have been conducted and a set of criteria for SA metric assessment has been outlined [10–12]. Petrochemical refining researchers will seek to compare SA measures to these suggested criteria. To this end, we conducted an experiment with the goal of investigating the applicability of SAGAT and SPAM in a petrochemical refining context.

In this experiment, operators at a petrochemical plant performed a process control task on a medium fidelity simulator while receiving SAGAT and SPAM probes. The goal was to look at several aspects of these measures' applicability, including criterion validity via association with performance, sensitivity to changes in operator SA as a result of changes in experienced workload, and potential intrusiveness to the primary operating task. Scenarios differed in terms of the probe measurement technique administered (SAGAT, SPAM, and a baseline without measurement) as well as the workload level imposed by the scenario (Low and High workload) itself. SAGAT accuracy, SPAM accuracy, SPAM response latency, and NASA TLX scores were collected for each scenario. In addition, measures of performance previously identified by subject matter experts (SME) were collected during the simulated process control task.

Performance measures of relevance to scenarios were regressed with collected SA measurements via simple regression for all significant correlation pairs. It was predicted that a good measure of SA would have significant correlations with performance, thus describing the measure's criterion validity.

In addition to criterion validity, the sensitivity of SAGAT and SPAM was examined. Scenario difficulty was manipulated to generate differences in workload experienced by operators. As workload is a factor known to influence the SA of operators, we examined which metric of SA was better able to detect fluctuations of SA as a function of changes in workload. A workload manipulation check was employed in addition to previous SME validation to ensure the effectiveness of the workload manipulation. A good measure of SA was expected to be sensitive to changes in SA resulting from this workload manipulation.

Finally, the intrusiveness of SAGAT and SPAM were assessed. SAGAT and SPAM measures were compared to a baseline in terms of their effect on performance and workload. A good measure of SA is one that would not significantly impact operator performance or workload compared to a baseline without probe measurement.

2 Methods

2.1 Participants

Eleven participants were recruited from a pool of qualified console operators at a petrochemical refinery. The participant console operators manage a hydroprocessing plant that uses hydrogen to break down chemicals in crude oil.

2.2 Materials

Simulator and Scenario. A medium fidelity process control simulator mirrored after the hydroprocessing plant currently managed by participant console operators was used. Six process displays which received input through three membrane keyboards were used by operators to manage a simulated process. The displays and keyboards had the ability to access all plant information and make process adjustments, mimicking the actual plant's control scheme. A master control station used to manage various simulator functions such as scenario selection, start, pause, and stop was situated behind the process displays (see Fig. 1). The simulator collected data on numerous process variables at a 1 min sampling rate as the scenario progressed.

The scenario chosen for this study was a loss of hydrogen feed to a simulated hydroprocessing plant. In this scenario, the plant supplying hydrogen to the hydroprocessor trips offline. The participant was required to respond to this loss of hydrogen in an appropriate manner in order to prevent damage to equipment and maintain product quality specifications. For the purposes of this study, all trials were limited to 33 min to balance data collection efforts and costs.

Subjective Measures. Two subjective questionnaire measures were collected from participants: NASA Task Load Index (NASA TLX; [13]) and a SA usability questionnaire developed for this experiment. The NASA TLX measures workload on six dimensions and has been shown to have acceptable levels of reliability and validity [14].

The SA usability scale consisted of six questions that asked participants about their experiences with both SA probe techniques. Participants were asked about the simplicity of answering questions during their task via a 7-point Likert scale ranging from 1 "Very Simple" to 7 "Very Difficult." Participants were also asked if they believed that answering questions during their task affected their ability to manage the scenario. Next, participants were asked to what degree their ability to manage the scenario was affected by answering questions with each technique, also

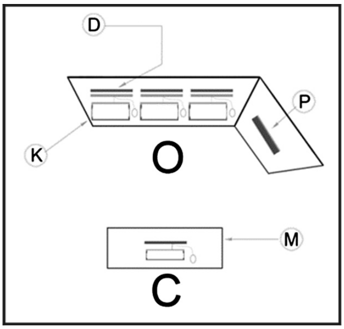

Fig. 1 Overhead view of the simulation setup. *O* is the participating console operator, *C* is the confederate administrator, *M* is the Master control station, *D* is a set of two displays (six total displays), *K* is a membrane keyboard (three total keyboards), and *P* is the SA probe station

on a 7-point scale ranging from 1 "Did not affect my ability at all" to 7 "Greatly affected my ability." Lastly, participants were asked which method they preferred overall.

SPAM and SAGAT. Both SAGAT and SPAM queries were administered using a touchscreen computer located adjacent to the participant control board. Recommendations for administering probe queries put forth by [12] were followed. For SPAM, a total of nine SA probe queries were presented to participants per trial; one about every 3 min starting at 6 min into each scenario. In addition to SA accuracies, the time it takes to first accept the probe question (ready latency) and to subsequently answer it (response latency) were collected but only the results of response latencies are discussed in this paper. SPAM proposes ready latency to be collected as a measure of workload while response latency is collected as an additional measure describing operator SA.

For SAGAT, probe queries were administered via a battery of three queries at three points in time during the scenario while the simulator screens were blanked and paused: first at about 7.5 min, second at about 15 min and lastly at about 22.5 min.

The time frames for the questions were not exact for either measure. This was done so that participants could not predict when a query would arise [3, 12]. Participants chose responses by pressing a button on the touchscreen. Situation awareness accuracies for both SPAM and SAGAT, and ready and response latencies for SPAM, were collected by the touchscreen computer.

Probe Queries. Probe queries and categories were created with guidance from a SAGAT-like measure of SA used in nuclear process control [9]. The three categories are defined as queries relating the recent past to the present, the present state to the normal state, and the present state to the near future. Methods for responding to queries were limited to "Yes" and "No" [7, 15] to simplify the process (See Appendix: Sample SA Queries).

2.3 Experimental Design

This experiment used a 2 (Task workload: less workload and more workload) by 3 (SA measurement: SPAM, SAGAT, and baseline) repeated measures design.

Independent Variables. Scenario difficulty was manipulated by increasing or decreasing certain feed rates to and within the plant to make the scenario more difficult to manage, but maintain similar participant responses to handle the situation. Within each level of difficulty, a SAGAT, SPAM, and baseline scenario was experienced by the participant, yielding a total of six trials.

Dependent Variables. Several dependent variables were collected during this study. These variables include performance, SA, and workload.

Performance. All performance variables were analyzed using numerical integration methodology and a performance threshold, thus lower performance variable values were indicative of better performance. See Table 1 for a complete list of performance variables used.

Table 1 List of performance variables used in current study

Performance variable	Abbreviation	Performance criteria
Hydrogen header pressure	HydPres	<750 PSIG counts against
Jet fuel freeze specification	JetSpec	Maintain −40 °F
Compressor temperature	ComTemp	>300 °F counts against
Distillation column temperature	CoTemp	<605 °F counts against
Distillation column level	CoLevel	Maintain between 40–70 % full
Reactor temperature	RxTemp	Maintain between 747–753 °F
Reactor 1 feed rate	Rx1Rate	Maintain between 32–42 MBPD
Reactor 2 feed rate	Rx2Rate	<46 MBPD counts against

Situation Awareness. The average percent accuracies to probe queries per participant were calculated for SAGAT and SPAM respective to their scenarios. SA was also measured by the average latencies to respond to correct probe queries during SPAM scenarios (response latency) [7].

Workload. Perceived workload was measured using the NASA TLX, which was administered immediately after each trial, including trials where SAGAT and SPAM were not administered. During SPAM scenarios, operators were instructed not to accept queries until their workload allowed. Therefore, workload was also measured using SPAM ready latencies [7], but ready latency results are not reported in this paper.

Usability. The participants' thoughts and preferences regarding the method of answering queries (SPAM or SAGAT) during their task were also collected through a usability questionnaire generated for this study.

2.4 Procedure

After informed consent, participants received three training blocks. The first block was meant to familiarize participants with operating the simulator. The remaining two blocks were administered to introduce and practice the two SA techniques independently.

Six 30 min experimental trials then proceeded. The six trials encompassed all combinations of task difficulty and SA probe technique and were counter-balanced. Ten minute breaks were allowed in between trials after having collected NASA TLX ratings. At the conclusion of the study, participants completed the SA usability questionnaire, were debriefed as to the purpose of the study, and dismissed. The complete study lasted roughly 8 h, which was well within the range of a 12 h shift that console operators are accustomed to.

3 Results

Separate analyses to address each of the goals of the experiment were conducted. An assessment of validity for each of the SA probe techniques was first conducted. Tests of sensitivity were also conducted on both techniques. An assessment of each probe techniques' potential intrusiveness was examined as well as a subjective appraisal of each technique by the operators. Finally, the results of a probe technique usability question administered to participants are described.

3.1 Validity Assessment

SAGAT and SPAM mean accuracies, and SPAM mean response latencies were regressed against the means of each of the performance variables with respect to each scenario type for all significant correlations. It was predicted that the SA measures that most strongly predict performance would have the strongest criterion related validity.

Regression analysis indicated that SAGAT accuracies explained 36.3 % of the variance in HydPres management performance in High workload conditions, $F(1, 9) = 5.127$, $p = 0.050$. Higher SAGAT accuracies were indicative of better performance with the management of HydPres, $\beta = -0.602$, $t(9) = -2.264$, $p = 0.050$.

SPAM accuracies predicted a significant amount of variance in performance on Jet Freeze Specification in Low workload conditions, $F(1, 9) = 6.685$, $p = 0.029$, $R^2 = 0.426$. Higher SPAM accuracies denoted better performance managing JetSpec, $\beta = -0.653$, $t(9) = -2.586$, $p = 0.029$.

Regression analysis also indicated SPAM accuracies significantly predicted Distillation Column Temperature in Low workload conditions, $F(1, 9) = 6.262$, $p = 0.034$, $R^2 = 0.410$. Higher accuracy on SPAM queries denoted better performance managing ColTemp, $\beta = -0.641$, $t(9) = -2.502$, $p = 0.034$.

Additional regression analysis results indicated that SAGAT accuracies explained 38.5 % of the variance in CoLevel management performance, $F(1, 9) = 5.629$, $p = 0.042$. Unlike with HydPres management, higher SAGAT accuracies were indicative of worse performance for CoLevel management, $\beta = 0.620$, $t(9) = 2.373$, $p = 0.042$.

Lastly, regression analysis found that SPAM accuracies accounted for 47.4 % of the variance in performance with managing Reactor Bed Temperatures in Low workload conditions, $F(1, 9) = 8.107$, $p = 0.019$. Higher SPAM accuracies denoted better performance managing RxTemp, $\beta = -0.688$ $t(9) = -2.847$, $p = 0.019$.

3.2 Sensitivity of SA Measures

Workload Manipulation Check. Before assessing the sensitivity of SA measures, the effectiveness of the workload manipulation was assessed (outside of the previous SME validation). To perform this manipulation check for scenario workload, paired-sample t-tests on participant's NASA TLX ratings between Low and High workload scenarios were run. As an additional manipulation check, paired-sample t-tests were also run on each of the eight participant performance variable outputs between Low and High workload scenarios. Finding differences in workload and performance between Low and High scenarios would indicate that the workload manipulation worked.

A difference in NASA TLX workload ratings was found between Low and High workload conditions in the SAGAT scenarios, $t(10) = -2.32, p = 0.042$. Differences in performance managing HydPres in SAGAT scenarios, $t(10) = -3.465, p = 0.006$, as well as a marginally significant difference in performance for HydPres in Baseline scenarios, $t(10) = -1.860$, $p = 0.093$, was observed. Differences in performance managing Rx1Rate between Low and High workload scenarios were found in all three SA measurement conditions, Baseline, $t(10) = -4.753$, $p = 0.001$, SAGAT, $t(10) = -3.962$, $p = 0.003$, and SPAM, $t(10) = -4.733$, $p = 0.001$. Differences in performance managing CoTemp in both Baseline, $t(10) = -2.951$, $p = 0.014$, and SPAM scenarios, $t(10) = -3.221, p = 0.009$ were also found. A difference was also found in managing RxTemp between Low and High workload conditions in the baseline conditions, $t(10) = -2.263$, $p = 0.047$. All significant findings were in the expected direction.

Sensitivity Assessment. Sensitivity of measures was assessed using three separate one-way, repeated measures Analyses of Variance (ANOVA), with scenario workload as the independent variable and measures of SA (SAGAT and SPAM response accuracies, SPAM response latencies) as the dependent measures. Sensitive measures of SA were predicted to detect differences in SA resulting from differences in task workload. However, no SA measurements were found to be sensitive to changes in workload for SAGAT Accuracies ($F(1, 10) = 0.01$, $p = 0.912$), SPAM Accuracies ($F(1, 10) = 0.06$, $p = 0.817$), and SPAM response latencies ($F(1, 10) = 0.81$, $p = 0.389$).

3.3 Intrusiveness Assessment

Intrusiveness was analyzed by running separate one-way repeated measures ANOVAs with three levels of measurement technique (SAGAT, SPAM, baseline) on each of the performance variables as well as NASA TLX ratings.

Only a marginal effect of SA measurement technique was found for NASA TLX ratings in Low workload scenarios, $F(2, 20) = 2.86$, $p = 0.081$, $\eta^2 = 0.22$. Operators rated their workload as lower during SAGAT scenarios ($M = 47.90$, *SE* = 5.01) compared to the Baseline scenarios ($M = 54.13$, $SE = 4.03$, $p = 0.043$), and SPAM scenarios ($M = 57.41$, $SE = 5.33$, $p = 0.062$). No differences were found between SPAM and Baseline scenarios, $p = 0.493$. No other effects of SA measurement technique were found for any other performance variables.

3.4 Usability Questionnaire

A questionnaire asking participants to rate their experiences with SA measurement techniques was included at the end of the experiment. There were a total of six questions within the questionnaire. The first pair of questions asked participants about their opinion of the measurement technique itself. The first question focused on participant experience with SAGAT: “How simple did you find answering questions while the screens were turned off and paused?” The question was formatted with a Likert scale ranging from 1–7 (1 was very simple, 7 was very difficult). Participants on average rated the SAGAT probe technique as 2.82 on this scale ($SD = 1.25$). The second question asked about participant experience with SPAM: “How simple did you find answering questions while the screens were on and active?” This question was presented with an identical 1–7 point Likert scale. Participants on average rated the SPAM probe technique as 3.36 on this Likert scale ($SD = 1.21$).

The next pair of questions directly addressed the influence of SA technique on scenario management. The question response format used a 1–7 point Likert scale ranging from “did not affect my ability at all” to “greatly affected my ability.” The third question focused on SAGAT, asking participants “how much did answering questions while the screens were turned off and paused affect your ability to manage the scenario?” Participants on average rated SAGAT as a 3.09 ($SD = 1.14$) on this scale. The fourth question asked about SPAM’s influence on scenario management “How much did answering questions while the screens were on and active affect your ability to manage the scenario?” Participants on average rated SPAM as a 3.55 ($SD = 1.44$) on this scale.

The last pair of questions asked about overall feelings toward probe measurement experienced in the experiment. Question five asked “In general, do you feel that answering questions during your task affected your ability to manage the scenario?” Six out of 11 participants (i.e., 55 %) stated they felt that probe techniques did not affect their ability to manage the scenario. Finally, the last question asked “Overall, which method did you prefer to use?” Participants could choose “screens on and active” (SPAM) or “screens off and paused” (SAGAT). Six out of 11 (55 %) stated they preferred “screens on and active.”

4 Discussion

This experiment was conducted with the goal of investigating the applicability of SAGAT and SPAM in a petrochemical refining context. Past research has described both the value and drawbacks of SA measures on a number of theoretical and objective criteria. Assessing SAGAT and SPAM probe techniques based on key psychometric properties such as validity, sensitivity, and intrusiveness [11, 12] was the manner that was sought to better describe which measure should be used in a petrochemical refining context to assess SA of plant operators. For the relationship between SA and performance, regression analyses showed that some measures of SA predicted performance, but only for certain performance variables. In the assessment of SA measure sensitivity, even though the workload manipulation was confirmed effective based on analysis and prior SME validation, measures of SA were not found to be sensitive to differences in scenario workload. Finally, SA measures were not found to impact performance or perceived workload of participants in this experiment, indicating SA measures were not significantly intrusive to the primary task.

Results of regression analyses conducted in this study align with past SA research suggesting relationships between performance and SA exist. While SAGAT Accuracies predicted HydPres and CoLevel performance, SPAM Accuracies predicted CoTemp, JetSpec, and RxTemp performance. No relationships, however, were found for SPAM response latencies. This exclusive relationship between performance variables and certain measures of SA show the importance of using different types of tools to measure SA. Future research should strive to identify why certain tools work better for certain aspects of performance.

An unexpected finding was that SA negatively predicted CoLevel performance, where the better a participant's awareness the worse their predicted performance. This unexpected result may speak to the nuances of process control. A possibility explained by SMEs is that the greater a participant's awareness of their situation the more they may have intentionally chose to disregard CoLevel management for processes deemed more important. This finding suggests the potential descriptive value of SA measures in understanding Operator management strategies.

Although a workload manipulation check implied workload was to an extent successfully manipulated, measures of SA used in this experiment did not show differences in measured SA across Low and High workload conditions. Higher levels of workload were expected to have created lower levels of SA, as well as lower workload levels expected to promote higher SA. A possible reason for the lack of difference in SA between conditions was that the workload manipulation was not strong enough. This is potentially supported by the finding that workload only affected some but not all variables.

A more likely factor influencing the sensitivity of SA measures used in this study is the SA probe administration itself. The SA probe query scheme used in this experiment was an adapted form of SACRI [9]. Additional pilot testing may have been warranted to tailor the question scheme to the petrochemical process operator's conceptualization of SA. A review of SA literature by [16] highlights the differences in conceptualization of SA by domain. While some domain practitioners may prioritize future state information, others may emphasize past state information, which is the case for process control operators that use awareness of the past state to investigate root causes of process disturbances.

Further testing of SA measurement techniques against the three major measurement criteria used in this study could lead to the improved applicability of these tools in domains such as petrochemical refining. As demonstrated by the findings of this study, SA measures did no intrude on primary task performance, which is promising for researchers seeking to apply SA measures in the domain of petrochemical refining.

Although low sample size was a limitation of this study, efforts had to be taken so that the study was monetarily and operationally feasible to complete since incumbent operators were used. Although not unusual in applied work, efforts should still be made to obtain a large enough sample size to draw reliable conclusions.

The results of this investigation could help guide researchers and practitioners in choosing the SA measurement technique that would be most appropriate for their domain. Measuring SA could inform decisions regarding system, interface, and training changes. Understanding and describing the SA of console operators is the first step to improving operator SA by guiding the construction and development of training programs to improve their awareness. Improving the SA of operators has positive implications for the petrochemical refining industry; more situation-aware operators would detect, diagnose, and resolve incidents sooner. This would lead to less frequent and critical events, fewer environmental impacts, higher safety of workers, and more monetary gains instead of losses.

Appendix: Sample SA Queries

Time relation	Query
Past with present	In the last 3 min, did [process] increase?
Past with present	In the last 3 min, did [process] decrease?
Present with ideal	In comparison with steady state operation, is [process] currently higher?
Present with ideal	In comparison with steady state operation, is [process] currently lower?
Present with future	In the next 3 min, will [process] increase?
Present with future	In the next 3 min, will [process] decrease?

References

1. Saleh, J.H., Haga, R.A., Favarò, F.M., Bakolas, E.: Texas city refinery accident: case study in breakdown of defense-in-depth and violation of the safety-diagnosability principle in design. Eng. Fail. Anal. **36**, 121–133 (2014)
2. Department of Industrial Relations Division of Occupational Safety and Health, http://www.dir.ca.gov/dosh/DoshReg/Process-Safety-Management-for-Refineries/PSM-Draft-Regulation.2014-10-31.pdf
3. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Factors: J. Hum. Factors Ergon. Soc. **37**, 32–64 (1995)
4. Kaber, D.B., Endsley, M.R.: Out of the loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. Process Saf. Prog. **16**, 126–131 (1997)
5. Kaber, D.B., Endsley, M.R.: The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. Theor. Issues in Ergon. Sci. **5**, 113–153 (2004)
6. Endsley, M.R.: Measurement of situation awareness in dynamic systems. Hum. Factors: J. Hum. Factors Ergon. Soc. **37**, 65–84 (1995)
7. Durso, F.T., Dattel, A.: SPAM: The real-time assessment of SA. In: Banbury, S., Trembley, S. (eds.) A Cognitive Approach to Situation Awareness: Theory, Measures and Application, pp. 137–154. New York, NY (2004)
8. Chiappe, D.L., Strybel, T.Z., Vu, K.P.L.: Mechanisms for the acquisition of situation awareness in situated agents. Theor. Issues in Ergon. Sci. **13**, 625–647 (2012)
9. Hogg, D.N., Folleso, K., Strand-Volden, F., Torralba, B.: Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. Ergonomics **38**, 2394–2413 (1995)
10. Durso, F.T., Bleckley, M.K., Dattel, A.R.: Does situation awareness add to the validity of cognitive tests? Hum. Factors: J. Hum. Factors Ergon. Soc. **48**, 721–733 (2006)
11. Salmon, P.M., Stanton, N.A., Walker, G., Green, D.: Situation awareness measurement: a review of applicability for C4i environments. Appl. Ergon. **37**, 225–238 (2006)
12. Strybel, T.Z., Vu, K.L., Chiappe, D., Battiste, V., Johnson, W., Dao, Q.V.: Metrics for situation awareness, workload, and performance: manual for online probe administration. (Report No. NASA/TM-2011-215991). NASA Ames, Mountain View, CA (2011)
13. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock P.A., Meshkati, N. (eds.) Human Mental Workload, pp. 139–183. Elsevier, Amsterdam, The Netherlands (1988)
14. Gawron, V.J.: Human Performance Measures Handbook. Lawrence Erlbaum, Mahwah (2000)
15. Morgan, C., Chiappe, D., Kraut, J., Strybel, T.Z., Vu, K.P.L.: An investigation into the potential sources of interference in SA probe techniques. In S. Landry (ed.), Advances in Human Aspects of Aviation, pp. 623–631. CRC Press, Boca Raton, FL (2012)
16. Lau, N., Jamieson, G.A., Skraaning, G.: Distinguishing three accounts of situation awareness based on their domains of origin. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 220–224. SAGE Thousand Oaks, CA (2013)

Full Scale Evaluation of How Task-Based Overview Displays Impact Operator Workload and Situation Awareness When in Emergency Procedure Space

Zachary Spielman, Rachael Hill, Katya LeBlanc, Brandon Rice, Gordon Bower, Jeffrey Joe and David Powers

Abstract Control room modernization is critical to extending the life of the 99 operating commercial nuclear power plants (NPP) within the United States. However, due to the lack of evidence demonstrating the efficiency and effectiveness of recent candidate technologies, current NPP control rooms operate without the benefit of various newer technologies now available. As nuclear power plants begin to extend their licenses to continue operating for another 20 years, there is increased interest in modernizing the control room and supplementing the existing control boards with advanced technologies. As part of a series of studies investigating the benefits of advanced control room technologies, the researchers conducted an experimental study to observe the effect of Task-Based Overview Displays (TODs) on operator workload and situation awareness (SA) while completing typical operating scenarios. Researchers employed the Situation Awareness Rating Technique (SART) and the NASA Task Load Index (TLX) as construct measures.

Keywords Nuclear power plant · Control room · HIS · Automation · Human-automation collaboration · Human performance · Situation awareness · Overview displays · Workload

---

Z. Spielman (✉) · R. Hill · K. LeBlanc · B. Rice · G. Bower · J. Joe · D. Powers
Idaho National Laboratory, 2525 Fremont Avenue, PO Box 1625, Idaho Falls, ID 83415-3818, USA
e-mail: zachary.spielman@inl.gov

R. Hill
e-mail: rachael.hill@inl.gov

K. LeBlanc
e-mail: katya.leblanc@inl.gov

B. Rice
e-mail: brandon.rice@inl.gov

G. Bower
e-mail: gordon.bower@inl.gov

J. Joe
e-mail: jeffrey.joe@inl.gov

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_2

1 Introduction

1.1 Background

The following research is part of the United States (U.S.) Department of Energy (DOE) sponsored Light Water Reactor Sustainability (LWRS) Program conducted at the Idaho National Laboratory (INL). The LWRS program is a collaborative effort with industry research and development (R&D) programs to establish the technical foundations for licensing and managing the long term, safety, and economical operation of current nuclear power plants (NPPs). One primary mission of the LWRS program is helping the U.S. nuclear industry modernize operating capabilities by adopting tested technologies and engineering solutions found to facilitate safe operation through a NPP's extended lifetime [1]. The Electrical Power Research Institute (EPRI) describes several potential drivers of control room modernization including [2]:

1. To address obsolescence and lack of spare parts
2. To meet the need for equipment replacement due to high maintenance cost or lack of vendor support for existing equipment
3. To implement new functionality necessary for adding beneficial capabilities
4. To improve plant performance, HSI functionality, and reliability
5. To enhance operator performance and reliability
6. To address the difficulties in finding young professionals with education and experience with older analog technology.

Items one and two both refer to the growing need that must be addressed for a NPP to continue operating while items three through six address possible improvements that should be considered during a modernization.

Despite the six drivers listed; none of the 99 currently operating NPPs within the US have undergone a full control room modernization [3]. Current control rooms are not built to facilitate a simple modernization effort. Between various regulatory, operations benchmarks, and licensing hurdles the control room requires careful deliberation as to what updates occur first. Note also that any physical changes to the current analog boards would be, for all practical purposes, permanent. Due to such difficulty a hesitation towards modernization exists because the evidence demonstrating the value of candidate technologies that goes beyond the original design concept (i.e. more than replacing analog alarm panels with identical panels) is lacking. Therefore strong evidence to validate the struggle and work required to modernize must be found before a plant can commit to upgrading or supplementing the control room. As a result current NPP control rooms operate without the benefit of various capabilities newer technologies offer.

However, as NPPs have been extending their licenses to continue operating for another 20 years the interest towards modernizing the control room and supplementing the existing control boards with advanced technologies has increased. The recent boost of interest has provided the opportunity to work with the industry to identify technology that improves upon current operating systems. The LWRS project aims to provide industry with tested solutions expected to exceed the performance of the current labor intensive model and address the latter items on EPRI's list.

1.2 Hybrid Approach

Hybrid control rooms with supplemental technology have been proposed as an initial solution to the complicated modernization process. As a stepping stone to full modernization, new technologies might supplement rather than replace portions of the control room [3]. Furthermore plants can make changes to the control room in a shorter amount of time, minimize permanent changes to the analog boards, and is generally less of an investment. Supplementing offers NPPs the opportunity to 'test drive' the technology as a potential replacement before committing to a physical redesign of the control room. To minimize the number of failed 'test drives', candidate technologies undergo testing to guide decisions towards the best options. However, the attempt to evaluate candidate technologies within a full-scale main control room operating context is accompanied by challenges in recreating realistic scenarios and testing real world operators. The challenges include:

- Finding available and experienced crews of operators
- Comparing candidate technology using unbiased expert operators
- Training operators on a generic operating system within a feasible time span
- Developing a facility flexible enough to mimic various NPP control rooms
- Measuring impact of candidate technology on operator-system interaction.

Due to the complexity of control room operations, real-world operators in a high-fidelity control room will be used to test candidate technologies in a control room setting to gain results generalizable to the industry as a whole. However, using real operators in a familiar control room may bring in operator bias towards the familiar. Instrumental in overcoming these challenges is the Human-System Simulation Laboratory (HSSL). The lab hosts a simulated operator control room environment equipped for testing candidate technologies. The flexible interface can simulate either a generic pressurized water reactor (gPWR) or the control boards of an industry control room depending on need. The lab is equipped with various recording technology for data collection and, later, analysis.

Recently a pilot study was launched to begin addressing these challenges and develop a protocol to achieve meaningful results without sacrificing fidelity. Researchers under the LWRS pathway at INL are investigating which near-term control room technologies may enhance operator performance in hybrid control rooms.

The advantage to using the gPWR is to reduce bias a simulated by creating even ground for all operators; no single operator will have more experience in the generic control room than another. The generic control room also supports the ability to generalize results across various plants. Once potential candidate technology is identified using the gPWR simulation, the HSSL can mimic the control rooms of industry partners interested in seeing how the technology affects their particular plant operations method.

Finding which candidate technologies improve operator performance is expected to lead to updating current control room technology as well as inform design concepts of future control rooms and continue evaluation of evolving or novel candidate technologies. The results described here are the beginning of an ongoing effort to verify and validate the impact candidate technologies have on operator and plant performance.

Researchers at the INL conducted a full-scale evaluation and began establishing a method of testing sensitive to changes in performance caused by candidate technologies. The study observed the effect of Task-based Overview Displays (TODs) on operator workload and SA while completing typical operating scenarios.

1.3 Task-Based Overview Displays

The TODs are expected to reduce operator workload and improve SA by using information-rich design principles to provide many variables and trends occurring in a NPP simultaneously at a given moment. The displays are "Task based" because the available content is dependent on current plant state. Content important for monitoring and maintaining a plant in normal operating procedure space is available during regular procedures. When the plant enters emergency procedure space post reactor trip, the available information is adapted to provide the crew with information relevant to the tasks involved in emergency procedure space [4].

Using trend displays and graphical representations of pumps, valves, and flow balance the TODs break from traditional analog displays to provide greater amounts of information with a single display element to increase the rate and ease of information gathering to improve situation awareness when a plant condition changes. Currently, most NPP control rooms are comprised of "boards" that contain the controls, alarms, and indicators, to specific aspects of the plant. Normally, operators must monitor a range of dials and indicators that requires them to move up and down the boards, or 'ping-pong' between boards to retrieve the information necessary to successfully complete their task. The TODs are intended to reduce the need to ping-pong by placing task relevant information for a single operator in a single location easily accessible from any position along the boards. For a detailed description of the design principles used in the TOD's refer to [4].

2 Methods

2.1 Participants

INL employees with nuclear operations experience were recruited to participate in this study. Seven operators participated serving the roles of Senior Reactor Operator (SRO), Reactor Operator (RO) and Balance-Of-Plant Operator (BOP). All the participants were male averaging 48.6 years of age. All of the participants have bachelor's degrees in nuclear/mechanical engineering and experience as Navy nuclear operators; however none of the participants currently work in operations.

2.2 Environment and Stimuli

The HSSL houses many integral components used to complete this study. Most notably is the full scale, full scope, and reconfigurable virtual NPP control room simulator. The simulator consists of fifteen bays each consisting of three, 47 inch LCD screens (measured diagonally). The bottom two LCDs have touch-screen capabilities via infrared overlays. A Dell OptiPlex desktop computer running Microsoft Windows 7 Professional is housed inside each of the bays, and acts as the client to the simulator software code running on a secure server. The server room houses backend servers that allow for rapid image deployment via Free Open Ghost (FOG), Windows Server 2008 R2 for different plant models and configurations, and Microsoft Hyper-V utilization to satisfy virtualization needs. The bays are mounted on frames with lockable wheels for mobility, maintainability and convenience. Because of these features, the control room is reconfigurable into almost any NPP control room layout.

Other resources used in the HSSL include virtual machines, an air-gapped network infrastructure, Foscam wireless Internet Protocol cameras for video capture, and Peavey wireless lavaliere microphones for audio capture. Blue Iris software is used to record and synchronize the audio and video feeds. There is also an observation room in the HSSL, in which resides a Dell OptiPlex computer that serves as the Instructor Station for the simulator. From there, all simulator activities are controlled, including: powering on the bays, starting the simulator, loading initial conditions, inserting a malfunction scenario, and powering down the bays. The core of the system is run by GSE Systems Java Application Development Environment (JADE) simulator platform (Fig. 1).

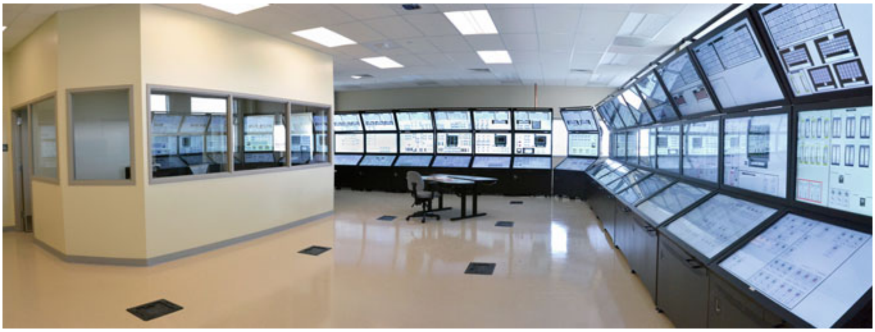

Fig. 1 The human system simulation laboratory (HSSL) with the generic pressurized water reactor control boards displayed on the three-screen bays

2.3 Scenarios

The scenarios were designed as two sets: each set with relative similarity in complexity and difficulty. Furthermore, scenarios were designed to move operators from normal procedure space to emergency procedure space. The movement through these procedure spaces is important as there are many decision gates the crew must accurately navigate. Decisions are easy targets when evaluating crew performance. Either the crew makes a correct or incorrect decision at each point providing a simple evaluation for non-expert observers. Scenario set B was the practice set used to familiarize the operating crews with the plant. The set of scenarios consisted of:

A. Scenario Set A: Loss of Coolant Accident (LOCA)
   1. Simple Case
   2. Complex Case
B. Scenario Set B: Faulted Steam Generator (SG)
   1. Simple Case
   2. Complex Case
C. Scenario Set C: Steam Generator Tube Rupture (SGTR)
   1. Simple Case
   2. Complex Case

A1: Simple LOCA. This scenario is a standard LOCA, a commonly trained scenario in commercial NPPs. The only complication was the leak occurring during a power ramp, meaning the crew will already be involved in a procedure as the break begins. The scenario progresses relatively slowly, giving the crew ample time to react and identify the issue before any critical issues occur. The scenario is ended once the affected leg is isolated and the crew is transitioning to “post-LOCA cooldown and depressurization”.

A2: Complex LOCA. Scenario A2 progresses quickly and has a masking fault making it the more complex of the two LOCA scenarios. Heater drain pumps are programmed to fail sequentially within a minute of each other causing alarms to trip and the crew to take action. As the crew is working to resolve the heater drain pump issue a large LOCA is inserted immediately tripping the reactor. Crews have to work quickly to stabilize the reactor and diagnose the issue. The scenario will end as the crew begins establishing alternate cooling routes.

B1: Simple Faulted Steam Generator. The scenario begins with a masking fault when charging pump ‘A’ experiences a shaft shear rendering the pump ineffective however indicators will not pick up the pump failing, only the flow ceasing. As the crews attempt to resolve the situation a steam generator experiences a main line break outside of containment. The resulting action moves the crew to emergency procedure space where they continue to stabilize the plant and diagnose the issue.

B2: Complex Faulted Steam Generator. The scenario begins with a masking fault when a charging line begins to slowly leak at 50 gallons per minute (gpm) moving the crew into abnormal procedure space. As the crew is determining if they can maintain the leak and still remain in operation the main steam line break is inserted automatically tripping the reactor. The reactor trip, as always, moves the crew into emergency procedure space. During this reactor trip however, the turbine does not automatically trip with the reactor as it is designed to do. The crew is monitored to be sure they detect the abnormality in system operation. The scenario is ended once the crew reestablishes continued Reactor Coolant System flow.

C1: Simple SGTR. The scenario begins with a masking fault, the failure of a pressurizer instrument expected to be identified before the SGTR begins and does not significantly affect diagnosis. The pressurizer indicator shows a sudden drop to the low end, causing alarms as well as three heater pumps to turn on as the system attempts to increase pressure. A redundant indicator that does not break should indicate to the crew that the indicator is malfunctioning. The crew is expected to appropriately handle the malfunction and at the resolution of which an SGTR in is injected. A slowly developing leak eventually leads to charging flow and pressurizer alarms. The size of the leak will exceed makeup capabilities and the crew will need to trip the reactor (an automatic trip would represent a failure condition). The scenario concludes when the crew has correctly identified the issue.

C2: Complex SGTR. The scenario begins normally with a masking fault, a tripped heat exchanger pump, inserted shortly after. The standby pump will start from low system pressure almost immediately, and the crew will progress through Abnormal Operation Procedure space (AOP) to confirm that this has happened and that parameters are returning to normal. As the crew is working through AOP, the SGTR is inserted. The team should eventually transition into the appropriate AOP once loss of primary side coolant is identified. This should progress similarly to the previous SGTR, with step 4 as a possible trip point depending on how quickly the crew has progressed but eventually the reactor will trip. The scenario concludes when the crew has correctly identified the issue.

2.4 Measures

Identifying differences in situation awareness and workload afforded by the TODs was the research objective. By counterbalancing which scenarios were accompanied by the TODs between crews and standardizing the questionnaires across conditions the researchers attempted to isolate how the displays impacted the two aspects of operator working space. The following describes the measures used and the experiment design.

Researchers used the Situation Awareness Rating Technique (SART) to determine the level of participant's situation awareness with and without the OVDs. Often utilized for aircrew studies, SART is a common questionnaire validated for measuring SA for the duration of a task [5]. The questionnaire is administered after a completed trial and presented in an ordinal format with nine items rated by the participants on a scale of one to seven (low to high).

- Instability of situation
- Variability of situation
- Complexity of situation
- Arousal
- Concentration of attention
- Division of attention
- Spare mental capacity
- Information quantity
- Familiarity with situation.

The NASA-TLX was used to measure the subjective workload of the participants after completing a scenario. This evaluation is a subjective six-item scale that is a widely used and validated scale for measuring workload after a task. It was developed specifically for the aviation industry [6], though it has been used in many studies in a wide variety of fields, including many NPP control room studies [7]. It is accepted as a reliable measure of workload differences between tasks. It measures six constructs on a scale of one to ten.

- Frustration Level
- Effort
- Performance
- Temporal demand
- Physical demand
- Mental demand.

The posttest questionnaire was used to gather additional data not already captured by NASA-TLX and SART. It includes four follow up items;

- How difficult was the scenario?
- How was your performance during the scenario?

- Were the overview displays available during the scenario?
- If the overview displays were available during the scenario, did you use them? Why or why not?

There were additional measures that were recorded but were not integral to the purposes of this report.

2.5 Procedure

Participants with a background in nuclear reactor operations were recruited for the study. They first provided informed consent to participate in the study acknowledging that they understood their rights during the study. Each crew worked together over three contiguous days to complete training and testing together. A day was considered an eight hour business day.

Day one began with training videos to familiarize the participants with the mechanical structure of gPWR and how to operate the simulated plant using a combination of video and referencing piping and instrumentation diagrams. The crews were allowed to ask questions and find what was being described on the video, on the simulator boards.

Day two continued the mechanical over view as well as the conduct of operations for the plant. Participants were given intermittent opportunities to "walk the boards" and interact with the simulator performing common actions associated with their role in the plant operations. Day two concluded with operators completing two formal training scenarios; a simple and a complex 'steam-generator tube rupture'.

Day three consisted of four testing scenarios used to evaluate how the candidate technologies impacted operator performance. The four scenarios were constructed in a 4 × 2 fashion. Two types of scenarios of two complexity levels were ran; a "Loss of Coolant Accident" and a "Steam Generator Failure". Each type came in two versions, with and without the TODs present. The order was counterbalanced across crews to account for learning effects.

Protocol was the same across every scenario. Crews were both video and audio recorded during all scenarios. The participants were handed over operation of the plant at full power with no known or occurring issue with exception to a single scenario requiring a power ramp. Once in control, the crews monitored the plant as faults were entered from the observation room by the simulator controller located behind the operators. All Scenario details were guarded from the crew beforehand to gain as naturalistic responses as possible. Operators were also not aware of which scenarios the TODs would be available. However the researchers had established a display order a priori. When the displays were available it was clear to the operators.

During the scenario crews were stopped at three different points to fill out a short freeze-probe questionnaire requiring only couple minutes at which point all simulator screens were blacked out to remove any hints or answers to the questions. The crews were allowed a quick brief to refocus on the task at hand right before resuming the simulation. At the conclusion of each scenario, each crew member filled out the NASA-TLX, the SART, and the post-test questionnaire. Note the post-test questionnaire asked if displays were present during a scenario to ensure the crews were aware of the TOD's provided them. During this time the simulator was reset, data exported, and procedure lists refreshed.

3 Results

All three crews were able to carry out each scenario to the pre-selected termination point. Crews were also able to answer the questionnaires easily with few questions themselves. The final day of testing went smooth for all crews with a small adjustment period for the final crew as one member cancelled last minute.

Using the Likert Scale values collected in the SART and NASA TLX, researchers made many comparisons to determine where meaningful differences occurred in the study. All the individual questions from the self-reporting measures were kept separated to search for differences within the questionnaires (i.e. in mental demand but not physical demand of the NASA TLX). Then researchers identified pair groupings that differed in only one respect. The average Likert rating for each question of one half a pair grouping was compared with the average of the other half. The result indicated if one pair was consistently rated greater than the other across the questionnaires. The comparisons were as follows:

1. TOD availability
   - On
   - Off
2. TOD on/off condition broken down by crew designation
   - RO
   - SRO
   - BOP
3. Scenario complexity
   - Simple
   - Complex
4. Scenario complexity by crew designation
   - RO
   - SRO
   - BOP
5. Task display by complexity
   - Simple
   - Complex.

The first comparison was to determine the candidate technology's impact on operator performance in general. No distinctions were made between crews or scenarios in case there was a main effect of display. However, the different crew designations have different tasks and different overview displays to use. Breaking down TODs by crew designation was a natural decision to understand the technology's impact on individual roles in the control room. Hence, the second grouping was compared. The results found no difference between having and not having the displays present during the scenario for any of the three roles.

The third comparison was helped verify if different scenario complexities had the intended impact on operator workload. Again, no significant differences were found to determine if simple scenarios were easier to complete than complex scenarios. The fourth comparison took a deeper look to see how scenario complexity affected individual roles. The result was no different than the second comparison.

The fifth and final comparison was used to evaluate how the presence of theoretically more accessible information made a difference when the need for information was increased. No difference was discovered within the fourth comparison either. Since a difference between the complex and simple scenarios was not confirmed this does not come as a surprise.

The post-test debriefing asked for operators subjective input regarding the use of the displays. Of 14 responses total one said the displays were not useful attributing their remark to the number of valve changes they were doing. The positive statements often contained similar content and more than one reason stating the TODs were helpful. The positive 13 responses included statements similar to the following:

- "Yes, useful for tracking all trends …" (4/13)
- "Ability to validate plant information" (4/13)
- "Yes, good indicators for checking [Steam Generators]" (3/13)
- "Checked pressurizer display" (5/13).

Thus the operator's majority responses were positive towards the presence of the TOD's.

4 Discussion

The researchers evaluated whether an environment was able to test crew performance while manipulating the various candidate technologies and observing the resulting effect on crew performance, SA, and workload. It was understood from the beginning that not every measure may produce a measurable result due to the small sample size. Such a result was acceptable as another goal was testing the facilities ability to collect measures, incorporate new technologies, and host operating crews able to perform scenarios at an acceptable level.

The results of the measures were non-descript as to whether the TODs made an impact on operator performance. It appears as though there was no difference made however with such a small sample size it is far more likely the measures were not sensitive to changes in performance. The finding exemplifies the need for multiple crews to perform the tasks in order to validate the testing measures and scenario differences. Due to the positive responses from all operators regarding the TOD's there is promise the displays may have a positive impact on performance.

The project was successful in addressing challenges associated with evaluating candidate technologies within a full-scale main control room operating context. While some limitations were present in reducing the number of crews the project could host, a viable source at the INL was found in previously experienced Navy Nuclear operators. Additionally, these participants were able to train and carry out tasks on the gPWR within a feasible time limit. Further attention towards the training regime would likely increase the positive transfer and reduce the time required for training allowing more time for testing. Using ex-operators in a generic control room leads us to believe there is reduced bias towards using the familiar systems but without sufficient sample size it is difficult to determine exactly.

The HSSL demonstrated a functional flexibility that provided researchers with the necessary tools to administer various performance measures collected during this study. The observation room allows for unobstructed visual access to all the actions being taken in the control room while all plant parameters are recorded and loaded to a safe directory for later viewing. The simulator demonstrated the capability to freeze, hide all screens, and quickly unfreeze again during a scenario for freeze probe questionnaires. The crews could operate the simulator with enough proficiency to complete the tasks without their inability hindering their task completion.

Overall, the study has demonstrated the HSSL has all the capability to perform full-scale evaluations of crew performance during control room scenarios. The laboratory has the flexibility to work with varying candidate technologies in an applied workspace. Furthermore multiple measures may be gathered at once and stored in a single place. The training developed for the study was sufficient in training operators to carry out scenarios on the generic plant in a feasible time period. Finally, to gain meaningful results the number of crews run and the resources to do so need to be considered when beginning such an endeavor to gain the most information from a single evaluation. The next task is to bring in a sufficient number of experienced operating crews and begin making decisions towards how the control room can best be upgraded.

References

1. United States of America, Department of Energy, Nuclear Energy: Light Water Reactor Sustainability Program Integrated Program Plan, revision 3 edn., vol. INL/EXT-11-23452, pp. 43–61 (2015)
2. Naser, J., Hanes, L., O'Hara, J., Fink, R., Hill, D., Morris, G.: Guidelines for Control Room Modernization as Part of Instrument and Control Modernization Programs (2004)

3. LeBlanc, K., Joe, J., Rice, B., Ulrich, T., & Boring, R.: Benefits of Advanced Control Room Technologies: Phase One Upgrades to the HSSL and Performance Measures, INL/EXT-15-3538 revision 0 (2015)
4. Braseth, A.O.: Information-Rich Design: A Concept for Large-Screen Display Graphics: Design Principles and Graphic Elements for Real-World Complex Processes (2015)
5. Taylor, R.M.: Situational awareness rating technique (SART): the development of a tool for aircrew systems design. AGARD, Situational Awareness in Aerospace Operations 17 p. (SEE N 90-28972 23-53) (1990)
6. Hart, S.G.: NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 50, no. 9, pp. 904–908, Sage Publications (2006)
7. LeBlanc, K., Powers, D., Joe, J., Spielman, Z., Rice, B., Fitzgerald, K.: A Pilot Study Investigating the Effects of Advanced Nuclear Power Plant Control Room Technologies: Methods and Qualitative Results, INL/EXT-15-36432 revision 0 (2015)

Identifying Macrocognitive Function Failures from Accident Reports: A Case Study

Peng Liu, Xi Lyu, Yongping Qiu, Juntao Hu, Jiejuan Tong and Zhizhong Li

Abstract Reliable macrocognitive functions are important for maintaining system safety. Few studies were conducted to investigate macrocognitive function failures in a complex system. NUREG-2114 proposes a cognitive framework connecting macrocognitive function failures, proximate causes, failure mechanisms, and performance influencing factor (PIFs). This model can serve as a model for analyzing human failure events in human reliability analysis (HRA). This study investigated macrocognitive function failures in a complex environment and also examined the usability of the cognitive framework in the HRA qualitative analysis. A total of 103 investigation reports of incidents and accidents from a petrochemical plant in China were involved. It was found that 35 % of the incidents and accidents could be attributed to human errors. Failures of action implementation and team coordination were the dominant failures. This study also gave the information of proximate causes, failure mechanisms, and PIFs for each macrocognitive function failure. The usability issue of the cognitive framework in NUREG-2114 was discussed. It seems that the current cognitive framework needs to be improved to inform HRA.

P. Liu (✉)
Department of Industrial Engineering, College of Management and Economics,
Tianjin University, Tianjin 300072, China
e-mail: pengliu@tju.edu.cn

X. Lyu · Z. Li
Department of Industrial Engineering, Tsinghua University, Beijing 10084, China
e-mail: lvx12@mails.tsinghua.edu.cn

Z. Li
e-mail: zzli@tsinghua.edu.cn

Y. Qiu · J. Hu
Shanghai Nuclear Engineering Research and Design Institute, Shanghai 200233, China
e-mail: qiuyp@snerdi.com.cn

J. Hu
e-mail: hujuntao@snerdi.com.cn

J. Tong
Institute of Nuclear and New Energy Technology, Tsinghua University, Beijing 10084, China
e-mail: tongjj@tsinghua.edu.cn

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_3

Keywords Macrocognitive function · Human reliability analysis · Human error · System safety · Accident investigation

1 Introduction

System safety is highly dependent on human performance. Human is widely acknowledged to be a major cause of incidents and accidents in complex systems [1, 2]. Convergent evidence shows that the contribution of human error to incidents and accidents is approximately 30–90 % in safety-critical systems [3]. From 1991 to 2011 more than 50 % of operating events were implicated with operator errors in Chinese nuclear power plants (NPPs) [4].

Several human error taxonomy systems have been proposed to classify human error types and identify error enforcing conditions, in order to understand which aspect of human information processing fails easily and why. For example, Wiegmann and Shappell [5] developed the Human Factors Analysis and Classification System (HFACS) and Embrey [6] developed the Systematic Human Error Reduction and Prediction Approach (SHERPA), in order to detect human errors in incidents and accidents. These approaches have been found to be able to effectively capture human errors from accident reports.

We concern about human errors in human macrocognitive functions. Macrocognitive functions (or macrocognitions), coined by Cacciabue and Hollnagel [7], are defined as the internalized and externalized high-level mental activities in a naturalistic environment [8], which are important for maintaining the safety of complex socio-technical systems. Recently, NUREG-2114 [9] proposed a cognitive framework to connect macrocognitive function failures with their proximate causes, failure mechanisms, and performance influencing factors (PIFs). This work is supposed to develop the cognitive foundation for the new NRC-sponsored HRA method (the Integrated Decision-Tree Human Event Analysis System, IDHEAS [10]). It will serve as a guidance for the HRA qualitative analysis and can be used to develop quantification methods for the HRA quantitative analysis in IDHEAS. Also it can be used as a human error taxonomy to investigate the error types and their causes.

Few studies have been done to examine the failures of macrocognitive functions in complex systems. This study will adopt the IDHEAS cognitive framework in NUREG-2114 [9] to detect the failures of macrocognitive functions in incidents and accidents. It will also examine the usability of the IDHEAS cognitive framework in NUREG-2114 in the HRA qualitative analysis. The next is organized as follows. Section 2 will describe the methodological issues including the incidents and accidents, the IDHEAS cognitive framework in NUREG-2114 [9], and the analysis procedure. Section 3 will present the results. Section 4 will discuss the results. Section 5 will conclude this study.

2 Methodology

2.1 Materials

A total of 103 short investigation reports of incidents and accidents from June 2005 to June 2007 in a petrochemical company in China were used as samples for this case study.

2.2 IDHEAS Cognitive Framework

The model of macrocognitive functions in IDHEAS [9] has five components: (1) detecting and noticing, (2) understanding and sensemaking, (3) decision making, (4) action implementation, and (5) team coordination. IDHEAS [9] suggests a generic cognitive framework of macrocognitive functions failures to describe its proximate causes (i.e., error type), failure mechanisms leading to proximate causes, and PIFs, as shown in Fig. 1. For example, for the failure of detecting and noticing, it has three proximate causes, i.e., cue/information not perceived, cue/information not attended to, and cue/information misperceived; for the proximate cause of cue/information not perceived, it has five failure mechanisms, i.e., cue content—cue salience is low and not detected, vigilance in monitoring—unable to maintain vigilance, attention—inattentional blindness, expectation—mismatch between

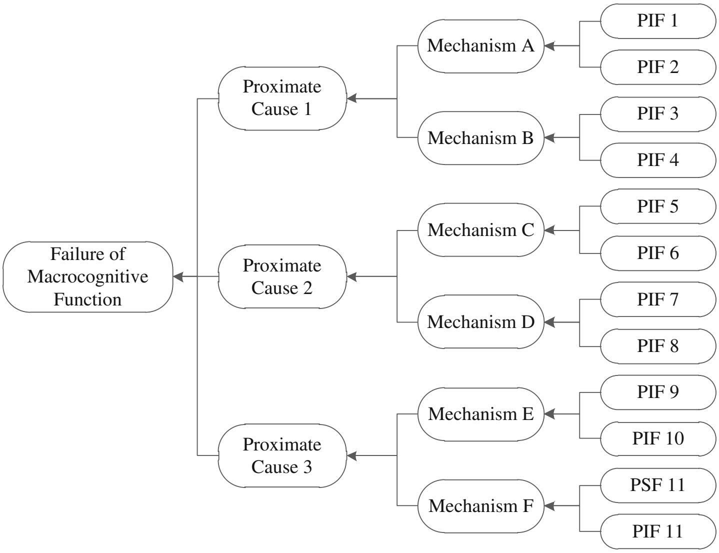

Fig. 1 IDHEAS cognitive framework of failures of macrocognitive functions [9]

expected and actual cue, and working memory capacity overload. For these failure mechanisms, there have dozens of PIFs to trigger them. This model of macrocognitive functions is expected to describe the logic relationship between failures of macrocognitive functions, proximate causes, failure mechanisms, and PIFs [9]. It provides a basis for HRA qualitative analysis in IDHEAS [9, 10].

2.3 Procedure

First, incidents and accidents were classified into two main types: equipment- and operator-related. Then, human errors in operator-related incidents and accidents were classified into three types: Type A (pre-initiators), Type B (initiators), and Type C (post-initiators). Finally, failures of macrocognitive functions in the operator-related incidents and accidents were identified according to the IDHEAS cognitive framework [9], including their proximate causes, failure mechanisms, and PIFs. For each failure mechanism, its major PSF was identified.

3 Results

3.1 General Results

Eighty-one incidents and accidents were equipment-related, 36 operator-related, and 14 related to both. About 35 % of incidents and accidents were implicated with operator errors.

A total of 74 failures of macrocognitive functions were identified, including 19 Type A errors (~26 %); 38 Type B errors (~51 %), and 17 Type C errors (~23 %). Figure 2 presents the distribution of the number of failures of macrocognitive functions. The number of failures of action implementation and team coordination was higher than that of other three macrocognitive functions. In addition, seven human errors, which were Type A errors, were hard to be classified into any types. The following will give the more detailed information for each macrocognitive function failure.

3.2 Failures of Detecting and Noticing

Among the reported 13 failures of detecting and noticing, 11 failures shared the common proximate cause "cue/information not perceived" (i.e., the cues or information may be missed, not seen, or not heard) and one was associated with the proximate cause "cue/information not attended to" (i.e., the cues or information

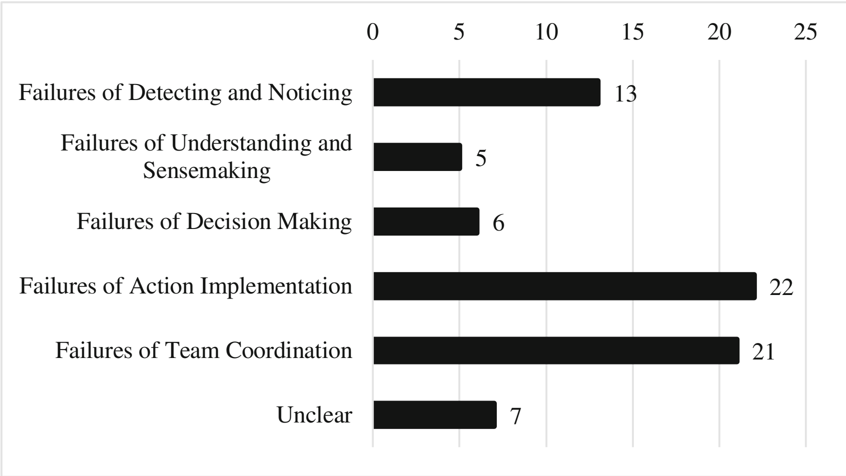

Fig. 2 Number of failures of macrocognitive functions

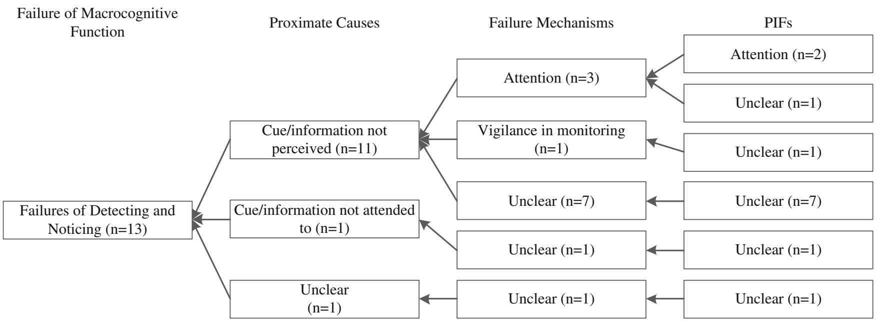

Fig. 3 Proximate causes, failure mechanisms, and PIFs in failures of detection and noticing

may be sensed and perceived but not attended to) (see Fig. 3). For one case, its proximate cause has not yet been defined.

For the causes “cue/information not perceived”, three of them were due to “attention—missing a change in cues” (i.e., change blindness) and one due to “vigilance in monitoring—unable to maintain vigilance”; however, for most of them, their failure mechanisms were not yet defined.

For the cause “cue/information not attended to”, its failure mechanism was not yet defined. For most of induced failure mechanisms, their PIFs were not yet identified.

3.3 Failures of Understanding and Sensemaking

Among the reported five failures of understanding and sensemaking, proximate causes of two failures were "incorrect integration of data, frames, or data with a frame" (e.g., operators do not properly integrate pieces of information together and improperly merge a frame of a system with the frame for the ongoing event) and the proximate causes of the other three were "incorrect frame used to understand the situation" (e.g., the frame or mental model to understand the situation is improper, incomplete, or insufficient) (see Fig. 4).

For the former two causes, one failure mechanism was "improper integration of information or frames" (e.g., operators may develop an incorrect mental model from the separate elements for the situation) and the other was "data not properly recognized, classified, or distinguished" (e.g., operators may misinterpret the situation based on existing knowledge). Both of them were induced by the lack of knowledge/experience/expertise.

For the latter three causes, they had the common failure mechanisms "incorrect or inadequate frame or mental model used to interpret or integrate information". Two of them were thought to be induced by the lack of knowledge/experience/expertise.

3.4 Failures of Decision Making

Among the six failures of decision making, five failures could be attributed to the proximate cause "incorrect mental simulation or evaluation of options" (e.g., operators may have incorrect prediction of a possible course of action or have unrealistically evaluations the options) and the other one, "incorrect goals or priorities set" (e.g., operators may have an inappropriate goal or prioritize goals improperly) (see Fig. 5).

For the former five causes, three failure mechanisms were "inaccurate portrayal of the system response to the proposed action" (e.g., operators may incorrectly predict how the system will respond to the proposed action), one was

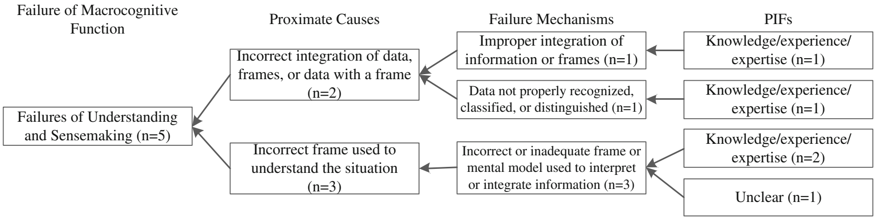

Fig. 4 Proximate causes, failure mechanisms, and PIFs in failures of understanding and sensemaking

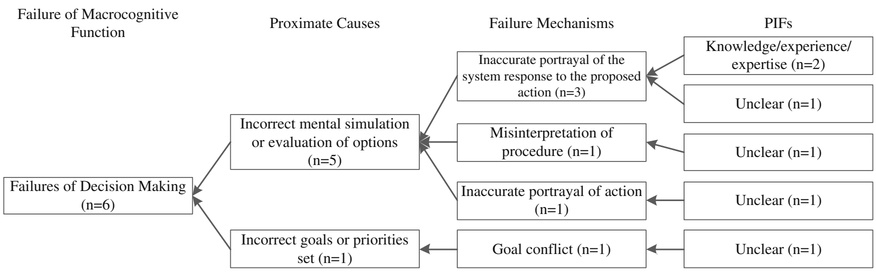

Fig. 5 Proximate causes, failure mechanisms, and PIFs in failures of decision making

"misinterpretation of procedure" (e.g., wrong procedures are used or the procedure makes operators difficult to use), and the left one was "inaccurate portrayal of action" (e.g., operators may incorrectly characterize the action or predict how the action will be performed). Two failure mechanisms were thought to be triggered by the lack of knowledge/experience/expertise. The PIFs for the other three failure mechanisms were not yet defined.

For the latter cause, its failure mechanism was "goal conflict". Its PIFs for this failure mechanism was unclear.

3.5 *Failures of Action Implementation*

Among the 22 failures of action implementation, the common proximate cause of four failures was "failure to take desired action (error of omission)" and that of the rest 18 failures, "execute desired action incorrectly (error of commission)" (see Fig. 6).

For the former four causes, the failure mechanisms were not yet defined.

For the latter 18 causes, nine were associated with the failure mechanism "continuous control deficiencies" and induced by PIF of system dynamics. The failure mechanisms and PIFs for the left nine causes were unclear.

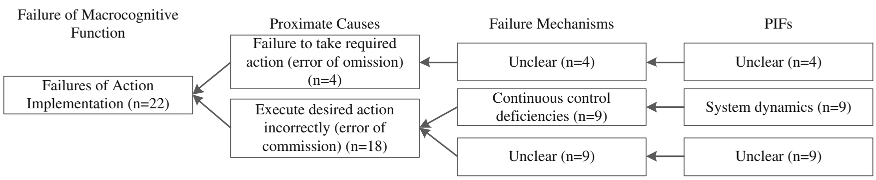

Fig. 6 Proximate causes, failure mechanisms, and PIFs in failures of action implementation

3.6 Failure of Team Coordination

Among the 21 failures of team coordination, 13 had the common proximate cause "failure of team communication" (e.g., crew members do not properly share or distribute information) and eight had "error in leadership/supervision" (e.g., the leader does not facilitate group discussion or failed to correct operator errors) (see Fig. 7).

For the former 13 causes, five causes could be attributed to the failure mechanism "source error of omission" (e.g., the source does not communicate information to the target), three "source error of commission" (e.g., the source transmits the wrong information to the target), two "target error of omission" (e.g., the target does not detect or notice the communicated information), and three "incorrect timing of communication" (e.g., the communication does not occur at the right time). Five PIFs were involved, i.e., risk perception, team cohesion, time pressure, role awareness.

For the latter eight causes, seven had the common failure mechanism "failure to verify that other operators have correctly performed their responsibilities" (e.g., the leader fails to the oversight duties) and one had "decision making failures" (e.g., bad decisions are made). Their corresponding PIFs were unclear.

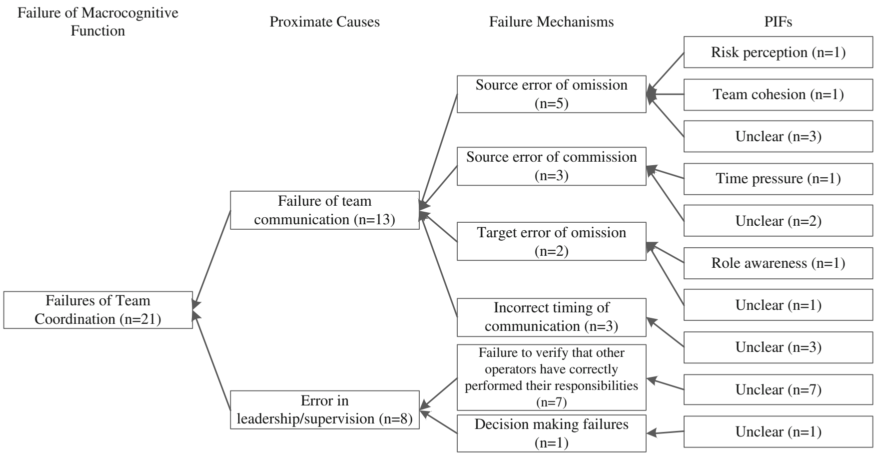

Fig. 7 Proximate causes, failure mechanisms, and PIFs in failures of team coordination

4 Discussion

4.1 Human Contribution to Incidents and Accidents

This study found that 35 % of incidents and accidents in this case of a petrochemical plant were related to human errors. This finding is consistent with the popular idea that 30–90 % of incidents and accidents connected to human errors [3, 11]. Nivolianitou et al. [12] reported that 40 % of the major accidents had immediate causes either exclusively (19 %) or partially (21 %) attributed to human factors in the petrochemical industry. Kidam and Hurme [13] observed that 19 % of contributors to chemical process accidents were human and organizational factors in operation. Baranzini and Christou [14] mentioned that the suspected cause of 43 % of major events in process industry in Europe was human and that plant or equipment accounted for the most major events. Our study and these aforementioned studies may imply that human may be not the most dominant contributor to accidents in the petrochemical plants. However, Zhang and Zheng [15] investigated 1632 hazardous chemical accidents (HCAs) occurring in China from 2006 to 2010 and reported that human factors accounted for the majority of HCAs. Zhang and Zheng's study obtained that 65.1 % of HCAs were due to human factors. It is a open question whether human or technology contribute more system failure events. It depends on how we treat human in the system failure events. Anyway, human factors have significant contributions to incidents and accidents in complex systems.

4.2 Failures of Macrocognitions

The dominant human errors were failures of action implementation and team coordination. Ghasemi et al. [16] also reported a similar finding that action errors had the highest frequency of occurrence in petrochemical plants. The significant contribution of action implementation and team coordination might be due to that operators have more macrocognitive activities of action implementation and team coordination or that operators have higher failure probability for performing those two macrocognitive activities. Currently, we are not sure whether the high frequency of occurrence or the high vulnerability of those two macrocognitive activities can explain their high impact on system incidents and accidents. A study by Liu and Li [17] on main control rooms in nuclear power plants found that complexity factors in action implementation and crew activity had the highest frequency of occurrence in nuclear power plants. This study may imply that the high frequency of occurrence of those two macrocognitive activities may explain their dominant role.

For the failure of team coordination, its main proximate cause was the failures of team communication, more than half of which were caused by the errors of source (i.e., who send the information). Ten of them were caused by information missing or inaccurate and three of them by the timing error. In time-critical situations (e.g., operating rooms), timing errors may be the dominant cause to communication failures (see [18]).

For the failure of action implementation, its main proximate cause was the error of commission (EOC) rather than the error of omission (EOO). Most of EOCs in this case were induced by continuous control deficiencies which were triggered by system dynamics. One previous study [19] in nuclear power plants observed that most of observed human factors events in control rooms were EOCs. Current probabilistic risk analysis (PRA) considers the effect of EOO on system performance, and however, ignores the effect of EOC in the HRA quantification process. Our present study points out that the number of incidents and accidents related to EOC was higher than that related to EOO, implying the potential higher negative effect of EOC on system performance. Thus, in PRA/HRA, EOC should be given more or at least equal attention. Fortunately, regulatory bodies have a growing interest to model the risk of errors of commission in PRA in nuclear power plants.

4.3 Usability of IDHEAS Cognitive Framework

IDHEAS (NUREG-2114) was expected to develop a solid theoretical foundation for HRA. Its first step was to build a model of macrocognitive functions to provide a technical basis combing the state-of-art knowledge of human performance from cognitive science, behavioral science, operational experience, etc. As it stated, it bridged HRA and psychology, which is expected to be a great step for the development of science-based HRA methods. It is supposed to clearly differentiate failures of macrocognitive functions, proximate causes, failure mechanisms, and PIFs, and most importantly, clearly describes the hypothetical cause-and-effect relationships between the four types of elements. It provides a comprehensive list of failure mechanisms.

The IDHEAS cognitive framework of failures of macrocognitive functions can be used to classify and analyze human errors in incidents and accidents in complex situations in the qualitative HRA. Regarding its usability, we perceived some difficulties to apply it in accident investigations. First, it might be not user-friendly to human reliability analysts and human factors analysts in practice. It adopts many terms (e.g., frame) from psychology and cognitive science, which might be not easily understood by those analysts and risk analysis engineers. Second, although it was relatively easy to classify human errors into different types of macrocognitive failures, it was a challenging task to analyze their proximate causes, failure mechanisms, and PIFs, especially for the latter two elements, which of course might also be due to the level of detail in the accident reports. Third, it was hard to distinguish several proximate causes/failure mechanisms. Regarding failures of understanding and sensemaking, for example, its three causes (i.e., incorrect data, incorrect integration of data, frames, or data with a frame, and incorrect frame) and

18 failure mechanisms were not easily to be distinguished. Several proximate causes and failure mechanisms may have some confusion. For example, the proximate cause "incorrect integration of data, frames, or data with a frame" is similar to the failure mechanism "improper integration of information or frames" in the terminology. In addition, sometimes it was hard to distinguish failure mechanisms and PIFs. For example, the concept "attention" was regarded as a failure mechanism and a PIF at the same time. Fourth, it was not easy to track back to the failure mechanisms and PIFs in a specific incident or accident, which might also be attributed to the reporting form in which context information related to human performance was not given too much attention. It seems that the current cognitive framework in IDHEAS [9] needs to be improved to inform HRA. We suggest to simplify the current IDHEAS cognitive framework [9] without sacrificing its theoretical soundness.

5 Conclusions

The current study investigated the failures of macrocognitive functions in a petrochemical plant, including their proximate causes, failure mechanisms, and possible performance influencing factors. It was found that human failures contributed to 35 % of incidents and accidents. Failures of action implementation and team coordination were the two main failures, highlighting the importance of correct actions and teamwork to system reliability. Regarding the failure of action implementation, it was found that commission errors had a higher contribution than omission errors did. This study also discussed the usability issues of IDHEAS (NUREG-2114) [9] as a human error taxonomy system in the qualitative human reliability analysis.

Acknowledgments This work was supported by the National Natural Science Foundation of China under Grant 71371104.

References

1. IAEA.: Managing Human Performance to Improve Nuclear Facility Operation. International Atomic Energy Agency, Vienna, Austria (2013)
2. Reason, J.: Human Error. Cambridge University Press, New York (1990)
3. Griffith, C.D., Mahadevan, S.: Inclusion of fatigue effects in human reliability analysis. Relib. Eng. Syst. Safe. **96**, 1437–1447 (2011)
4. Wang, Z., Li, Z.: Summary Reports of Operating Events in Chinese Nuclear Power Plants. China Environmental Science Press, Beijing, China (2012). (in Chinese)
5. Wiegmann, D.A., Shappell, S.: A Human Error Approach to Aviation Accident Analysis: The Human Factors Analysis and Classification System. Ashgate, Burlington, VT (2003)

6. Embrey, D.: SHERPA: a systematic human error reduction and prediction Approach. In: International Topical Meeting on Advances in Human Factors in Nuclear Power System, pp. 184–193. American Nuclear Society, Knoxville (1986)
7. Cacciabue, P.C., Hollnagel, E.: Simulation of cognition: applications. In: Hoc, J.M., Cacciabue, P.C., Hollnagel, E. (eds.) Expertise and Technology: Cognition and Human-Computer Cooperation, pp. 55–73. Lawrence Erlbaum Associates, Hillsdale, NJ (1995)
8. Klein, G.A., Ross, K.G., Moon, B.M., Klein, D.E., Hoffman, R.R., Hollnagel, E.: Macrocognition. IEEE Intell. Syst. **18**, 81–85 (2003)
9. Whaley, A.M., Xing, J., Boring, R.L., Hendrickson, S.M.L., Jeffrey, C.J., Le Blanc K.L.: Cognitive basis for human reliability analysis. NUREG-2114. U.S. Nuclear Regulatory Commission, Washington D.C. (2016)
10. Xing, J., Presley, M., Parry, G., Forester, J., Hendrickson, S., Dang, V.: An integrated decision-tree human event analysis system (IDHEAS) method for NPP internal at-power operation. Draft Report. U.S. Nuclear Regulatory Commission, Washington, D.C. (in press)
11. Liu, P., Li, Z.: Human error data collection and comparison with predictions by SPAR-H. Risk Anal. **34**, 1706–1719 (2014)
12. Nivolianitou, Z., Konstandinidou, M., Michalis, C.: Statistical analysis of major accidents in petrochemical industry notified to the major accident reporting system (MARS). J. Hazard. Mater. **137**, 1–7 (2006)
13. Kidam, K., Hurme, M.: Statistical analysis of contributors to chemical process accidents. Chem. Eng. Technol. **36**, 167–176 (2013)
14. Baranzini, D., Christou, M.D.: Human factors data traceability and analysis in the European community's major accident reporting system. Cogn. Tech. Work **12**, 1–12 (2010)
15. Zhang, H.-D., Zheng, X.-P.: Characteristics of hazardous chemical accidents in china: a statistical investigation. J. Loss Prevent. Proc. **25**, 686–693 (2012)
16. Ghasemi, M., Nasleseraji, J., Hoseinabadi, S., Zare, M.: Application of SHERPA to identify and prevent human errors in control units of petrochemical industry. Int. J. Occup. Saf. Ergo. **19**, 203–209 (2013)
17. Liu, P., Li, Z.: Comparison between conventional and digital nuclear power plant main control rooms: a task complexity perspective, part i: overall results and analysis. Int. J. Ind. Ergonom. **51**, 2–9 (2016)
18. Lingard, L., Espin, S., Whyte, S., Regehr, G., Baker, G.R., Reznick, R., Bohnen, J., Orser, B., Doran, D., Grober, E.: Communication failures in the operating room: an observational classification of recurrent types and effect. Qual. Saf. Health. Care **13**, 330–334 (2004)
19. Dai, L., Zhang, L., Li, P., Hu, H., Zou, Y.: Study on human errors in DCS of a nuclear power plant. In: Proceedings of the 12th Probabilistic Safety Assessment and Analysis (PSAM12). Honolulu, Hawaii (2014)

A Literature Review on Human Reliability Analysis Techniques Applied for Probabilistic Risk Assessment in the Nuclear Industry

Ninochka Dsouza and Lixuan Lu

Abstract A literature review of a number of Human Reliability Analysis (HRA) methodologies is carried out in this paper. The focus of the paper is on the use of an HRA method to quantify the probability of a human error that can then be plugged into the overall plant Probabilistic Risk Assessment (PRA), specifically in the nuclear industry. In keeping with this criterion, the modeling techniques selected for review are Technique for Human Error Rate Prediction (THERP), Accident Sequence Evaluation Program (ASEP), Cause-Based Decision Tree Method (CBDTM), Human Cognitive Reliability/Operator Reliability Experiments (HCR/ORE), Simplified Plant Analysis Risk-Human (SPAR-H) reliability assessment, Justified Human Error Data Information (JHEDI), Cognitive Reliability and Error Analysis Method (CREAM), A Technique for Human Error Analysis (ATHEANA), Nuclear Action Reliability Assessment (NARA) and Human Error Assessment and Reduction Technique (HEART). It is concluded that while no one methodology can cover all aspects of HRA due to each having their own set of limitations, a combination of certain methodologies can provide a more accurate Human Error Probability (HEP) for input into the PRA.

Keywords Human reliability analysis · Nuclear engineering · Probabilistic risk assessment · Technique for human error rate prediction · Accident sequence evaluation program · Cause-Based decision tree method · Human cognitive reliability/operator reliability experiments · Simplified plant analysis risk-human (SPAR-H) reliability assessment · Justified human error data information · Cognitive reliability and error analysis method · A technique for human error analysis · Nuclear action reliability assessment

N. Dsouza · L. Lu (✉)
Faculty of Energy Systems and Nuclear Science, University of Ontario Institute of Technology, Oshawa, ON, Canada
e-mail: Lixuan.Lu@uoit.ca

N. Dsouza
e-mail: Ninochka.Dsouza@uoit.ca

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_4

1 Introduction

Risk is inherently present in all facets of human endeavor. There are two main aspects to risk: the likelihood for things to change and the resulting magnitude of the consequences when said changes occur. The risk identification process helps to avoid the unintentional retention of risk that can take place when a source of performance variability remains unknown. Once identified, the risk can be characterized using deterministic and probabilistic risk assessment methods [1]. Probabilistic Risk Assessments (PRA) are used to express uncertainty of possible future negative consequences of complex systems such as a chemical or nuclear plant, in terms of probabilities. This methodology is used most extensively in Nuclear Power Plants (NPP) to numerically quantify risk measures [2].

PRA in NPP primarily emerged in the years following World War II, as a technique was sought that could estimate the likelihood of an accident, using mathematical calculations of the probability that systems and subsystems would fail [3]. The Rasmussen Report (WASH-1400) produced in 1975 for the Nuclear Regulatory Commission (NRC) used linked fault tree technique to evaluate the probability of a number of accident sequences that might lead to a core melt accident in the reactor. This study along with other German investigations done at the time provided a more realistic assessment of the risks associated with the operation of commercial NPPs. These studies also showed that human errors could be a major contributor to reactor accidents [2]. While actual values are difficult to attain, human errors are estimated to be responsible for 60–90 % of industrial accidents, while the remainder of accidents are attributed to technical deficiencies including equipment failures due to environmental degradation. There is therefore a need to assess human reliability with the aim of reducing the likely cause of errors.

There are many factors that impact human performance in a NPP. These Performance Shaping Factors (PSFs) illustrate the many facets of human error and are used to provide a numerical basis for modifying nominal human error probability (HEP) levels, identify contributors to human performance as well as enhance performance [4, 8]. As PSFs have the potential to positively or negatively affect an individual's performance, identifying and quantifying the effects of a PSF is one of the key steps in the HRA process [4]. PSFs can be internal or external to a person. External PSFs involve the entire work environment especially equipment design, written procedures or oral instruction whereas internal PSFs involve the individual characteristics of the person such as skills, motivation, and expectations that influence the persons performance. Another classification of PSFs involves psychological and physiological stresses resulting from a work environment where the demands placed on an individual result in poor performance.

HRA has always been closely associated to risk analysis and was developed to predict potential human errors as an enhancement to risk analysis of systems and equipment. While human factors are part of many industry applications (e.g. chemical and refining plants, aviation industry, etc.) the main focus of HRA in this paper will be in safety-critical applications for which human error has the potential for severe consequences, i.e. most specifically in the nuclear industry [5]. These unwanted consequences include, but are not limited to, degraded safety, production, maintenance, operation, or performance of a plants system [6]. Incidents at Three Mile Island and Brown's Ferry have also clearly shown humans to act as accident initiators, propagators and mitigators in nuclear power plants.

Human errors are typically divided into 4 categories [7]: Type A: Pre-initiator Errors (Latent); Type B: Errors Causing Initiating Events; Type C: Post-initiator Errors (Dynamic) and Type D: Cognitive Errors of Commission (this type of error is typically not in the scope of internal event PRAs). Type A, B and C as mentioned above are typically evaluated based on the following HRA process steps [7]:

1. Identification and Categorization of HFE
2. Screening (done for Type A only)
3. Definition (for Qualitative data)
4. Quantification (starting with a qualitative analysis and then selecting a HRA method to quantify the Human Error Probability (HEP) for an HFE)
5. Documentation
6. HRA to PRA Model Incorporation.

This paper is focused primarily on step 4 of the above stated process, i.e. the use of an HRA method to quantify the probability of a human error that can then be plugged into the overall plant PRA.

There is no specific number or types of PSFs that must be used in an HRA method though the range used, lies between the extremes of 1 to 60 PSFs. Table 1 [8] shows a comparison of the number and types of PSFs used in a few HRA methodologies where it can be see that there is considerable overlap in the PSFs. The variety in PSFs however, depicts the vastness of factors that can influence human performance. Boring, in [8], concludes that the number of PSFs is not as crucial as the use for which the PSFs are engaged. Depending on the sensitivity of the HRA application, the number of PSFs used should be adjusted accordingly [8].

As mentioned in [10], PSF measures can be categorized as direct or indirect. An example of a direct measure would be time taken to finish a task whereas an example of indirect measure would be fitness for duty as it would require measure of a different PSF such as fatigue.

2 Three HRA Generations

PSFs are used in almost all methods and phases of a HRA. While most HRA methods can generically be said to have three common phases i.e. identification, modeling and quantifying of human error, they are more broadly categorized into qualitative methods, quantitative methods and into methods that encompass the complete spectrum of a HRA [8, 11].

Table 1 Comparison of PSFs among the good practices, SPAR-H, CREAM and the 9-Factor model [8]

Good practices	SPAR-H	CREAM	9-Factor model
Training and experience	Experience/training	Adequacy of training and preparation	Training knowledge
Procedures and administrative controls	Procedures	Availability of procedures/plans	Resources
Instrumentation	Ergonomics/HMI	Adequacy of HMI and operational support	Machine
Time available	Available time	Available time	Loads/perceptions
Complexity	Complexity	Number of simultaneous goals	Complexity
Workload/time/pressure/Stress	Stress/stressors	Number of simultaneous goals	Loads/perceptions
Team/crew dynamics	Work processes	Crew collaboration quality	Team
Available staffing	Work processes	Adequacy of organization	Resources
Human-system interface	Ergonomics/HMI	Adequacy of HMI and operational support	Machine
Environment	Stress/Stressors	Working conditions	Complexity
Accessibility/operability of equipment	Ergonomics/HMI	Adequacy of HMI and operational support	Machine
Need for special tools	Ergonomics/HMI	Adequacy of HMI and operational support	Resources
Communications	Work processes	Crew collaboration quality	Team
Special [Equipment] fitness needs	Ergonomics/HMI	Adequacy of HMI and operational support	Resources

(continued)

Table 1 (continued)

Good practices	SPAR-H	CREAM	9-Factor model
Consideration of "realistic" accident sequences diversions and deviations	–	–	–
–	Fitness for duty	Time of day	–
–	Work processes	Adequacy of organization	Organizational culture
–	–	–	Attitude

2.1 First Generation HRA

Quantitative HRA includes the estimation of time-dependent and time-independent HEPs for errors impacting proper system function. Human performance data, models and analytical methods are used in conjunction with one another to calculate these estimates [6]. In the First Generation HRA methods, HEPs were considered to be a major factor whereas PSFs were considered to be a minor factor in the estimation of the probability of human failure. The generation concentrated on quantification, paying greater attention to the success/failure of an action rather than the depth of the causes/reasons of human behavior. The approaches taken encourage the analyst to divide a task into sub components and then consider the potential impact of modifying factors like time, pressure, equipment design and stress. Combining these elements, a nominal HEP can be determined. Cognitive modeling in these methods followed the Rasmussen skill-rule-knowledge (SKR) model which is without adequate human and psychological realism. Due to the ease of use, these quantitative methods are regularly used in the industry although they are often criticized for failure to consider the impacts of context, environment, organizational factors, and errors of commission. This results in analysts being more conservative and assigning higher estimates of HEPs [4, 12].

2.2 Second Generation HRA

The qualitative aspect of HRA uses task analysis to identify situations with potential for human error and requires an understanding of the cognitive situation people are in, rather than the physical arrangement of things [5, 6]. The task analysis is an iterative and interactive process whereby specific human behaviors are determined and analyzed by observation, interviews, talk-through, error records, etc. Incorrect inputs that can occur during human-system interaction are then related to one or more PSFs. The selected PSFs in turn provide the basis for the quantitative aspect of HRA [6]. In the past, HRA analysts neglected the development and inclusion of the qualitative phase and things like cognitive science in their analysis however it has slowly been integrated back into HRA [5]. While various methods have distinct approaches to the qualitative and quantitative parts of a HRA, some methods are primarily qualitative e.g. root-cause analysis, while others only provide a simplistic means to quantify the HFE. In the simplistic quantitative approach if an analyst were to assume two different qualitative methods, there would be considerable variability in the end results of the analysis. Hence both qualitative and quantitative elements should be included for a more complete understanding of the strength and weaknesses of an HRA methodology [13].

2.3 Third Generation HRA

Based on the earlier limitations and shortcomings of the first and second generation methods, new methods are now being developed and are being identified as the Third Generation HRA methods. An example of this is NARA, Nuclear action reliability assessment, which is an advanced version of a quantitative method known as HEART that is used in the first generations HRA methods in the nuclear industry [4].

3 HRA Methods

In this paper, the primary focus of the literature review will be on HRA methodologies that are currently being used to support NPP probabilistic risk assessments. In keeping with this criterion, the modeling techniques selected for review are as follows.

3.1 Technique for Human Error Rate Prediction (THERP)

As an HRA founding method and as well as a first generation method, THERP, which was authored by Swain and Guttman, was initially developed and used in 1961 and has been in constant development and use ever since. In late 1972, WASH-1400-Reactor Safety Study [14] employed this method to assess the impact of human errors in a NPP. This led to the development of a HRA handbook and workbook which along with the THERP internal data bank and its other sources of data in still in use at many NPPs, military systems, oil production and refining plants and chemical plants to date [6]. This handbook presents methods, models and estimated HEPs that enables quantitative or qualitative assessments to be carried out by the analyst. The HEP tables presented in the handbook can be modified by the effects of the PSFs chosen using other tables. This method deals with all aspects of human reliability assessments from: task analysis, human failure events and PSFs in error identification and representation, to nominal, basic and conditional HEP in quantification of the human error probabilities followed by its integration into plant

PRA [12, 15, 16]. Kirwan et al. in [17] describes the six major process steps of this method as decomposition of tasks into elements, assignment of nominal HEPs to each element, determination of effects of PSF on each element, calculation of the effects of dependence between tasks, modeling in an HRA tree and finally quantification of total task HEP [12].

Although THERP is largely used by the industry, it still has many short comings and limitations. When compared to the set of good practices to be fulfilled by HRA methods as set out in [18], a number of technical deficiencies are noted in the THERP methodology [19, 20] Some of the limitations include it being resource intensive, time consuming, excessive level of detail, insufficient guidance on modeling scenarios and impact of a limited number of PSFs on performance, non-provision of a cognitive model for human beings whereby the THERP quantification tables of human errors are based on a taxonomy that does not take into account the human error mechanism [12, 21]. Alvarenga and Fonseca in [19, 21] provide possible solutions to eliminate these deficiencies with the use of second generation methodologies such a CREAM, ATHEANA, etc.

3.2 Accident Sequence Evaluation Program (ASEP)

Due to the resource intensive nature of THERP, the Nuclear Regulatory Commission (NRC) expressed a need for an HRA method that would provide estimates of HEPs and response time for tasks performed during normal operating conditions and post-accident operating conditions that would be both sufficiently accurate for PRA and require minimal expenditure of time and other resources. With these requirements in mind, ASEP was developed based heavily on the THERP/Handbook method for HRA while incorporating many simplifications of the human performance models and HRA methodology in NUREG/CR-1278 [9, 22]. It was also developed solely for the nuclear industry.

ASEP HRA procedure is divided into procedures for pre-accidents tasks, post-accident tasks, screening HRAs and nominal HRAs. Pre-accident tasks are those tasks which if performed incorrectly could result in the unavailability of a safety system or component thereby preventing an appropriate response to an accident. Post-accident tasks are those tasks that are intended to assist in returning the plants systems to a safe condition during an abnormal condition. A nominal HRA is the HRA applied to tasks which survive the screening analysis that is part of the systems analysis. It is more judgment based and hence more conservative. Screening HRAs involves the screening of probabilities and response times assigned to each task as an initial sensitivity analysis thereby reducing the amount of detailed analysis to be performed. The four procedures that the ASEP HRA procedure consists of are therefore Pre-Accident Screening HRA, Post-Accident Screening HRA, Pre-accident Nominal HRA and Post-Accident Nominal HRA. It is based on fixed combinations of recovery and dependency factors and on adjustments to a constant basic human error probability [7, 12, 22].

3.3 Human Cognitive Reliability (HCR)/Operator Reliability Experiments (ORE) Method

In the 1980s, EPRI amongst other organizations, in support of a NRC program on non-training simulator use, carried out a series of data collection projects with respect to the role of operators performing safety functions. The results led to the use of time reliability curves (TRCs) to the development of the HCR method [23]. The HCR method was required to quantify the numerical relationship between non-response probability and response time based on human-machine interfaces, human cognition and mean response time [24–26]. This quantification technique was developed by the Electric Power Research Institute (EPRI) for estimating non-response probability of post-initiator actions only. Although part of the first generation HRA methodology set, it uses cognitive psychology to study dynamic cognitive processes and explore the mechanisms of human error [27]. In NPPs, the HCR correlation is an analytical method that enables the quantification of reliability of control room personnel with respect to their response to an abnormal plant condition. It is a time-reliability correlation which takes into account a number of factors such as expected operator stress level, type of human-machine interface, time for task completion, etc. This model is used to estimate the cognitive reliability for skill, rule and knowledge based behaviors [27]. It is a fairly quick technique and has great ease of use [26].

The HCR/ORE technique is found to not explicitly address potential causes of human errors in diagnosis. It also requires a relatively significant number of simulator exercises to produce reasonable results. The tables used in this methodology are also not available for public scrutiny [28].

3.4 Cause-Based Decision Trees Method (CBDTM)

This first generation, task related, HRA quantification method is a derivative of THERP. It came about due to a need to incorporate the data collected during the ORE set of Operator experiments into something that could be used in HRA. With this objective, a decision tree approach to try to incorporate ORE and THERP was created, known as the Cause-Based Decision Tree Method (CBDT) [23]. Originally developed by EPRI, it addresses very low probability values produced by the HCR/ORE method (<1E-02) and actions with longer time frames where “extrapolation of HCR/ORE TRC can be viewed as extremely optimistic” [28]. Similar to HCR/ORE, it is a quantification technique used for estimation of non-response probabilities of post-initiator human actions only. CBDT uses expert judgment and data extrapolated from the THERP method. It considers 8 potential error mechanisms and factors that could contribute to those failures through the use of the decision trees. While this methodology can now be used as a stand-alone method, it has a number of limitations such as it provides no guidance for use under time-limited conditions and one would need to be quite experienced in the application of HRA concepts and have some experience of how control room crews operate to use this technique [23, 28].

3.5 Simplified Plant Analysis Risk Human Reliability Assessment (SPAR-H)

The U.S. NRC in the early 1990s identified a need for an improved, traceable, easy to use HRA method for use with analytical models associated with the NRCs Accident Sequence Precursor (ASP) program [29]. As a quantification method it does not provide any of the qualitative aspects such as identification or modeling of HFEs that is needed to support quantification [30]. The basic framework of this technique: decomposes probability into contributions from diagnosis failures and action failures; accounts for the context associated with HFEs by using PSFs and dependency assignment to adjust a base-case HEP; uses pre-defined base-case HEPs and PSFs, together with guidance on how to assign the appropriate value of the PSF; employs a beta distribution for uncertainty analysis; uses designated worksheets to ensure analyst consistency [29].

A major component of this method is the simplification of the estimation procedure using the SPAR-H worksheet. Amongst others, one of the positive aspects of this worksheet is that it includes an adjustment factor to avoid probability estimates greater than one. The nominal HEPs in this method are assigned to both action and diagnosis failures that can be adjusted to reflect the impact of all eight PSFs. More on the steps involved in this methodology is available in [30].

3.6 Human Error Assessment and Reduction Technique (HEART)

This technique was first outlined in 1985 in a conference paper by Williams [12]. The main elements of the HEART process are: Classification of tasks into one of the Generic Categories, Assigning a nominal HEP to the task, Determining which Error Producing Conditions (EPCs or PSFs) may affect task reliability, Determining the Assessed Proportion of Affect (APOA) for each EPC and calculating the task HEP [15]. The premises this method is based on include: basic human reliability is dependent upon the generic nature of the task to be performed; in 'perfect condition', the level of reliability will be achieved consistently with a given nominal likelihood within probabilistic limits and; if said 'perfect' conditions were non-existent, the human reliability may degrade as a function of the extent to which identified PSFs may apply. Further information on the steps involved in this process can be taken from [31].

3.7 Justified Human Error Data Information (JHEDI)

This technique was developed to provide a faster screening technique than its 'parent' Human Reliability Management System (HRMS) approach. The HRMS approach is a fully computerized HRA system that contains a human error identification module that can be used by the assessor on a previously prepared and computerized task analysis. This process was created to inform the design process of the British Nuclear Fuels Ltd, Thermal Oxide Reprocessing Plant [12, 31]. JHEDI starts from a set of basic error descriptors and empirically derived error probabilities, and uses a set of PSF questions that have been answered by the assessor to determine the HEP value [15]. Both JHEDI and HRMS can be used for task and error analysis and PSF based quantification but JHEDI involves a lot less detailed assessment than HRMS. The JHEDI quantification system has been simplified and its HEP values have been made more conservative to allow for simplicity. It also requires less PSF questions to be answered in comparison to HRMS. In comparison to THERP and HEART to which JHEDI is conceptually similar to, the PSF rating process in this method is more straightforward due to the factual rather than subjective questions asked that can be substantiated. The multipliers used in this technique are fewer than those used in the HRMS method and some of the extrapolation rules have been omitted from this method as well, thereby making the JHEDI system relatively rapid. The system uses actual industry data and requires very little training for use by the analyst [12, 15, 31].

3.8 Cognitive Reliability and Error Analysis Method (CREAM)

A second generation HRA technique, CREAM was introduced by Hollnagel in 1998 [32]. The method is designed to take better account of context than the earlier first generation techniques. It is different from other second generation methods in that it sets aside the errors of commission/errors of omission categorization [33].

CREAM has both a basic and extended version thereby allowing for a preliminary analysis to take place using the basic version, prior to deciding to continue on with a detailed analysis. The basic version can therefore be used as a screening tool [33]. As described in [12, 32], CREAM is a fully bidirectional process where its principles can be applied in both retrospective and performance analysis. As described in [33], the process of CREAM involves a task analysis from which a list of operator actions is produced. The likelihood of failure in these tasks is impacted by Common Performance Conditions (CPCs). There are 9 CPCs considered and depending on the choice made, the CPC will improve, reduce or not significantly change the task failure probability. For each activity, the nine CPC scores are assessed to give a combined CPC score [33]. A result of [9, 0, 0] would indicate the least desirable situation while a result of [0, 2, 7] would be a more desirable situation. These steps consist of the Basic CREAM.

3.9 A Technique for Human Event Analysis (ATHEANA)

ATHEANA is a second generation HRA technique that was developed for the U.S. NRC in 1996. ATHEANA is a method capable of obtaining both qualitative and quantitative HRA results for pre and post initiators. It was created on the premise that significant human errors occur as a result of “error forcing contexts” (EFCs). EFCs are defined as combinations of PSFs, plant conditions and other influences that make an operator error more likely. These EFCs are identified using four related search schemes. There are 10 steps in this process but the main ones are: integration of the issues of concern into the ATHEANA HRA/PRA perspective; identification of human failure events and unsafe actions that are relevant to the issue of concern; identification of the reasons why such events occur of the same; quantification of EFCs and the probability of each unsafe action given its context and; evaluation of the results of the analysis in terms of the issue for which the analysis was performed [12, 34–36].

The technique uses a quantification model for HFE probabilities based on estimates of EFC frequency. It is seen that there are three basic elements in the quantification process i.e. the probability of the EFC, the probability of unsafe human action (UA) and, the probability of not recovering from the initial UA. Following this an expert elicitation approach for performing ATHEANA quantification was developed. However no empirical validation of this technique has been carried out [12, 34–36].

3.10 Nuclear Action Reliability Assessment (NARA)

NARA is a third generation HRA technique developed for British Energy in 2005. Therefore this methodology is more tailored to the UK NPPs PRAs and HRAs. NARA was created with more recent data but using the HEART methodology as its basis i.e. the same formula is used for deriving the HEP. Similar to HEART, NARA consists of 14 GTTs and 18 EPCs. The process includes: classification of the task for analysis into a GTT and assigning it a nominal HEP; deciding which EPCs may affect task reliability and; consideration of the APOA for each EPC. Once completed the task HEP is calculated. The new features of NARA are: an approach to quantifying operator reliability in relation to long time-scale events; a prototype approach to EOC quantification; more guidance for use of the APOA process and; dependence approaches for NARA applications which are currently in the process of being determined [12, 37]. While a positive aspect of NARA is that it is based on HEART which is a validated and established method, as seen in [38], NARA was found to not provide an explicit process for error identification and screening. Additionally, NARA provides uncertainty information only for a limited range of GTTs and EPCs. This lack of consistent uncertainty information may limit the seamless integration of NARA into some PRA models.

4 Conclusion

In this study, only a handful of HRA methods were touched upon. Many more HRA methods do exist that have not been included in this paper. In the current nuclear industry, one of the most used HRA software tools is the EPRI HRA calculator that automates HCR/ORE, CBDT, THERP, SPAR-H and the ASEP method to diagnose and quantify pre and post initiator HFEs. This not only encourages consistency in the HRA results but also provides for a consistent framework to analyze human actions in PRAs. Hence it can be concluded that while no one methodology can cover all aspects of HRA due to each having their own set of limitations, a combination of methodologies can provide a more accurate HEP for input into the PRA.

References

1. IAEA: Risk management: A tool for improving nuclear power plant performance (2001)
2. Nusbaumer, O.: Introduction to probabilistic safety assessments (PSA). Leibstadt NPP. http://nusbaumer.tripod.com/resources/publications/nusbaumer_introduction_to_probabilistic_safety_assessments.pdf (2016). Accessed on April 26, 2016
3. Carlisle, R.P.: Probabilistic risk assessment in nuclear reactors: engineering success, public relations failure. Technol. Cult. **38**(4) (1997)
4. Pasquale, V.D., Iannone, R., Miranda, S., Riemma, S.: An overview of human reliability analysis techniques in manufacturing operations, operations management. InTech (2013). doi:10.5772/55065
5. Oxstrand, J.: Human Reliability Guidance- How to Increase the Synergies between Human Reliability, Human Factors, and System Design & Engineering. Phase 2: The American Point of View—Insights of how the US Nuclear Industry works with Human Reliability Analysis. NKS_R-2009_77 Interim Report. (2010)
6. Swain, A.D.: Human reliability analysis: need, status, trends and limitations. Reliab. Eng. Syst. Safety. **29**(3), 301–313 (1990)
7. Education of Risk Professionals: PRA 105-106 Data Analysis/HRA. Charlotte, North Carolina (2013)
8. Boring, R.L.: How many performance shaping factors are necessary for human reliability analysis? In: 10th International Probabilistic Safety Assessment and Management Conference (2010)
9. Swain, A.D., Guttmann, H.E.: Handbook of Human Reliability with Emphasis on Nuclear Power Plant Applications Final Report. NUREG/Cr-1278. US NRC (1983)
10. Boring, R.L., Griffith, C.D., Joe, J.C.: Measure of human error: direct and indirect performance shaping factors. In: 8th IEEE Conference on Human factors and Power Plants (2007)

11. Bell, J., Swain, A.: A procedure for conducting a human reliability analysis for nuclear power plants. Final Report. U.S. NRC. NUREG/CR-2254. (1983)
12. Bell, J., Holroyd, J.: Review of Human Reliability Assessment Methods. Health and Safety Executive Research Report RR679 (2009)
13. Boring, R.L., Forested, J. A., Bye, A., Dang, V.N., Lois, E.: Lessons learned on benchmarking from the international human reliability analysis empirical study. In: The 10th Probabilistic Safety Assessment and Management Conference (2010)
14. U.S. NRC: Reactor Safety Study- An assessment of Accident Risks in US Commercial Nuclear Power Plants. NUREG-75/014, WASH-1400 (1975)
15. Kirwan, B.: The validation of three human reliability quantification techniques, THERP, HEART and JHEDI: Part I- technique descriptions and validation issues. Appl. Ergonomics **27** (6), 359–373 (1996)
16. Boring, R.L.: Fifty Years of THERP and Human Reliability Analysis. Department of Energy (2012)
17. Kirwan, B., Kennedy, R., Taylor-Adams, S., Lambert, B.: The validation of three human reliability quantification techniques, THERP, HEART and JHEDI: Part II- results of validation exercise. Appl. Ergonomics **28**(1), 17–25 (1997)
18. Kolaczkowski, A., Forester, J., Lois, E., Cooper, S.: Good Practices for Implementing Human Reliability Analysis (HRA). NUREG-1792 (2005)
19. Alvarenga, M.A.B., Fonseca, R.A.: On the use of the THERP methodology in the human reliability analysis of nuclear power plants—compliance with the USNRC good practices criteria. In: Proceedings of the International Nuclear Atlantic Conference (2007)
20. Jeong, W.D., Kim, J.H., Jang, S.C., Ha, J.J.: Abstract of comprehensive analysis to identify the limitations of ASEP/THERP methods based on HRA experiences. In: Proceedings of the Korean Nuclear Society Autumn Meeting (1998)
21. Alvarenga, M.A.B., Fonseca, R.A.: Comparison of the THERP quantitative tables with the human reliability analysis techniques of second generation. In: Proceedings of the International Nuclear Atlantic Conference (2009)
22. Swain, A.D.: Accident Sequence Evaluation Program Human Reliability Analysis Procedure. NUREG/CR-4772. SAND86-1996 (1987)
23. Spurgin, A.J.: Human Reliability Assessment Theory and Practice. CRC Press (2010)
24. Hannaman, G.W., Spurgin, A.J., Lukic, Y.: Human Cognitive Reliability Model for PRA Analysis: Draft report (NUS-4531, EPRI Project RP2170-3). San Diego, CA (1985)
25. Humphreys, P.: Human Reliability Assessor's Guide: A Report by the Human Factors in Reliability Group. United Kingdom Atomic Energy Authority Safety and Reliability Directorate (1995)
26. Sun, R., Chen, Y., Liu, X., Peng, T., Liu, L.: A method of analysis integrating HCR and ETA modeling for determining risks associated with inadequate flight separation events. J. Aviation Technol. Eng. **1**(1), 19–27 (2011)
27. Varde, P.V., Lee, D.Y., Han, J.B.: Human error probability evaluation as part of reliability analysis of digital protection system of advanced pressurized water reactor- APR1400. Korea Atomic Energy Research Institute. INIS vol. 36, Issue 19. (2003)
28. Lois, E., Forester, J., Kolaczkowski, A.: Presentation for Public Meeting on Evaluation of Human Reliability Analysis (HRA) Methods Against HRA Good Practices (NUREG-1842-Draft for Public Comment). Sandia National Laboratories (2006)
29. Gertman, D.I. et al.: The SPAR-H Human Reliability Analysis Method. U.S. NRC. Idaho National Laboratory. NUREG/CR-6883 (2005)
30. Whaley, A.M., Kelly, D.L., Boring, R.L., Galyean, W.J.: SPAR-H Step-by-Step guidance. In: 11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference (2012)
31. Kirwan, B.: A Guide to Practical Human Reliability Assessment. Taylor & Francis, London (1994)
32. Hollnagel, E.: Cognitive Reliability and Error Analysis Method-CREAM. Elsevier Science, Oxford (1998)

33. Bedford, T., Bayley, C., Revie, M.: Screening, sensitivity, and uncertainty for the CREAM method of human reliability analysis. Reliab. Eng. Syst. Saf. **115**, 100–110 (2013)
34. Cooler, S.E., Ramey-Smith, A.M., Wreathall, J., Parry, G.W., Bley, D.C., Luckas, W.J.: A Technique for Human Error Analysis (ATHEANA). Technical Basis and Methodology Description. U.S. NRC. NUREG/CR-6350. BNL-NUREG-52467 (1996)
35. Barriere, M.T., et al.: Technical Basis and Implementation Guidelines for a Technique for Human Event Analysis (ATHEANA). U.S. NRC. NUREG-1624, Rev. 1 (2000)
36. Forester, J., Kolaczkowski, A., Cooper, S., Bley, D., Lois, E.: ATHEANA User's Guide. Final Report U.S. NRC. NUREG-1880 (2007)
37. Kirwan, B., Gibson, H., Kennedy, R., Edmunds, J., Cooksley, G., Umbers, I.: Nuclear action reliability assessment (NARA): A data-based HRA tool. Saf. Reliab. **25**(2), 38–45 (2005)
38. Chandler, F.T., Chang, Y.H.J., Mosleh, A., Marble, J.L., Boring, R. L., Gertman, D.I.: Human Reliability Analysis Methods- Selection Guidance for NASA. National Aeronautics and Space Administration. NASA/ OSMA Technical Report (2006)

Part II
Applying Human Factors: Building Better Processes, Procedures and Organizations in Energy

Developing a Human Factors Engineering Process for Control Room Upgrades

Hanna Koskinen, Jari Laarni, Leena Salo and Paula Savioja

Abstract There is a lot of evidence that complex engineering projects do not always proceed as have been planned. One of the reasons is the lack of socio-technical systemic approach to systems design. The aim of this paper is to present a review of basic principles of HFE in the nuclear domain. Our findings and practical experiences suggest that there are several challenges for successful implementation of HFE work, such as the proper timing of HFE activities, the appropriate sharing of knowledge among designers, HFE experts, and management, and the integration of HFE into the systems engineering process. Some recommendations are offered for a more systematic application of HFE practices in the nuclear domain.

Keywords Human factors engineering · HFE process · Systems engineering · Control room design

1 Introduction

There is a lot of evidence that complex engineering projects do not always proceed as has been planned. For example, in the nuclear domain project delays and budget overruns seem to be more a rule than an exception. One of the reasons for the

H. Koskinen (✉) · J. Laarni · P. Savioja
VTT Technical Research Centre of Finland, Vuorimiehentie 3,
02150 Espoo, Finland
e-mail: hanna.koskinen@vtt.fi

J. Laarni
e-mail: jari.laarni@vtt.fi

P. Savioja
e-mail: paula.savioja@vtt.fi

L. Salo
Fortum Power and Heat, Keilaniementie 1, 00048 Fortum, P.O. Box 1000,
02150 Espoo, Finland
e-mail: leena.salo@fortum.com

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_5

design of large systems not to meet their deadlines and costs and expectations is the lack of socio-technical systemic approach to systems design, that is, the social and organizational complexity of the environment in which the systems are used are not recognized [1]. Similarly, experiences from several design projects have shown that the urge to deliver systems within budget and on schedule has resulted in insufficient attention to human factors considerations in system engineering [2]. There is also the potential for the organization to drift away from the systems engineering and human factors standards it still believes it is achieving [1, 3]. We propose that a Human Factors Engineering (HFE) program which is tightly integrated to systems engineering processes is an effective cure for these problems and challenges.

A starting point for HFE is that the way the design work is carried out and how the human and organizational issues of the design are analyzed during the design process has implications in the human performance in the future [4]. According to Booher [5], HFE does not only help us to avoid accidents and to increase efficiency of poorly performing systems, but it also helps us to reduce both project costs and to produce improvements in performance and productivity. Bruseberg [6, 7] has also noticed that by following a standard HFE process the likelihood of financial benefits are higher.

Experiences from several projects suggest that HFE input is needed throughout all phases of the product/system life cycle [8]. It is also important that safety implications are considered as early as possible in the design. If the design flaws are corrected later on in the process the costs are higher and the solutions are often less optimal and user-friendly [7, 9]. A problem regarding the time span of design decisions is that there are some decisions considering safety that have to be made immediately in the beginning of the project, whereas the disadvantages that are caused by these decisions are actualized much later in time.

1.1 Systems Engineering Process in the Nuclear Domain

Designing a new nuclear power plant or modernizing and upgrading the automation systems and control rooms (CR) of existing ones is a complex engineering design activity. To better manage the complexity of the NPP systems and the risks associated with the design of these systems, a systems approach and systems engineering practices are needed in their design.

A typical system engineering process covers four main activities: procurement, analysis, construction and operation, but more than four main stages or activities have often been identified in more detailed categorizations described in standards and guides. Quite often, the procurement phase is not included in the design process and only the following main phases are listed: requirements specification, overall architecture design, detailed design and analysis, installation, acceptance testing and operation (e.g., [10]).

Systems are increasingly considered from the life-cycle perspective, that is, it is followed the system evolution from the statement of a stakeholder need through the disposal of the system [11]. Typically, the acronym CADMID is used when referring to the six phases of the system lifecycle, Concept, Assessment, Demonstration, Manufacture, In-service and Disposal.

2 Human Factors Engineering in the Nuclear Power Plant Life Cycle

According to a review of related literature and studies on the field, the following quite an extensive list of main HFE activities have been identified in the engineering design of CR systems: screening, management and planning, operating experience review, development of concept of operations, function allocation and analysis, task analysis, human reliability and failure analysis, manning, personnel selection and qualifications, system design and development, design of maintenance activities, procedure design, training design, verification and validation, and commissioning and in-time monitoring.

NUREG-0711 [12] identifies and addresses most of the aforementioned activities, however, not all of them. Three activities that are not mentioned above are described below. The first one of them is screening. The screening refers to the fact that HFE activities need to be tailored and individually modified for each I&C and CR project separately. The scale of needed activities is defined based on the complexity and safety impact of the proposed change [13]. As a result, more effort is put on those activities, which have a larger impact on safe and efficient human tasks. The second activity not addressed by NUREG-0711 is the drafting of an overall Concept of Operations (ConOps). According to EPRI [13], ConOps for the new CR describes on how the operating crew is organized and how it monitors and controls the plant systems under different plant conditions. Thus it describes the operation of the system from users/usage point of view. Considerations on ConOps should be done early on, at the beginning of a development project. Lastly, the design of maintenance activities is also typically unheeded from the listings of needed HFE activities. Maintenance activities include all activities that are carried out to prevent equipment failures and repair out of order ones. According to ETSON [4], specific requirements should address also the HF aspects of the maintenance activities and those should be paid attention to as early as possible during the design process. The objective is to facilitate future maintenance by anticipating possible future errors and by trying to make the systems more "maintenance-error tolerant".

To really have an impact and to be able to address appropriately all the activities that needs HF attention in complex new buildings and upgrade projects a more integrated and holistic approach to HFE is called for. This importance of HFE integration is recognized and discussed widely by many authors and organizations (e.g., [9, 14, 15]). For example, according to NEA/CSNI/R [15], a systematic and integrated HFE program is needed, and there should be a continuous, timely and effective dialogue between HF experts and technical specialists. Open and continuous communication is required between designers and HF specialists so that critical human-system interactions have been specified and standard methods and valid analyses have been applied [15].

According to Madni [2], two kinds of integration of HFE with systems engineering are needed: Since the HFE area is quite fragmented itself, it has to be internally integrated first; only after that the HFE activities can be meaningfully integrated with system engineering domains. For example, there are complex interactions between HF activities that have to be specified before it can be started to think about their integration with systems engineering domains and activities. IEEE 1023-2004 [16] among others emphasizes that HFE should be planned as an integral part of the lifecycle activities in NPPs, and HFE activities should be implemented under dedicated HFE programs. According to IAEA N-PT-3.10 [10], human factors should be considered at all levels of engineering design, and it should be integrated into the plant change process at various levels, from project management, to verification and validation, and to end users.

There are several clear benefits of HFE in systems design [15]. HFE is considered as an essential support function throughout the project lifecycle, through which HF problems can be avoided and mitigated (e.g., [6]). HFE activities provide knowledge to inform the design of complex systems, and HFE provides quality control at different stages of the design process. HFE experts play a mediating role between end-users and designers and in establishing a relation between different parties. But despite all this the value of HF considerations has not been always fully acknowledged [17, 18]. In general, the success and impact of HFE activities has shown to be difficult to evaluate. HFE is considered a costly process, and there is a general lack of understanding of how HF could produce cost benefits [6]. Because of big investments without certainty of later return, there are a lot of reservations against HFE [19].

3 HFE Process for Control Room Upgrades

To address the challenge of systematically describing an HFE design process and integrating it to the systems engineering process, an example of HFE process model that was developed as a part of an industrial partner's HFE program is described in the following. In the example company they are having an ongoing I&C modernization project aiming at a partial renewal of the I&C and CR systems at their operating plant. The long traditions and experience in CR and HSI design are displayed also in the present modernization project as in the project a strong emphasis is placed on conceptual and safety design as well as human factors engineering. This includes human factors analyses, user-centered design, and use of ergonomics standards being integral part of the HFE design process. Due to the large scale of the ongoing modernization project and the new regulatory requirements for HFE, the need to describe the HFE knowledge and processes in a form of a more formal HFE program was recognized. The aim of the HFE process development was to create an HFE process and procedures for different kinds of projects grounded on company's own practices and HFE standards and guidelines. Other pursued goals include efficiency in conducting HFE in projects, standardization and high quality of HFE practices, and improved integration of HFE to the engineering process. In the development of the HFE process and procedures international standards and guidelines, such as ISO 11064-1 [20] and NUREG-0711 [12], have been compared to the company's own HFE practices. It has also been seen important to take the internal project guidelines as basis for HFE program development in order to ensure that HFE would be integrated with ease to the current engineering processes and practices.

3.1 HFE Process Model

As often recognized, it is challenging to get HFE activities involved in projects from the start when initial strategies and concepts are created. Many models of ergonomic design only describe the HFE activities during the actual design phases, whereas some sources (e.g., [20]) describe also the tasks beginning from the start of a project. The ISO 11064-1 standard was taken as the basis for the HFE process development, since it covers all the phases of an upgrade process from initial planning to commissioning. Moreover, we found it to portray the timely progress of a design process better than, for example, the NUREG-0711. The developed HFE process model is presented in Fig. 1. The HFE process covers an upgrade project from the initial clarifications and project planning to the implementation and commissioning. In each of the phases the main HFE tasks and their inputs and outputs are depicted.

The recognized need for change triggers the design process and actualize the first phase of the HFE process model i.e., the initial clarification phase. In the starting phase it is useful to have a checklist of HFE issues and tasks that should be taken into account when first plans and strategies are made. In addition, the impacts of the change ahead and the needs and constraints of the proposed new design solution is analyzed and evaluated against the existing concept of operations and operational experiences. This phase includes also a screening process where the scope of HFE activities is estimated. The scope needs to be determined based on the safety criticality and the amount and novelty of changes to the existing control room concept and operator work. The HFE goals and requirements as well as the preliminary solutions are described and serve as the HFE input to the project proposal and plan. The initial clarification phase also produces the HFE program in which the HFE tasks, HFE organization and responsibilities regards to the specified tasks are defined. The HFE program should also clarify the links between the HFE tasks and the other fields of design engineering and relevant stakeholder groups.

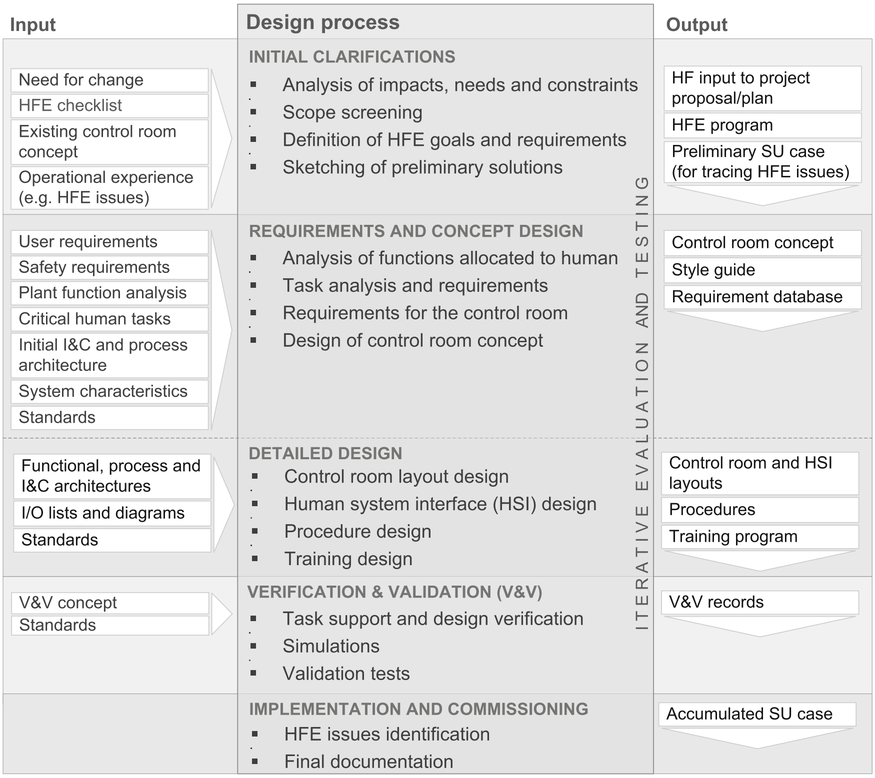

Fig. 1 HFE tasks, inputs and outputs in the HFE process

In addition, the preliminary systems usability case is drafted in which the HFE issues of importance for the project are initially recorded.

After the initial clarifications the project shifts to actual design phase which starts with the requirements specification and concept design. Requirements are derived from identified user needs, regulatory guidelines, ergonomics standards, and safety design principles. In this process also task analyses are performed. Other essential inputs in requirements and concept design phase are plant function analysis, specification of critical human tasks, the existing plant configuration and the new I&C and process architectures. Furthermore, the characteristics of the selected technologies set limits for the possible solutions. The main outputs from this phase are the description of the control room concept, the style guide and setting up the requirements database for design continuation. The control room concept describes what tools, procedures and organization are used in different operational conditions and degraded situations, and it also describes the human-system interface and procedure concepts.

Following the requirement specification and concept design, the detailed design can take place. In this phase the control room layout, human-system interfaces, procedures and the training program are designed. Inputs from the I&C and process design include e.g., I/O lists and functional and process diagrams. Now the design of different sub-systems progress side by side. In this phase challenges arise from the coordination and information exchange between the different design tasks and establishing proper feedback loops from the users to the design (in accordance with user-centered design).

In the verification and validation phase the ergonomics of the design solutions are evaluated and their acceptance assessed with task support and design verification, simulations and validation tests. Validation of design should be done early on before costly commitments and the configurations of the design solutions have been set. Validation throughout the design provide in-depth insight into possible emerging design problems/HFE issues and allow design decisions to be made on an informed basis. In practice, this means subsequent validations of control room systems to be planned and carried out when reasonable. The integrated system validation functions as the final validation test of the whole control room and its ergonomic and safety acceptance. The implementation and commissioning phase progresses partly at the same time as the verification and validation phase. If HF issues are identified, they are recorded, evaluated and fed back to the design process when possible. During the implementation phase the final documentation of the design is produced.

The HFE process model portrays the tasks of which the HFE organization is of main responsibility in design projects. The model shown above includes the input and output documentations of the process but lacks the clear links between the HFE tasks and other design engineering disciplines. The HFE organization plays a mediating role between end-users, different technical design disciplines, and the supplier. Therefore, it is important to recognize the interfaces of the different HFE tasks with the other project stakeholders. In addition to the input and output documents the boundary objects common to all design parties and the important information exchange points should be identified. Moreover, the common baselines where the design is aligned within HFE and with other more technical design disciplines should be identified. In HFE program all individual HFE tasks and tasks that require collaborative and multidisciplinary effort are stated and documented. Also, the common and shared tools are important to describe in the program.

3.2 HFE Process and the Engineering Process

One of the main goals of the development of the HFE process was to enable better integration of the HFE process to the systems engineering process.

Figure 2 depicts the overall system engineering process and HFE and the recognized links between the two processes. In addition, for each of the main phases and relative boundary objects connected to them, an example of what the specific HFE contributions may be are highlighted and described.

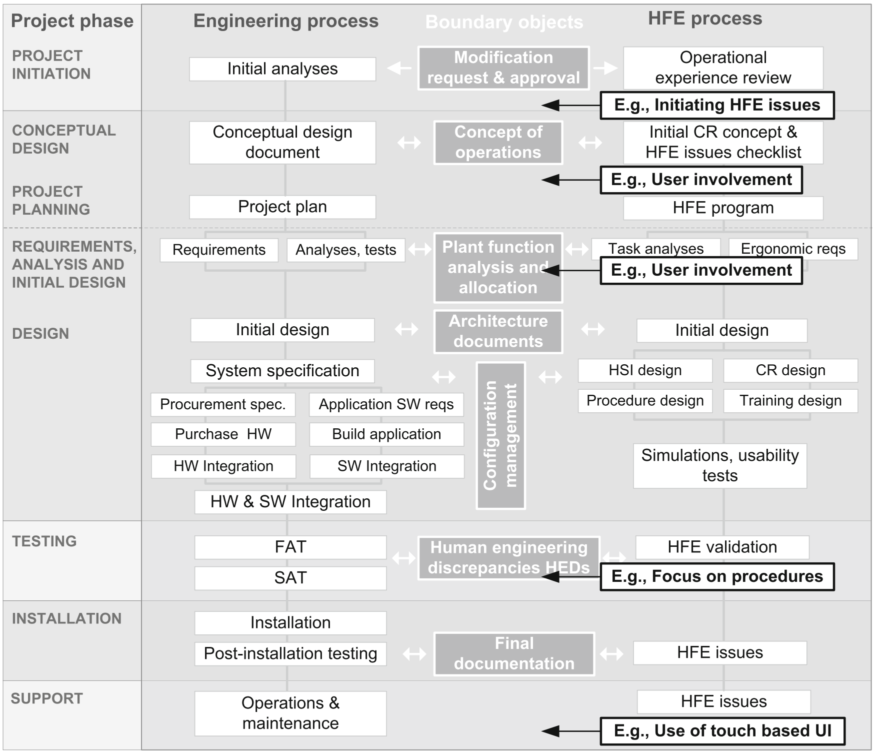

Fig. 2 Links between the systems engineering process [14] and the HFE process with highlighted examples of individual HFE contributions

The aim of illustrating the systems engineering process and the HFE process side by side is to create a visual representation that could help understand the interdependencies between the two processes and tasks carried out within them. It is not possible to capture comprehensively all these links in single figure, but some important information exchange points are defined for which common procedures for handling information and making design decisions need to be specified.

Already at the very first moment when the “initial need” for the change or development project is recognized the HFE process is launched. The operational experience reviews carried out by HFE may even initiate these development projects. In either way the HFE plays an important role in capturing the background of the developments and identifying and recording the HFE issues to be concerned in the design. After project initiation, in the concept design and project planning phase, the concept of operations is a boundary object that describes the common design principles and the pursued future operation after reaching the project completion. A real life example of the HFE contribution made in the conceptual phase of the modernization project can be found from the end-user involvement in the development of large screen display concept for main CR. HFE played as mediator between the end-users and the design engineering and recorded essential HFE issues (e.g., management of the content) to be dealt with in the creation of the overall concept of operation. In the following more detailed design phase the boundary objects include plant function analysis and architectural level documents. Plant function analysis is a multidisciplinary task in which the functions that are changed during the upgrade project are analysed and allocated either to human or machine. In this phase, end-user involvement may also play an important role. The allocated functions and tasks are further analysed by the two processes. The architecture documents describe the basic structure of solutions to which more detailed designs are based.

During the design many parallel design activities may advance the detailing of the different sub-systems, whereas later in the testing, implementation and commissioning phases these sub-systems are put together and viewed as one integrated whole. To promote the integration of the parallel design processes of sub-systems and different fields of design, the points of information exchange and design freezing need to be identified and common procedures and tools defined, e.g. configuration management databases. Also the points where the individual HFE tasks are mutually integrated should be defined. Such points may include e.g., architectural documents, usability testing at simulators, HFE validation, and updating the HFE issues register. In the testing phase the design is validated from the operation and usability point of view. An example of this can be drawn from the graded sub-system validation approach developed for the validation needs of complex and stepwise realized upgrades. Grading enables also identification of and emphasis on critical focus areas (i.e., HFE issues) regarding each sub-system to be tested. For example, it was recognised critical to validate the use of operating procedures exploiting and introducing totally new information presentation format i.e., flow-charts if compared to the earlier used procedures. As a result of the kind of validation test, human engineering discrepancies (HEDs) may be identified. The HEDs function as boundary objects when discussed about the handling of them with e.g., the technical design and training personnel. In implementation and commissioning phases the HFE process collects and records operation experiences of new CR and HSIs. For example, in one project it was found important to pay particular attention to the implementation and commissioning of new touch-screen based safety UI because of the identified HFE issues during its design. The contribution of HFE was to recognize the emerging issues early enough so that they could still be appropriately addressed without danger to safety of operation.

Further development of the above described HFE process model and showcase for HFE contributions would require more detailed consideration of the integration of HFE to systems engineering, not only on the level of the design and development but also on the levels of project management and licensing. The model should also better address the lifecycle perspective of NPPs.

4 Conclusions and Recommendations

Based on the literature review, the HFE process development and our earlier experiences of carrying out variety of HFE tasks we propose that an established Human Factors Engineering program can be an effective cure for the problems and challenges that may rise in complex NPP design processes. The HFE tasks need to be first mutually integrated and secondly integrated to systems engineering processes for HFE to have effective impact on the designed solutions. In the following some recommendations are given that hopefully open up a route for systematic application of HFE practices:

- Since NUREG-0711 is a review guide, not a design process model, it does not provide definitive guidance on modernization/new build projects. Additional guidance has to be derived from other sources.
- HFE tasks have to be mutually integrated through feed-forward-feedback links to achieve a well-functioning system of activities. The top-down approach for a safety evaluation advocated by NUREG-0711 affords a good basis for the mutual integration of HFE activities.
- HFE process has to be started as early as possible in the design process, and it should continue throughout the life-cycle of the target system. Also in the new build projects, HFE considerations should be started from the beginning of the design work. The HFE process models should identify all HFE tasks from the initiation of the project until commissioning, operations and maintenance.
- HFE process has to be fully integrated into an iterative systems engineering process throughout the lifecycle. There is little existing guidance on how to make the HFE process more unified and systematic, however, some novel frameworks such as the Incremental Commitment Model [21] might provide some help in this respect. Visualizing the systems engineering process and the HFE process together may help understanding the interdependencies between the processes.
- Common boundary objects and tools should be identified and developed in projects to promote integration of design disciplines. The concept of operations that describes an end point vision of the project may be one example of common boundary objects between the engineering and HFE processes.
- Operator involvement is required from the early stages through the end of the project. Users and other important stakeholder groups have to have opportunities to provide input to all HFE activities. The HFE organization has a mediating role between end-users and different design engineering disciplines, and the supplier.
- The verification and validation process must be iterative by nature, and there must be a systematic procedure for aggregating and systematizing V&V data.
- Grading process is important in determining the emphasis and amount of effort allocated to HFE tasks and activities. It is important to concentrate on the most critical aspects of the design (e.g., identified HFE issues).

- HFE's role in safety management needs clarification, and a special plan has to be prepared for facilitating a dialogue between HFE and HRA experts.
- More detailed analyses of HFE costs and benefits are needed, and they should be adjusted iteratively throughout the planning process.

References

1. Mumford, E.: The story of socio-technical design: reflections in its successes, failures and potential. Inf. Syst. J. **16**, 317–342 (2006)
2. Madni, A.M.: Integrating humans with software and systems: Technical challenges and a research agenda. Syst. Eng. **13**, 232–245 (2010)
3. Healey, A.N., Cacciabue, P.C., Berman, J.: Human factors in nuclear safety: present and Future. Cogn. Technol. Work **15**, 1–3 (2013)
4. ETSON—European Technical Safety Organisations Network. Technical Safety Assessment Guide: Human and Organisational Factors in Nuclear Facilities Design and Modification Processes (2013)
5. Booher, H.R.: Introduction: Human systems integration. In: Booher, H.R. (ed.) Handbook of Human Systems Integration. Wiley, Hoboken, NJ (2003)
6. Bruseberg, A.: Presenting the value of human factors integration: guidance, arguments and evidence. Cogn. Technol. Work **10**, 181–189 (2008)
7. Bruseberg, A.: Cost-Benefit Analysis for Human Factors Integration: A Practical Guide. HFI DTC/WP 2.7.2/3. BAE Systems (2009)
8. Samaras, G.M.: Human-Centered systems engineering: managing stakeholder dissonance in healthcare delivery. In: Kolker, A. (ed.) Management Engineering for Effective Healthcare Delivery Principles and Applications. IGI Global, Hershey, PA (2012)
9. Hale, A., Kirwan, B., Kjellén, U.: Safe by design: where are we now? Saf. Sci. **45**, 305–327 (2007)
10. International Atomic Energy Agency (IAEA).: Integration of Analog and Digital Instrumentation and Control Systems in Hybrid Control Rooms. IAEA Nuclear Energy Series, No. NP-T-3.10. International Atomic Energy Agency, Vienna (2000)
11. IEEE 1220–2005.: IEEE Standard for Application and Management of the Systems Engineering Process. The Institute of Electrical and Electronics Engineers, Inc., New York, NY
12. NUREG-0711.: Human Factors Engineering Program Review Model. United States Nuclear Regulatory Commission, Washington, D.C. (2012)
13. EPRI Human Factors Guidance for Control Room and Digital Human-System Interface Design and Modification—Guidelines for Planning, Specification, Design, Licensing, Implementation, Training, Operation, and Maintenance. Electric Power Research Institute, EPRI 1008122, Nov (2004)
14. Hugo, J.: Towards a unified HFE process for the nuclear industry. In: Proceedings of NPIC & HMIT, on CD-ROM. San Diego, CA, 22–26 July. American Nuclear Society: American Nuclear Society, LaGrange Park, IL (2000)
15. Nuclear Energy Agency. Organisation for Economic Co-operation and Development (OECD), CSNI Technical Opinion Papers, No. 10.: The Role of Human and Organisational Factors in Nuclear Power Plant Modifications. Organisation for Economic Co-operation and Development (OECD) (2009)

16. IEEE 1023–2004.: IEEE Recommended Practice for the Application of Human Factors Engineering to Systems, Equipment, and Facilities of Nuclear Power Generating Stations and Other Nuclear Facilities. The Institute of Electrical and Electronics Engineers, Inc., New York, NY (2004)
17. Holmström, C., Green, M.: Human factors and control centre design issues. What lessons have we learned. In: Fostervold, K.I., Endestad, T. (eds.) Ved Innegangen til Cyberspace—ergonomisk teenkning inn i et nytt årtusen. Nordiska Ergonomisälskapet (2000)
18. Waterson, P., Kolose, S.L.: Exploring the social and organisational aspects of human factors integration: a framework and case study. Saf. Sci. **48**, 482–490 (2010)
19. Kirwan, B.: An overview of a nuclear reprocessing plant Human Factors programme. Appl. Ergonomics **34**, 441–452 (2003)
20. ISO 11064–1.: Ergonomic Design of Control Centres. Part 1: Principles for the Design of Control Centres. The International Organization for Standardization (2000)
21. Pew, R.W., Mavor, A.S.: Human-system Integration in the System Development Process: A New Look. National Academy Press, Washington, DC (2007)

Defining Expertise in the Electric Grid Control Room

Susan Stevens Adams and Francis P. Hannigan

Abstract Electric distribution utilities are on the brink of a paradigm shift to smart grids, which will incorporate new technologies and fundamentally change control room operations. Expertise in the control room, which has never been well defined, must be characterized in order to understand how this shift will impact control room operations and operator performance. In this study, the authors collaborated with a utility company in Vermont to define and understand expertise in distribution control room operations. The authors interviewed distribution control room operators, HR personnel, and managers and concluded that a control room expert is someone who has 7–9 years' experience in the control room and possesses certain traits, such as the ability to remain calm under pressure, effectively multi-task and quickly synthesize large amounts of data. This work has implications for control room operator training and how expertise is defined in the control room domain.

Keywords Human factors · Expertise · Power grid · Control room

1 Introduction

Distribution electric grid centers, considered a critical infrastructure in the United States [1], play a vital role in ensuring that electricity reaches the end user. Distribution control room operators must monitor grid load and coordinate with field crews for both routine switching tasks (such as scheduled maintenance on a particular line) and unplanned switching tasks (such as outages due to weather). These distribution utilities are on the brink of a paradigm shift from a more traditional grid model, with a one-way flow of information from the utility to the end

S.S. Adams (✉)
Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM, USA
e-mail: smsteve@sandia.gov

F.P. Hannigan
Quantum Improvements Consulting, 3259 Progress Drive, Orlando, FL, USA
e-mail: fhannigan@quantumimprovements.net

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_6

user, to smart grids, which will incorporate new technologies available to the home user—such as solar panels and electric cars—and will result in a bi-directional flow of energy and information. In addition, increases in automation will result in a 'self-healing grid' in which grid equipment will take action without the input of the control room operator. These changes to the grid model will fundamentally change control room operations and decision-making. As found in one of the author's past research [2], this increase in automation is likely to change the control room operator's communication with the field crew and, by extension, their situation awareness of what is happening on the grid.

Another factor that has not yet been considered with the changes in automation and the grid model is that of expertise. Expertise in the control room has never been well defined and, given the impending change to the grid model, it is important to understand how operator expertise, combined with the changes in automation may, or may not have an impact on the grid. What are the characteristics of an expert in the control room? Will the characteristics of what defines an expert have to change? These questions have not thoroughly been studied, despite the drastic changes that are about to take place.

In this study, the authors collaborated with a utility company in the state of Vermont (the first state to be 100 % smart grid) to define and understand expertise in distribution control room operations in order to better understand how these changes in automation will influence operator performance. The authors conducted structured interviews with distribution control room operators, HR personnel, and managers to identify the current characteristics and traits of an expert in the distribution control room.

2 Expertise

Expertise has been studied in numerous domains using a wide variety of tasks, from chess to air traffic control tasks [3, 4], physicians and clinical diagnosis tasks [5, 6], music [7], and weather forecasting [8] among many others. Expertise research can be categorized by skill acquisition [9] or knowledge acquisition [10] or whether it is important to differentiate between several levels of expertise [11] or whether the differentiation between expert and non-expert is acceptable. In trying to define who or what an expert is, the theories are wide and varied. Ericsson suggests an expert must be able to select superior actions, generate rapid reactions, and control movement production [12]. Weiss and Shanteau claim there are 4 different categories of expertise, including expert judges, experts in prediction, expert instructors, and expert performers, yet each of these types of experts are bound the fundamental cognitive ability of evaluation [13]. Finally, the Dreyfus model of skill acquisition [11] defines 5 different levels of expertise including novice, advanced beginner, competent, proficient, and expert. While dissecting the nuances of the complete expertise literature is beyond the scope of this effort, it was the authors' goal to understand the main characteristics that define an expert, agnostic of domain, and how best to define what makes an expert in the electric power grid distribution control room.

The fact that expert reasoning is specific to a domain [14] is a widely accepted statement regarding expertise and speaks to the importance of the current study. Certainly an airline pilot with thousands of hours of flight time in a particular aircraft would be considered an expert in that aircraft; however, that same individual would not be considered an expert if he suddenly found himself doing brain surgery. While that is a drastic example, other attributes that are theorized to define what an expert is are more nuanced. For example, while extensive experience of activities in a particular domain is necessary to reach exceptional levels of performance that experience does not necessarily translate to expert levels of achievement [9]. In fact, it appears as though one's ability to reach expert levels of achievement are constrained by individual characteristics such as abilities, mental capacities, and innate talents [9] and more specifically constrained by information processing ability and working memory capacity [15] of the individual. Furthermore, some research shows "knowledge gained through extensive experience of activities in a domain" is not a differentiator between experts and novices [16, 17]. In other words, while expertise can clearly be defined as being domain-specific, simply time spent working in a domain is not the sole factor in determining expertise.

Also core to expertise are the constructs of discrimination and evaluation [13, 18–21]. Discrimination refers to the ability of the expert to discern between the relevant and important pieces of information and those that are not, given the task or scenario. An expert should be able to perceive subtle differences in given situations that non-experts would typically not notice [19]. One step further than discrimination or recognition of a stimulus is the evaluation of those critical details. Appropriate evaluation of the task or scenario at hand allows for proper actions to be contemplated and acted upon. Without proper evaluation, an incorrect course of action is likely to be taken.

The way experts store and recall knowledge is critically different from that of novices as well. For instance, Munshi et al. postulates there are three stages of experience. Individuals in a less than expert state use deductive reasoning, simply testing individual hypotheses in an attempt to find the correct action or decision. In the second stage, as might be expected, increased knowledge and learned experiences allow the individual to form causal relationships between data points. Finally, in the 'scheme' stage, knowledge is hierarchically organized in which pattern recognition and inductive reasoning is used to come to conclusions [22]. Anderson et al. also proposed three similar stages of expert knowledge acquisition: (1) The cognitive stage in which there is limited application of textbook learning to the real world. (2) The associative stage where individuals begin applying their knowledge from the cognitive stage. (3) The autonomous stage where knowledge where knowledge is automatic rather than deliberate [10]. These states and the similar knowledge acquisition models [11] show some consensus around the cognitive changes and processes involved with knowledge acquisition. It is these changes in representation of knowledge that novices experience in their pursuit of expertise [23] which leads to a better functional and behavioral understanding of complex systems [24].

Given this understanding of what it means to be an expert, how does one go about determining who is an expert in a particular organization or domain? Experts must have extensive experience in a domain, though this is not predictive of expert levels of achievement [9]. Social acclimation [25] is the agreement of professionals regarding who is an expert in the domain. Typically, multiple individuals will not elect the same individual who is not an expert. However, it is also true that occasionally a nominated individual's performance is found to be lacking [26, 27]. Additionally, amount of overall experience in a domain needs to be distinguished from deliberate practice and the number of years spent on relevant activities is not strongly related to performance [28].

Given this murky determination of expertise, the authors collected information pertaining to years in the field, years in the job, academic experience, training experience, other relevant experience for control room operators. In addition, the authors discussed with each participant who he believed was the most expert person at his work location, and why.

3 Methodology

3.1 Participants

The authors interviewed 13 control room operators, 3 managers, and 1 human resources (HR) personnel from a Vermont utility company across 2 weeks in December, 2015.

With respect to the control room operators, each had varying degrees of experience ranging from just 2 months on the job to more than 37 years. Ordered by this utility's job roles, researchers interviewed 1 Apprentice, 1 3rd Class (3C) operator, 2 2nd Class (2C) operators, and 9 1st Class (1C) operators. An 'Apprentice' is defined as a new-hire who is learning and executing basic operations under supervision. As an operator progresses through each class, they are expected to learn more complex tasks and complete them under decreasing amounts of supervision, until reaching the 1C level at which time an operator is expected to be able to operate independently.

While the focus of this step in the research is to determine what control room operator expertise looks like, management and human resources personnel were interviewed to gain perspective on the qualities and attributes that are rewarded in the workplace as well as to gain a reference for how hiring and promotion are conducted for control room operator positions. Furthermore, job descriptions, performance reviews, and promotion assessments provided by the human resources department, provided the basis for some of the interview questions.

3.2 Structured Interviews

The authors conducted structured interviews with the managers, human resources personnel, and control room operators. Interviews were conducted on utility property either in a conference room or the control room itself and participants were interviewed individually. Interviews conducted in the control room took place during a 'quiet' time and did not impact the operator's job performance. Each interview was scheduled for 1 h although some interviews took less time. Interviews were structured in such a way as to incorporate the expertise literature and to investigate how an expert would be defined within the power distribution control room domain. The authors asked questions pertaining to the importance of experience in the control room, peer assessments of who is considered an expert, how experts are different from non-experts and what attributes an expert in the control room possesses. In addition, the authors interviewed operators to better understand how they currently perform system restoration tasks and how these tasks might change as the result of increasing automated components on the grid.

The following interview questions were asked of the control room operators, and a sub-set of questions or modified questions were asked of the HR and management personnel:

1. How would you define an expert in your domain?
2. What are the top tasks that an expert typically performs?
3. What are the top qualities or characteristics of an expert?
4. Who among your peers do you consider to be an expert and why?
5. How comfortable would you be sketching an area of the sub-transmission lines?
6. Generally, what is SCADA (Supervisory Control And Data Acquisition) data used for and how would a novice interact with the data differently from an expert?
7. Describe the testing/assessment required for progressing to the 1C level.
8. When is someone considered to be an expert with respect to the defined classes and years of experience?
9. Which skills would and would not transfer to a different control room?
10. Does management ask the input of 1C operators to determine the readiness of an Apprentice/3C/2C to advance to the next class?
11. What is the difference between a 'basic' understanding vs. an 'advanced' understanding of system knowledge?
12. Is there anything else you think is pertinent to defining an expert control room operator that we have not covered?

4 Results

The participant responses to the interview questions were consolidated and are outlined below.

1. How would you define an expert in your domain?
   An expert was said to be someone who had previous field knowledge, an in-depth knowledge of the system, exposure to a variety of events and had a great deal of hands on experience.
2. What are the top tasks that an expert typically performs?
   The top task that an expert performs (that a non-expert does not) is complex switching.
3. What are the top qualities or characteristics of an expert?
   An expert was said to be able to synthesize data, to respond to a large volume of data, to remain cool, calm and collected under pressure, to be adaptable, to multi-task effectively and quickly, and to efficiently use and navigate the system.
4. Who among your peers do you consider to be and expert and why?
   The experts who were named had combination of strong academic or practical experience along with significant time in control room.
5. How comfortable would you be sketching an area of the sub-transmission lines?
   A 1C would be very comfortable sketching the area; a 2C would be comfortable but the sketch would not be as detailed as a 1C's; a 3C would not be able to generate a sketch.
6. Generally, what is SCADA data used for and how would a novice interact with the data differently from an expert?
   It was emphasized that all operators use SCADA data but how quickly an operator could process the SCADA data was the important factor. A 1C should be able to synthesize SCADA data quickly but a 3C will take longer to understand SCADA data.
7. Describe the testing/assessment required for progressing to the 1C level.
   The test involves both oral and written portions, both of which consist of multiple choice, problem solving and schematic drawings.
8. When is someone considered to be an expert with respect to the defined classes and years of experience?
   Taking an average of the provided responses, an expert is someone who is a 1C with an additional 4.7 years of experience.
9. Which skills would and would not transfer to a different control room?
   The skills that would transfer are those related to customer service, fundamentals and principles of electricity, safety knowledge and equipment type knowledge. The skills that would not transfer are the specific system details, such as knowledge of the transmission and sub-transmission systems.
10. Does management ask the input of 1C operators to determine the readiness of an Apprentice/3C/2C to advance to the next class?

There were a variety of responses leading the authors to conclude that this activity was dependent on the relationship between the particular management personnel and the 1C operator.

11. What is the difference between a ‘basic’ understanding vs. an ‘advanced’ understanding of system knowledge (as it is referred to in the class job descriptions)?
    A basic understanding includes knowledge of electricity fundamentals and basic knowledge of system. An advanced understanding includes detailed knowledge of system (such as power flows, voltage levels), the ability to make independent decisions and perform actions without supervision, and writing and executing switching plans on the fly in an emergency situation.
12. Is there anything else you think is pertinent to defining an expert control room operator that we have not covered?
    None of the participants had anything to add.

Based on the responses to these questions, the authors determined that a control room ‘expert’ is someone who typically has 7–9 years’ experience in the control room. In addition, distribution control room experts were said to possess certain traits, such as the ability to remain calm, cool, and collected under pressure, being adaptable and able to effectively multi-task, quickly synthesize large amounts of data, quickly and efficiently using and navigating the system and having exposure to lots of different types of events. In line with the previously discussed literature, the results of the interviews point toward experiential learning and qualities and attributes of successful individuals in the control room rather than simply time on task.

5 Conclusions and Future Steps

Control room operators, managers and human resources personnel from a distribution utility company in Vermont were interviewed to define and understand expertise in distribution control room operations in order to better understand how changes in automation will influence operator performance. An expert in the control room was defined as someone having 7–9 years’ experience in the control room as well as possessing certain traits, such as the ability to remain calm under pressure, effectively multi-tasking and quickly synthesize large amounts of data. This is the first study of its kind to characterize expertise in the distribution control room.

For the next stage in this project, expertise and automation will be investigated in an empirical study at the same Vermont distribution utility. Operators will complete power outage scenarios in a simulator and response time and decision accuracy will be assessed. The level of automation in each scenario will be manipulated and operators will be classified as an expert or a non-expert based on the findings from the current study.

The results from this study will help inform tools and strategies that will help control room operators adapt to a changing grid, respond to critical incidents and maintain critical performance skills.

Acknowledgments The authors would like to acknowledge Green Mountain Power for their participation and involvement with this study.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2016-2219 C.

References

1. The Department of Homeland Security: www.dhs.gov (2015). Accessed 27 March 2015
2. Stevens-Adams, S., Cole, K., Haass, M., Warrender, C., Jeffers, R., Burnham, L., Forsythe, C.: Situation awareness and automation in the electric grid control room. Procedia Manufact. **3**, 5277–5284 (2015)
3. Chi, M.T.H.: In: Siegler, R. (ed.) Children's Thinking: What develops?, pp. 73–96. Erlbaum, Hillsdale, NJ (1978)
4. De Groot, a D.: Thought and choice in chess. Am J Psychol. **79** (1965)
5. Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., Wang, Y.: Expertise in a complex skill: diagnosing X-ray pictures. In: Chi, M.T.H., Glaser, R., Farr, M.J. (eds.) The Nature of Expertise, pp. 311–342. Erlbaum, Hillsdale, NJ (1988)
6. Feltovich, P.J., Barrows, H.S.: Issues of generality in medical problem solving. In: Schmidt, H.G., de Volder, M.L. (eds.) Tutorials in Problem-Based Learning, pp. 128–142. Van Gorcum, Maastricht, Netherlands (1984)
7. Glenberg, A.M., Epstein, W.: Inexpert calibration of comprehension. Memory Cogn. **15**, 84–93 (1987)
8. Hoffman, R.R., Trafton, G., Roebber, P.: Minding the Weather. How Expert Forecasters Think. MIT Press, Cambridge, MA (2005)
9. Ericsson, K.A.: The Influence of experience and deliberate practice on the development of superior expert performance. The Cambridge Handbook of Expertise and Expert Performance, pp. 685–705 (2006)
10. Anderson, J.R., Corbett, A.T., Koedinger, K.R., Pelletier, R.: Cognitive tutors: lessons learned. J. Learn. Sci. **4**(2), 167–207 (1995)
11. Lester, S.: Novice to expert: The Dreyfus model of skill acquisition. **44**(1984), 1980–1982 (2005)
12. Ericsson, K.A.: The search for general abilities and basic capacities. The Psychology of Abilities, Competencies, and Expertise, pp. 93–125 (2003)
13. Weiss, D.J., Shanteau, J.: Empirical assessment of expertise. Hum. Factors **45**(1), 104–116 (2003)
14. Voss, J.F., Greene, T.R., Post, T.A. Penner, B.C.: Problem-solving skills in the social sciences. In: Bower G.H. (ed.) The Psychology of Learning and Motivation: Advances in Research and Theory, vol. 17, pp. 165–213. Academic, New York (1983)
15. Jipp, M.: Expertise development with different types of automation: A function of different cognitive abilities. Hum. Factors **58**(1), 92–106 (2016)
16. Camerer, C.F., Johnson, E.J.: The process-performance paradox in expert judgment: how can the experts know so much and predict so badly? In: Ericsson, K.A., Smith, J. (eds.) Towards a General Theory of Expertise: Prospects and Limits (pp. 195–217). Cambridge: Cambridge University Press. Retrieved from: https://www0.gsb.columbia.edu/mygsb/faculty/research/pubfiles/1152/Process_and_performance_of_experts.pdf (1991)

17. Shanteau, J., Stewart, T.R.: Why study expert decision making? Some historical perspectives and comments. Organ. Behav. Human Decis. Processes **53**, 95–106 (1992)
18. Lipshitz, R., Klein, G., Orasanu, J., Salas, E.: Taking stock of naturalistic decision making. J. Behav. Decis. Making **14**(5), 331–352 (2001)
19. Hammond, K.R.: Human Judgment and Social Policy: Irreducible Uncertainty, Inevitable Error. Oxford University Press, Unavoidable Injustice (1996)
20. Thomas, R.P., Willem, B., Shanteau, J., Raacke, J., Friel, B.: CWS Applied to Controllers in a High Fidelity Simulation of ATC. International Symposium on Aviation Psychology, Columbus, Ohio USA (2001)
21. Shanteau, J., Weiss, D.J., Thomas, R.P., Pounds, J.C.: Performance-based assessment of expertise: how to decide if someone is an expert or not. Eur. J. Oper. Res. **136**(2), 253–263 (2002)
22. Munshi, F.M., AlJarallah, B.M., Harasym, P.H.: Problem solving strategy and diagnostic performance at three expertise levels. Biosaf. Health Educ. **1**(4) (2013)
23. Bedard, J., Chi, M.T.H.: Expertise. Curr. Dir. Psychol. Sci. **1**(4), 135–139 (1992)
24. Hmelo-Silver, C.: Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. Cogn. Sci. **28**(1), 127–138 (2004)
25. Shanteau, J., Weiss, D.J., Thomas, R.P., Pounds, J.C.: Performance-based assessment of expertise: how to decide if someone is an expert or not. Eur. J. Oper. Res. **136**(2), 253–263 (2002)
26. Ericsson, K.A., Smith, J.: Prospects and limits of empirical study of expertise: an introduction. In: Ericsson, K.A., Smith, J. (eds.) Toward a General Theory of Expertise: Prospects and Limits, pp. 1–38. Cambridge University Press, Cambridge (1991)
27. Vicente, K.J., Wang, J.H.: An ecological theory of expertise effects in memory recall. Psychol. Rev. **1**(105), 33–57 (1998)
28. Ericsson, K.A., Krampe, R.T., Tesch-Romer, C.: The role of deliberate practice in acquisition of expert performance. Psychol. Rev. **100**, 363–406 (1993)

Measurement Sufficiency Versus Completeness: Integrating Safety Cases into Verification and Validation in Nuclear Control Room Modernization

Ronald Boring and Nathan Lau

Abstract This paper reviews verification and validation (V&V) as applied in the context of nuclear power plant control room modernization. A common approach for V&V is summative or late-stage evaluation of the finalized design through a process called integrated system validation. Yet, common practice in user-centered design is to conduct evaluations early on in-progress system prototypes. Iterative, early-stage evaluation can form the basis of a safety case argument to ensure the regulatory acceptability of the new human-machine interface in the control room. It is argued that a series of formative evaluations provide more complete evidence of the safety of the new system than does a single summative evaluation.

Keywords Verification · Validation · Integrated system validation · Safety case · Control room · Nuclear power plant · Human factors

1 Introduction

Nuclear power plants in the United States (U.S.) contain predominantly legacy control rooms comprised of analog or mechanical instrumentation and controls (I&C). Yet, new digital control systems and displays are readily available and have been extensively implemented in other process control industries [1, 2]. As noted across several research reports [3–5], barriers to control room upgrades in the nuclear industry are multifold—from regulatory, to know-how, to plant downtime, to cost. Despite such barriers, there is a desire on behalf of many plants to move forward with control room modernization. Reliability issues of aging I&C, the cost of maintaining obsolete systems, training requirements for new operators, and

R. Boring (✉)
Idaho National Laboratory, Idaho Falls, ID, USA
e-mail: ronald.boring@inl.gov

N. Lau
Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
e-mail: nathan.lau@vt.edu

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_7

successful international examples of control room modernization make a compelling case for upgrades, slowly overriding those barriers.

The Human Systems Simulation Laboratory (HSSL) [6] at Idaho National Laboratory (INL) serves as a testbed for operator-in-the-loop studies on modernization, where simulator studies have been carried out using licensed reactor operators using new digital human-system interfaces (HSIs) as part of industry upgrades. A key lesson learned from these operator-in-the-loop studies is the importance of human factors engineering (HFE) to the overall project outcome. The chief barrier to upgrades is not technological, as noted, and I&C replacement and upgrades are not simply an engineering problem. Rather, the systems being upgraded should result in improvements. Improvements may be marked by engineering metrics like reliability, but legacy systems have proved exceptionally reliable during their useful life. In fact, while digital I&C may improve reliability compared to its analog antecedents, its effective lifespan may prove considerably shorter. So, digital technology may introduce the need for more frequent and costly replacement to maintain its performance advantage.

The performance improvements from digital upgrades in the control room are likely to be found in terms of the reactor operators. Well-designed digital islands in the control room promise shorter training cycles, increased operator situational awareness, shorter response times in the face of plant upsets, and decreased human error. These goals cannot be achieved without careful consideration of the reactor operators. HFE serves as the bridge between technological solutions in the control room and the operators of those solutions. A poorly engineered solution is unlikely to yield significant operator performance improvements; in fact, it may actually introduce new error traps and decrease operational efficiencies.

The hallmark of HFE in the nuclear industry is compliance with applicable guidelines, particularly regulatory guidance such as NUREG-0711, *Human Factors Engineering Program Review Model* [7]. As noted in [8], NUREG-0711 is written primarily for use by the U.S. Nuclear Regulatory Commission (NRC) in reviewing the HFE activities undertaken by the licensee. As such, the level of explanation provided in the guideline is not specifically calibrated to the needs of industry; rather, it serves as a quality check on the HFE process that industry should follow to achieve successful HSIs in the control room. This seeming disconnect between the purpose of the guideline and industry need for additional process guidance should not be seen as a deficiency of NUREG-0711. Additional guidance is readily available in supplemental reports commissioned by the U.S. NRC. For example, NUREG/CR-6393, *Integrated System Validation*: *Methodology and Review Criteria* [9], outlines the details pertinent to one of the major HFE processes that the U.S. NRC expects industry to follow. Additional guidance is also available from the Electric Power Research Institute (EPRI), e.g., EPRI 3002002770, *Guidance for Developing a Human Factors Engineering Program for an Operating Nuclear Power Plant* [10], or even through HFE standards such as ISO 9241-210, *Human Centered Design for Interactive Systems* [11].

The problem is not the lack of relevant guidance on using HFE to support control room modernization. Instead, the problem is actually an overabundance of

HFE guidelines, methods, and processes from which to choose. This richness can in itself become another barrier to control room modernization by obfuscating the HFE process. Rather than becoming an enabling process, HFE risks becoming murky or overwhelming to the plant design engineer who is planning the upgrade. What is the best process for ensuring operator needs and wishes are met in the modernization process? What are the measures of operator performance in this process? What are the success criteria for HFE?

One goal of research in the HSSL is to demystify the HFE process and provide concise guidance to utilities and vendors to enable them to design and validate new control systems for the main control rooms of nuclear power plants. The present paper focuses on one element of the HFE process, namely verification and validation (V&V) of the new system. V&V is the evaluation of the HSI according to operators or HFE standards. As noted in this paper, there exists a fine line between sufficiency and completeness in V&V.

2 V&V of Control Rooms in the Nuclear Industry

V&V as applied in the nuclear power industry has tended to focus on final evaluation in the licensing applications to the U.S. NRC. In the context of licensing review, NUREG-0711 [7] explicitly states that V&V "is considered a test that final design requirements are met" (p. 74), although HSI tests and evaluation are recommended during the entire design process. Consequently, both research and practice of V&V in the nuclear industry have emphasized final or summative evaluation of the HSI design. This late-stage evaluation is called integrated system validation (ISV) and is comparable to a factory acceptance test (FAT), except the acceptance criteria center on operator performance while using the system in ISV versus software or hardware reliability in the FAT. The U.S. NRC prescribes four key V&V activities:

1. *Sampling of the Operating Conditions*: ensures that licensees identify the environment and potential situations that may arise during the actual operation of the plant, reflect system performance under those varying conditions, and examine significance of HSIs in those operating conditions. Effective sampling of operating conditions ensures that system safety inferred from subsequent V&V activities can generalize to the entire operating life of the plant.
2. *Design Verification*: ensures that licensees design HSIs to support operators for the full range of operating conditions (i.e., sampled scenarios). This includes analytical evaluation of the HSIs using task analysis. Effective design verification analytically identifies HSI deficiencies or Human Engineering Discrepancies (HED) that must be addressed prior to ISV or plant commissioning.
3. *Integrated System Validation*: ensures that licensees validate the system performance necessary for safe plant operations over a range of operating conditions.

ISV typically involves human-in-the-loop studies recruiting full-scope simulators and professional reactor operators to provide empirical, performance-based measurements. Effective ISV empirically identifies HSI deficiencies or HEDs that are missed in design verification and must be addressed prior to plant commissioning.

4. *Human Engineering Discrepancy Resolution*: ensures that licensees resolve any design HSI deficiencies identified in the V&V process. Effective HED resolution ensures that the HEDs are eliminated prior to plant or system commissioning, eliminating risks of unsafe operations.

This paper examines the limitations of adopting the conventional perspective that V&V should emphasize final evaluation, especially for the practitioners who must manage the engineering design and licensing process as well as the V&V process. The conventional perspective of V&V as ISV may be less suitable for the nuclear industry engaging in step-wise modernization projects that involve ongoing or gradual modifications to existing HSIs for plants in operation. Further, there are theoretical and practical limitations in obtaining best available empirical evidence for generalizing plant safety during operations from ISV results only.

3 Evaluation Theory Versus Practice

The majority of data collection and analysis methods in science are developed (initially) with considerations of neither the industrial purpose (and safety implications) of ISV nor inherent constraints of the nuclear domain. Scientific research methods for performance assessment originating in psychology and physiology rely on a large participant sample size to investigate many narrow research questions (cf., scenario types) and thereby produce knowledge or generate discussion for further testing and validation. In addition, methodological limitations often become impetus for further research. For instance, qualitative methods can focus on single participants to explore details and contexts with limited emphasis on generalization. Quantitative methods can focus on strict statistical or other criteria with large samples for validation. From this perspective, science is a continual process, readily accommodating *half-answers* to a research topic in anticipation of future studies.

The nuclear industry cannot accommodate *half-answers* to ISV or safety assessment. For instance, regulators cannot grant licenses on the basis of perfectly validated safety performance for only half the operating conditions. However, science classically produces research methods and study designs that produce conclusive or highly confident narrow findings (i.e., *half-answers*). This approach is impractical for ISV given the constraints of the nuclear industry. Consequently, both researchers and practitioners raise questions on classic research methods for ISV. Within the topic of ISV, the methodological discussions range from meaningfulness of inferential statistics and relevance of qualitative measurements to requirements for follow-up evaluation studies for validation of the main control room. The nuclear industry must address the applied research issues of ISV, since the direct application of classic research methods may not practically provide the necessary evidence and confidence in the assessment of plant safety.

It is difficult to acquire sufficient evidence and reasonable confidence in the results of V&V for plant safety assessment. Confidence in the late-stage V&V process, analysis, and results has major implications in the engineering design process as well as regulatory licensing decisions. Interestingly, confidence in performance testing connects closely with the longstanding basic research on test validity and validity generalization. Validity research provides a new perspective to revisit the purpose of V&V in the nuclear industry that may simplify the discussion and provide directions for future research.

V&V is concerned with predictive validity. Thus, broadly speaking, all qualitative or quantitative data are collected and analyzed to make (or to become confident in making) inferences on future performance. Though the traditional assessment criteria may be impractical, many principles in inferential statistics (e.g., avoiding Type I errors—false positives—and Type II errors—false negatives) remain essential for establishing confidence in V&V performance conclusions. More importantly, this perspective encourages all evidence that could support the prediction to be admitted for plant safety assessment, even though individual pieces of evidence carry different merits that must be carefully weighted for achieving valid conclusions on the integrated operational performance. This approach of emphasizing predictive evidence stands in contrast to treating the V&V activity and results solely based on testing the final design for a highly defined set of human performance metrics. Instead, V&V should steer toward establishing confidence in plant safety over the licensing period. This approach suggests the value of many types of evidence to establish the trajectory rather than a single all-inclusive snapshot of performance through ISV.

4 Consequential Validity

The paucity of early design evaluation and the limitations of final performance testing motivate an investigation into new approaches and methods for V&V in the U.S. nuclear industry, especially for those plants undergoing modernization. In particular, it is important that any new approach to V&V must encapsulate the perspective of *consequential validity*.

Consequential validity [12] emphasizes the implications of the decision made as a result of the outcome of the evaluation method (i.e., V&V). For the practitioners of the nuclear industry (e.g., vendors, utilities, and regulators) who focus on outcome of a specific instance of V&V, consequential validity may be viewed as predictive validity—validity of the empirical evidence for predicting safety for the licensing period of the specific plant. For V&V researchers developing generalizable evaluation methods relevant to multiple modernization and construction projects, consequential validity extends beyond predictive validity in that the developed V&V method can have major decision implications on licensing, plant operations, and ultimately public safety. The focus on consequential validity ensures an emphasis on evidence predictive of safe plant operations over evidence centered on performance testing a high-fidelity representation. This perspective reconciles the need for early evaluation and importance of final integrated testing by weighing the relative merits of evaluation at different stages and the quality for predicting operational safety in the future. That is, empirical evidence at early or formative evaluation likely qualifies less than final or summative evaluation for predicting safety. Nevertheless, evidence from formative evaluation may be sufficient to predict many aspects of safety or system performance, including establishing the trend toward improved performance.

Thus, this paper presents a safety-predictive approach to V&V that accommodates evidences at different stages of evaluation. This emphasis on safety prediction should improve the licensing process for the utilities and the confidence in the licensing decisions for the regulators. ISV has tended to be treated as a sufficient measurement for V&V, but evidence gathered over the design lifecycle may actually be more complete and better targeted in predicting operational safety—the consequences of deciding on the adequacy of the control room design.

To support the nuclear industry in adopting a safety-predictive approach that leaves behind the notion of V&V concerning only the final design, four methodological areas require substantial development:

1. Structuring the evidence, especially the information gathered outside of the traditional V&V process.
2. Identifying the appropriate evidence to gather with respect to different stages of design and evaluation.
3. Assessing the merits of various evidence for predicting plant safety.
4. Integrating all of the evidence to provide a final safety assessment of integrated operations in the main control room.

Clear and effective methods in these four areas can ensure that the process of gathering, assessing, and presenting evidence would lead to products that could satisfy regulatory concerns on public safety and meet objectives of complete and targeted measurements in V&V.

5 The Safety Case

The *safety case* is a meaningful starting point for structuring evidence to encapsulate the concept of consequential validity. According to Kelly [13] (p. 22), "A safety case should communicate a clear, comprehensive, and defensible argument that a system is acceptably safe to operate in a particular context." The three elements of a safety case as described by Kelly are:

- *Requirements*—the safety objectives for the system.
- *Argument*—the mapping of the evidence to the requirements.
- *Evidence*—supporting safety documentation such as risk assessments.

The exact nature of the evidence and the way in which arguments are most effectively conveyed is the subject of ongoing discussion [14]. Nonetheless, they remain the cornerstone of safety regulation, for example, in the United Kingdom (UK) defense [15] and nuclear power sectors [16]. Rather than rely on a single source of evidence for the safety of a system, safety case regulation requires a body of evidence that clearly argues for meeting safety requirements. This approach is not unlike the judicial trial system, where evidence must be argued to influence the verdict. The verdict, in this case, is the safety of the system.

Kelly [17] (p. 31) notes, "Safety cases have little hope of adding value if they are impotent in their influence on the design and operation of the system in question. Safety cases shouldn't be produced after the system design has been finalized." Thus, the argument can be made that safety cases should not rely on late-stage evidence but rather on evidence derived early in the design stage and actually used to shape the design of the system. This approach seems to contradict the late-stage emphasis of ISV. That is not to say that ISV is an unimportant piece of evidence in the case for safety; however, final stage evidence is rarely complete and shouldn't be the only evidence used. The evidence that best predicts safety of operations at any stage of V&V should be considered for licensing.

Figure 1 provides one example of evidence in the form of preliminary usability evaluations during the design phase. Recall that NUREG-0711 emphasizes the primary form of evaluation as a standalone ISV in the V&V phase. In the figure, currently being used to support control room modernization activities at a fleet of U.S. nuclear plants, there are actually three rounds of assessment that occur prior to the ISV. Each round consists of operator-in-the-loop studies and expert evaluation by HFE professionals. The three phases correspond to the 30, 70, and 100 % system completion milestones, resulting in increased fidelities of the system being tested. However, each phase of evaluation serves as input for the next design and development activities for the system. As such, the system undergoes an iterative design-evaluation cycle leading up to the completion of the system.

The process outlined in Fig. 1 illustrates a systematic design process by the licensee and vendor, one that takes operator and HFE input at several junctures

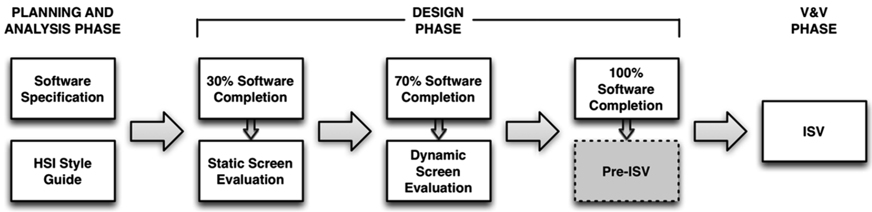

Fig. 1 Design phase evaluation (from [5])

during the design and uses them to refine and optimize the design. This process might even be said to exemplify a user-centered design best practice. Yet, these iterative pieces of evidence do not have a clear placeholder in the NUREG-0711 framework. They are supplemental evaluations leading up to the ISV, which is the truly meaningful evaluation in the common interpretation of NUREG-0711.

What if these design-phase evaluations were not just supplemental steps in the process? What if they were framed as a type of safety case—evidences that build the argument for meeting safety objectives for the design of the new system? Surely there is value in the design-phase evaluations beyond guaranteeing the system will pass the ISV. The design-phase evaluations are more than a dry-run for the final evaluation, because they are actually shaping the design of the system. These evaluations become evidence for the veracity of design decisions and the quality of the process of the design. In other words, they provide data to show why one design decision was selected over another but also build confidence that the final design represented the convergence of a vetted process. By using iterative data, it becomes clear that passing the ISV is not a matter of luck; it follows a traceable path since the design inception. Unfortunately, when the regulatory process emphasizes evidence coming from the ISV, there is no clear guidance for the licensee and vendor to build a complete safety case with evidence across the design lifecycle.

If the case has been made for the value of early-stage evaluation, is there still a need for late-stage ISV? There are several key differences that help delineate early-versus late-stage evidence:

- *Scenarios*—the situations against which the system is tested.
- *Participants*—the representative sample of operators who will interact with the system.
- *Measures*—the reflection of operator performance and preferences using the system.

Table 1 summarizes these considerations. Essentially, the distinction between early- and late-stage evaluation can be understood in terms of completeness and conclusiveness. Late-stage evaluation, as a type of final evaluation before licensing, will seek to have comprehensive scenarios against which the system is tested, a representative and statistically valid sample of operators, and definitive performance measures. Early-stage evaluation may feature a subset of these tailored toward gathering sufficient data for meeting design objectives. The goal of early-stage evaluation is design input, while the goal of late-stage evaluation is sufficient assessment to meet certification for regulatory safety requirements. The fact that early-stage evaluation may not be as rigorous as late-stage evaluation does not diminish its value as evidence toward completing a safe final design. There emerges a paradox: Late-stage evaluation seeks to be as complete as possible, while early-stage evaluation may avoid such goals. Yet, successive early-stage evaluations may actually paint a more complete picture of system use than is possible with late-stage evaluation.

Table 1 Considerations of early- and late-stage evaluations

		Evaluation stage	
		Early stage	Late stage
Evaluation considerations	Scenarios	Generally limited to test the functionality and operator interaction with specific aspects or features of the system	Comprehensive across the range the system will encounter relative to real world situations, including safety-critical situations
	Participants	Limited number of operators and process control experts needed to test the evolving system design and provide feedback to the design team	Ideally, a large enough sample size to be statistically significant, covering a range of operators (e.g., different experience levels) representative of the operator population
	Measures	Suitable for design decisions, akin to discount usability, with consideration of subjective preference data to drive the design	Suitable for safety compliance decisions with emphasis on objective measures of performance

These considerations apply only to operator-in-the-loop studies. The need for clear appreciation of different types of evidence becomes more importunate when the evidence is not strictly numeric. Other forms of evaluation may not produce quantitative assessments. The safety case argument is crucial for incorporating qualitative evidence that may result from the V&V [18].

Despite its prevalence in the European regulatory community and others, the safety case has not been widely adopted in the U.S. Recently, there has been consideration of safety cases to help regulate the U.S. chemical industry by the U.S. Chemical Safety Board [19], and the U.S. Food and Drug Administration has introduced safety assurance cases (which are in most cases synonymous with safety cases) to minimize risk in the use of medical devices [20]. These industries are considering safety cases as part of a risk regulatory framework. There has not been a parallel adoption of safety cases by the nuclear industry or regulator.

It is important to note that although NUREG-0711 emphasizes ISV, it does not prohibit other stages of evaluation. Licensees and vendors will need to provide examples of such evaluations in order for the regulator to determine their effectiveness. Such examples must not be fragments that confuse the merits of the safety case. They should proceed in a systematic manner that builds an effective argument. Although the U.S. nuclear regulatory framework does not currently require safety cases, this framework certainly does not discount the value of well-argued safety evidence.

6 Conclusion

Every system that is designed for use by operators should be validated and verified to ensure it meets safety and usability objectives. Despite the importance of V&V, it is an activity that is often relegated to the late stages of the design process in safety-critical domains like control room modernization. In this paper, we've made the case for using V&V across the design lifecycle, especially at early stages when V&V can positively shape the design of the system. The advantages of early-stage evaluation include not only the ability to improve the design but also to ensure operator buy-in and to avoid potential reworks of the system that might be necessary when issues are first discovered late in the design and development process.

Despite these advantages, the true value of early-stage V&V should also be understood in terms of building the case for the safety of the system. Evidence for the safety of the system should not be limited just to ISV. Adoption of iterative design and evaluation demonstrates a solid HFE process and should serve to establish confidence by the utility and the regulator that error traps have been eliminated as the design has matured. Early-stage V&V coupled with late-stage ISV forms a complete—not just a sufficient—picture of the safety of the system. Although there is no requirement for V&V outside ISV in NUREG-0711, the process outlined in this paper fully supports regulatory goals of the new design.

We close this paper with two final considerations about V&V as it pertains to control room modernization. First, distributed control systems used for control room modernization are much less static than their analog predecessors. They can be fine-tuned over time for better performance, better display presentation, or better alarm management. These digital HSIs run on standard operating systems and will need continuous maintenance. Ongoing changes to the digital architecture and feature set suggest the need for additional V&V, even after the ISV is completed. The scope of gradual HSI changes to control systems may not warrant a large-scale design activity that steps through all stages of NUREG-0711. The use of early-stage V&V methods may translate into a sustainable approach for ensuring safety of digital systems as they are gradually and necessarily upgraded. The model of a large evaluation for a large change in the control room may not hold when the changes become more nuanced. Without a suite of methods to assess these changes quickly and cost effectively, there is the risk that small changes may not take advantage of V&V. Small-scale evaluations, not ISV, may be the key to ongoing V&V that mirrors the natural evolution of digital HSIs.

Second, it should be remembered that V&V is a confirmatory approach. It is only intended to show that operators can use the system for prescribed conditions. As such, HFE, operations, and engineering need to be diligent in casting a wide net in selecting scenarios. Still, it is never possible to anticipate all possible scenarios. The role of ISV is to test the integration of the tested new system against the other systems with which it interacts. In this manner, system dependencies and common cause failures can be identified. Early-stage V&V presents a different type of confirmation. Early-stage evaluation tends to be more informal and open-ended, exploring operators' first interactions with the system across unscripted activities. In many cases, early-stage V&V precedes operating procedures, thereby necessitating a degree of discovery by the operators. This discovery may actually be seen as a type of stress test of the system as operators familiarize themselves with the system interface and its strengths and weaknesses. The opportunity to gather performance data on first and unconventional use scenarios can actually instill confidence in the robustness of the system. Insights from early-stage evaluation are crucial in establishing a pattern of interaction that can be extrapolated to novel and even unanticipated scenarios. The safety of the system is not just proved through carefully considered scenarios; it is ultimately demonstrated through the system's resilience across diverse uses including those that are unforeseen. Early-stage V&V represents an ideal test case for the system outside normal operations.

Control room modernization has only begun to realize the benefits of V&V. V&V is often considered within a narrowly defined function to support ISV. Expanding the application of V&V promises to create an integrated design process that can become the backbone of plant safety assurance. V&V should become a continuous process as plants modernize, providing graded levels of evaluation suitable to support both small and large upgrades at various stages of design and deployment. V&V is complete when it covers all stages of design. The safety case approach can help ensure that the measures of V&V are not simply or just sufficient.

Acknowledgments and Disclaimer The authors wish to thank Dr. Roger Lew of the University of Idaho for his insightful review and suggestions for an earlier version of this paper. This work of authorship was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights. Idaho National Laboratory is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517.

References

1. Hollifield, B., Oliver, D., Nimmo, I., Habibi, E.: The High Performance HMI Handbook: A Comprehensive Guide to Designing, Implementing and Maintaining Effective HMIs for Industrial Plant Operations. PAS, Houston (2008)
2. Strobhar, D.A.: Human Factors in Process Plant Operation. Momentum Press, New York City (2013)
3. Ulrich, T., Boring, R., Phoenix, W., DeHority, E., Whiting, T., Morrell, J., Backstrom, R.: Applying Human Factors Evaluation and Design Guidance to a Nuclear Power Plant Digital Control System, INL/EXT-12-26797, Idaho National Laboratory (2012)
4. Hugo, J., Boring, R., Hanes, L., Thomas, K.: A Reference Plan for Control Room Modernization: Planning and Analysis Phase, INL/EXT-13-30109, Idaho National Laboratory (2013)

5. Boring, R., Joe, J., Ulrich, T., Lew, R.: Operator Performance Metrics for Control Room Modernization: A Practical Guide for Early Design Evaluation, INL/EXT-14-31511, Rev. 1, Idaho National Laboratory (2015)
6. Boring, R., Agarwal, V., Fitzgerald, K., Hugo, J., Hallbert, B.: Digital Full-Scope Simulation of a Conventional Nuclear Power Plant Control Room, Phase 2: Installation of a Reconfigurable Simulator to Support Nuclear Plant Sustainability, INL/EXT-13-28432, Idaho National Laboratory (2013)
7. O'Hara, J.M., Higgins, J.C., Fleger, S.A., Pieringer, P.A.: Human Factors Engineering Program Review Model, NUREG-0711, Rev. 3, U.S. Nuclear Regulatory Commission (2012)
8. Boring, R.L., Joe, J.C.: Baseline Evaluations to Support Control Room Modernization at Nuclear Power Plants. In: 9th International Topical Meeting on Nuclear Power Plant Instrumentation, Control, and Human-Machine Interface Technologies (NPIC&HMIT), pp. 2155–2165 (2015)
9. O'Hara, J.M., Stubler, W., Higgins, J., Brown, W.: Integrated System Validation: Methodology and Review Criteria, NUREG/CR-6393, U.S. Nuclear Regulatory Commission (1995)
10. Electric Power Research Institute: Human Factors Guidance for Control Room and Digital Human-System Interface Design and Modification: Guidelines for Planning, Specification, Design, Licensing, Implementation, Training, Operation, and Maintenance, EPRI TR-1010042 (2005)
11. International Standards Organization: Ergonomics of Human-System Interaction-Part 210: Human Centered Design for Interactive (2010)
12. Messick, S.: Standards of validity and the validity of standards in performance assessment. Educ. Meas. Issues Pract. **14**(4), 5–8 (1995)
13. Kelly, T.P.: Arguing safety—a systematic approach to managing safety cases. Ph.D. Dissertation, University of York Department of Computer Science (1998)
14. Haddon-Cave, C.: The Nimrod Review: An Independent Review into the Broader Issues Surrounding the Loss of the RAF Nimrod MR2 Aircraft XV230 in Afghanistan in 2006. Office of Public Sector Information, London (2009)
15. UK Ministry of Defence: Safety Management Requirements for Defence Systems, Part 1: Requirements and Guidance, Defence Standard 00-56 Part 1, Issue 6, UK Ministry of Defence (2015)
16. UK Office for Nuclear Regulation: The Purpose, Scope, and Content of Safety Cases, NS-TAST-GD-051 Rev. 3, UK Office for Nuclear Regulation, Health and Safety Executive (2013)
17. Kelly, T.: Are 'Safety Cases' working? Saf. Crit. Syst. Club Newsl. **17**(2), 31–33 (2008)
18. Boring, R.L., Ulrich, T.A., Joe, J.C., Lew, R.T.: Guideline for Operational Nuclear Usability and Knowledge Elicitations (GONUKE). Procedia Manuf. **3**, 1327–1334 (2015)
19. Hopkins, A.: The Cost-Benefit Hurdle for Safety Case Regulation, U.S. Chemical Safety Board (2013)
20. U.S. Food and Drug Administration: Infusion Pumps Total Product Life Cycle: Guidance for Industry and FDA Staff, OMB 0910-0766, Center for Devices and Radiological Health, U.S. Department of Health and Human Services, U.S. Food and Drug Administration (2010)

Part III
Human Factors in Energy

Consumer Electric Energy Management Strategies and Preferences in Emergency Demand Response: Results from a Survey

Huiyang Li, Haya Salah and Ziang Zhang

Abstract Demand response has not been widely accepted by the residential consumer, in part due to the lack of understanding of how average residential energy consumers will behave under different scenarios. The objective of this study is to examine consumer energy management strategies in emergency demand response using a survey. Participants were given a scenario where they were the owner of a single-family house and had participated in an emergency demand response program, and answered questions related to electricity usage. Results showed that between 27 and 39 % of the participants were willing to turn off the AC in the emergency demand response. More than 86 % of the participants were willing to change the AC setting to some extent. The higher the bonus was, the more participants were willing to do so. Willingness of postponing the use of other appliances highly depends on the category of the appliances.

Keywords Energy management · Demand response · Smart grid · Consumer behavior

The original version of this chapter was revised: The last name of the chapter author salah was corrected. The erratum to this chapter is available at DOI 10.1007/978-3-319-41950-3_18

H. Li (✉) · H. Salah · Z. Zhang
State University of New York at Binghamton, Binghamton, NY, USA
e-mail: hli@binghamton.edu

H. Salah
e-mail: hsalah1@binghamton.edu

Z. Zhang
e-mail: zhangzia@binghamton.edu

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_8

1 Introduction

Power systems are becoming more complex in part due to the use of smart infrastructures, such as the Advanced Metering Infrastructure (AMI). As AMI data becomes more prevalent, the utility will be capable of using pricing mechanism to improve the efficiency of the power grid by reducing the peak load. The demand response technology has been around since the 1980s [1] and has been extensively studied in the past several years [2]. Existing researches conducted by researchers from the power systems community has been focused on using optimization theory and control theory to develop the optimal operation schedule [3, 4]. There are many variations of demand response program, and it has been widely applied to factories, commercial buildings and university campuses [5] (NYISO 2014).

Demand response has not been widely accepted by the residential consumer due to two barriers: (1) how to effectively present the real-time energy usage as well as the payback of energy conservation behavior to a large-scale consumer group, (2) lack of understanding of how average residential energy consumers will behave under different scenarios. The current development on the Internet-of-Things (IoT) [6] (Wu et al. 2011) and the Advanced Metering Infrastructure (AMI) [7] (Sui et al. 2011) from the power grid modernization movement can address some of the challenges of the first barrier.

To date, few empirical studies have investigated the energy management strategies used by residential end-users. The focus of this study is to address the second barrier by empirical data analysis. The objective of this survey is to reveal strategies and preferences that residential users hold in energy management. The lessons learned from this study will be used to guide us to design a more comprehensive platform to address both barriers.

2 Methods

2.1 Participants

The study is based on survey data collected from residents in the state of New York. Participants were recruited from students enrolled in an undergraduate human factors course and their family and/or friends. The eligibility requirement is that the participants must be responsible for paying a separate electricity bill at the place they reside. With this criterion, students who live in on-campus residence halls and some off-campus apartments are not eligible for participation, because the utility expense was included in the rent. Each student was asked to recruit three qualified participants as part of a course assignment. Students and their family were all residents of New York State.

2.2 Questionnaire

The first set of questions was based on a scenario where the participant was the owner of a single-family house and had participated in a voluntary emergency demand response program. In an emergency situation on a hot summer evening, the utility requested the house owners to reduce electricity consumption between 6−8 p.m., a time when system-level peaks normally occur in the Northeast.

The participants were asked how much change they were willing to make to the temperature setting of their air conditioner: turn off, turn up by 3 °F, turn up by 2 °F, turn up by 1 °F, or keep the same temperature. They were asked to answer questions under three conditions: receiving $3 bonus per kWh electricity saved, $2/kWh, and $4/kWh. This bonus was in addition to the savings from using less electricity. The estimated temperature change in the house and the estimated bonus that the participant may receive were given for each condition. For each condition, two questions were asked: (1) which option would they most likely to choose to do? (2) which option is the maximum discomfort they can tolerate? Table 1 shows the option list for Case 1 ($3/kWh) in the paper version of the questionnaire. The explanation of the temperature calculation is in Appendix 1.

The participants were also asked about whether they were willing to postpone the use of other appliances, such as washer, dryer, dishwasher, coffee maker, computer, microwave oven, and range (stove and oven). These appliances were selected because the estimated bonus upon postponing the use was more than $1 between 6 and 8 pm.

Demographic data of the participants as well as their self-reported daily electricity management strategies were also collected. Questions included age, gender, monthly electricity bill, monthly household income, percentage of electricity bill in overall income, type and size of the property they normally reside, and temperature settings of AC in their property in the summer.

2.3 Survey Administration

Two versions of the survey were available to the participants: an online Google form and a paper-based questionnaire. The paper-based questionnaire was provided as a .pdf file and can be printed out by the student or the participant.

Table 1 Options for Case 1 ($3 bonus per kWh)

	AC setting change	Estimated bonus	Estimated temperature change
A	Turn off AC	$12	Up by 4° in 2 h
B	Turn up by 3°	$9	Up by 3° in the first 1.5 h, then remain the same
C	Turn up by 2°	$6	Up by 2° in the first hour, then remain the same
D	Turn up by 1°	$3	Up by 1° in the first 0.5 h, then remain the same
E	Keep the same	$0	Same temperature as you usually set

3 Results

3.1 Respondents

Two hundred and forty-four responses were received. Table 2 shows the characteristics of the participants.

Table 2 Participants characteristics

Gender, % Female	39.3
Age (%)	
18–24	23.4
25–34	10.2
35–44	13.1
45–54	25.8
55–64	17.2
>65	2.5
Monthly electricity bill ($)	
<20	4.9
20–39	7.0
40–59	15.2
60–79	13.9
80–99	18.9
100–120	20.5
Other	12.7
Monthly household income (%)	
<2000	20.5
2–4 k	10.7
4.1–6 k	15.6
6.1–8 k	16.0
8.1–10 k	12.3
10.1–12.1 k	12.7
Other	5.7
Percentage of electricity bill in overall income	
Type of property (%)	
Single-family house	61.1
Town house	3.7
Condo/apartment	23.0
Mobile house	0
Others	2.5
AC temperature setting in summer, Mean (SD)	70.06 (4.36) range 50–80° 11 responses were N/A

Table 3 Preferences for AC setting under three bonus conditions

	Case 1			Case 2			Case 3		
	Bonus	Q1 (%)	Q2 (%)	Bonus	Q1 (%)	Q2 (%)	Bonus	Q1 (%)	Q2 (%)
A. Turn off AC	$12	36.1	54.5	$8	27.5	49.6	$16	38.9	54.1
B. Turn up by 3 °F	$9	24.6	20.9	$6	25.8	24.4	$12	26.2	24.2
C. Turn up by 2 °F	$6	21.3	12.7	$4	25.8	13.1	$8	14.8	10.2
D. Turn up by 1 °F	$3	6.6	5.7	$2	7.0	4.9	$4	10.2	4.9
E. Keep the same	$0	11.5	5.7	$0	13.9	8.2	$0	9.8	6.6

3.2 Preferences for AC Setting

Descriptive statistics was conducted to analyze the preferences for AC settings in emergency response scenarios. Table 3 shows the percentage of participants who chose each of the five options when answering the two questions for each condition. Question 1 was about the option the participant most likely to choose, while question 2 was about the maximum tolerance.

3.3 Preferences for Postponing the Use of Appliances

Table 4 shows the descriptive statistics on participants' preference for postponing the use of appliances.

In addition, correlation analysis was used to explore the relationship between the amount of total bonus and willingness to postpone. Chi-square analysis was used to examine the relationship between AC setting preference, i.e. participant's answer

Table 4 Preferences for postponing the use of appliances

Appliance	Bonus ($)	Purpose of use	% Yes	% No
Washer	1.5	Clothes	75.0	25.0
Dryer	9	Clothes	78.7	21.3
Coffee maker	1.5	Food/drink	65.2	34.8
Microwave	1.8	Food/drink	59.0	41.0
Range (stove and oven)	3.9	Food/drink	55.7	43
Dishwasher	3.6	Cleaning	77.5	22.5
Desktop computer	3	Business or entertainment	48	52

(A-E) to Case 1 Question 1 and their preferences for postponing the use of appliances. Results showed that there was no correlation between expected bonus and the willingness of postponing the usage of the appliances. No significant difference was found on the answers to Case 1 Question 1 and any of other appliances.

4 Discussion and Conclusions

The use of Advanced Metering Infrastructure allows utility companies to apply demand response programs using pricing mechanisms. Demand response, however, has not been widely accepted by the residential consumer, in part due to the lack of understanding of how average residential energy consumers will behave under different scenarios. The objective of this study is to examine consumer electricity energy management strategies in emergency demand response using a survey. Participants were given a scenario where they were the owner of a single-family house and had participated in a voluntary demand response program. They then answered questions related to electricity usage, primarily about setting the AC temperature and postponing the use of other appliances, under different bonus conditions.

Results showed that between 27 and 39 % of the participants were willing to turn off the AC in the emergency demand response. The higher the bonus was, the more participants were willing to do so. The percentages of participants who were not willing to do anything are surprisingly low, ranging from about 10 % to about 14 %. Consistent with the trend for choosing to turn off the AC completely, the higher the incentive, the fewer people chose to do nothing. Responses to the question “the maximum temperature change one can tolerate” were similar across all levels of bonus. About 50–55 % of the participants indicated that they could tolerate the temperature change while turning off the AC completely, i.e. the temperature rising 1 °F every half an hour. The results imply that pricing policies may be effective in achieving goals of reducing consumer energy consumption in emergency demand response scenarios. Higher incentives may lead to higher rate of compliance. In extreme cases, force shut-down of cooling system may be an option.

Willingness of postponing the use of other appliances highly depends on the category of the appliances. More than 75 % participants were willing to postpone the use of clothes washer, clothes dryer, and dishwasher, appliances for clothing and household cleaning. However, less than 60 % of the participants were willing to postpone the use of microwave oven and stove, appliances for preparing food. Only 50 % of the participants were willing to postpone using a desktop computer, which can be used for either business purposes or entertainment purposes. These results suggest decision support tools for house owners could manipulate energy consumption and savings by rescheduling less urgent tasks. Suggestions on postponing time-sensitive tasks such as cooking may not be adopted by the user.

The results from this survey can be used in the design of pricing policies as well as the design of energy management support tools for house owners. In the future, we will expand the survey to increase the sample size and the participants' representation of geographic areas. We will also conduct more detailed analyses to examine the factors that influence consumer electricity energy management strategies, such as social economic status, education, location, temperature, and perception of coldness/warmth. Future research will also focus on the use of high-fidelity simulated environment, such as a test room/house, to observe consumer manage electricity consumption when they can actually feel the consequences of the strategies they applied. The results will then be used in the development of a Human-and-Hardware-In-the-Loop Smart Energy Testbed (H^2IL-SET) for the simulation of future smart grids.

Appendix: Energy Consumption and Temperature Change Calculation

The following assumptions have been made to simplify the problem:

- Without AC, indoor temperature will raise 2° per hour.
- The AC needs 2 kWh of energy per hour to maintain the indoor temperature.

Here is a numerical example, assume compensation is \$3 per kWh, initial indoor temperature is 75 °F.

- If the AC is off during 6–8 PM:
  - The AC is off for 2 h. You will save 4 kWh of energy, compare with your regular energy usage
  - You will receive 4 × 3 = \$12 of compensation, in addition to the regular energy charge on your utility bill.
  - Your room temperature will be 77 °F at 7 PM, and 79 °F and 8 PM
- If the AC is set on 76 °F, (x + 1):
  - The AC is off for 0.5 h, you will save 1 kWh of energy, compare with your regular energy usage
  - You will receive 1 × 3 = \$3 of compensation, in addition to the regular energy charge on your utility bill.
  - Your room temperature will be 76 °F at 7 PM, and 76 °F and 8 PM
- …
- If the AC is set on 75 °F, (do nothing):
  - The AC is always on, no energy consumption reduction.
  - Your room temperature is always 75 °F.

References

1. Schweppe, F.C., Tabors, R.D., Kirtley, J.L., Outhred, H.R., Pickel, F.H., Cox, A.J.: Homeostatic utility control. IEEE Trans. Power Apparatus Syst. PAS-99 (3), 1151–1163 (1980). doi:10.1109/TPAS.1980.319745
2. Palensky, P., Dietrich, D.: Demand side management: demand response, intelligent energy systems, and smart loads. IEEE Trans. Ind. Inform. **7**(3), 381–388 (2011). doi:10.1109/TII.2011.2158841
3. Li, N., Chen, L., Low, S.H.: Optimal demand response based on utility maximization in power networks. In: Power and Energy Society General Meeting, pp. 1–8 (2011). doi:10.1109/PES.2011.6039082
4. Albadi, M.H, El-Saadany E.F.: A summary of demand response in electricity markets. Electr. Power Syst. Res. **78**(11), 1989–1996 (2008). doi:http://dx.doi.org/10.1016/j.epsr.2008.04.002
5. NYISO. 2014. Power Trends 2014. http://www.nyiso.com/public/media_room/publications_presentations/index.jsp
6. Wu, G., Talwar, S., Johnsson, K., Himayat, N., Johnson, K.D.: M2M: from mobile to embedded internet. IEEE Commun. Mag. (2011) doi:10.1109/MCOM.2011.5741144
7. Sui, H., Sun, Y., Lee, W.J.: A demand side management model based on advanced metering infrastructure. In: 2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT). (2011) doi:10.1109/DRPT.2011.5994150

Exploring Observed Cognitive Error Types in Teams Working in Simulated Drilling Environments

Margaret T. Crichton, Scott Moffat and Lauren M. Crichton

Abstract Errors made by drilling teams were observed during simulator-based exercises that formed part of a well control training course. Each course lasted four days and comprises both non-technical and technical theory with five simulator-based exercises. The exercises are observed for key team non-technical skills (NTS). In feedback sessions, the observers debriefed the team members about their areas of effective NTS performance and also where improvements could be made. This paper will specifically focus on the errors (i.e. performance that was classified as either 'marginally below' or 'well below' expectations) made by the teams during 105 observed exercises. An understanding of such errors will allow future training programs to focus on areas for improvement and designing training that transfers into the real rig-site.

Keywords Training · Non-technical skills · Team performance

1 Introduction

The environment in which drilling teams work can be characterized as dynamic, uncertain, high-risk, and involves multiple players in often geographically distributed locations. Such conditions therefore require the team to perform safely and effectively to achieve their objectives. Other high hazard industries, such as aviation, healthcare, and maritime, where teams must function in similar circumstances to drilling teams, have increasingly acknowledged the importance of non-technical skills (NTS) [1] and have taken action to investigate, identify, train and assess NTS.

Behavioural markers systems include NTS categories and behaviours, and provide a framework that can be used to observe non-technical behaviours that contribute to superior or substandard performance within a work environment [2]. Behavioural markers are typically identified specifically for the role or context under examination. For example, behavioural marker systems have been developed

M.T. Crichton (✉) · S. Moffat · L.M. Crichton
People Factor Consultants Ltd., Aberdeen, UK

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_9

in aviation, such as the NOTECHS system for air crew [3], and in medicine, such as ANTS for anesthetists [4] and NOTSS for surgeons [5]. Behavioural marker systems can provide a basis for designing a curriculum for training NTS, for assessment, and for providing feedback.

In the drilling context, NTS have been receiving relatively recent attention. Following a number of accidents and incidents in the oil and gas industry over the past decade, for example the Deep Water Horizon tragedy in 2010 [6], there is growing awareness of the impact of NTS on safety and performance. The Energy Institute [7] and also the International Association of Oil and Gas Producers (IOGP) [8] have published guidance on the implementation of Crew Resource Management (CRM) with the aim of providing an overview and learning objectives for non-technical skills training delivery and assessment.

This paper presents the results of an investigation into the types of errors commonly made by drilling teams during simulator-based exercises. By identifying these errors, behavioural marker systems can be developed for team performance and can support training interventions, such as CRM, to specifically address the causes of such errors in the attempt to reduce the potential for these errors to occur.

2 Team Behavioural Markers

A drilling team comprises members who have different experiences and expertise, who may work for different organisations (operators, contractors, service companies) who may have different agendas, goals, and objectives, and may only join the actual team for short periods of time dependent on the requirements of the plan of work. Figure 1 illustrates the interactions between various members of a drilling team, with the role of the driller at its center.

In order to perform safely and effectively, a drilling team therefore depends on the competencies of individuals as well as the team's [9]. Team competencies refer to the knowledge, skills and attitudes (KSAs) of the team as a whole, that is: what team members think; what team members do, and what team members feel [10]. These KSAs form the basis of the team non-technical skills.

As well as reinforcing the NTS of individual team members through training interventions such as CRM, other high hazard industries, particularly aviation and healthcare, have recognized that the behaviours of the team strongly influence outcomes and therefore safety. Training interventions, especially simulator-based team training, have been developed to stimulate team behaviours in inter-disciplinary teams, which can then be debriefed to improve safe and efficient performance [11]. Team members are encouraged to reflect on their own behaviours as well as those of the team [12], in particular, team non-technical skills such as team communication, co-operation and co-ordination, situational awareness, and decision making.

Moffat and Crichton [13] described a pilot of a project undertaken to identify the non-technical skills required by drilling teams. Simulator-based team training

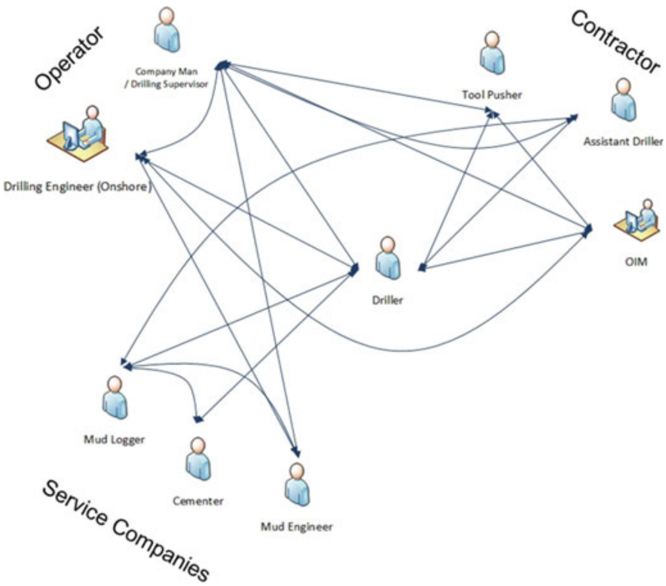

Fig. 1 Interactions between members of a drilling team

sessions were observed and the recorded observations used for feedback purposes to highlight examples of expected performance as well as areas for improvement.

A preliminary taxonomy of team behavioural markers was then developed as shown in Table 1 along with the categories and definitions of the observation rating criteria.

A four-point observation rating scale was created ranging from exceeded expectations to well below expectations, and examples of the behaviours were included in the team behavioural marker system. The observation rating scale

Table 1 Matrix of team behavioural markers and rating descriptions

Team behavioural marker	Exceeded expectations	Met expectations	Marginally below expectations	Well below expectations
Situation awareness	*Observed behaviour* was of a consistently high standard enhancing safety and could be used as a positive example for others	*Observed behaviour* was of a satisfactory standard but could still be improved	*Observed behaviour* indicated some cause for concern and improvements are needed	*Observed behaviour* considerable improvements required
Teamwork decision making				
Teamwork and communication				
Team workload and stress management				

allowed observers to note extremely positive behaviours as well as those behaviours which, although might not be defined as potentially causing an incident, raised concerns in terms of safety. For the purposes of improving performance it is essential to specify the types of errors made in drilling teams to raise awareness of their causes and effects.

3 Typology of Errors Made by Drilling Teams

3.1 Identifying Error Types

Human error is defined by Reason [14] as being all occasions in which a planned sequence of mental or physical activities fails to achieve its intended outcome, and when these failures cannot be attributed to the intervention of some chance agency. He proposed a description of error classifications including slips and lapses, mistakes, and violations. A number of different systems exist, as discussed by Stanton and colleagues, to identify types of human errors and their causes [15]. These include techniques such as Cognitive Reliability and Error Analysis Method (CREAM), Human Error Identification in Systems Tool (HEIST), or Human Error Assessment (THEA). However, these techniques can often be context-specific, rely on probabilities, or be more focused on the predictive aspects of error and reliability.

It is valuable to look beyond the label “human error” to identify more precisely what the error was and why it occurred [16], particularly related to errors in cognition which then affect behaviours. Cognitive system factors, according to Woods et al. [16], refers to bringing knowledge to bear, changing mindset as situations and priorities change, and managing goal conflicts.

In order to identify examples of ineffective performance, or errors, affecting drilling team interactions, a more explanatory error identification typology was developed. The aim of this drilling team-specific error typology is to present the errors in terms that would be more readily understood by drilling team members, and to guide them when debriefing themselves on their performance. The typology could also then form an integral of a simulator-based training intervention focusing on non-technical skills for drilling teams.

3.2 Data Collection and Analysis

Following the process described by Moffat and Crichton [13], observations took place during simulator-based exercises using the team behavioural marker system previously developed. A total of 21 sessions were undertaken with drilling teams (n = 8−16), each of whom took part in 5 simulator-based exercises resulting in 105 observation reports. The drilling teams included roles from Mud Logger, Assistant Driller, Driller, Tool Pusher, and Drilling Supervisor. The resultant data were analyzed into the four categories and observation rating criteria in order to create lists of the positive behaviours as well as areas for improvement.

3.3 Error Types in Team Behaviours

The team behavioural markers, as shown in Table 1 above, includes four categories. Examples of the areas of concern, as noted under the ratings of being marginally below expectations and well below expectations, were collated and analyzed to identify the main factors that either caused or were caused by ineffective behaviours.

4 Results

The results are presented here in terms of total numbers of observed cognitive errors, then by category (Team Situation Awareness, Team Decision Making, Team Work and Communication). The category of Team Workload and Stress Management was not included in the analysis as few, if any, observations falling under this category were recorded during these simulator-based exercises. The final analysis presented refers to the types of errors observed within each category.

4.1 Total Number of Observed Cognitive Errors

A total of 721 errors affecting team performance were observed during the simulator-based exercises. Of those, as demonstrated in Table 2, 486 errors were made in the category of marginally below expectations (67 %), with 235 errors recorded as well below expectations (33 %).

4.2 Frequency of Observed Cognitive Errors by Category

The frequency of errors by category is shown in Fig. 2. More errors were observed in the Team Work and Communication category than Team Situation Awareness or Team Decision Making. Although the frequency of errors classed as marginally below expectations for Team Decision Making was greater than the other two categories. The highest number of errors classed as well below expectations were recorded for Team Work and Communication.

Table 2 Total frequency of observed cognitive errors by category and rating (Percentages of total observations shown in parentheses)

Rating	Team behavioural marker category			Total
	Team situation awareness	Team decision making	Team work and communication	
Marginally below expectations	143 (20 %)	142 (20 %)	201 (28 %)	486 (67 %)
Well below expectations	57 (8 %)	66 (9 %)	112 (16 %)	235 (33 %)
Total	200 (28 %)	208 (29 %)	313 (43 %)	721 (100 %)

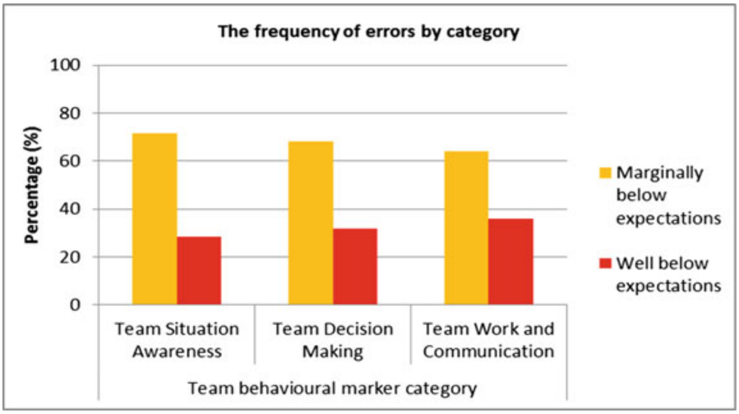

Fig. 2 Frequency of observed cognitive errors by category of marginally below expectation and well below expectations

4.3 *Frequency of Observed Cognitive Errors by Type of Error*

The observed cognitive errors were analyzed to identify the key types of errors made under each category of Team Situation Awareness, Team Decision Making, and Team Work and Communication. Table 3 presents a typology of the frequently observed cognitive errors with the combined ratings of marginally below and well below expectations by category. A full list of the typology of errors with frequencies and percentages separated into marginally below expectations and well below expectations is presented in the Appendix.

As illustrated in Table 3, similarities emerge between the frequency of observed cognitive errors in both marginally below expectations and well below expectations. For example, in Team Situation Awareness, distractions during operations scores highly under both categories (13 and 21 % respectively). However, there are some differences in that the error of actions being carried out without other team

Table 3 Typology of observed types of observed cognitive errors with combined count of marginally below and well below expectations by category

Team behavioural marker	Type of error	Combined count	
		Count	%
Team situation awareness (Total number of errors: 200)	Information not shared with relevant other team members	31	16
	Distractions during operations (non-sterile cockpit)	31	16
	Too much information available but not being paid attention to	15	8
	Losing track of the current situation	17	9
	Information present but missed	14	7
	Accept only one assessment of situation (Tunnel vision)	17	9
	Data misinterpreted	12	6
	Fixated on non-existing problems	11	6
	Failing to challenge the assumptions made by other team members	11	6
	Poor structure to Tool Box Talk	8	4
	Important information disregarded	15	8
	Initial parameters not recorded	6	3
	Confirmation bias (looking for what expect to see)	3	2
	Actions carried out without other team members being advised	9	5
Team decision making (Total number of errors: 208)	Team members not contributing to decision making process	49	24
	Choosing first option discussed and no contingencies considered	48	23
	No-one taking responsibility for making decisions	24	12
	Options discussed but selected option not articulated	20	10
	Team dividing into separate groups and making different decisions	18	9
	Team members not assertively presenting their option(s)	15	7
	Individual team members making decisions without involving other team members	12	6
	Lots of discussion but no actual decision made	15	7
	Inexperienced team members not being invited to contribute during decision making process	7	3

(continued)

Table 3 (continued)

Team behavioural marker	Type of error	Combined count	
		Count	%
Team work and communication (Total number of errors: 313)	Not verifying information being exchanged	72	23
	Incorrect use of open or closed questioning	48	15
	Unclear/non-specific instructions passed on	42	13
	Use of leading questions	19	6
	Multiple conversations taking place simultaneously	22	7
	Roles not allocated	33	11
	Team members not listening or paying attention	18	6
	Team members not participating in team discussions	26	8
	Missing or poor handover	7	2
	Interruptions while others speaking	14	4
	Team members ignoring input from others during discussions	3	1

members being advised is observed highly under the well below expectations category (11 %) but much less on the marginally below expectations (2 %). Observations of important information being disregarded occurred 18 % of the time under the well below expectations category but only 3 % in marginally below.

Under the category of Team Decision Making, both errors of not contributing to the decision making process, and choosing the first option discussed and no contingencies being considered were frequently observed. A count of 26 out of 66 occasions (39 %) were attributed to choosing the first option discussed and no contingencies being considered. More errors were observed as being no-one taking responsibility for making a decision (15 %) under marginally below expectations than in well below expectations (5 %).

Under Team Work and Communication, a similar picture emerged with four errors being observed and explaining 60 and 54 % respectively of the total errors being rated as marginally below expectations or well below expectations. Errors explained as roles not being allocated were observed as 17 % of well below expectation but only 7 % of marginally below expectations.

5 Discussion

The purpose of this project was to identify the types of errors made by drilling teams as exhibited during simulator-based exercises. The simulator was a high-fidelity simulator which encouraged the team members to become immersed in managing scenarios and solving problems collectively. This allowed the team behaviours to be observed and analyzed.

Through identifying this set of ineffective team behaviours, it is anticipated that training interventions specifically addressing team performance can be developed or existing interventions modified to reinforce effective behaviours. Trainers, or facilitators, can incorporate the error typology in a training curriculum and to discuss in the classroom what errors occur and why. Experiential learning opportunities in a simulator—whether low or high fidelity—can be designed to allow the behaviours to be demonstrated and targeted during debriefs.

Specific and directed debriefing and feedback is considered essential to improving performance [17] but requires a tool or technique to act as a basis for the debrief. The feedback should attempt to strengthen effective habits and behaviours and to modify ineffective ones. Using an error typology which describes the errors in drilling team terms can enhance the learning opportunities from the exercise by highlighting the types of errors that can occur and raising awareness of their impact. Team members can reflect on what they did, how they did it, and why they did it that way, and can generate mitigations for themselves or for the team.

In total, the majority of cognitive errors were observed as being marginally below expectations (67 %) than well below expectations (33 %). This implies that, with focused training objectives, team performance could be improved to avoid or mitigate errors in all three categories. Errors made under Team Work and Communication were rated almost the same in terms of being well below expectations (16 %) as the other two categories combined (8 and 9 % respectively). This substantiates other research showing that communication failures affecting team functioning are a major contribution to incidents and near misses [18]. Training interventions need to target the types of errors categorized as being well below expectations to improve communication within a drilling team. The errors made under each category were similar but occasional differences did emerge.

The effect of stressors such as uncertainty, time pressure, dynamic events, and multiple non-co-located players were seldom observed during these simulator-based exercises. Although it is not possible to categorically describe why this would be, one potential explanation could be that the exercises were not designed to create stressful situations. These observations took place during a training intervention developed to allow trainees to demonstrate their knowledge and skills, rather than being assessed. In order to observe examples of errors due to team workload and stress management the scenarios would need to be modified to include triggers to increase stress.

6 Conclusion

Drilling teams operate in environments characterized by high risk, uncertainty, and multiple players, and where the consequences of errors can be catastrophic. There is growing recognition in the oil and gas industry of the importance of NTS, initially related to well control, and that these skills require specific focus and training. Errors made by drilling teams that could have an impact on safety and efficiency can be identified during simulator-based exercises, leading to directed training interventions to avoid or reduce the effects of those errors.

Appendix: Full Typology of Observed Cognitive Errors by Category and Rating

Team behavioural marker	Type of error	Marginally below expectations		Well below expectations	
		Count	%	Count	%
Team situation awareness	Information not shared with relevant other team members	25	17	6	11
	Distractions during operations (no sterile cockpit)	19	13	12	21
	Too much information available but not being paid attention too	15	10	0	0
	Losing track of the current situation	11	8	6	11
	Information present but missed	11	8	3	5
	Accept only one assessment of situation (Tunnel vision)	10	7	7	12
	Data misinterpreted	10	7	2	4
	Fixated on non-existing problems	9	6	2	4
	Failing to challenge the assumptions made by other team members	9	6	2	4
	Poor structure to tool box talk	8	6	0	0
	Important information disregarded	5	3	10	18
	Initial parameters not recorded	5	3	1	2
	Confirmation bias (looking for what they expected to see)	3	2	0	0
	Actions carried out without other team members being advised	3	2	6	11
	Total number of errors	143		57	
Team decision Making	Team members not contributing to decision making process	30	21	19	29
	Choosing first option discussed and no contingencies considered	22	15	26	39
	No-one taking responsibility for making decisions	21	15	3	5
	Options discussed but selected option not articulated	17	12	3	5

(continued)

(continued)

Team behavioural marker	Type of error	Marginally below expectations		Well below expectations	
		Count	%	Count	%
	Team dividing into separate groups and making different decisions	14	10	4	6
	Team members not assertively presenting their opinions or option(s)	12	8	3	5
	Individual team members making decisions without involving other team members	10	7	2	3
	Lots of discussion but no actual decision made	9	6	6	9
	Inexperienced team members not being invited to contribute during decision making process	7	5	0	0
	Total number of errors	142		66	
	Not verifying information being exchanged	44	22	28	25
	Incorrect use of open or closed questioning	34	17	14	13
	Unclear/non-specific instructions given or comments made	24	12	18	16
	Use of leading questions	19	9	0	0
	Multiple conversations taking place simultaneously	16	8	6	5
	Roles not confirmed or allocated	14	7	19	17
	Team members not listening or paying attention	13	6	5	4
	Team members not participating in team discussions	13	6	13	12
	Missing or poor handover	7	3	0	0
	Interruptions while others speaking	7	3	7	6
	Interference when other team members are performing tasks	3	1	0	0
	Team members ignoring input from others during discussions	7	3	2	2
	Total number of errors	201		112	

References

1. Flin, R., O'Connor, P., Crichton, M.: Safety at the Sharp End: A Guide to Non-technical Skills. Ashgate, Aldershot (2008)
2. GIHRE (Group Interaction in High Risk Environments): Enhancing performance in high risk environments. Recommendations for the use of behavioural markers. Berlin (2001)

3. van Avermaete, J.A.G., Kruijsen, E.A.C.: NOTECHS: The Evaluation of Non-technical Skills of Multi-pilot Aircrew in Relation to the JAR-FCL Requirements (Project Report CR 98443). NLR, Amsterdam (1998)
4. Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N., Patey, R.: Anaesthetists' non-technical skills (ANTS): Evaluation of a behavioural marker system. Br. J. Anaesth. **90**, 580–588 (2003)
5. Yule, S., Flin, R., Maran, N., Rowley, D., Youngson, G., Paterson-Brown, S.: Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. World J. Surg. **32**(4), 548–556 (2008)
6. Hopkins, A.: Disastrous Decisions: The Human and Organisational Causes of the Gulf of Mexico Blowout. CCH, Australia (2012)
7. Energy Institute: Guidance on crew resource management (CRM) and non-technical skills training programmes. Report prepared by Energy Institute, 1st edn, London (2014)
8. IOGP: Guidelines for implementing well operations crew resource management training. Report no 502. International Association of Oil and Gas Producers, London (2014)
9. Salas, E., Cannon-Bowers, J., Weaver, J.: Command and control teams: Principles for training and assessment. In: Flin, R., Arbuthnot, K. (eds.) Incident Command: Tales from the Hot Seat. Ashgate, Aldershot, UK (2002)
10. Salas, E., Cannon-Bowers, J.A.: The anatomy of team training. In: Tobias, L., Fletcher, D. (eds.) Handbook on Research in Training. MacMillan, New York (1997)
11. Hamman, W.R.: The complexity of team training: what we have learned from aviation and its applications to medicine. Qual. Saf. Health Care **13**(1), 72–79 (2004)
12. Weaver, S.J., Salas, E., Lyons, R., Lazzara, E.H., Rosen, M.A., DiazFranados, D., Grim, J.G., Augenstein, J.S., Birnbach, D.J., King, H.: Simulation-based team training at the sharp end: A qualitative study of simulation-based team training design, implementation, and evaluation in healthcare. J. Emergencies, Trauma Shock **3**(4), 369–377 (2010)
13. Moffat, S., Crichton, M.: Investigation non-technical skills through team behavioral markers in oil and gas simulation-based exercises. In: Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015, Las Vegas (2015)
14. Reason, J.: Human Error. Cambridge University Press, Cambridge (1990)
15. Stanton, N.A., Salmon, P.M., Rafferty, L.A., Walker, G.H., Baber, C., Jenkins, D.P.: Human Factors Methods: A Practical Guide for Engineering and Design, 2nd edn. Ashgate, Aldershot (2013)
16. Woods, D.D., Dekker, S., Cook, R., Johannesen, L., Sarter, N.: Behind Human Error. Ashgate, UK (2010)
17. Fanning, R.M., Gaba, D.M.: The role of debriefing in simulation-based learning. Simul. Healthcare **2**(2), 115–125 (2007)
18. Flin, R., O'Connor, P., Crichton, M.: Safety at the Sharp End: A Guide to Non-technical Skills. Ashgate, Aldershot (2008)

The Human Factor Role in the Reducing of the Power Supply Reliability

Irina Semykina and Eugene Skrebneva

Abstract The article describes the problem of power supply reliability of hazardous production facilities for Russian coalmines. The authors analyze the accidents in the coalmine external power supply networks for the period from 2014 to 2015 and reveal the weak points in the power supply organization. It is proved that the main cause of power supply violations is the human factor. As a way to increase the power supply reliability the authors propose the algorithm based on the objective quantitative adjectives that eliminates the subjective opinion of the staff and the management of both company-consumers and electric grid companies.

Keywords Power supply reliability · Category of reliability · Coalmine · Hazardous production facilities · Utility connection · Power supply network

1 Introduction

The power supply reliability of hazardous production facilities is inextricably connected with not only the technical and economic performance of the organizations but also with its security, that involves both the social security and the environmental safety. Currently it is physically impossible to guarantee an uninterrupted supply of electric power because of too many uncontrollable factors influences on the supply process, such as technical, natural, anthropogenic and

I. Semykina (✉) · E. Skrebneva
Power Institute, T.F. Gorbachev Kuzbass State Technical University, Vesennyaya 28, 650000 Kemerovo, Russia
e-mail: siyu.eav@kuzstu.ru

E. Skrebneva
e-mail: evgeniyas77@rambler.ru

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_10

others. The human factor is contained in most major accidents in the energy sector that have occurred over the last 60 years:

- Design errors in the electric power supply;
- Insufficient attention to the safety standards in the power system;
- Insufficient staff qualifications;
- Incorrect operations in whether normal or emergency modes, which are realized in the company-consumers of electricity, in the power companies and in the centralized power dispatching authority;
- The emotional and psychophysiological staff oppression.

The authors investigate the impact of human factor on the power supply reliability in the case study of Kuzbass. This region is the largest coal deposit in Russia. The coal industry development requires a continuous increase in coal production that entails the intensive increase in equipment power and, as a consequence, the increase in the consumption of electric energy and power. The reliability and service continuity of power supply have the top priority by virtue of the significant harm that may occur even at short-term discontinuation of the power supply. The harm involves not only the economic losses, but also the life and health staff hazard, as well as environmental safety violation in the region.

2 Causes of Power Supply Violations

The analysis of accidents in the external power supply networks of Kuzbass coalmines for the period from 2014 to 2015 reveals the weak points in the power supply organization of hazardous production facilities.

First, the electric grid companies that provide electricity for coalmines have the electric grid facilities with high-life on the balance sheet, which reduces the level of the power supply reliability. A lack of available financial resources does not allow the regional electric grid companies to carry out the global reconstruction of power supply networks. This company have to make arrangements of maintaining equipment such as maintenance and service, control over the width of forest corridor, well-timed and regular diagnostics and recovery et cetera. However this arrangements not always be executed in full so it may become a cause of accident.

Moreover, the construction of new transmission lines and the reconstruction of operating ones are carried out with a principle of cost minimization to date without reference to cost savings following the service of new transmission lines and generally within disregard for the reliability. For example, an emergency shut down of transmission line located in a forest area have a larger probability than in a steppe area. Therefore, if design engineers change the transmission line routing to increase the power supply reliability, the emergency downtime cases will be avoid and the electric grid company will have long-term-profit even the first costs will be larger. However, the dependence between reliability and routing is commonly neglected [1].

Second, the coalmine utility connections are often not compatible with the demanded category of the power supply reliability. In Kuzbass coalmines, a part of equipment are always the power-consuming equipment with the first special category of the power supply reliability. However, the management of coalmines makes applications for utility connection with a lower category, as it is cheaper. This is possible because in line with the Decree of the Russian Government dated 27.12.2004 No. 861 the company-consumers have the prerogative rights to choose the category of power supply reliability. Consequently, the electric grid companies do not have the possibility to refuse such coalmine utility connection and under-reporting of the category of the power supply reliability leads the probability that electric power supply of hazardous production facilities at coalmines will be stopped for 24 h in a row.

In addition, it is the common situation when category of the power supply reliability for the coalmine utility connection is specified adequately but the procedures for documenting are not complied. For instance, the capacity of the equipment with the first special category of the power supply reliability must be given in the Certificate of delineation of balance sheet attribution, however it is often be signing without referring to this capacity. Furthermore, the Acts of coordination of emergency and technological reserved quota are often inaccurate. On these grounds, in the cases of power supply violations in the networks the electric grid company has the rights to power cutoff of coalmine.

Third, the existing schemes of external power supply networks of Kuzbass coalmines had designed in the middle of the last century. Since then, the composition and location of company-consumers and energy usage level in the region have dramatically changed, which affects the power supply reliability. Today, most Kuzbass coalmines have a dead-end power layout and the substations that provide mine feeding are located outside the responsibility area of the regional dispatch control.

Besides, there is no the Single regional electric utility capacity expansion planning system as it existed in USSR [2]. The development of power supply network issue in the requests of technological connections from company-consumers. It means firstly that the electric grid companies set going construction and reconstruction of power supply equipment for external power supply networks of coalmines only if the power supply demand has already arisen and secondly that one electric grid company will not make reconciliation its network development with whether another electric grid companies in the region or power plants connected to the network.

Fourth, despite the heavy accidents that occurred in the external power supply networks of Kuzbass coalmines, which number exceed 10, the mines' management does not consider the ensure of power supply reliability as a problem. For example, let us look at the external power supply network of coalmine in Fig. 1.

Two high-voltage dual circuit electric transmission lines make connection between the coalmine and two high-voltage substations with mutual redundancy. The requirements to guarantee the power supply according with the first special category of the power supply reliability are protocolary fulfil, there are two

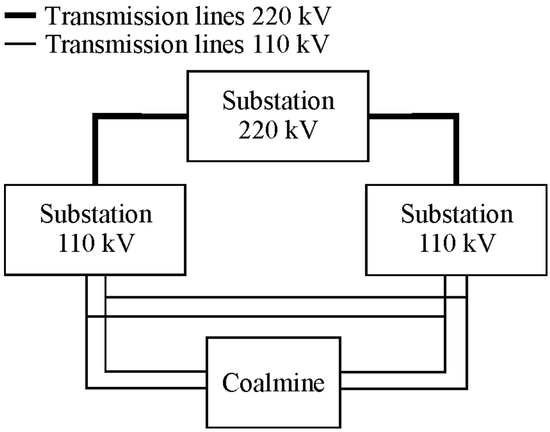

Fig. 1 The typical external power supply network of coalmine

independent power sources and the redundancy of dual circuit lines. This is the coalmines' management point of view but reality is the regional power supply network has the limited number of main substations so with a high probability the considered substations are supplied from one substation with a higher voltage level. It means the accident in this substation lead to power cutoff of considered coalmine and its category of the power supply reliability is not be achieved.

3 Technical Measures of Reliability Growth

There are three general kinds of technical measures to increase the power supply reliability of coalmine. All of them specify a construction of new elements in external power supply network. The choice between the technical measures is made by management of coalmine in the basis on a cost-benefit analysis, so there is also the human factor in this choice.

3.1 Independent Power Supply Source

In line with the Decree of the Russian Government dated 27.12.2004 No. 861 the company-consumers must equip with the independent power supply source having the required capacity all its hazardous production facilities classed as first special category of the power supply reliability. In the case of Kuzbass, this independent source can be the next types [2]:

- Coal burning combined heat and power station with low power, it capacity might be from 5 to 15 MW. However, this type of stations is very expensive as in construction as in operation and it is unprofitable for coalmine.

- Diesel power station. The problem is diesel power stations with required capacity costs more when 1 million dollars per 1 MW installed capacity.
- Gas-producer plant that operate on the degassing-extracted mining methane. Even if the methane is evolving in each Kuzbass coalmine, it is impossible to guarantee steady gas input for a gas generator.

That is why the Decree No. 861 notwithstanding, there are only a few independent sources in Kuzbass.

3.2 Construction the Third Transmission Line

Another ability to increase the power supply reliability is construction of the third transmission line which connect the coalmine with another substation or even with a power station (Fig. 2). This solution makes the redundancy of coalmine power supply.

3.3 Reconstruction of Power Supply Network

The construction of the third transmission line is the responsibility of coalmine, but the electric greed companies into its area of responsibility may also increase the

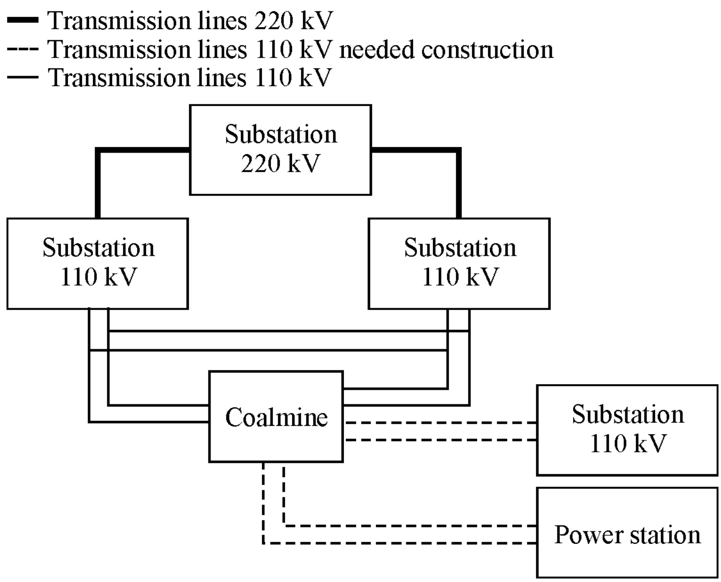

Fig. 2 The coalmine utility connections to the redundant power sources

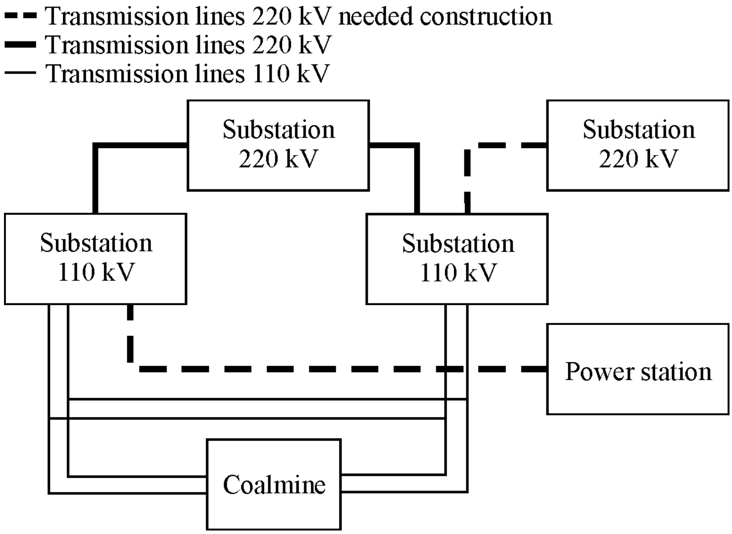

Fig. 3 The redundancy of substations power supply

power supply reliability. It means reconstruction of power supply network to make connection of substation which supply the coalmine with another one substation with a higher voltage level or with the power station (Fig. 3).

4 Conclusions

Thus, the human factor in one way or another is present in supply process of hazardous production facilities, and it often adversely affects the level of reliability. It once in a way happens the human factor increase the reliability. For example, the coalmine utility connection has the third category of the power supply reliability but the mine's chief power engineer press the point of using the modern commutation switches and power technologies that make emergency power switching and so the uninterrupted power supply is achieved and required level of power supply reliability is provided [3]. However, the human factor frequently decreases the power supply reliability.

As a way to increase the power supply reliability the authors propose the algorithm based on the objective quantitative adjectives that eliminates the subjective opinion of the staff and the management of both company-consumers and electric grid companies. This algorithm provides not only the calculation and the statement of the existing level of power supply reliability but also offers the advices intended to ensure the reliability with minimal financial investment:

- Change the emergency power switching of the coalmine;
- Change the utility connection scheme of coalmine;
- Adjust the category of the power supply reliability to the composition of power-consuming equipment;
- Implement a system of monitoring and emergency response to the power supply network of coalmine;
- Reconstruct the electric grid, optimized by the coast-reliability criterion.

It can help mostly eliminate the human factor from the providing of the power supply reliability.

Acknowledgments The authors wish to express appreciation to Dmitry Kudryashov, the head of Kuzbass Regional Approval Authority Department of Electric Power Industry, for invaluable help in data acquisition about the accidents in the external power supply networks of Kuzbass coalmines.

References

1. Goldshtein, V.G., Kubarkov, Y.P., Revyakina, K.V., Rygalov, A.Y.: The application of information technologies to assess the level of reliability and risk of power facilities. In: Proceedings of the Kola Science Centre RAS, no. 5, pp. 150–154. (2011) (in Russian)
2. Kudryashov, D.S.: The law has lagged behind the industry. Some problematic issues of power supply of Kuzbass coalmines. In: Kuzbass coal, no. 1. (2015). http://www.uk42.ru/index.php?id=426 (in Russian)
3. Vorotnitsky, V.V.: The power supply reliability as the regulatory tool of interconnection between electric greed companies and company-consumers. In: Energy and Management, no. 3 (2009). http://www.web-energo.by/EiM_3_2009_art3.html (in Russian)

Part IV
Simulation and Interface Design for Safety Focused Research

An Overview of the IEEE Human Factors Standard Development Activities—2016

David R. Desaulniers and Stephen Fleger

Abstract Since 1980 the Institute of Electrical and Electronics Engineers (IEEE) has supported development of human factors (HF) standards. Within IEEE, Subcommittee 5 (SC5) of the Nuclear Power Engineering Committee develops and maintains HF standards applicable to nuclear facilities. These standards are structured in a hierarchical fashion. The top-level standard defines the HF tasks required to support the integration of human performance into the design process. Five lower tier documents expand upon the upper tier standard. Presently, two new HF standards projects are underway; one to provide HF guidance for the validation of the system interface design and integrated systems operation and another for designing and developing computer-based displays for monitoring and control of nuclear facilities. In addition to producing and maintaining HF standards, SC5 is also involved in outreach activities, including sponsorship of a series of conferences on human factors and nuclear power plants.

Keywords Human factors guidelines · Nuclear power plants · Control rooms · Standards

1 Introduction

The Institute of Electrical and Electronics Engineers (IEEE) standards documents are developed within the IEEE Societies and the Standards Coordinating Committees of the IEEE Standards Association (SA) Standards Board. The IEEE develops its standards through a consensus development process approved by the American National Standards Institute. The process engages volunteers representing varied

D.R. Desaulniers (✉) · S. Fleger
U.S. Nuclear Regulatory Commission, Washington, DC 20555, USA
e-mail: David.Desaulniers@NRC.gov

S. Fleger
e-mail: Stephen.Fleger@NRC.gov

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_11

disciplines, agencies, and organizations to ensure that a range of viewpoints are considered in developing the institute's guidance documents.

Since 1980 the IEEE has supported development of human factors standards. Within the Nuclear Power Engineering Committee (NPEC) of the IEEE Power and Energy Society, Subcommittee 5 (SC5) develops and maintains human factors standards applicable to nuclear facilities. SC5 members have extensive expertise in application of human factors engineering principles and human reliability analysis for nuclear power plants and other nuclear facilities. NPEC manages human factors standards development and maintenance by overseeing the activities of SC5, and approving new or revised draft standards for ballot by members of the IEEE SA.

Currently, the subcommittee has nine active/attending members and approximately thirty contributing/corresponding members. The members are employed by the electric power industry, including utilities, the Nuclear Regulatory Commission, the Department of Energy (DOE), DOE laboratories, architect/engineering firms, consulting companies, and related industries. For budgetary or other reasons, contributing/corresponding members typically do not attend meetings but keep abreast of the subcommittee's activities through meeting minutes. As matters arise that align with their individual areas of interest and expertise, these individuals take a more active role by supporting the review of draft standards under development, and existing standards under revision, by providing comments and participating as members of the ballot pool.

SC5 meets three times per year, usually in January, July, and October. The first two meetings of the year are planned to coincide with NPEC's biannual winter and summer meeting, while the fall meeting is SC5 only. To make best use of their time and resources during the meetings, the subcommittee usually splits into working groups. Members may participate in more than one working group. All of the meetings last 3 days, which includes 1 day for the NPEC meetings in January and July. While a great deal of technical work is accomplished during the meetings, many members accept writing and review assignments outside of the meetings. In addition, to manage the workload and maintain steady progress on standard development and revision initiatives, many of the working groups elect to conduct their business remotely through conference calls and web conferencing services.

SC5 has structured its standards in a hierarchical fashion. The top-level SC5 guidance document is IEEE Std 1023 [1]. IEEE Std 1023 is an upper tier document that defines the human factors tasks required to support the integration of human performance into the design process and promotes the systematic integration of human performance considerations in the life cycle of commercial nuclear power facilities. Presently, five lower tier documents expand upon the upper tier standard. IEEE Std 845 presents methods for measurement of human performance [2]. IEEE Std 1082 contains guidance for incorporating human reliability analysis (HRA) into probabilistic risk assessments of nuclear generating facilities [3]. IEEE Std 1289 presents HF guidance for design of computer-based display systems [4]. IEEE Std 1786 is a guide for the design and use of computerized operating procedure systems at nuclear generating stations [5]. Finally, IEEE Std 1707 provides a recommended practice for the investigation of events at nuclear facilities [6].

In addition to these existing standards, two new human factors projects are currently underway. IEEE P2411 will be a guide to provide HF engineering guidance for the validation of the system interface design and integrated systems operation. IEEE P2421 will be an HF guide for designing and developing computer-based displays (CBDs) for monitoring and control of nuclear facilities. IEEE P2421 will differ from IEEE Std 1082, which has a similar scope, in that it will focus on the process of designing and developing computer-based displays rather than specific criteria for design elements.

2 The Standards Development Process

IEEE standards are developed and maintained through the standards development lifecycle process depicted in Fig. 1.

2.1 Initiating the Project

The development of a new standard or revision of an existing one is initiated by a Project Authorization Request (PAR) submitted to the IEEE Standards Association (IEEE SA). The request is developed by a sponsoring body, which is an individual

Fig. 1 IEEE standards development lifecycle

or entity, such as an SC5 Working Group (WG) in charge of the standard. The WG completes a PAR form, which includes the standard's title, scope, purpose, brief description of the need for the project, key dates, and contact information for the sponsoring society (Sponsor) and committee, the WG chair and the standards representative. The IEEE SA mandates, oversees, and helps facilitate the process for standards development. The Sponsor for the standards project assumes responsibility for the respective area of standards development, including the organization of the standards development team and its activities.

2.2 Mobilizing the Working Group

Once the IEEE SA approves the request for a new standard development project, the sponsor follows the rules and processes to recruit and assemble a collaborative team (Working Group) to actively engage in standard development. Working Groups are comprised of individuals and entities (e.g., companies, organizations (including non-profits), and government agencies) who volunteer to support the development of standards. Working Group officers may either be elected by the Working Group members or appointed by the Sponsor and they oversee the standard development project.

The IEEE Standards Association (IEEE-SA) has established rules related to membership and participation to ensure that highly dedicated individuals lead participation and no single interest dominates the standards development process. Working Groups establish their own organizational, communications, and meeting structures to support standard development and to address matters such as consensus building, decision making and balloting.

2.3 Drafting the Standard

Upon approval of the PAR by IEEE, an outline is produced and writing assignments are agreed upon. Each standard is assigned a project champion who assumes a leadership role by working with the WG Chair to ensure the standard progresses in a collegial and timely fashion through the writing process. Typically the WG members meet on a bimonthly basis to provide review and comment on each other's work, which is carried out independently between WG meetings. The standards development process is conducted in an iterative manner, in which sections of the standard are written, drafts are submitted for review and comment, and sections are revised to address comments that represent a consensus view of the subject matter experts. Eventually the standard is ready for review by the entire subcommittee.

2.4 Balloting the Standard

Following SC5 approval, the draft standard is ready for presentation to NPEC. The standard is previewed at an NPEC meeting to see if it is acceptable for IEEE to ballot. If judged ready to ballot, a ballot pool is formed and the standard is distributed to those members of the IEEE SA who expressed an interest and willingness to review and vote on the standard. It is important to note that the composition of the ballot pool must meet IEEE SA guidelines to ensure that a diversity of organizations and associated interests/views are represented in the ballot pool. A successful ballot requires a minimum 75 % response rate by the ballot pool and approval votes by at least 75 % of the individuals casting ballots. Ballots may include comments and in such cases these comments must be reviewed and dispositioned. If substantive changes are necessary to resolve comments a recirculation ballot may be required.

2.5 Gaining Final Approval

Upon successful completion of the Sponsor ballot, the draft is submitted to the Review Committee (RevCom). The balloted draft is reviewed by RevCom and then submitted to the Standards Board for approval. After submission, review, and acceptance, the approved standard is published and issued for 10 years, after which time it is automatically withdrawn if no further action is taken to reaffirm or revise the standard. Development of a standard typically takes 5 years from project inception to issuance by IEEE.

2.6 Maintaining the Standard

It is important to remember that standards are "living documents," which may be iteratively modified, corrected, adjusted, and updated based on lessons learned from operating experience and advances in methods, tools, and technologies, among other factors.

The remainder of this paper provides a high level overview of the current status of SC5 standards activities, and introduces additional SC5 business by order of the working groups that are responsible for the activity.

3 SC5 Working Groups and Activities

IEEE SC5 comprises four working groups that are charged with overseeing the development and maintenance of human performance standards for the nuclear industry.

3.1 WG 5.1—Human Factors Applications and Methods

WG5.1 is responsible for maintaining four existing standards, and currently has a fifth standard under development.

Human Factors Analysis Standard. IEEE Std 1023-2004 is SC5's upper tier or "mother" standard. It was originally published in 1988 and underwent a major revision in 2004 to provide more comprehensive guidance and to improve its coordination with the lower tier standards, especially the performance measurement standard. This document provides recommended practices to engineering personnel for development of integrated programs for applying HF engineering to the design, operation, and maintenance of nuclear power generating stations and other nuclear facilities. It contains guidance for program organization, the design aspects to consider, the HF methodologies and tools to apply, and for developing an HF program plan. By following the standard, the diverse activities of design, construction, and procedures development can be integrated to improve human-system performance. The standard was reaffirmed in 2010, which approved use of the standard until 2020.

Human Performance Measures Standard. IEEE Std 845-1999 was originally published in 1988 and underwent a complete revision in 1999. This document provides guidance for evaluating human-system performance related to systems, equipment, and facilities in nuclear power generating stations. It summarizes specific evaluation techniques and presents a rationale for their application within the integrated systems approach to design. This document provides guidance for the selection and application of human performance evaluation techniques and presents recommendations for their application. The standard was reaffirmed in 2011 with use approved until 2021.

Computer-Based Control and Displays Standard. IEEE Std 1289-1998 was originally published in 1998 and reaffirmed in 2004. The standard is intended for use by managers and engineers who must replace, modify, or design instrumentation and control (I&C) systems. The standard provides system design considerations, identifies information display and control techniques for use with computer-based displays, and provides HF engineering guidance for the use of these techniques in nuclear power generating stations.

The original standard was created when utilities were first beginning to consider using computerized graphical user interfaces (GUIs) in nuclear plants with broader capabilities than those found in the simple plant process computers or safety parameter display systems available at that time. Subsequently, much of the basic technical guidance contained in the standard became available in later revisions of NUREG-0700 [7] and other industry standards such as ANSI/HFES-100 [8] and ANSI/HFES-200 [9]. In addition, the nuclear industry has gained a great deal of experience developing GUI-based I&C systems in control room modernization projects involving digital system upgrades, as well as for the new generation of NPP.

In response to these developments, SC5 gained approval in March 2010 for a project to update IEEE Std 1289. Although the WG initially made good progress with the revision effort, it became necessary to suspend work while SC5 focused on other projects. In July 2014, SC5 voted to allow the existing PAR to expire in 2016 and create a new PAR when adequate resources could be dedicated to the update. In November 2015, SC5 considered options for future guidance efforts related to computer-based displays and opted to request a PAR for a new standard focused on the process designing and developing computer-based displays. IEEE/NPEC approved this request in February 2016. The new standard development effort (P2421) will be a principal work activity of SC5 over the next several years. In the interim, SC5 is continuing to assess the need and path forward for an update to IEEE Std. 1289. The next revision is required by 2018 or IEEE Std 1289 will become inactive.

Computer-Based Procedures Standard. IEEE 1786 was published in 2011 as a standard for computerized operating procedure systems (COPS). The purpose of the standard is to provide application guidance, based on current industry experience, for the design and use of COPS in nuclear power generating stations and other nuclear facilities. The guide supports developers, users, and reviewers of COPS, and identifies acceptable practices and important considerations for applying COPS to facility operations within the control room. The existing revision is valid until 2021.

In June 2010, the IEEE-SA Standards Board approved the joint development by IEC and IEEE of a Computer Based Procedure (CBP) standard. IEC/SC45A took the lead role in writing this standard, with IEEE/SC5/WG5.1 primarily providing review and comment. During the process it was agreed not to issue a co-logo standard at that time, and IEC issued edition 1 of IEC 62646 [10] in September 2012. The IEC standard is broader in scope than IEEE P1786, including guidance for implementing all types of procedures that a utility may decide to computerize. The scope includes procedures for use outside the main control room, as well as guidance for formulating a utility policy about which procedures to computerize and to what extent. Edition 2 of IEC 62646 is under development at this time, and SC5 is part of the review process. Edition 2 is scheduled to be issued in 2017.

Validation of Systems Design and Integrated Systems Operations. With the exception of the limited guidance on validation that is presented in IEEE-Std-1023-2004, there are currently no dedicated industry-based consensus standards governing the conduct of validation, including integrated system validation, for nuclear power generating stations and other nuclear facilities. Yet there is a need for such guidance to support design certifications and combined operating license applications given the regulatory expectations set forth in NUREG-0800 [11] and NUREG-0711 [12].

To address this need SC5 is currently developing human factors engineering guidance for the validation of the system interface design and integrated system operation. The PAR for this effort, P2411, was approved by the IEEE-SA in March 2014. The project will provide guidance to be used by nuclear facility designers, applicants, licensees, architect/engineers and regulators to assure reasonable confidence that the integrated system can be safely operated by personnel during a representative set of operating conditions that could be encountered during the facility's operation. This guidance will provide acceptable means to: (1) identify performance criteria, (2) collect sufficient evidence of performance, (3) plan and conduct validation tests, and (4) analyze and resolve validation results.

3.2 WG 5.2—Human Factors International Conference

Under the leadership of WG5.2, SC5 sponsors and supports international conferences on human factors in nuclear power plants. A series of IEEE-led conferences were held approximately once every 5 years since 1980, with the last one held in conjunction with the Human Performance, Root Cause and Trending (HPRCT) conference. In more recent years, WG5.2 has collaborated with the American Nuclear Society to support their international conferences on Nuclear Power Instrumentation Control (NPIC) and Human Machine Interface Technologies (HMIT).

3.3 WG 5.4—Human Reliability Analysis

Human Reliability Analysis (HRA) Standard. First published in 1997, IEEE Std 1082-1997 provides an orderly process framework for the inclusion of human-system interactions in probabilistic risk assessments (PRAs). The document is intended to improve the analysis of human-system interactions in PRAs, to help ensure that conclusions are reproducible, and to standardize the documentation of such analyses. Rather than describing a specific method for doing HRA, the standard presents a method for integrating the HRA process into PRA, including a systematic technique for structuring, conducting, and documenting the results of an HRA. The standard was reaffirmed in 2003 and again in 2010. In parallel with the recent reaffirmation ballot, working group A8 to IEC Subcommittee 45A conducted a review of the standard to determine its suitability for adoption as a dual logo standard. Based on the comments received by IEC and additional comments made by the IEEE WG 5.4, the decision was made to revise the standard. The draft revision, now largely complete, includes editorial enhancements (e.g., clarification of existing terminology, new definitions, updated bibliography), as well as substantive changes to reflect the updated practices that are influencing contemporary HRA as part of PRAs. In January 2016, IEEE/NPEC granted SC5's request to ballot the draft revision of this standard. SC5 anticipates balloting the draft revision in the summer of 2016, resolving comments in the fall, and publishing the revision by the close of 2016.

3.4 WG 5.5—Lessons Learned

Event Investigation Standard. Until recently, the nuclear industry lacked a standard that establishes a common practice for event investigations and a common language (terminology) for communicating the tools, methods, and results of such investigations. To address this need, SC5 formed a working group in early 2006 to begin development of a recommended practice for event investigations. After a period of early development a variety of challenges slowed progress on the standard until 2013, at which point SC5 was able to focus its efforts on completion of the standard. In December 2015, the results of this standard development effort were published as IEEE Std 1707.

The document is intended for use by staff and management at nuclear facilities and those tasked with evaluating event investigation reports. The standard is broad in scope, and provides considerations for preserving data for analysis while implementing early actions to manage the event, along with recommended practices to assist analysts in planning the scope, team composition, and timeline for conducting an investigation. The standard defines key terms for communicating the methods and results of event investigations and described methods for data gathering and analysis, as well as cause determination and corrective action identification. The guidance addresses analysis of not only narrow issues and failures, but also the entire organizational infrastructure. Organizational aspects needed to support the investigation, including management oversight, training, record keeping, and roles and responsibilities, are described along with the recommended attributes of a report on the investigation.

4 Summary

SC5 is the subcommittee within NPEC that is concerned with the analysis of the human performance aspects of systems and equipment, the development of control facilities criteria, and the treatment of all matters relating to human reliability analysis for nuclear facilities. Included are the development of human factors criteria for systems; equipment and facility design; operation, maintenance, and testing; and the development of methodologies for human performance data collection, modeling, model evaluation, and model validation.

SC5 prepares and reviews technical papers, disseminates information to industry on new developments, and responds to requests for interpretation. The subcommittee has responsibility for coordination with other groups with respect to the acquisition, evaluation, and application of human factors data, control facilities criteria, human reliability data, and the coordination of nuclear standards. Lastly, SC-5 supports or sponsors technical conference sessions and educational courses, including international conferences on human factors and power plants.

Currently, SC5 is responsible for maintaining six IEEE human factors standards and is in process of developing two new standards. SC5 is also working in partnership with IEC to revise and update the existing IEEE HRA standard for adoption as a co-logo standard by IEC. SC5 is always open to suggestions for new standards.

Authors' Disclaimer The views presented in this paper represent those of the authors alone and not necessarily those of the U.S. Nuclear Regulatory Commission.

References

1. IEEE Std 1023-2004: IEEE recommended practice for the application of human factors engineering to systems, equipment, and facilities of nuclear power generating stations and other nuclear facilities, Institute of Electrical and Electronic Engineers, New York
2. IEEE Std 845-1999: IEEE guide for the evaluation of human-system performance in nuclear power generating stations, Institute of Electrical and Electronic Engineers, New York
3. IEEE Std 1082-1997: IEEE guide for incorporating human action reliability analysis for nuclear power generation stations, Institute of Electrical and Electronic Engineers, New York
4. IEEE Std 1289-1998: IEEE guide for the application of human factors engineering in the design of computer-based monitoring and control displays for nuclear power generating stations, Institute of Electrical and Electronic Engineers, New York
5. IEEE 1786-2011: IEEE human factors guide for applications of computerized operating procedure systems at nuclear power generating stations and other nuclear facilities, Institute of Electrical and Electronic Engineers, New York
6. IEEE 1707-2015: IEEE recommended practice for the investigation of events at nuclear power generating stations and other nuclear facilities, Institute of Electrical and Electronic Engineers, New York
7. NRC, Human-System Interface Design Review Guidelines. (NUREG-0700, Rev. 2): U.S. Nuclear Regulatory Commission, Washington, D.C. (2002)
8. ANSI/HFES 100-2007: Human factors engineering of computer workstations, Human Factors and Ergonomics Society, Santa Monica
9. ANSI/HFES 200-2008: Human factors engineering of software user interfaces, Human Factors and Ergonomics Society, Santa Monica
10. IEC 62646: 2012, nuclear power plants—control rooms—computer based procedures, International Electrotechnical Commission, Geneva, Switzerland
11. NRC, Standard Review Plan for the Review of Safety Analysis Reports for Nuclear Power Plants: LWR Edition. (NUREG-0800), U.S. Nuclear Regulatory Commission, Washington
12. NRC, Human Factors Engineering Program Review Model. (NUREG-0711, Rev. 3), U.S. Nuclear Regulatory Commission, Washington (2012)

The Impact of Simulation Display on Nuclear Power Plant Task Error Frequencies

Jonathan Harris, Lauren Reinerman-Jones and Grace Teo

Abstract This research investigated the impact physical fidelity has on error frequencies when operating a simulated nuclear power plant (NPP) main control room (MCR). The simulated environment used in this study uses dual 24″ monitors and a mouse as its interface, which represents the interface of digital power plants planned to come online. The simulator models an NPP MCR using scroll/pan/zoom (SPZ) to navigate Instrumentation and Control (I&C) panels housed along MCR walls. However, touchscreens can be used to display entire I&C panels, which represents the legacy plants in use today and requires participants to stand to operate in the MCR. A between-subjects experiment was conducted to evaluate desktop and touchscreen interfaces for their impact on response times and error rates when performing reactor operator (RO) tasks. While increased physical fidelity through larger field of view did help reduce response times, using touch induced more miss touch errors than the mouse.

Keywords Nuclear power plant control room · Simulator interface · Touchscreen · Human error

1 Introduction

The role of a nuclear power plant (NPP) reactor operator (RO) is an important one. The RO is responsible for the safe operation of the plant and energy production for the utility company, hence, much effort has been made to understand the tasks of

J. Harris (✉) · L. Reinerman-Jones · G. Teo
Institute for Simulation and Training, University of Central Florida,
Orlando, FL 32826, USA
e-mail: jharris@ist.ucf.edu

L. Reinerman-Jones
e-mail: lreinerm@ist.ucf.edu

G. Teo
e-mail: gteo@ist.ucf.edu

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_12

ROs, which are performed on the Human System Interface (HSI) located in the main control room (MCR) where they work.

2 Background

2.1 NPP Control Room Simulator and Tasks

The commercial nuclear industry in the United States is in a transition where new digital power plants like the AP1000 are coming on line in the next few years [1] and some legacy analog plants from the 1960s and 1970s are licensed to operate into the 2040s [2]. An HSI interface in an NPP control room comprises instrumentation and controls (e.g., gauges, switches, valves, light box status indicators, etc.) that are mounted along the MCR walls. The role of the RO is to monitor and manipulate these controls as needed to operate the NPP for safe and efficient power generation. The tasks performed by RO with the interface can be categorized into knowledge-based, skill-based, and rule-based tasks. The present study focuses on rule-based tasks for the purposes of experimentation and investigates three task types: (i) checking, (ii) detection, and (iii) response implementation. These three tasks were modeled after research findings in [3] and are representative of tasks performed by ROs in an NPP MCR.

3 Experimentation and Simulation

Studies have to be conducted to understand how the RO performs his task, i.e., the workload experienced, the types and causes of errors committed, effectiveness of training methodologies, etc. However, these studies are typically not permitted in operational MCRs as they disrupt daily operations, are costly to perform on operational equipment, and jeopardize plant safety. Furthermore, these studies are constrained by RO availability for experimentation. One solution is to use a simulation of the MCR coupled with tasks that do not require specialized NPP training. Since rule-based tasks are procedural and well defined, a novice population can perform these with limited training. The generic pressurized water reactor (GPWR) MCR simulator [4] from GSE systems has enabled such research and experimentation to be conducted.

3.1 Simulation Fidelity

Simulations offers a controlled, safe, and repeatable experimentation and training environment where management of dangerous conditions can be rehearsed [5]. One of the most important aspect of simulations is its fidelity. Fidelity is the extent to which a simulator emulates its real world equivalent. Simulation fidelity has been shown to affect training effectiveness [5] as greater transfer of training has been associated with higher fidelity simulations. Fidelity is commonly decomposed into three dimensions: physical, functional and psychological [6]. Simulator interface design has the most impact on physical fidelity, and is the focus of the present study.

3.2 Simulator Interfaces

Using the GPWR simulator with two different interfaces, a series of rule-based scenarios were developed. One interface used desktop monitors with a mouse and the other used touchscreens. The differences between the simulator interfaces (desktop vs. touchscreen) were isolated to the user interface (UI) and mostly impacted the physical fidelity of the simulation. The touchscreen interface provides a richer UI experience by utilizing a more natural direct interface compared to the computer mouse [7], and the technologies such as smart phones, tablets, and touch enabled laptops as well as public kiosks and automatic teller machines (ATMs) all increase familiarity of touch in human-machine interfaces. However, for most people, the computer mouse is still the most often used input device for interacting with a computer [7].

Desktop Interface. The simulator configuration with the desktop interface modeled an RO station in an NPP MCR using photo-realistic light boxes, gauges, and valves found in operational control rooms. It utilized dual 24″ monitors and a mouse input. The desktop-based simulator modeled the controls along the MCR walls of an operational NPP control room using a virtual environment. The simulator uses scroll/pan/zoom (SPZ) functionality with the mouse as the navigation method to move around the virtual control room wall. The SPZ functionality of this simulator imposes a non-value added step to the training task. It also reduces visual complexity of a control room wall by limiting the field of view (FOV) to a sub section of an entire I&C panel. By isolating the view, the simulator allows distracting events to be removed from the field of view. The desktop configuration limits the field of view to a sub-section of an I&C panel when zoomed to normal size (i.e., zoom level is at a 1:1 scale to the real control). This limited view required the participant to use the mouse to SPZ in order to navigate around the panel to locate controls. The desktop configuration, because of the mouse input, required users of the simulator to perform the tasks seated (Fig. 1).

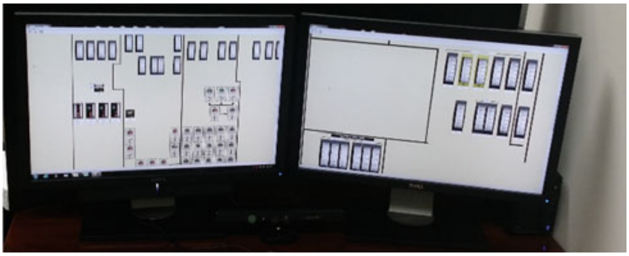

Fig. 1 GPWR simulator with the desktop configuration

Touchscreen Interface. The simulator configuration with the touchscreen interface was developed to compare represent the legacy plants in operation, contrasted with the future desktop MCR planned to come online. In place of the dual 24″ monitors and mouse inputs of the desktop interface, the touchscreen interface comprised eight 27″ touchscreen surfaces arranged in a grid (two high by four wide). This interface allowed direct touch manipulation of valves and eliminated SPZ by displaying an entire I&C panel on a large flat touchscreen surface. The touchscreen configuration spans about 100″ wide and 30″ tall. The field of view for the touchscreen configuration is over $4\frac{1}{2}$ times larger than the desktop configuration at the default desktop zoom level. Although the large size of the touchscreen allows an entire I&C panel to be displayed at once, its large size requires the participant to stand and physically move in order to scan and locate a control (Fig. 2).

3.3 Performing Tasks on the Different Interfaces

The two interfaces differ on physical fidelity. The touchscreen interface has a higher level of the physical fidelity compared to the desktop (with mouse input) interface as it utilizes a direct manipulation input (with finger) on a control rather than a translated input through a mouse. Also, in having a larger high resolution display, the touchscreen interface enhances physical fidelity by increasing the field of view closer to an operational NPP MCR, but also increases visual complexity.

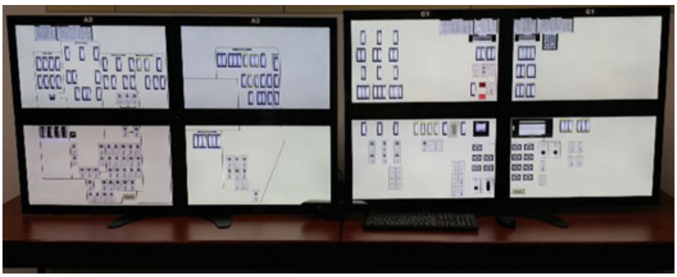

Fig. 2 GPWR simulator with the touchscreen configuration

In all three tasks (checking, detection, response implementation), users were told to locate a control on an I&C panel. Due to the differences in field of view and SPZ functionality between the interfaces, the time it takes to visually scan and locate a specific control may be affected. In addition to the response time for locating a control, the different interfaces are also likely to differ in the frequency of accidental clicks/touches during the detection task. The detection task requires users to click or touch on an acknowledge button located directly beneath the gauge they are monitoring each time they detected a change in the gauge readout. This requirement was for experimentation purposes as there was the need to record when users directed their attention to the control for the purpose of detecting changes. The button's dimensions are 3/4″ wide by 1/4″ tall for the both the touchscreen and desktop interfaces (when desktop is at a zoom of 1:1 scale). For the desktop interface, this entails moving the mouse cursor over the acknowledge button and clicking, and for the touchscreen interface, users had to touch the button with their finger. Interface-induced errors of accidental clicks/touches are committed when users click or touch anywhere but the acknowledge button.

During the response implementation task, users were instructed to open or shut specific valves. In order to open or shut a valve, users were required to click and hold (desktop interface), or touch and hold (touchscreen interface), a valve's handle and then rotate it at least 45° with a drag gesture to the left or right (left to shut, right to open) then release. If a user releases the handle before reaching the 45° threshold, the gauge snaps back to its current state. The differences in the interface between using a mouse and touch to perform a valve manipulation gesture may affect the ease with which valves are manipulated. These interface errors are reflected in the number of unsuccessful valve manipulations and repeated attempts to open/shut the same valve.

4 Hypotheses

4.1 Hypothesis 1

Although the touchscreen interface has a higher level of visual complexity from having the entire I&C panels displayed, it did not require SPZ. This should allow for quicker visual scanning of the panel resulting in a quicker response time for locating a specific control.

4.2 Hypothesis 2

For the desktop interface, the mouse provided a visual indicator of its exact location in the form of a pointer cursor. In contrast, the touchscreen interface provided no indication of the user's finger location with respect to the touchscreen until the screen is touched. It is expected that this would adversely impact the accuracy of clicks/touches for the touchscreen interface.

4.3 Hypothesis 3

The drag-and-release mouse gesture is quite common in desktop environments and relatively easy to perform. On the other hand, touch-and-drag gestures are less common, and may be harder to perform. Hence, in the response implementation task, users of the touchscreen interface may experience more unsuccessful valve manipulations than users of the desktop interface.

5 Methodology

5.1 Participants

One hundred forty-seven students participated in the study with ages ranging from 18 to 40 (M = 20.56, SD = 3.45). There were 83 males and 64 females. Participants were undergraduate student volunteers recruited from the psychology research participation system at the University of Central Florida (UCF). They were all unfamiliar with NPP control room operations. Participants were given class credit for their participation. All participants had normal or corrected-to-normal vision and were not colorblind. Participants were also required to abstain from alcohol and/or sedative medications for at least a 24-hour period and nicotine for two hours prior to participating in the experiment.

5.2 Materials

Two display configurations of the same adapted GSE GPWR simulator were used to collect data for this experiment. The simulator used for both configurations consisted of a standard desktop computer (6.4GT/s, Intel XeonTM 5600 series processor) running windows 7 professional [4]. The software runs on java JVM 1.7. For the desktop interface, the participant interacted with the simulator using two 24″ UXGA monitors with a total resolution of 3600 by 1200px, a USB 3-button laser mouse, and a 104-key Windows keyboard. For the touchscreen interface, the participant interacted with the simulator using an eight touchscreen monitor grid (two high by four wide) of 27″ WQHD monitors with a total resolution of 10,240 by 2880px, and a 104-key Windows keyboard.

5.3 Experimental Tasks

Participants were trained for 2 h on how to perform the role of a RO in an NPP MCR. This included several practice scenarios where they had to show proficiency before they were permitted to proceed. If they were unable to show proficiency with the task within two attempts, the participants were dismissed. The experimenter played the role of Senior Reactor Operator (SRO) and tasked the participants with three different task types (checking, detection, and response implementation). The three tasks were partially counterbalanced across participants.

Checking. During the checking task, the SRO verbally instructed the participant to correctly identify the state of a control located somewhere on the I&C panel. The participant was then required to locate the control of interest and click/touch the control to signal they have identified the control. After the participant clicked/touched the control, they were then required to verbally notify the SRO if the control was or was not in the desired sate. During this task, measures of response time were collected. The response time was calculated as the time between when the SRO completed the checking instruction to the time the participant clicked/touched the correct control on the I&C panel.

Detection. During the detection task, the SRO verbally instructed the participant to locate a specific gauge and then to notify the SRO once the gauge reached or crossed a certain threshold value. The participant then had to locate the control of interest and click/touch the control to signal they have identified the control. The click/touch event triggered a time-based scripted change in the gauge's values over a 5 min period. The gauge value changed 60 times during the 5 min period before reaching the threshold for reporting to the SRO. Each time the gauge event changed, the participant was required to click the acknowledge button below the gauge to signal they noticed the change. The detection task was repeated with four different gauges for a total of 240 acknowledge clicks/touches the participant was to report. During this task, response time measures were collected, as well as accidental background clicks. The response time was calculated using an identical method to that used in the checking task. The frequency of accidental background clicks corresponded to the number of times the background, instead of the control's acknowledge button, was clicked during the 5 min task.

Response Implementation. During the response implementation task, the SRO verbally instructed the participant to open or shut a specific valve. The participant then had to locate the control of interest and click/touch the control to signal they have identified the control. Once the control was identified, participants were required to shut or open (left to shut, right to open) the valve based on the SROs instruction. To shut a valve, participants using the desktop interface had to left click on the valve's handle, hold down the mouse button, and drag the mouse to the left until the handle rotated to the left by at least 45°. Participants using the touchscreen interface were required to perform the equivalent gesture using their finger. This task was conducted four times with different valves each time. During this task, response time measures were collected as well as the frequency of unsuccessful manipulations due to under rotating gestures. Response time was calculated using the identical method used in both the checking and detection tasks. Under-rotating gestures are calculated as the number of times the participant started to open/shut a valve but did not rotate the valve far enough to change the valves state.

5.4 Procedure

After giving their informed consent, participants were administered the Ishihara color-blind test and the demographics questionnaire, then they trained for two hours on their role as an RO working with an SRO. The researcher conducting the study performed the role of the SRO. The training included a PowerPoint presentation that provided a brief of the procedures and protocols for NPP simulation research, as well as given practice on the tasks on the simulator. In addition, participants were trained on the 3-way communication procedure to ensure that critical information were accurately conveyed to and from the SRO. They were also trained on how to navigate within the simulator. One group of participants was assigned to perform the tasks with the desktop interface, while the other group was assigned to use the touchscreen interface. Each group was trained on how to navigate within the assigned interface to locate controls, read status indicators, and implement the commands given by the SRO. Each aspect of the tasks was trained separately, and all aspects were combined in the practice session. Participants were also given feedback and assessed on their proficiency with the tasks. They had to attain a minimum level of proficiency (score 80 % and above) to proceed with the experiment. When training concluded, participants were given a short break. Experimental scenarios with the tasks were administered. For each task, the three-way communication procedure was carried out with the researcher playing the role of the SRO. The duration of the entire experiment session was about two hours.

6 Results

Independent samples *t*-tests were conducted to determine the extent to which the different interfaces affected the amount of errors. We examined the effect of the interfaces on (i) the time it took to locate the correct control, (ii) the number of times the 'background' rather than the control was clicked on, and (iii) the number of unsuccessful attempts at manipulating the control (i.e. not rotating the valve handle enough to trigger an open or shut event on the valve).

6.1 Time to Locate Correct Control

Findings showed that there was a significant difference between the groups in the average response time it took to locate a control, $t(112.504) = 6.103$, $p < 0.001$. The group in the Desktop condition took much more time to locate the controls compared to the Touchscreen group (Desktop: $M = 39.380$, $SD = 27.081$, Touchscreen: $M = 18.871$, $SD = 11.907$; $d = 0.980$), as shown in Fig. 3.

6.2 Accidental Background Clicks/Touches

Examination of the data showed that all the accidental background clicks/touch observed during the Detection task were committed by the Touchscreen group (Desktop: M = 0.00, SD = 0.00, Touchscreen: M = 3.35, SD = 4.525), as shown in Fig. 4. One participants data was removed from the touchscreen group whose unusually high frequency of 59 background touches seemed to indicate non-compliance with instructions rather than difficulties with the interface.

6.3 Unsuccessful Control Manipulations

Although there were incidences where users had difficulty manipulating the controls (i.e., valves) during the Response Implementation task, these were relatively

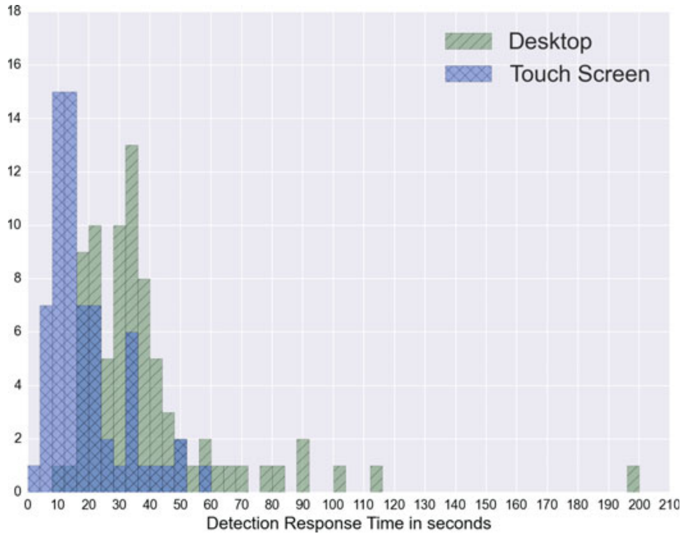

Fig. 3 Graph showing response times in seconds for participants to correctly locate a control

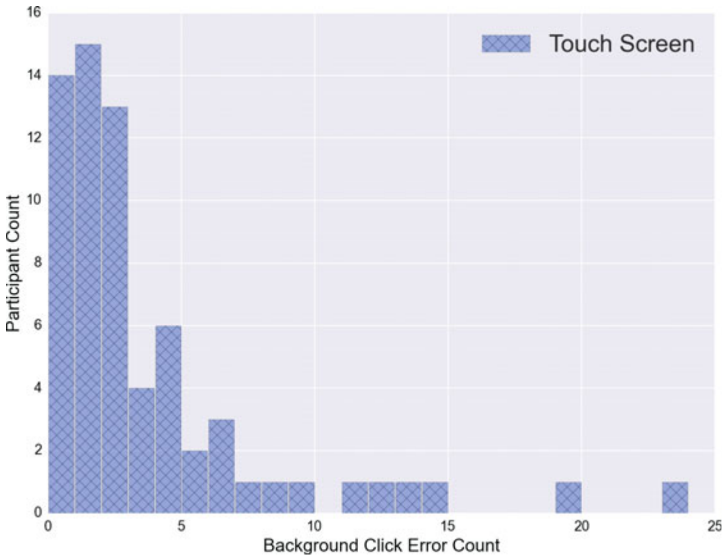

Fig. 4 Graph showing erroneous background click counts made by the Touchscreen group. No accidental background clicks occurred with the desktop group

infrequent (16.32 %) and most were successful after the second try (see Fig. 5). However, the difficulties encountered were not likely to be due to any particular interface as the two groups did not differ significantly on the frequency of unsuccessful manipulations, $p = 0.481$.

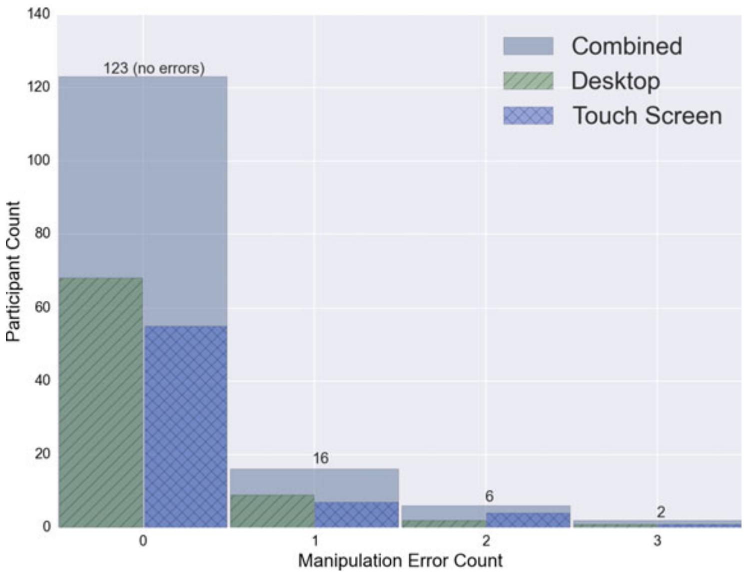

Fig. 5 Graph showing unsuccessful valve manipulation counts for each group

7 Discussion

In order to collect meaningful measures during experimentation, the simulator imposes artificial tasks not performed in an operational NPP MCR. These task include pressing an acknowledge button to signal a detected change in a gauge readout. In an MCR an RO would not be tasked with pressing an acknowledge button each time a value changes on a gauge but rather just report back to the SRO when the gauge value crosses some threshold. However, this simulator imposed task of physically acknowledging a change allows for the objective evaluation of metrics such as human error frequencies, attention control, and workload. Therefore, while the simulator's imposed task is artificial, it is an informative task for assessment and training.

This study investigated the differences between two interfaces using the same simulator software baseline. Because both systems used identical software, which was designed to function as a digital MCR (i.e., desktop and mouse), the touchscreen interface was not able to leverage common software UI practice when designing for touchscreens. However, buttons sizes on the touchscreen were larger than the 30×30 pixels recommended design guidance [8] for touchscreen displays. While the touchscreen was found to induce accidental background touches with the current software baseline, software modifications that leverage best practices for touchscreen design could ameliorate this issue.

Lastly, it was shown that response time is significantly faster when the entire I&C panel is displayed on one large display matrix. Field of view, at least for a novice population unfamiliar with the I&C panel layouts, was beneficial for locating controls.

8 Conclusions

Increasing physical fidelity through a larger field of view and utilizing touch over a mouse might lead to a more realistic training experience of a legacy plant; however, it might also lead to an increase in interface-induced errors. In the NPP domain, errors can be costly. This research illustrates the importance of understanding the impact a UI can have on error rates. While it is desirable to have a quick response time when responding to events, in the NPP domain, safety is of the highest importance. Therefore, the desire to minimizing errors trumps speedy responses. The main findings of this study are as follows:

(i) On average, the time it takes for a participant to locate a specific control on a display that is large enough to show an entire I&C panel is about half as long when compared to a participant with a smaller display which requires SPZ for searching the I&C panel.
(ii) Acknowledge events are better when performed with a mouse click gesture than when using a finger to perform a touch gesture on a button that is 3/4 in

wide and a 1/4 in tall. Not a single participant in the desktop group had an issue clicking the button with a mouse, while over half the touchscreen group had at least one incident where they clicked background adjacent to the button rather than on the button. Button size and placement need be considered in the system design if touch gestures are going to be a critical factor in safety systems.

(iii) There is no statistical difference for performing click-hold-rotate or touch-hold-rotate gestures for manipulating a valve open or shut. Both interfaces had several unsuccessful valve manipulation gestures, but the majority of participants had no issues performing either gesture.

Acknowledgmentss This report was prepared as an account of work sponsored in part by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product, or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights. The views expressed in this paper are not necessarily those of the U.S. Nuclear Regulatory Commission.

References

1. Company, G.P.: Thirteenth Semi-annual Vogtle Construction Monitoring Report. 13(29849), https://georgiapower.com/docs/about-energy/13th-VCM-Report-Final.pdf?hp=lnau_box2 (2015)
2. Slater-Thompson, N.: Nuclear Regulatory Commission resumes license renewals for nuclear power plants (2014)
3. Reinerman-Jones, L., Guznov, S., Mercado, J., D'Agostino, A.: Developing methodology for experimentation using a nuclear power plant simulator, in foundations of augmented cognition, pp. 181–188. Springer. Berlin (2013)
4. Corporate, Generic PWR: Control Room Simulator, GSE Systems: http://www.gses.com/wp-content/uploads/GSE-GPWR-brochure.pdf. pp. 1–4 (2015)
5. Ragan, E.D., Bowman, D.A., Kopper, R., Stinson, C., Scerbo, S., McMahan, R.P.: Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. Vis. Comput. Graph. IEEE Trans. on. **21**(7), 794–807 (2015)
6. Roberts, A., Stanton, N., Fay, D.: The command team experimental test-bed stage 1: Design and build of a submarine command room simulator. Procedia Manuf. **3**, 2800–2807 (2015)
7. Wu, F.G., Lin, H., You, M.: Direct-touch vs. mouse input for navigation modes of the web map. Displays **32**(5), 261–267 (2011)
8. Sun, X., Plocher, T., Qu, W.: An empirical study on the smallest comfortable button/icon size on touch screen, in Usability and internationalization. HCI and Culture, Springer. Berlin. pp. 615–621 (2007)

Neuropsychological Aspects Observed in a Nuclear Plant Simulator and Its Relation to Human Reliability Analysis

André Ricardo Mendonça Pinheiro, Eugênio Anselmo Pessoa do Prado, Marcelo Ramos Martins, July Silveira Gomes and Enrique López Droguett

Abstract This paper will discuss preliminary results of an evaluation methodology for the analysis and quantification of errors in manual (human) operation by training cognitive parameters and skill levels in the complex control system operation based on Neuropsychology and Psychophysiology approaches. The research was conducted using a game (nuclear power plant simulator) that simulates concepts of operation of a nuclear plant with a split sample evaluating aspects of learning and knowledge in the nuclear context. Operators were monitored using biomarkers (ECG, EEG, GSR, face detection and eye tracking) and the results were analyzed by statistical multivariate techniques. The experiments aimed at observing state change situations such as shutdowns and planned matches, incidents assumptions and ordinary features of operation. The preliminary findings of this research effort indicate that neuropsychological aspects can contribute to improve the available human reliability techniques by making them more realistic both in the context of quantitative approaches for regulatory purposes as well as in reducing the incidence of human error.

A.R.M. Pinheiro (✉) · E.A.P. do Prado · M.R. Martins
Analysis, Evaluation and Risk Management Laboratory,
Departamento de Engenharia Naval, University of São Paulo,
São Paulo, SP, Brazil
e-mail: andrericardopinheiro@usp.br

E.A.P. do Prado
e-mail: eugenio.prado@labrisco.usp.br

M.R. Martins
e-mail: mrmatins@usp.br

J.S. Gomes
Laboratório Interdisicplinar de Neurociências Clínicas,
Federal University of São Paulo, São Paulo, SP, Brazil
e-mail: july.flp@gmail.com

E.L. Droguett
Department of Mechanical Engineering, University of Chile, Santiago, Chile
e-mail: elopezdroguett@ing.uchile.cl

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_13

Keywords Human factors · Human-systems integration · Neuropsychology · Psychophysiology · Human reliability analysis · Incidence of human error

1 Introduction

The neuro scientific study of human behavior has advanced greatly in recent decades and today is an invaluable tool for studying human behavior in various situations. It also allows the dealt with limitations of available methods for analysis of human factor contribution to complex control system operation by identifying and evaluating factors involved in decision making, such as the influence of ergonomic elements in the criteria of acquired skills (training) and cognitive load.

An important component in the evaluation of complex systems is the human reliability during operation. Human reliability refers to the probability of the human element of performing the scheduled tasks during a defined period for system operation when tested under specified environmental conditions.

Cognitive ergonomics refers to mental processes such as perception, memory, reasoning and motor response, affecting the interactions among humans and other elements of a system. Relevant topics include the study of mental workload, decision making, specialized performance, man-machine interaction, stress and training as well as correlation between designs involving complex operating systems and human operator.

Taking into account the difficulties imposed by the human profile, the use of cognitive monitoring equipment is an interesting option for the full assessment of training and operating procedures, as it is possible to identify and record the patterns of cognitive skills in each operator as face attention, reaction ability, level of knowledge and motor actions, which may be later assessed by a monitoring group composed of the most experienced operators, psychologists and engineers linked to the process.

After evaluating operators via the application of the proposed methodology, the collected information can be used in a Human Reliability Analysis. In particular, through the analysis of Eye Tracking, EEG (electroencephalogram), ECG (Cardiac Monitoring) and GSR (Galvanic Skin Response) data, a model of the operator flow experiences will be developed allowing us to increase the operator performance to a higher level of human reliability.

To reach this end, it is necessary to observe moments of high workload, when there is a higher probability of micro incidents.

This research was conducted using a game (Power Plant Simulator) that simulates concepts of operation of a nuclear power plant with a split sample to evaluate aspects of learning and knowledge in the nuclear field. Operators were monitored using biofeedbacks (ECG, EEG, GSR, and eye tracking), and the results were analyzed by multivariate statistical techniques. The research has two main objectives:

1. Identify biomarkers (cognitive and psychophysiological variables) that influence the behavior during the decision-making process on tasks with situations of risk and uncertainty involving a group of operators in a virtual control room of a nuclear power plant simulator;
2. Establish preliminary protocol Neuropsychology and Psychophysiology assessment to be used in studies of human reliability for operations in complex hybrid systems.

2 Methodology

For decision-making experiments, we used a game that partially reproduces the control room of a simulated nuclear power plant in 2D computing environment. Game's basic idea is to produce sufficient electricity to lite a whole city without causing a Nuclear disaster. The procedure is performed by the increase or decrease of the control bar to start the nuclear reaction in the reactor. It is crucial to find the right combination of settings to produce energy not damaging the reactor components for exceeding its limits of operation, knowing that all reactor components have their pre-determined limits. If you exceed them and not perform contra measures, the reactor will be damaged, which will appear in the Repair Facility section of the game. If any of the components of the reactor oscillate the condition 'Warning' should rapidly reduce the slider that caused the condition. While the "warnings" (Warning) in oscillation, will damage the proper components. In the repair installation, this damage is indicated in two forms. Beyond ways to produce electricity to supply the energy demand, it should also generate profit. There are various expenses acquired during the game, especially early in the game. These costs are called Aux Power and appear in the financial section.

For the experiment the operators were divided in four groups:

- G1—High levels operators. Individuals that know nuclear aspects of a power plant and that get used to deal with IHM (Human Machine Interface);
- G2—Individuals that know nuclear aspects of a power plant but do not get used to deal with IHM (human machine interface);
- G3—Individuals that do not know nuclear aspects of a power plant but have a good ability with IHM (human machine interface)—game simulation;
- G4—Individuals that do not know nuclear aspects of a power plant and do not have a good ability with IHM (human machine interface).

The grouping was based in classical technique of questionnaire and formal test with a follow up of psychologist and a nuclear engineer with experience in nuclear power plant operation (Fig. 1).

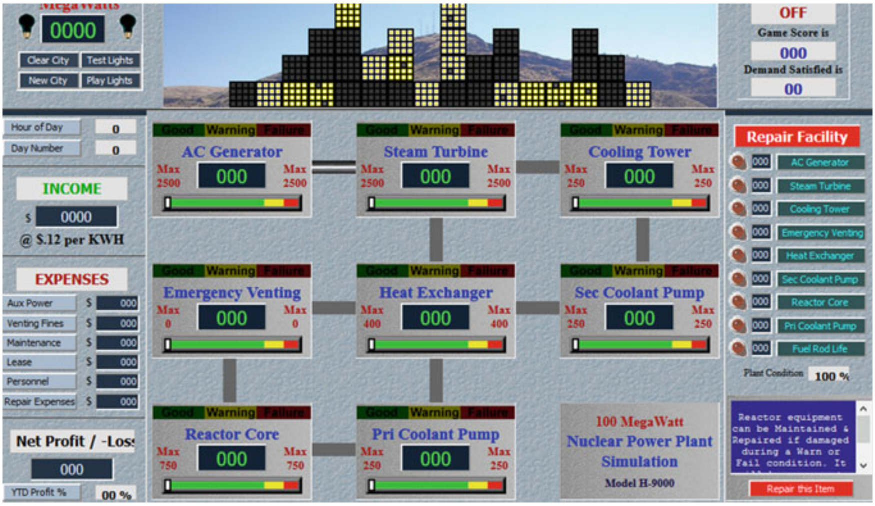

Fig. 1 The display with the variables utilized to control the nuclear reactor. *Source* http://www.nuclearpowersimulator.com/ [1]

The operator was subjected to three conditions in sequence, with sub-conditions, as described below:

1. Baseline: eyes closed and eyes open;
2. Cognitive tasks: selective attention, visual-spatial working memory and arithmetic evaluated from the cognitive assessment battery ProA;
3. Nuclear power plant simulator in a 2D computing environment (operation—events triggers).

All operators played the game two times in a continuous test and the following bio signals were recorded: GSR (Galvanic Skin Response); HRV (Heart Rate Variability); Eye Tracking, and EEG (Electroencephalography).

The individuals were grouped based on their performance, and it was constructed a matrix in which the following were analyzed on a multivariate regression:

- Rule–Based Learning (RBL)
- Knowledge in the nuclear area (KNA)
- Skills on HMI or game (SBG).

According Reason [2], the error is directly associated with performance levels as shown in Fig. 2.

Aiming to merge and link the various levels of knowledge, skill and rule with the performance and mistakes, the groups were analyzed individually comparing individuals within each control group.

After the identification of differences and similarities in operation (action) between the operators of each control group and the identification of possible

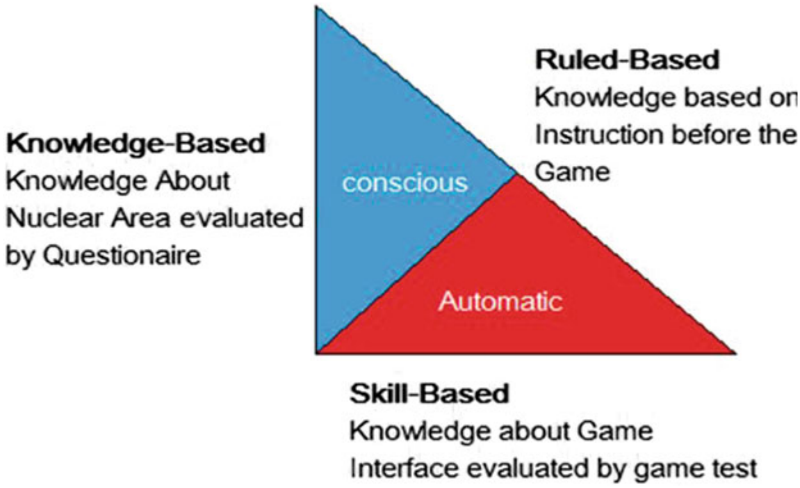

Fig. 2 The continuum between conscious and automatic behavior (Adapted from [2])

Table 1 Performance numerical classification

Performance	Numerical classification
Rule (R)	1-Yes/0-No
Knowledge (K)	1-Yes/0-No
Skill (S)	1-Yes/0-No

patterns of each level of performance, it was provided a binary sequence (RKS) based on Table 1. This sequence indicates the level of performance of each operator in its characteristics.

How operators receive the same set of instructions before the game, the basement rule was considered equal to 1 for all groups. Leaving assess only performance with bases in the knowledge and skills of operators. To characterize the groups and classify each individual by areas of knowledge and skill it was used a questionnaire with questions about the professional and academic activities of each individual, as well as an interview with each individual before the collection procedure. Therefore, individual performance is related to the foundation as a rule, specific knowledge and skills linking stressors moments that can compromise the operation to change patterns provided by bio-signals. Subsequently the operation, the following equations were developed to identify the influences based on rule (RBL), performance based on knowledge in the nuclear field (KNA), and skill based on knowledge of HMI or game (SBG), considering each operator. To treat it numerically, It was used a multivariate regression based on least squares method.

$$X\beta + \varepsilon = Y \tag{1}$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} RBL_{G1} & RBL_{G2} & RBL_{G3} & RBL_{G4} \\ KNA_{G1} & KNA_{G2} & KNA_{G3} & KNA_{G4} \\ SBG_{G1} & SBG_{G2} & SBG_{G3} & SBG_{G4} \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} R_{G1} & 0 & 0 & 0 \\ 0 & R_{G2} & 0 & 0 \\ 0 & 0 & R_{G3} & 0 \\ 0 & 0 & 0 & R_{G4} \end{bmatrix} \tag{2}$$

Based on the results, this study aims to quantify the percentage of evolution between operations (1) and (2) and their relationship with each of the biological signals, characterizing operators and groups. Performance was measured by an in game function that displays the profit generated by the operator over the task.

3 Results

3.1 Regression Based on Performance Data

The results presented in Table 2 can identify the weight of each level of performance by group. Equation (1) was developed by multivariate regression analyzing the result of the β coefficient that was normalized in percentage. The evolution of an operation was calculated by comparing game tasks 1 and 2 total profit.

The above results show the performance variation in both game activities, considering the groups division proposed by knowledge, skill and rule.

Table 2 Result from multivariate regression based on operator performance divided per group

Game 1				Game 2			
Group	RBL (%)	KNA (%)	SBG (%)	Group	RBL (%)	KNA (%)	SBG (%)
1	10	60	30	1	18	50	32
2	20	80	0	2	25	75	0
3	50	0	50	3	30	0	70
4	100	0	0	4	100	0	0

Performance improvement game 1–2	
Group	Improvement (%)
1	+16
2	+60
3	+95
4	+65

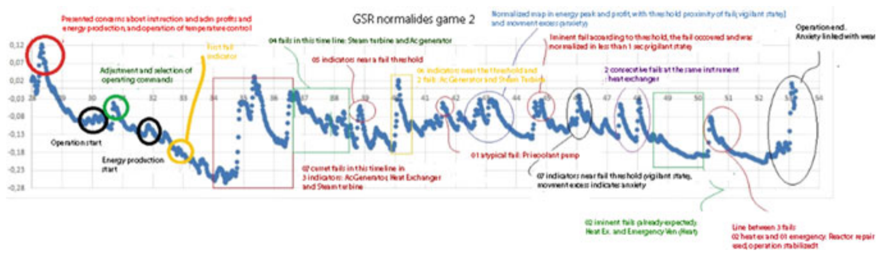

Fig. 3 Example of normalized GSR data from an operator during a game

The methodology presented can be applied in an operator training to be a tool to identify relevant aspects in a training individualization that depends on the level initially presented by the individual. Therefore, an individualized standard in performance based on training could reduce or increase the hours of training in certain groups depending on the knowledge of the individual.

The percentage of each performance level when analyzed by the control group has close patterns in bio signals among its individuals during operation, providing evidence on the mistakes and missteps committed by systemic pattern of bio signals per group when analyzed a specific error type. In following results will be better explained.

Even comparing directly related individuals between groups and individually, according to the performance levels we can classify the influence of each of the decision-making standards, including the technical and behavioral analysis of each individual.

3.2 Results from Bio Signals Data

During the experiments, the bio signals (GSR, temperature and HRV) data from each individual was collected from J&J hardware [3] analyzed with the bio explorer package.[1] In order to allow the comparison among different individuals in different groups, the data was normalized using the individual base line.

The signal obtained from the GSR sensor (Galvanic Skin Response) can be a very useful data to identify error during the operation. Figure 3 shows GSR variation from the operator during game.

The game it is possible to correlate warnings, errors or missteps committed by operator with peak of normalized GSR response.

The observation considers a peak as a variation at least 15 % greater than the previously value observed. This criteria were considered to analyze the collect data

---

[1]Bio-Explorer is a Windows program for real-time biophysical data acquisition, processing, and display. It is intended for personal use in entertainment, education, and experimentation.

and to provide a correlation with stress condition during an operation. Other important result is the duration of peak; The failure event usually ends when the problem observed after the warning, or the observation (of fail eminence) of an operator results in a counter measure adopted by the operator itself that normalizes the system. After the operator solved the problem, the stress stimulation disappears and the GSR signal reduces to the previous level.

The use of HRV (Heart Rate Variability) is not direct as GSR. The possibility to use HRV with others sensor is aim of futures studies. Prospective analyses to correlate EEG with HRV have been conducted by the authors.

3.3 Results from Eye Tracking

The principle of the eye tracker is the analysis of eye movement to assess subconscious cognitive processes. The method of analysis is done with the monitoring of eye movements using an infrared reflected light in the eye and then through a geometric model, it is determined the exact look of the attachment point. To assess the patterns of eye movement and fixation of the look, can be used the "heat map" that shows how much attention is directed to the menu, or the "plot gaze" which shows the pattern of visual user research before making a decision.

The principle of the eye tracker is the analysis of eye movement to demonstrate subconscious cognitive processes. The analysis method is done with the tracking of eye movements via infrared sensors, then applies a geometric model, given exactly the visual fixation point. To assess the movement patterns of the eyes and look fixation, its used a "heat map" or gaze plot that indicates how much attention is directed to activity, indicating the standard visual user review before making a decision. In developing the methodology proposed by the research, the use of this visual focus monitoring equipment allows the extraction of a large amount of relevant information about the error (failure) as levels and focus of attention, time and error types (failure), operating standards.

Another important point is the ability to quantify the level rule (reading time and steady focus on the task) based on the instruction provided to all operators equally as a factor to quantify performance levels, on closer analysis the individual error of investigation SKR following the criteria developed by Reason [2] used in the methodology.

Considering qualitative analysis of data provided by eye tracking, the same group individuals have a small variation of visual focus on operating instruments. This variation has shown the weight of each performance level in the decision-making cycle in relation to the focus of attention in control instruments according to each operators group. For example:

The control group 111 (level of rule-knowledge-skill) showed no significant change of focus in the instruments remaining faithful to the operation pattern based on their performance levels, varying only in intensity in a different game. Below the example of high-level operation pattern that can determine a standard to compare others operators. That pattern is a result of good rule-knowledge-skill based learning (Fig. 4).

Analyzing the group 110 (level of rule-knowledge), it presented a unique variance focus on instruments just by changing the operation pattern of trying to fix problems encountered by lack of skill (simulation and games) considering the user level of performance (SKR), thereby undermining the effectiveness of the group control system operation (Fig. 5).

Considering the group 101 (level of rule-skill), its operator showed significant variations focus on instruments ranging around 2–3 instruments for the operation in order to remedy problems caused by lack of performance level knowledge in the nuclear field (Fig. 6).

Fig. 4 Gaze plot control group 111-Pattern game 1 and game 2

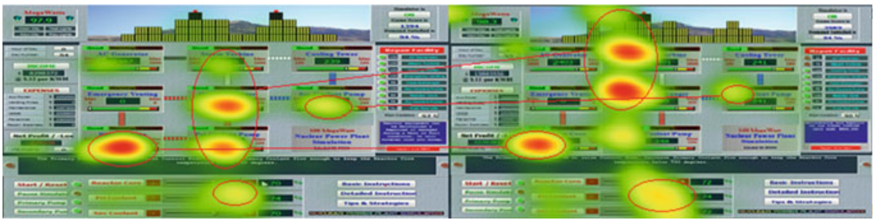

Fig. 5 Gaze plot control group individual 110-Pattern game 1 and game 2

Fig. 6 Gaze plot control group 101-Pattern game 1 and game 2

Fig. 7 Gaze plot control group 100

Finally, the group 100 (level of rule) showed the worst result among the four control groups, ranging several times the visual focus on the control instruments. This is due to lack of basis of performance levels (Fig. 7).

The Eye tracking was a useful tool to identify of operation based on gaze plots pattern of performance. There was a decreasing focus point when analyzing from groups 1–4. It is possible to quantify how much a lack of one level of performance impact in operation performance crossing the data of Table 2 and results from eye tracking. In this game, a good performer has 6 focus points in correct place; the knowledge level guide to identify and the skill level guide to transform this knowledge in correct action.

3.4 Results from EEG

The electroencephalogram (EEG) was continuous recorded by a 19 channels elec-trode cap (Electro-Cap International Inc.), according to the 10–20 International Elec-trode Placement System on a 24 channel MITSAR 202 EEG machine [4] in a monopolar montage referenced to linked ears. The EEG was recorded for 5 min for either eyes open and eyes closed condition and continuously for all the cognitive tasks and during the simulation as well. The recording bandwidth was 0.5–70 Hz with a notch filter from 55 to 65 and sampling rate of 250 Hz. The data was analyzed using WINEEG software, and the off-line artifacts' removal procedures was followed. Artifact correction was done using independent component analyses tool (ICA) to correct eyes blinks and eyes movement artifacts, followed by a visual inspection of the EEG signals to remove remaining artifacts. A Fast Fourier transform was used to separate the frequency bands and the absolute power (AP = μV2) was calculated for the following bands: Theta: 4–8 Hz, Alpha: 8–12 Hz, and Beta: 12–21 Hz) due to their importance when investigating cognitive load.

For this paper purposes, i.e. setting up a methodology to verify EEG changes associated to operator failure, the analyses were run individually for each subject comparing the mean difference on spectra activation from 30 s before an error occurrence to 30 s following it. The error occurrence presented here was Primary Coolant Pump flashing light alarm. As an example, Fig. 8 shows the difference in the activation pattern of four subjects (G1, G2, G3 and G4), one from each group. Each subject has 3 maps, representing the 3 frequency bands analyzed.

From Fig. 8, it is possible to see different EEG activation patterns associate to the same error at the same task: while some operator increase theta (G1 and G3), which reflects information processing usually associated to encoding new information, G2 didn't show differences in this pattern, due less information process. In the other hand, a decrease in alpha activity, observed in the G2 operator, is usually associated to in-crease cognitive load and task difficulty. Considering that G2 was from the group of individuals with a higher knowledge on nuclear power plant, but not on IHM, this pre-analysis draws attention to the importance of the subjective and individual variables associated to the operator performance, instead of relying solely on knowledge.

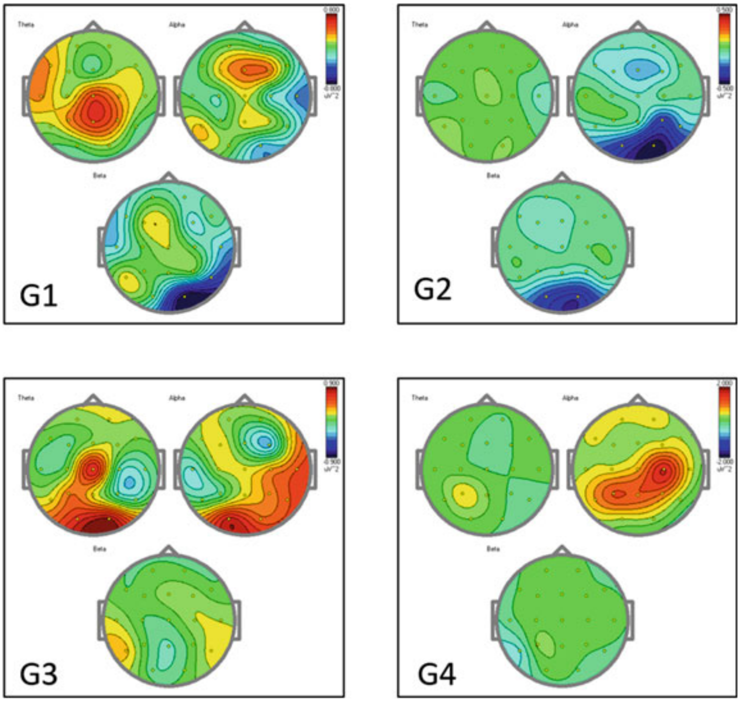

Fig. 8 Scale ranges from *blue* to *red* meaning respectively decreasing or increasing in individual EEG brain pattern (deactivation or activation) from 30 s pre event to 30 s post event. Note that *G1*, *G2*, *G3* and *G4* have different power scale, based on the individual magnitude of the changes. *Green* colors represent non-significant differences

4 Conclusion

This research that aims to correlate the bio signals with error analysis is only in its begging stage. The proposal tries to establish a classification standard based on performance to figure out which characteristic of a group of operator are relevant to find the better characteristic to development the pattern that are suitable to operation.

Based on this initial issues are possible to qualify good information provided by eye tracking. Comparing the results, it was possible to establish a pattern of good operation.

The use of GSR allows the identification of stress condition and how each operator handles the situation when it is necessary an intervention. The current study found a peak variation on GSR signal during an individual stress condition that suggests the needs of special attention to solve the problem and further studies will try to correlate response time with normalized GSR values to better understand this psychology behavior.

Finally, the prospective use of EEG shows a brain activation area array per group of operators and allowed linking the measurement with proposed classification for a specific error (primary coolant bomb) for all operators.

References

1. Nuclear Power Simulator-Game, Majik Mike Simulators. Information, http://www.nuclearpowersimulator.com
2. Reason, J. (1990). Human Error. Cambridge University Press
3. Software J&J Engineering, J&J Engineering. Inc. Information, http://jjengineering.com
4. Software Mitsar - EEG systems, Mitsar Medical, Petersburg, Russia. Inc. Information http://mitsar-medical.com

Framework for Evaluating the Impact of Environmental Conditions on Manual Actions

Garill A. Coles, Angela C. Dalton, Kristi M. Branch, Rajiv Prasad, Nancy P. Kohn and Alvah C. Bittner

Abstract Ensuring successful protection from and mitigation of external floods at nuclear power plants (NPPs) has received increasing attention in the wake of the Fukushima nuclear accident. Following the incident, the U.S. Nuclear Regulatory Commission (NRC) required all operating U.S. NPPs to identify nonconforming conditions and to verify the adequacy of monitoring and response procedures. Additional NRC initiatives aim to ensure that manual actions, i.e. actions taken outside of the main control room for flood protection and mitigation, are both feasible and reliable. We developed a framework to identify the key components and relationships required for an analytical approach or model to assess the impacts of environmental conditions (ECs) on the ability of individuals to perform flood protection and mitigation manual actions.

Keywords Human performance · Manual actions · Environmental conditions · Flood protection and mitigation · Nuclear power plants

G.A. Coles · A.C. Dalton · R. Prasad (✉)
Pacific Northwest National Laboratory, P.O. Box 999, Richland, WA 99352, USA
e-mail: Rajiv.Prasad@PNNL.gov

G.A. Coles
e-mail: Garill.Coles@PNNL.gov

A.C. Dalton
e-mail: Angela.Dalton@PNNL.gov

K.M. Branch
PNNL, 100 Dexter Ave. N #400, Seattle, WA 98109, USA
e-mail: Kristi.Branch@PNNL.gov

N.P. Kohn
PNNL, 1529 W Sequim Bay Rd., Sequim, WA 98382, USA
e-mail: Nancy.Kohn@PNNL.gov

A.C. Bittner
Bittner and Associates, 13839 SE, 260th St., Kent, WA 98042, USA
e-mail: DrBittner@comcast.net

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_14

1 Introduction

Following Japan's March 11, 2011, Tohoku earthquake and subsequent tsunami, which led to the Fukushima nuclear accident, the U.S. Nuclear Regulatory Commission (NRC) required all operating U.S. nuclear power plants (NPPs) to conduct flooding walkdowns aimed at identifying and addressing degraded, nonconforming, or unanalyzed conditions, and verifying the adequacy of their monitoring and response procedures. The NRC also undertook several initiatives to improve understanding of and to ensure the feasibility and reliability of required actions taken outside of the control room for flood protection and mitigation. Such actions are referred to in this paper as "manual actions." Another initiative was to assemble information about the environmental conditions (ECs) accompanying a flooding event and the impact of those ECs on the performance of those manual actions. This work would update and expand on *NUREG/CR-5680*, *The Impact of Environmental Conditions on Human Performance*, 1994 (henceforth *NUREG/CR-5680*), which reviewed research completed prior to the mid-1990s on the impact of ECs on human performance [1]. The purpose of *NUREG/CR-5680* was to provide information and technical guidance on how exposure to certain ECs could affect human performance; however, its scope was limited to consideration of the impact of heat, cold, noise, vibration, and lighting. This new work will assist NRC staff who review and evaluate plans for events in which exposure to ECs may result in considerable stress on those attempting to perform the manual actions. In support of this effort, we at Pacific Northwest National Laboratory (PNNL) developed a conceptual framework that identifies the underlying concepts and information required to assess the impacts of ECs on performance of flood protection and mitigation actions.

Our conceptual framework is based on a review of literature on ECs associated with external flooding, actions taken for flood protection and mitigation, and human factors research on ECs and performance stressors. This framework will facilitate development of an analytical approach or model for assessing the effects of ECs on human actions and guide identification of the concepts, analytical tools, and data required for site-specific NPP applications. We are actively working out the details of a site-specific application; however, the framework remains to be tested on actual NPPs. Consequently, some aspects of the framework presented here may be subject to modification based on the information obtained and experience gained during exercises conducted at specific NPPs.

2 Conceptual Framework

Figure 1 illustrates our framework, which has three basic elements: (1) characterization of manual actions, (2) characterization of potential ECs, and (3) characterization of potential impacts of ECs on manual actions. The purple box on the left

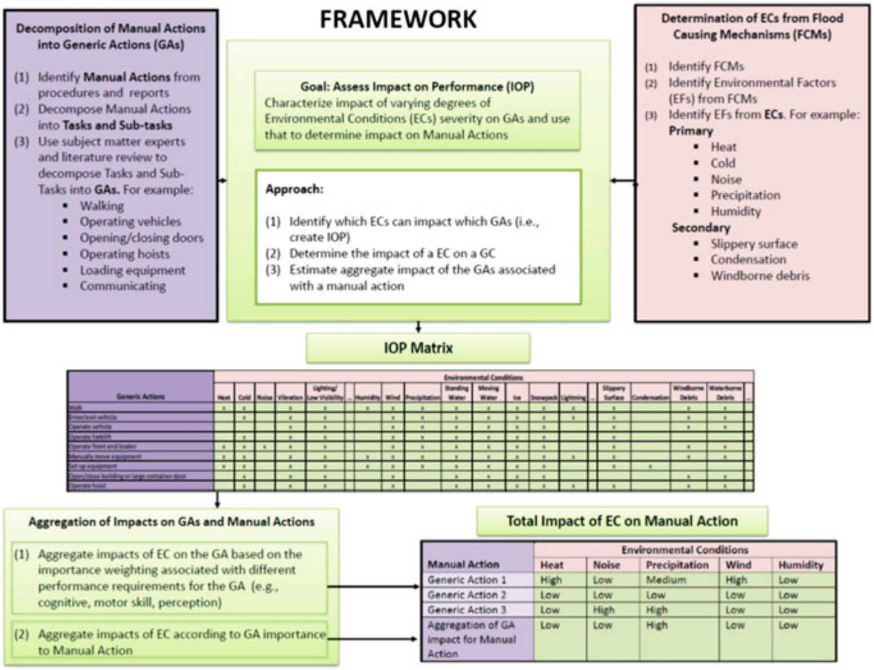

Fig. 1 Conceptual framework for assessing impact of ECs on operator actions

side of this figure depicts the characterization of manual actions. Specifically, manual actions are decomposed into elements such as tasks and subtasks, which are further decomposed into generic actions (GAs). Characterization of potential ECs, including how they are identified, is shown in the pink box on the right side of the figure. In turn, characterization of the potential impacts of ECs on manual actions is shown in the four green boxes in the center of the figure. The impact of a specific EC on a GA is estimated based on extant research, and impacts are aggregated from the level of GAs to the level of manual action following the pathway established by the decomposition process.

The first element of our framework is the identification and decomposition of manual actions (e.g., installing sump pumps) into GAs (e.g., lifting, walking, loading equipment) and the assessment of ECs on the manual actions. In this process, manual actions, tasks, subtasks, and GAs are described in terms of their cognitive and non-cognitive, gross and fine motor movement, and communication and coordination demands. Because much of the extant research addresses EC impact on performance (IOP) at a level of detail that is similar to GAs and their performance demands rather than manual actions, the decomposition of manual actions into GAs provides a more direct way to apply information about the impact of ECs from the literature. Decomposing manual actions into GAs also provides the flexibility needed to address variability in the implementation of manual actions subject to site-specific terrains and layouts.

The second element—identification of ECs associated with flood-causing mechanisms (e.g., hydrometeorological or geoseismic)—characterizes ECs that may prevail during flooding events. Our literature search and review focused on these ECs. We identified the ECs by examining each flood-causing mechanism (and combinations) for ECs that would occur concurrently with the mechanisms. The geographic area of focus was the United States.

The third framework element treats GAs as building blocks on which impacts of ECs are evaluated. In an NPP site-specific application, the analyst would develop a matrix showing where site-specific exposure-effect relationships exist, as illustrated in Fig. 1 in the box under the IOP Matrix label. After applying research findings on the impact of the ECs on the GAs, the manual actions would be recomposed, yielding an estimate of the impact of ECs on the manual action as a whole. Manual actions may fail if one or more of their component GAs cannot be completed or if the cumulative impact of ECs prevents completion of the action in the available time. The framework also accommodates consideration of variation in ECs due to changes in the severity of ECs over time and space, and the moderating effects of sheltering and/or use of protective equipment during manual action performance. In the remainder of this paper, we detail aspects of the approach and information used to develop each element of the framework.

3 Characterization of Manual Actions

Manual actions designed for flood protection and mitigation are defined as those actions that are being credited by NPPs and are taken or directed by plant staff to mitigate or prevent impacts associated with an external flooding event. Typical manual actions might include installing sump pumps and flood barriers. Though "manual action" might imply a set of simple movements, in the nuclear industry and our framework, it refers to sets of actions that are typically much more complex. Specifically, manual actions often involve a set of related tasks or sequence of steps that may be performed at different locations and may require communication and coordination among operators. Consequently, clear understanding of such manual actions and the ECs influencing them requires their decomposition into constitutive elements (e.g., tasks, subtasks, and GAs). This characterization must be detailed enough to determine the impact of ECs on the constitutive elements.

3.1 Identify Manual Actions from Available NPP Reports

As a consequence of the Tohoku earthquake and tsunami "lessons-learned," the NRC asked U.S. NPPs to conduct flooding walkdowns aimed at identifying and addressing degraded, nonconforming, or unanalyzed conditions, and verifying the adequacy of their monitoring and response procedures. We reviewed the NPP licensees' flooding walkdown reports, along with NRC staff assessment of those reports, to (1) obtain a high-level understanding of the manual actions and (2) identify those that occurred at multiple NPPs. We subsequently analyzed the identified manual actions and grouped them both by type and whether those performing them would be wholly or partially exposed to ECs (e.g., strong wind and/or precipitation).

We also identified manual actions by reviewing available site-specific NPP flood protection and mitigation procedures. Though our access to these documents was limited, we identified and reviewed several procedures that relied on complex and resource-intensive manual actions designed to respond to a variety of flooding events (e.g., dam failure, plant flooding, turbine building flooding, containment flooding, and site flooding). These procedures also included guidance for operating under severe weather conditions and natural disasters (e.g., hurricanes, tornadoes, tropical storms, severe thunderstorms, and seismic events). We found that the majority of these procedure-specific manual actions fit well into the categories of manual actions from the flooding walkdown reports. This two-pronged approach confirmed that we had identified a reasonably representative set of manual actions, which consisted of the following (many of which occur outdoors fully exposed to several ECs):

- deploy sandbags and build berms
- place flood barriers
- close doors, gates, hatches, and manhole covers
- secure drains, close valves, and seal openings
- setup and operate portable pump and sumps
- setup and operate diesel generators and stage fuel
- equalize pressure (open doors, weight floor)
- seal fuel vents and cover air intakes
- monitor leakage, hazards, weather, and debris
- clear debris from intake structure or haul path
- move equipment to higher elevations
- de-energize and adjust electrical power
- operate installed plant sump or pump systems
- connect piping spool to alternate cooling source
- connect electrical jumper to alternate power source
- monitor leakage, hazards, weather, and debris.

3.2 Identify Tasks, Subtasks, and Generic Actions

Manual actions are often complex. They may consist of multiple steps, involve sequential movements, a combination of motor and cognitive functions and

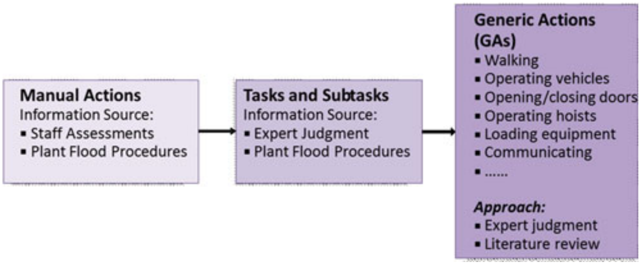

Fig. 2 Decomposition of manual actions

processes, be implemented in more than one location, and employ varying levels of automation and/or tools or equipment. To identify how ECs impact the performance of manual actions, we concluded that it was necessary to decompose them into discrete, lower-level actions. Figure 2 illustrates the hierarchical decomposition of manual actions into progressively finer-grained units of analysis. The assumptions underlying this process are that (1) an EC may impact different elements differentially and (2) the impact of different ECs on a given element may be similar or markedly differ.

Equipped with an understanding of what typical manual actions might be required at NPPs, we decomposed the manual actions into tasks and subtasks based on a combination of reference document and expert judgment sources. Informed by research on task analysis [2, 3], we defined tasks as "consisting of the actions required or believed to be necessary for an individual or a team of individuals to achieve a pre-determined system goal, using appropriate devices as needed." Tasks and any constituent subtasks typically consist of a mixture of both cognitive and non-cognitive elements, adding complexity that can affect both operator speed and reliability [4].

We further decomposed the tasks and subtasks associated with a manual action into GAs (e.g., walking, lifting, opening doors, and operating hoists). The analytical objective was to identify a representative set of GAs, which could function as somewhat independent "building blocks" that could be aggregated as needed to represent a broad range of manual actions. The impacts of ECs on constituent GAs would then be aggregated to estimate the impact of various ECs on the manual actions. If different GAs are shown to be impacted similarly by an EC or if different ECs are shown to impact a GA similarly, the analysis of impacts can be simplified by grouping the GA and ECs according to these similarities. To ensure that the GAs were characterized and represented consistently, we developed and applied the following criteria to guide the decomposition of manual actions into GAs.

1. GAs should be general in nature and function as "building blocks" that can be used for recomposing or decomposing other manual actions
2. GAs should be associated with accomplishing a functional objective (e.g., a step in a procedure)

3. GAs performed in an unsheltered location should be distinguished from those performed in a sheltered or semi-sheltered location
4. GAs requiring a higher level of cognitive support should be distinguished from those not requiring a high level of cognitive support (though perhaps requiring fine or gross motor skills)
5. GAs should be defined at a level consistent with efficient applications of findings from the research and modeling literature addressing EC effects on human performance.

Using these criteria, we have decomposed a limited number of typical manual actions into GAs, presented here as examples:

- walk
- enter or exit vehicles (transport vehicles, light and heavy equipment)
- operate a transport vehicle
- operate forklift
- operate frontend loader
- manually move equipment
- setup equipment
- open building or large container door
- operate powered hoist
- manually lift and move heavy materials or equipment
- work manually with simple equipment
- use hand tools
- communicate electronically
- communicate non-electronically.

4 Environmental Conditions

At the beginning of the project, we identified both the relevant set of external flood-causing mechanisms and the range of ECs that could occur during a flood of interest at U.S. NPP sites. Floods of interest are those that trigger flood protection and mitigation procedures at NPPs. The external flood-causing mechanisms evaluated for NPP safety include [5–7]:

- floods from local intense precipitation
- floods in rivers and streams
- floods from breaches of dams or failure of water-storage structures
- floods from storm surges and seiches
- floods from tsunamis
- floods from failures of ice dams or backwater effects from ice jams
- floods from geomorphic changes to river or stream channels.

Table 1 Identification of ECs associated with floods of interest

Flood-causing mechanisms and contributing events considered in combination	ECs that could affect manual actions
Local intense precipitation	Primary ECs (at the time the manual action is being performed)
Streams and rivers	Heat (high temperature)[a]
Dam/structure failures	Cold (low temperature)[a]
Storm surges	Relative humidity
Seiches	Precipitation type and intensity
Tsunamis	Wind velocity
Ice dams/ice jams	Noise level[a]
Channel diversion/migration	Water depth
Contributors to Event Combinations	Water velocity
Concurrent wind-induced wave activity	Vibration frequency and intensity[a]
Antecedent or subsequent precipitation	Lighting level[a]/low visibility
Snowpack, snowmelt, rain-on-snow	Presence of ice
Dam failure concurrent with riverine flood	Snow depth
Earthquakes	Presence of lightning
Concurrent high tides	Secondary ECs
	Slippery or muddy surfaces
	Condensation
	Windborne debris
	Waterborne debris

[a]ECs reviewed in *NUREG/CR-5680* [1]. Lighting was retained, but expanded to include the "outdoor light level" change that could occur with several flood-causing mechanisms. Secondary ECs occur in presence of an initiating or enabling primary EC

Table 1 shows the varied hydrometeorological and geoseismic processes that comprise flood-causing mechanisms and the varied hydrometeorological and geomorphic NPP settings result in a range of ECs that could co-occur with a flood. The ECs considered in NUREG/CR-5680 [1]—heat, cold, noise, vibration, and lighting—could be present with a number of flood-causing mechanisms. However, NUREG/CR-5680 did not address either the range of severity of those conditions or the other ECs that could be encountered outdoors during an external flooding event. In Table 1, we list the more comprehensive set of mechanisms and ECs included in our literature search.

5 Characterization of Impacts of ECs on Manual Actions

With an understanding of how manual actions and ECs may be characterized, we describe the third element of our framework in this section. Typical impacts from primary ECs could include, for example, (1) the impact of high wind on the ability to work manually with simple equipment, (2) the impact of precipitation on operating vehicles and light or heavy equipment, and (3) the impact of noise on electronic and non-electronic communication. Typical impacts from secondary ECs could include, for example, (1) the impact of waterborne debris on manually moving equipment and (2) the impact of slippery surfaces on walking. Some manual actions may be performed partially or wholly in sheltered locations. Although operators performing these manual actions (or portions of them) may therefore not be exposed to conditions such as wind or precipitation, they may be exposed to other conditions that affect performance, such as increased humidity or heat.

Our approach to assessing how ECs impact manual actions starts with an investigation of the extent to which operators performing the GAs comprising the manual action could be impacted by the ECs. We treat GAs as building blocks on which impacts of ECs are evaluated. Analyses can be complex because (1) the same GAs may occur multiple times during the performance of a manual action, (2) ECs may vary during the course of manual action performance due to changes in EC severity, (3) variations from the moderating effects of sheltering, and/or (4) use of protective equipment. Research has shown that different GAs place different performance demands on operators and that ECs may affect these performance demands differently. Therefore, to support the analysis, the performance demands of the GAs must be systematically categorized.

To operationalize the notion of performance demands, we drew principally on three sources: (1) *NUREG/CR-5680* Volume 2 [1] on performance abilities; (2) work by O'Brien et al. [8] and others on human characteristics important to evaluating task performance; and (3) macro-cognitive functions identified by *NUREG-2114* [8] that provide the basis for human reliability analysis (HRA) of NPP operator actions.

NUREG/CR-5680 Volume 2 defines "performance abilities" as the human capabilities, such as perception or psychomotor skills that are necessary to perform tasks. It also identifies a spectrum of performance abilities applicable to tasks associated with NPP operations and maintenance that are known to be sensitive to exposure to ECs. These performance abilities are: (1) attention, (2) vision, (3) perception, (4) psychomotor skill, (5) manual dexterity, (6) cognitive functions, and (7) mood and comfort.

O'Brien et al. [8] describes another taxonomy, which was drawn from his and previous personnel selection research [10]. It is explicitly focused on prediction of operator performance under variable conditions, such as exposure to different ECs. The characteristics in O'Brien's task taxonomy, which he referred to as "taxons," are: (1) perception, (2) cognition–numerical analysis, (3) cognition–information processing and problem solving, (4) motor—fine motor discrete, (5) motor—fine motor continuous, (6) motor—gross motor light, (7) motor—gross motor heavy, (8) communication—oral, and (9) communication—reading and writing.

Whaley et al., in *NUREG-2114* [9], provide another taxonomy of interest. They describe a unified approach to HRA that quantitatively determines the failure probability of actions required by NPP operators to put the plant in a safe state. Their broad review of literature addressing cognitive psychology, behavioral psychology, neuropsychology, human factors, and human performance identified five cognitive functions that form the basis of their HRA approach, the Integrated

Decision-tree Human Error Analysis System (IDHEAS): (1) detecting and noticing, (2) understanding and sense making, (3) decision making, (4) action, and (5) teamwork.

Based on these three sources, we developed the following consolidated classification scheme of performance demands:

1. Detecting and Noticing—Attention, memory, vigilance, switching, acuity, perception, and threshold perception, Sensation and visual recognition
2. Understanding and Sensemaking—Pattern recognition, discrimination, understanding, evaluating, hypothesizing, diagnosing, and integrating
3. Decision Making—Reasoning, computation, interpreting, classifying, goal setting, planning, adapting, and evaluating and selecting options
4. Action—Fine motor skills—discrete and motor continuous, and manual dexterity, Gross motor skills—heavy and light, Other neuropsychological functions
5. Teamwork—Reading and writing, Oral face-to-face and electronic communication, Cooperation, crew interaction, and command and control.

To estimate the impact of ECs on GAs and provide a basis for aggregating impacts to the manual action level, it is also necessary to evaluate and quantify the cognitive and non-cognitive demands associated with performing each GA, using the consolidated classification scheme. Performing different GAs, such as operating a vehicle or setting up equipment, requires meeting a mixture of cognitive (e.g., attention and memory that consume mental energy) and non-cognitive demands (e.g., gross and fine motor movements that consume physical energy) [11, 12] that will vary by GA. We are currently working to develop a method for quantifying performance demands in relation to ECs. Here we provide a preliminary discussion of our initial conceptualization of a proportional approach that might be serve this purpose, acknowledging that additional technical details remain to be developed (e.g., the dimensions and scale needed for fleshing out this approach).

With the proposed proportional approach, GAs would be characterized in terms of the combination of relative performance demands that must be met for their completion. This conceptualization can be visualized as pie charts, as illustrated for two different GAs in Fig. 3. Such percentage information provides a basis for weighting the impact of an EC on performance demands commensurate with their importance to successful completion of the GA. For example, one of the GAs associated with a manual action (i.e., Building a Sandbag Berm) is “work with simple equipment” to take filled sandbags off the rack. This GA has a low level of cognitive performance demand. The relative performance demands for this GA might be similar to that shown Fig. 3, in which the dominant requirement is for gross motor skills. This contrasts with another GA (i.e., Installation of a Portable Pump) to “operate a powered hoist” to lift the portable pump into place, which might be similar to that shown in Fig. 3 because the dominant performance demands are Detecting and Noticing, but also include Gross Motor Skills-Light and Fine Motor Skills.

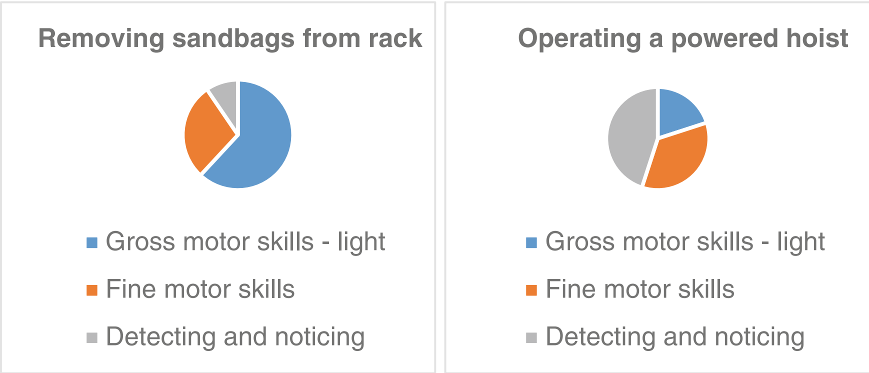

Fig. 3 Nominal illustration of a GA requiring a relatively low level of cognitive performance versus requiring a higher level of cognitive performance

A fundamental premise of the concept illustrated in Fig. 1 is that the process of decomposition and recomposition or aggregation provides a useful approach for assessing the impact of individual and/or combinations of ECs. A method for recomposing the GAs back into the manual actions in a process that yields an estimate of the impact of ECs on the manual actions as a whole is consequently an important requirement of our framework. The methodology for this aggregation must account for the relative importance of each GA to the subtask, task, and manual action. We are postulating that this can be accomplished by developing a method for weighting the GAs in terms of their relative contribution to the performance of the subtask, task, and manual action.

We anticipate that the impacts of ECs will be measured primarily as an increase in the time it takes to perform a manual action. This type of impact seems to be well addressed in the literature. The literature also shows that an increased likelihood of an undetected error (e.g., a step is not performed that has the effect of failing the manual action) can sometimes be the most significant impact from ECs. However, the relatively protracted time most flood protection and mitigation actions appear to take (e.g., building berms and placing barriers) may provide the opportunity to detect and correct many of these types of errors, thus reducing their IOP.

The complexities associated with applying the information about EC impacts from literature to GAs remains to be fully addressed in our research. For example, the framework—as displayed in Fig. 1—currently shows a two-dimensional IOP matrix, indicating cases (i.e., with "X") at intersections where GAs could be impacted by the different ECs. Our ongoing research will address (1) the degree of impact on the GA from the EC, (2) the degree of impact given the severity of the EC, (3) the degree of impact given the type of manual action being considered, and (4) the impact from multiple ECs in the IOP matrix.

Toward addressing these issues, we are conducting—as noted earlier—an ongoing literature review focused on the impacts of salient ECs on human performance, especially at the GA level. This literature search was initiated by three avenues of search: (1) effects of ECs (Table 1) on basic manual handling tasks (i.e., lifting, lowering, pushing, pulling, manual manipulation, and carrying [walking with/without loads]), (2) models of EC effects on manual handling and capabilities identified in *NUREG/CR-5680*, and (3) consultation with cognizant researchers. Citations to the initially identified "tentative key" literature items (e.g., journal or proceeding articles and agency reports) were then "forward searched" for more recent tentative key items, especially integrated reviews, meta-analyses, and/or highly cited items. In some cases, tentative key items were eliminated when found to be dominated by another identified item (e.g., researcher update of earlier review). Those surviving were also successively forward searched until the process became unproductive (typically with final key items relatively contemporaneous). Not surprisingly, an early 2016 snapshot of 237 items in the literature base found >30 % published between 2011 and 2016 (and ~60 % last decade). Thus far, this process has already identified studies on the impact of ECs on performance in which the degree of impact is well-characterized as a function of the severity of the EC (e.g., temperature or noise level) for a given activity (e.g., walking or driving).

6 Conclusions

Our research on the impact of ECs associated with external flooding on NPP flood protection and mitigation manual actions is still underway; however, we have made significant inroads by developing a conceptual framework based on extensive review of literature on ECs associated with external flooding, flood protection and mitigation actions, and human factors studies on ECs and performance stressors. This framework may be used to inform an NPP-specific evaluation of the effects of the ECs associated with external flooding on manual actions. The framework likewise provides guidance on (1) characterization of manual actions, (2) characterization of potential ECs, and (3) characterization of potential impacts of ECs on manual actions. Thus far, we have demonstrated how a manual action may be decomposed into GAs, how impacts to GAs can be determined, and how aggregation of the impact of ECs on GAs can be performed. We are in the process of completing our review of the literature on the impact of ECs on manual actions, further developing the methodology for applying the research findings to the performance of manual actions, and developing and implementing a test case or demonstration of the application of our methodology to a manual action exposed to a set of ECs.

In future work, the framework will be contextualized to guide the identification of the concepts and data required for site-specific evaluations. The concepts incorporated into this framework have broad applicability beyond the nuclear power industry. This work will also contribute to the research on human performance and HRAs by updating and extending the NRC's research in *NUREG/CR-5680*, to address a broader range of ECs.

Acknowledgments We performed this work at PNNL funded by the NRC under Contract Number NRC-HQ-60-14-D-0028. We gratefully acknowledge NRC's support for this work. The opinions expressed in this paper are those of the authors only and should not be interpreted as official NRC positions. PNNL is operated for the U.S. Department of Energy by Battelle under Contract DE-AC05-76RL01830.

References

1. Echeverria, D., Barnes, V., Bittner, A., Durbin, N., Fawcett-Long, J., Moore, C., Slavich, A., Terrill, B., Westra, C., Wieringa, D., Wilson, R., Draper, D., Morisseau, D., Persensky, J.: The impact of environmental conditions on human performance: A handbook of environmental exposure. NUREG/CR-5680 Vol. 1 and 2; PNL-7675. Pacific Northwest Laboratory, Seattle, Washington (1994)
2. Preece, J., Rodgers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T.: Human-Computer Interaction. Addison-Wesley, Harlow, England (1994)
3. Kirwan, B., Ainsworth, L.K. (eds.): A Guide to Task Analysis. Taylor & Francis, London, UK (1992)
4. Bedny, G., Karwowski, W., Bedny, I.: Task and Its Complexity Assessment. In: Marek, T., Karwowski, W., Frankowicz, M., Kantola, J., Zgaga, P. (eds.) Human Factors of Global Society: A System of Systems Perspective, pp. 203–210. CRC Press, Boca Raton, FL (2014)
5. NRC (U.S. Nuclear Regulatory Commission).: Design Basis Flood for Nuclear Power Plants. Regulatory Guide 1.59, Rev. 2. Washington, D.C. (1977)
6. NRC (U.S. Nuclear Regulatory Commission).: Standard Review Plan for the Review of Safety Analysis Reports for Nuclear Power Plants: LWR Edition, NUREG-0800, Revision 3, Section 2.4. Available at http://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr0800/ (2007)
7. Prasad, R., Hibler, L.F., Coleman, A.M., Ward, D.L.: Design-basis flood estimation for site characterization at nuclear power plants in the United States of America. NUREG/CR-7046, U.S. Nuclear Regulatory Commission, Washington, D.C. Available at http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr7046/ (2011)
8. O'Brien, L.H., Simon, R., Swaminathan, H.: Development of the personal-based system evaluation aid (PER-SEVAL) performance shaping functions. ARI Research Note 92-50, United States Army Research Institute for the Behavioral and Social Sciences (1992)
9. Whaley, A., Xing, J., Boring, R., Hendrickson, S., Joe, J., Le Blanc, K., Morrow, S.: Cognitive basis for human reliability analysis. NUREG-2114. Nuclear Regulatory Commission (NRC) Office of Nuclear Regulatory Research (2015)
10. Fleishman, E., Quaintance, M.: Taxonomies of Human Performance: The Description of Human Tasks. Academic Press, San Diego, CA (1984)
11. Genaidy, A., Karwoski, W., Succop, P., Kwon, Y., Alhemoud, A., Goyal, D.: Classification system for characterization of physical and non-physical work factors. Int. J. Occup. Saf. Ergon. **6**(4), 535–555 (2012)
12. Maras, W., Hancock, P.: Putting mind and body back together: A human-systems approach to the integration of the physical and cognitive dimensions of task design and operations. Appl. Ergon **45**, 55–60 (2014)

A Virtual Reality Based Approach to Improve Human Performance and to Minimize Safety Risks When Operating Power Electric Systems

Alexandre Cardoso, Paulo R. Prado, Gerson F.M. Lima and Edgard Lamounier

Abstract Power systems require continuous operation for reasons of public safety, emergency management, national security and business continuity. Companies today control an electric system by means of 2D line diagrams, whereas a substation in the field is a 3D space. There exist situations where new control center operators have never been immersed into a real substation environment. When these operators visit a real electric substation, the environment is at minimum 'strange'. This fact unquestionably reduces human performance when it comes to operating the electrical system, since a great deal of mental effort is required by the operator to associate both 2D and 3D worlds. There are situations where some modifications and replacements have to be executed within the real substation environment. Hence, to design such procedures on the 2D line diagram does not adequately reflect the reality of the field. For example, it is impossible, in this 2D scenario, to design the route taken by a truck carrying a huge electric component. In this case, safety factors also arise and need to be given due attention. It is important to seek new alternatives to ensure that systems are designed in a manner as to optimize human performance and minimizes risks, thus producing higher productivity, health and safety in the work place and safety in work processes. On the other hand, Virtual Reality (VR) is known as providing "the feeling of being there". With the features provided by VR, it is possible to simulate all real operations of an electric substation with such precision that it has bearing on real world environments. For this reason, this paper proposes a Virtual Reality approach for the simulation,

A. Cardoso (✉) · G.F.M. Lima · E. Lamounier
UFU-Federal University of Uberlândia, Uberlandia, Minas Gerais, Brazil
e-mail: Alexandre@ufu.br

G.F.M. Lima
e-mail: gersonlima@ieee.org

E. Lamounier
e-mail: lamounier@ufu.br

P.R. Prado
CEMIG-Electrical Company of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
e-mail: prmprado@cemig.com.br

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_15

training and control of electric substations. In this approach, a virtual substation is realistically replicated according to its dimensions, using electric component data sheets, pictures, videos and floor plans. This is relevant as safety rules state that the distance between electrical components must be taken into account. Next, by means of a web service, data from a supervisory system is allocated to each component in the virtual substation, so the operator can attain access to all the information required for possible intervention, as is the case in real life. It is believed that all the features explored in this work have the capacity to increase human performance when operating a power electric substation.

Keywords Virtual reality · Power transmission substations · Time-to-Market · CAD-to-VR · Safety risks operating

1 Introduction

The recent use of Virtual Reality techniques in power systems operations (generation and transmission of electric energy) has identified new paradigms for the routine monitoring and control activities concerning critical engineering systems. The existence of a three-dimensional model of a substation's electrical equipment, which is true to its actual structure, minimizes the differences in the operational mental model for those professionals that work in remote control rooms. Besides, studies have shown that this approach facilitates team communication [1–3].

The term Virtual Reality (VR) has many definitions, due to its interdisciplinary nature and evolution. According to [1], Virtual Reality can be defined as a way for users to visualize, manipulate and interact with computers and extremely complex data, in which ideas such as immersion, interaction and involvement with the virtual environment are considered imperative. One of the main advantages of this technology is the broad involvement of the human senses in man-machine interaction, with impacts of improved visualization components and assimilation of content (learning and training). Training procedure applications in Virtual Reality can be used for several purposes, among which are, teaching procedures for assembly, disassembly, maintenance and operation of complex machinery or performing activities of risk to both the user and equipment. In these cases, the utilization of VR allows the user to receive training regardless of their location or the availability of the training in usage of equipment. Moreover, the virtual environment also allows the user to explore the possibilities without compromising their safety or the operation of virtual peripherals.

The advantages associated with VR training have increased the demand for this type of application in different segments. Despite this, few studies have been developed involving the simulation and training related to the control of Electric Power Systems. The purpose of this work is to investigate efforts to develop computational solutions based upon Virtual Reality techniques for the control and operation of electric power systems, in particular, transmission substations, power plants and their components. The goal is to design and develop a Virtual Reality system (software and hardware specific suitability, based on components available on the market) integrated with the supervision and control system of the Cemig Control Center, which supports simulation, training and control of substations and power plants, with navigation requirements, immersion and interaction.

Specific objectives for this research are:

1. To train control center operators', involving the simulation commands and operation elements of substation control during routine operations and during restoration processes;
2. To train field operators' on the commands and operations of elements of substations and power plants;
3. To operate the substation and power plants;
4. To plan maintenance and outages;
5. To use real-time simulation for greater realism increasing the communication between field and control center operators.

2 Related Work

Galvan-Bobadilla et al. [4] presents architecture and the development methodology of a non-immersive VR training system for power line operators. The system incorporates an easy-to-use graphical interface and terminology that refer to real-world tasks. The developed software guides the employees, on a step-by-step basis using 44 different methods, containing background information on standards and useful tips written by experts. The system also offers a learning management module, which allows instructors to plan when participants can access the training module or the evaluation module.

This training system offers uniformity as well as a low cost for training programs in electric utilities, in addition ensures safety during training operations. Currently, it is being used to train thousands of live line operators across 13 divisions of an energy utility company in Mexico.

The interaction between user and system is performed by means of a control module called orchestrator, which provides the fundamental communication between the interface, the logic module (business logic), and the repository multimedia elements. The orchestrator module incorporates callback functions, determines which request is being made by the user and responds appropriately firing different methods for data manipulation. Upon receipt of the information, the orchestrator transfers it to the display module. The events are synchronized with the controller through sounds and animations on a 3D screen. Still, the controller is responsible for initializing and pausing events depending on user actions.

Wang and Li [5] propose a simulation system in Virtual Reality for training in electric energy substations, using multimedia technology and databases. According to the author, the system can simulate all types of substations, requiring only the uploading of the 3D model for the proposed substation. The simulated functions are common training operations and preparing drill simulations for solving problems concerning power equipment and treatment of accidents.

Using the system, operators do not only learn the correct operation of power equipment under different conditions, but also improve effectively the skills for emergency treatment and maintaining energy levels when a system failure occurs. Moreover, there is a reduction in the loss of energy supply and an increase in social and economic benefits. The proposed system provides 3D modelling needed for electrical equipment and real-time rendering of the virtual scene. However, limitations include the impossibility of establishing corresponding mathematical models and dynamic simulations in real time. Moreover, there is no mention of how the interaction between user and system is handled.

The objective of the project presented in [6] was to create a platform for providing training in the maintenance of complex machines in electric power systems, in order to understand and practice operational procedures. It illustrates how Virtual Reality techniques can be used in electric power systems by improving the effectiveness of vocational training in the field of energy industries.

The system provides an authoring tool used for input and setting data editing, which requires essential knowledge in computer science and experience in design and analysis of transformer structures. The user can interact with the virtual environment via an infrared optical tracking system, changing position and orientation. Experience in teaching was used to combine learning methods to the field of Virtual Reality in order to improve the educational process and knowledge transfer in training and education.

3 Proposed Methodology

The methodology proposed in this paper is composed of the following stages:

(a) Acquisition of information regarding the features of the substation (CAD plans, photos, videos and equipment catalogues), by means of a standardized protocol;
(b) Definition of techniques to model three-dimensional components of a substation contemplating its constructive and necessary information for the purposes of simulation control and maintenance (3D Library creation);
(c) Automatic generation of the three-dimensional environment (automatic positioning of equipment from the library, reuse of electrical arrangements, topology for equipment start-up);
(d) Standardized interface templates for best navigation control, reading of electric component information and command sending.

3.1 Flowchart for Construction Process

Figure 1 presents the flow process referring to the proposed methodology.

From the photos and construction documents (CAD, component files, manufacturer documentation etc.), one initiates the construction of 3D models that will go on to construct the virtual environment (physical modeling). Each model is validated and inserted into a Model Library, which further groups together photos and other documents.

Thus, the system allows for the converting of CAD models into virtual environments (semi-automatic generation for virtual environments, by means of the 2D CAD project).

By using such a mechanism, an incomplete VR environment is generated, without cables and connections between the distinct virtual objects. In order to insert the connections an algorithm was developed, for the connecting of cables between components automatically, by taking into consideration the topology of the electrical connections (electric circuit). At the end of this process, the virtual environment resembles the actual arrangement in the field (particular to each substation) and is evaluated by the operators. In turn, the design is sent forward to the association stage. Here, each element from the virtual model is associated to an

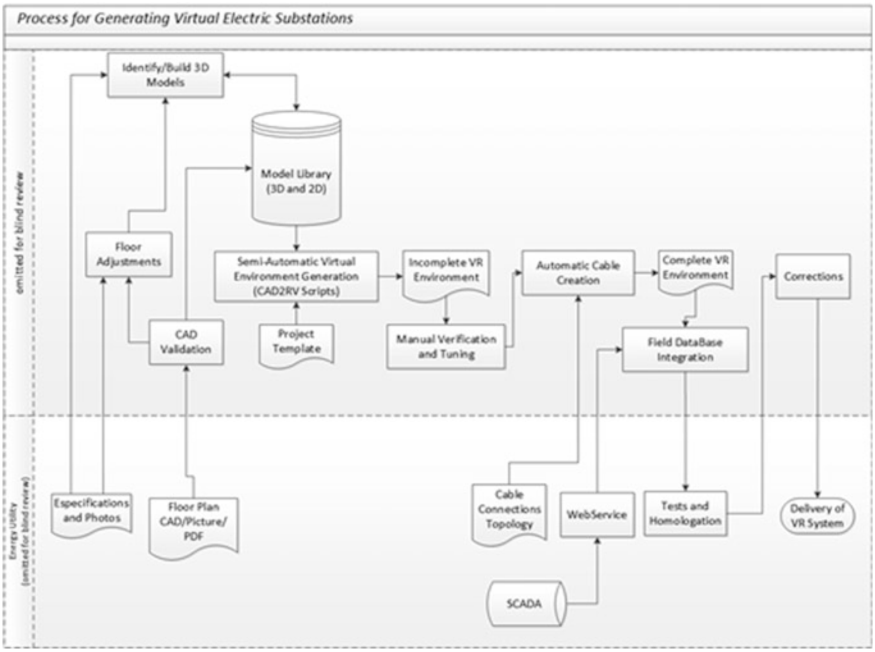

Fig. 1 Process for the generation of the substations

identifier within the SCADA system [7], thus generating conditions in the virtual environment, for presenting the state of each monitored component.

Finally, through consideration of the authoritative policies, permission of the active control is associated by means of the interface template.

3.2 Acquisition Protocol

A site survey at the substation location is necessary for obtaining the wealth of pictures essential to providing the real profile and connection between its elements (e.g. cables, electric bus, etc.). This procedure was used as a cheaper and faster alternative to 3D Scan, guarantying precision and similarity. The digitalization process is the first challenge [8] when it comes to obtaining the 3D model of each component of the power substations. Hence, one of the main causes of delay encountered in the execution of projects that require 3D modeling is the lack of the necessary documentation for carrying out the entire process. Elements that make up this documentation include CAD plant design, photos and videos with a substation and equipment field survey, plus technical equipment catalogues (data sheets), such as circuit breakers, disconnecting switches, and other substation components, all of which are paper based documentation, due to the fact that many power substation have been in operation for more than 10 years.

3.3 VR Environment

The developed VR environment allows for the simulation of a complete scenario of the power substation, presented in Fig. 2, where in the upper right corner a general plan view was created (minimap) that presents the position (red point) of the viewer in the 3D environment. In real scale, it is possible to walk inside the substation and simulate many scenarios. The interaction can be in first person, third person and "God view" modes.

3.4 2D Control Interface for the 3D Environment

In order to support the operator in monitoring and controlling the substation equipment, it is necessary to elaborate two-dimensional control interfaces for performing these activities.

In this context, a control interface and selection (menu, Windows, panels and icons) template has been elaborated, which attends to every task demanded that is

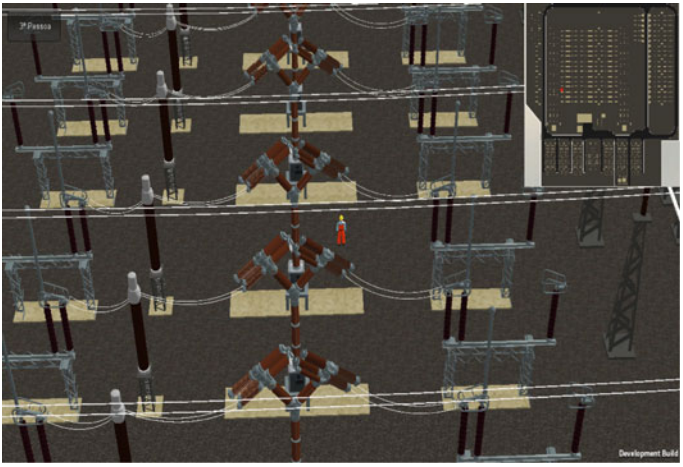

Fig. 2 Complete substation—(Third person view)

pertinent to the substation management. Therefore, after the generation of the virtual substation environment, the template containing the full interface description is added to it. This, in turn, will be incorporated across all substations, thus reducing the time spent for adjusting these components.

Through the means of an internal architecture, data referring to the state of the electrical equipment (turned on or off, electrical measurements) is received and processed in real time via WebService. These data are made available by the SCADA system at the Cemig Control Center. Therefore, with this information at hand, the virtual environment is updated with a true representation of the state of the equipment and the control components content.

In this context, it is possible to provide a new approach for controlling and operating power substation devices, by means of Virtual Reality techniques, which offer greater immersion and more intuitive interactions.

Another associated aspect is that the operators can navigate in the most diverse ways, exploring and viewing the conditions of the electric components that control the substation with greater safety.

Figure 3 presents the system during the consulting operation of the state of a substation component. Note that there are no environmental changes for reading component data. Therefore, the user does not lose the sensation of realism when immersed into the virtual environment, which imitates the workplace with a greater sense of reality. Thus, it is believed that a Virtual Reality based approach for the improvement to human performance and the minimization of health and safety risks when operating power electric systems in their everyday operations, provides the operator the means to produce easily well-constructed mental models. It is relevant, in the training case procedures to allow the user to view the exact situation of the real power substation, inside a virtual environment, via Web Service.

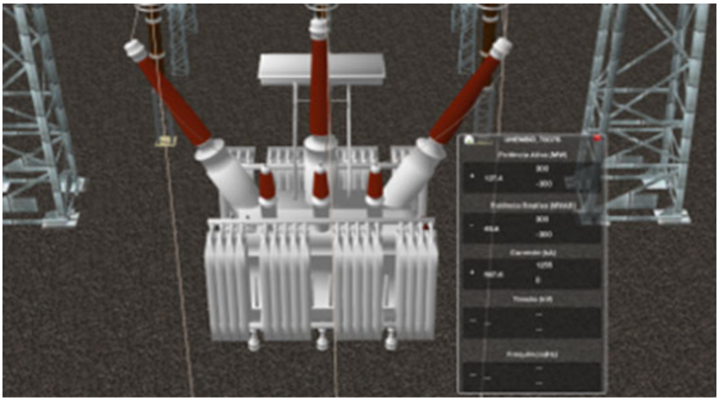

Fig. 3 Consultation of information concerning the state of a transformer

To prepare this control interface template, a number of requirements were defined, considering usability, layout and design aspects:

(a) The two-dimensional control interface should be integrated into the three-dimensional virtual environment.
(b) The control interface should promote mechanisms that allow for their quick response use.
(c) The control mechanisms should be presented only when necessary in the context of interaction or when activated by the user.
(d) Production of alternative interfaces for control that possess mechanisms that allow the user activation and deactivation functionality, along with the option of relocation to any desired space within the virtual environment. In the following, the principal components for this template are presented.

Selection Bar—Menu. This strategy contains a single bar for the selection of control options (menu) located on the left-hand side of the environment (Fig. 4).

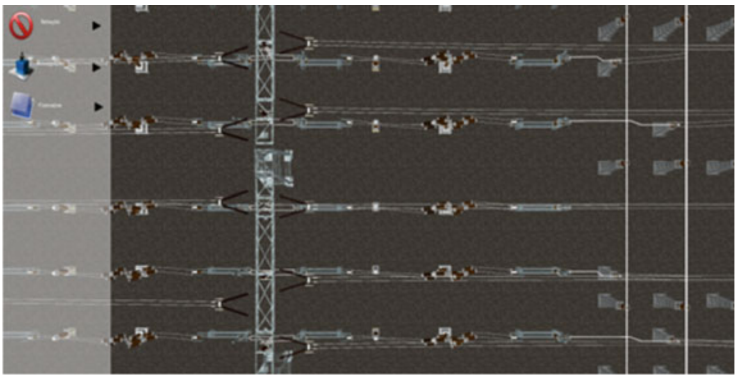

Fig. 4 Active selection bar

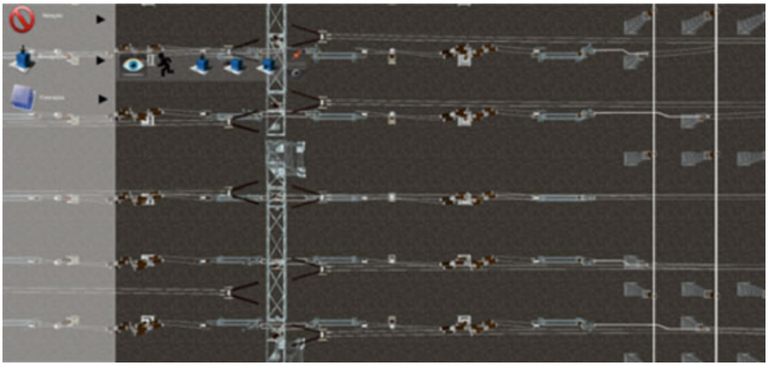

Fig. 5 Action control selection panel

The area occupied by the menu, in relation to the vertical direction of the screen is 100 %. In relation to the horizontal direction, in moments of inactivity, only 1.2 % of the screen area is occupied and does not present any reference to the selection options contained therein. However, in moments of activity, this space is altered occupying an area totaling 15 %. Figure 4 presents the item with an active selection bar. This activation is performed by placing the mouse pointer over the bar. One notes that the control interface is inserted in the context of a 3D application, with transparency of 50 %.

Action Control Panels. Each item situated in the side bar contains a panel with options relating to certain actions. In order to visualize this panel, it is necessary only to place the mouse over the opening icon. Thus, it appears only when the action involved is being requested. When the mouse pointer is removed from the control interface, the panel disappears and is no longer presented. At this moment, the side bar returns to a state of inactivity. In Fig. 5, a part of the side bar and the panel referring to an action contained in the virtual environment is opened.

3.5 SCADA Integration

The following architecture concept for supporting the integration of the virtual environment with the supervisory system (SCADA) is presented in Fig. 6.

The Virtual Reality system is integrated with the SCADA (supervisory control and data acquisition) [7] through interface integration. This interface, besides making the connection between the SCADA and the RV Cemig, behaves as a layer of compatibility between the two systems and permitting the continual flow of information related to training activities. The intention behind this product is to simplify the internal logic of reading data from the Virtual Reality system. It prevents the direct connection to SCADA, and ensures access to various data sources, such as the names and properties of the elements of substations and power

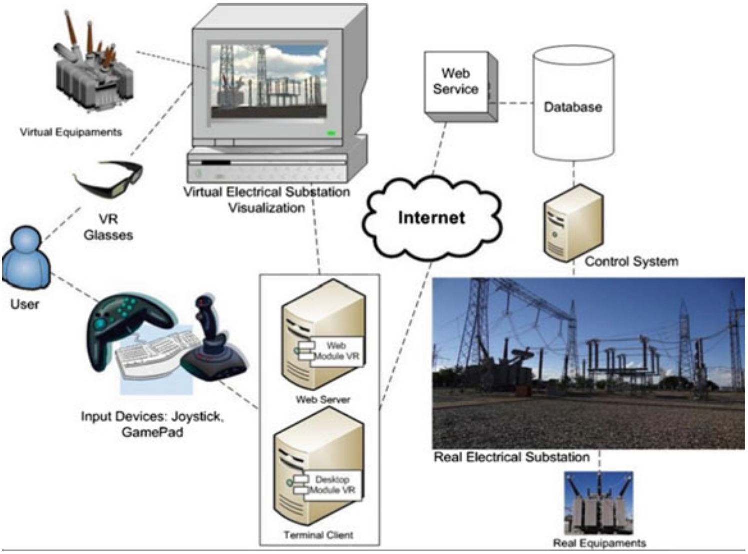

Fig. 6 SCADA environment architecture integration

plants, as well as the parameters and measurement values of each element. The Communication interface SCADA lays down the Virtual Reality parameters, which is also known as integration interface or interface layer. The interface then delivers to the VR environment, the status (measured values, on/off, etc.) of the information requested and sends the SCADA commands to be performed on the elements (on/off, open/close, etc.). The reading of the element states in SCADA is performed from the SQL Server database, and changes to these states will be carried out through CGI's that currently use HMI. An interface between the system and the SCADA CEMIG Virtual Reality environment has already been implemented.

4 Results

Through the use of this methodology, until the present moment, 24 real transmission substations have been reproduced virtually. It is important to highlight that the first substation was constructed without the adoption of the presented methodology. In fact, it was due to the time spent on the construction of the referred to substation (420 h) that our group highlighted the need for the creation of a methodology for perfecting the generation process of more than 50 substations. By way of example, only the task of positioning virtual components and the generation of conducting cables, consumed around 70 % of the overall time spent.

By using the new approach, through use of the elaborated methodology, only 71 h were spent for reproducing the same substation, which presents a time

Fig. 7 The photorealism observed by the operators

reduction of 83 %. The system developed herein was presented to its target audience—the system operators. The photorealism observed by the operators was classified as a great level or realism in relation to the field, where the respective substation can be found (Fig. 7). Besides, professionals working on the operation of the electrical system identified a great reduction in the mental effort and time needed to produce the interfaces. Figure 8 presents results using a Video Wall at CEMIG COS (Operational Center) in Belo Horizonte—Minas Gerais, Brazil.

Fig. 8 Video wall VR usage at Cemig's operational center-Belo Horizonte—MG—Brazil

5 Conclusions and Future Work

The methodology described herein was applied successfully to the production of Virtual Reality environments for power utility substations. The process proposed reduced both time and costs. Through the creation and application of the protocol for acquiring information from substations, it was possible to manage in an effective manner the issues related to the data necessary for initiating the construction process, which also made it an effective instrument for the validation of this material.

Regarding the convention rules for modeling, it was possible to identify that these are fundamental to the process that is associated with automation tasks, besides providing improved performance and fluidity while navigating through the system. Concerning the automation stage during the construction of the scene, one can conclude it demonstrates high efficiency based on the high percentage in time reduction offered. In the near future, it will be desirable to include a greater number of CAD applications from engineering and architecture, as well as VR engines, already available on the market.

Acknowledgments The authors would like to thank CEMIG GT411 Research and Development Program, along with CAPES (Coordination for the Improvement of Higher Education Personnel), CNPq (National Development Council Science and Technology) and FAPEMIG (Support Foundation Research of the state of Minas Gerais) for all their financial support given to this research project.

References

1. Okapuu-Von Veh, A., et al.: Design and operation of a virtual reality operator-training system. Power Systems, IEEE Transactions on, agosto **1585**, 1591 (1996)
2. Fanqi, M., Yunqi, K.: An improved virtual reality engine for substation simulation. 2nd International Conference on Future Computer and Communication (ICFCC), 2010. [S.l.]: [s.n.], p. V1-846, V1-849, 21–24 (2010)
3. Guangwei, Y., Zhitao, G.: Scene graph organization and rendering in 3D substation simulation system. Asia-Pacific Power and Energy Engineering Conference (APPEEC), 2009. [S.l.]: [s.n.], p. 1, 4, 27–31 (2009)
4. Galvan-Bobadilla, I. et al.: Virtual reality training system for the maintenance of underground lines in power distribution system. In: Third International Conference on Innovative Computing Technology (INTECH), 2013, pp. 199–204. IEEE (2013)
5. Wang, C., Li, P.: Development and challenges of distributed generation, the micro-grid and smart distribution system. Autom. Electr. Power. Syst. **2**, 004 (2010)
6. Arendarski, B., Termath, W., Mecking, P.: Maintenance of complex machines in electric power systems using virtual reality techniques. In: Conference Record of the 2008 IEEE International Symposium on Electrical Insulation (ISEI), pp. 483–487. IEEE (2008)
7. Boyer, S.A.: SCADA: supervisory control and data acquisition. International Society of Automation (2009)
8. Simoes, F. et al.: Challenges in 3D reconstruction from images for difficult large-scale objects: A study on the modeling of electrical substations. 14th Symposium on Virtual and Augmented Reality (SVR), 2012. [S.l.]: [s.n.], p. 74, 83, 28–31 (2012)

The NRC Human Performance Test Facility: An Approach to Data Collection Using Novices and a Simplified Environment

Niav Hughes, Amy D'Agostino and Lauren Reinerman-Jones

Abstract In the spring of 2012, as part of a 'hub and spoke' model of research to address the human performance concerns related to current as well as new and advanced control room designs and operations, the U.S. Nuclear Regulatory Commission (NRC) sponsored a project to procure a low cost simulator to empirically measure and study human performance aspects of control room operations. Using this simulator, the Human Factors and Reliability Branch (HFRB) in the Office of Nuclear Regulatory Commission (NRC) began a program of research known as the NRC Human Performance Test Facility (HPTF) to collect empirical human performance data with the purpose of measuring and ultimately better understanding the various cognitive and physical elements that support safe control room operation. To accomplish this, HFRB first procured two 3-loop Westinghouse pressurized water reactor simulators with the capability to run a full range of power operation scenarios. HFRB staff work as co-investigators along with a team of researchers at the University of Central Florida (UCF) to design and carry-out a series of experiments aimed at measuring and understanding the human performance aspects of common control room tasks through the use of a variety of physiological and self-report metrics. The intent was to design experiments that balanced domain realism and laboratory control sufficiently to collect systematic, yet meaningful human performance data related to execution of common main control room (MCR) tasks. Investigators identified and defined three types of tasks that are examined in the present project: Checking, Detection, and Response Implementation. Task type presentation was partially counterbalanced to maintain ecologic validity with experimental control. A variety of subjective and physiological measures were used to understand performance of those tasks in terms of

N. Hughes (✉) · A. D'Agostino
U.S. Nuclear Regulatory Commission, Rockville, USA
e-mail: Niav.hughes@nrc.gov

A. D'Agostino
e-mail: Amy.Dagostino@nrc.gov

L. Reinerman-Jones
University of Central Florida, Orlando, FL, USA
e-mail: lreinerm@ist.ucf.edu

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_16

workload. The simulator used to collect these data was a digital representation of a generic analog NPP MCR interface. The data resulting from this experimentation enhances the current information gathering process, allowing for more robust technical bases to support regulatory guidance development and decision making. The present paper describes the approach behind this research effort.

Keywords Nuclear energy · Main control room (MCR) tasks · Simulators · Human performance · Decision-Making

1 Introduction

The staff of the U.S. Nuclear Regulatory Commission (NRC) is responsible for reviewing and determining the acceptability of new designs to ensure they support safe plant operations. The human operator is a vital part of plant safety, thus, the NRC staff must understand the potential impact of new designs on human performance in order to make sound regulatory decisions. Much of the basis for current NRC Human Factors Engineering (HFE) guidance comes from data from research in other domains (e.g. aviation, defense), qualitative data from operational experience in Nuclear Power Plants (NPPs), and a limited amount from empirical studies in a nuclear environment. Unfortunately for new designs, technologies, and concepts of operations, there may be a lack of operational experience and a dearth of research literature. To address this, the commission in SECY-08-0195 directed the staff to consider using generic simulator platforms for addressing human performance issues. A simulator could provide a tool to gather more empirical nuclear specific human performance data. These data would enhance the current information gathering process thus providing stronger technical bases and guidance to support regulatory decision making.

Although this may seem like a simple undertaking, there are two primary challenges: (1) NPP simulators have historically been very costly to purchase, house, and maintain and, (2) recruiting trained operators for human-performance research is very difficult. When a simulator and operators can be secured for human performance research, the operator sample tends to be quite small, often allowing for only qualitative analysis or limited quantitative analysis which makes drawing conclusions difficult.

2 Overcoming the Challenges

To collect empirical nuclear specific human performance data, the two challenges outlined in Sect. 1 had to be addressed. The resources for this project were limited and building a large simulator facility for human performance research that would require staff and a long-term agency commitment was not a feasible option. It was determined that in order for this project to be successful, the staff had to find a low-cost simulator option that would allow for collection of meaningful quantitative human performance data to help answer research questions of interest to the NRC. In order to gather enough data for quantitative analysis, the staff concluded that it would be necessary to utilize a non-operator population for at least a portion of the research.

The long-term research vision for this project was to conduct human performance studies in two steps. The first step would involve testing many non-operator participants with various combinations of scenarios, system conditions, and new technologies. The results would allow researchers to identify safety-critical or error-prone contexts as well as identify measures most sensitive to changes within this environmental context. Using the insights from the first step, the second step would test a limited number of operators for those error-prone scenarios to further inform us about the potential human factors issues.

2.1 Procuring a Low Cost Simulator

As mentioned above, historically, purchasing an NPP MCR simulator has necessitated having a facility where all of the "hard" analog panels can be staged, trained operations staff and IT staff to use and maintain the simulator, and large start-up budget to either have a custom simulator built or purchase an already built simulator. As this was not an option for the Office of Regulatory Research (RES), the staff pursued several alternatives including:

1. Collecting human performance data in the simulators at the NRC Technical Training Center
2. Partnering with a utility to collect data in their simulator
3. Exploring availability of "soft" simulators (i.e. runs on computer, no "hard" panels)

Options 1 and 2 were quickly ruled out for several reasons. First, getting access to either the TTC simulators or a utility simulator is very difficult as they are often in use for training purposes. Second, to operate a full simulator, trained operators must be used. This is a problem, as mentioned above, because the number of trained operators are limited, hence, their ability to be available for research is very restricted. Thus, option 3 was determined to be the most reasonable path.

The staff determined requirements to facilitate the simulator search which included:

1. Must be a generic (pre-built) model
2. Must model primary and secondary systems
3. Must include basic process models of reactor physics, thermo-hydraulics, and control systems
4. Must allow for full-range of power operations
5. Must have straightforward method to configure the simulator to run in several modes (e.g. fully-simulated mode or a semi-manual mode)

6. Must allow the NRC to conduct real time, human-in-the-loop simulations so that operator responses can be observed and assessed during scenarios of various initial conditions, plant behaviors, malfunctions, and transients
7. Must have graphic tools to modify interfaces, as well as the ability to build additional graphic displays to study the impacts of new interface features or modifications on human performance
8. Interface configuration must be flexible so that the simulator allows one individual or a team of personnel to perform tasks
9. Must provide ways to allow for non-operator participants to perform simplified tasks or parts of the tasks in scenarios
10. Must operate on desktop computers under a Microsoft Windows environment
11. Fidelity of the simulator must be high enough not to mislead an experienced operator into error in actions
12. HSI must either simulate current control-room panels or advanced control room displays
13. Must include an instructor station capable of simulation control, monitoring, and data visualization activities
14. Must have a data-logging system to collect real-time plant parameter process values and be capable of exporting data to files in a format readable by Microsoft Excel.

After an open competitive bidding process and assessment of a variety of simulator options, ultimately, the simulator that best fit the needs of the NRC was determined to be the GSE Generic Pressurized Water Reactor (GPWR).

The GSE GPWR included the following features:

- Generic 3-loop Westinghouse PWR
- RETACT thermal hydraulics code
- Runs on eight 24 in LCD screens, 4 Dell Precision Workstations with Single Quad CPU
- Software includes a graphics tool, an instructor station, and a real time executive program
- System update time of at least 2 times per second
- Capability to run full range of power operations
- Allows for instrumentation failure
- Graphics development tool allows for drag and drop user interface
- HSI is hard panel mimics
- Each operator station can access entire control room soft panels
- Operator stations can be preconfigured to display specific panel sections
- Contains real time trending for data capture and logging
- Data logs can be exported to Excel
- Over twenty initial conditions (can add up to 200)
- Simulator is pre-loaded with 100 s of malfunctions
- Includes operating procedures for full range of operations, plant operating “curve book,” and technical specifications.

2.2 Finding Participants

As discussed in the previous section, access to operators is a major challenge to the use of simulation studies to understand human performance in the nuclear domain. Drawing substantial conclusions from experimental data requires a large sample size which is difficult and costly using the trained operator population.

Thus, in order to gain access to more potential research participants, the NRC determined that partnering with a university was the best course of action. Universities typically have access to a pool of students required to participate in research for class credit and often have ties to the community as a means of recruiting research participants as well. Partnering with a university was beneficial to the project in several additional ways including (1) their expertise in experimental design, simulation engineering, the use of state-of-the-art human performance measurement tools, and the collection and analysis of large quantities of data and (2) ensured NRC adherence to proper guidelines for conducting human subjects research by going through the university's established internal review board (IRB) process to ensure the ethical treatment of human subjects.

Access to a larger population from which to collect data was critical for the project's success, however, the staff realized that specific limitations had to be addressed when utilizing a novice population. In order to collect meaningful data from novices, we proposed that the environment needed to induce participants to experience both the complexity and cognitive requirements incurred by trained operators without requiring them to have all the knowledge and skills of a trained operator [1, 2]. In other words, the methodological approach adhered to the principal of different but equal; the environment (e.g. interface, task) is different, but in such a way that is controlled and meant to induce the same type of cognition and level of workload that would be experienced by trained operators. Underlying all human cognition, there are various cognitive mechanisms and performance influencing factors that ultimately impact human performance [3]. It is on this premise that we base our rationale for the use of a novice population as proxy for an expert operator population as a means to investigate the more generically human aspects of cognition associated with task performance within an NPP MCR environment. For instance, we know that operators have many parameters that they are required to monitor. A novice population can be used as a surrogate to understand what types of displays might cause more monitoring errors.

3 Proof of Concept

In order to have a successful program of research, the "different but equal" philosophy described in Sect. 2.2 had to be tested and validated. As a first step in this effort we needed to create an ecologically valid environment from which to conduct our research.

3.1 Creating an Ecologically Valid Environment

The challenge was to develop an experimental platform that was ecologically valid, but could also be systematically controlled and operated by a novice population. It was necessary to ensure that cognitive demands would be comparable to that experienced by trained operators, but the physical environment would be calibrated to accommodate the skill-level of the novice population.

3.2 Experimental Design and Defining the Tasks

In order to maintain a cognitively simplified yet similar environment for novice participants, it was determined that novice participants would need to complete realistic NPP operator tasks while still allowing for experimental control and performance measurement. In order to develop the experimental design and define the tasks to be measured, the research team collaborated with a NPP operations Subject Matter Expert (SME).

NPP MCRs are managed by teams or "crews" of professional operators; a minimum MCR crew is composed of a Senior Reactor Operator (SRO) who directs two Reactor Operators (ROs). The crew uses Emergency Operating Procedures (EOPs) to bring the plant to a safe state during emergencies. The use of (EOPs) is standard across U.S. control rooms. Thusly, our equal but different approach led us to use tasks derived from various EOPs and discussions with a domain Subject Matter Expert and to adopt an experimental paradigm that included an SRO (played by the experimenter), RO1 (played by a confederate) and RO2 (participant). The use of realistic tasks along with the team dynamic created by the use of the roles of SRO-RO1-RO2 allowed for a cognitively similar environment.

Several methodological steps were taken in order to arrive at the three types of NPP MCR tasks that participants would be asked to complete: checking, detection, and response implementation. We began by first considering all the possible tasks performed by trained NPP MCR operators. O'Hara et al.'s model [4–6] describes the following as the generic primary tasks involved in MCR operations: monitoring and detection, situation assessment, response planning, and response implementation. As we pre-determined the tasks participants would be asked to complete, we ascertained that situation assessment and response planning were outside the scope of the present work.[1] We therefore focused on *monitoring and detection* and *response implementation* as they could be defined and controlled simply and

[1]Situational assessment tasks consist of evaluating current state of NPP systems to determine whether they are within required parameters. Response planning refers to deciding upon a course of action to address the plant's current situation [4]. The use of an EOP and the SRO to direct participant actions remove the cognitive activity associated with situational assessment and response planning and therefore, determined to be outside the scope of the present work.

sufficiently for measurement using a novice population. Through team discussions with a SME, we further delineated O'Hara's task hierarchy to conclude that within the monitoring and detection[2] activity as described by O'Hara, there actually exists two distinct activities: (1) monitoring and detection and (2) checking.

The *checking* task type consisted of a one-time inspection of an instrument or control to verify that it was in the state that the EOP calls for it to be (e.g., open or shut). Participants were required to locate various instrumentation and controls by clicking on the correct control. The *detection* task type required participants to correctly locate a control and continuously monitor it for identification of change. Participants were required to monitor the gauge for five minutes and detect changes by clicking on a button located at the bottom of the display. Twelve changes per minute occurred, totaling 60 changes per detection task. The *response implementation* task type required participants to correctly locate a control and manipulate it in the required direction. Each task type consisted of four steps that were executed using three-way communication led by the experimenter acting as the SRO.

Task types were presented in partially counterbalanced blocks of four. Meaning, one block consisted of four checking tasks, four detection tasks, or four response implementation tasks. The purpose of the blocking method was to control the presentation of the tasks such that the resulting performance and workload results could be statistically analyzed. The partial counterbalancing of the blocks was an effort to balance ecological validity with laboratory control as the checking task type always preceded the response implementation task type because, in a real operating scenario, an operator would never implement a response prior to checking the state of the instrumentation first.

3.3 Modifying the Simulator

In order to create a cognitively similar environment for novice participants, the interface also needed to be simplified. Thusly, the control panels were modified in various ways to reduce complexity. The first reduction to complexity is that the experimental scenario only required the use of two control panels. Next, each panel was reduced in visual complexity. Specifically, the panels were modified by reducing the amount of instrumentation and controls (I&C) contained on each panel and changing the naming convention of the I&C. The names of the gauges and switches were modified to reduce the memory burden to maintain the short-term memory principal of seven plus or minus two items [2, 7]. These changes were made

[2]O'Hara et al. [5] identify monitoring and detection as one task, but their definition of the two tasks are separate. Monitoring requires checking the plant to determine whether it is functioning properly by verifying parameters indicated on the control panels, observing the readings displayed on screens, and obtaining verbal reports from other personnel. Detection occurs when the operator recognizes that the state of the plant has changed. Through discussions with a SME, the team separated and defined the *checking* task described in the text.

consistently to the instructions used in the experiment as well as to the panel interfaces. In order to systematically reduce the amount of I&C on each panel, the original panel with the least amount of controls was identified—in this case, panel C1. Next, a systematic reduction of the amount of instrumentation and controls on the A2 panel occurred based upon a calculated percentage to equal the amount of controls on panel C1, which had 113 I&C elements. In particular, the instrumentation and controls were categorized into five groups including gauges, switches, light boxes, and status boxes. Participants interacted with gauges, switches, and light boxes. Each type of I&C was reduced by the previously calculated percentage, thus leaving the ratio of I&C types the same on each panel. This systematic approach ensured the complexity of the original panel remained. In other words, the ratio of I&C on the modified panel remained intact to those of the original panel. In addition to enabling a novice population to interact at an appropriate level of complexity, the reduction of the amount of controls in panel A2 to equal the amount of controls in panel C1 balanced complexity between panels, thereby removing potential confounds. For further detail on these modifications, see Reinerman-Jones et al. [2].

After a series of pilot tests using the modified panels, we determined that having the simulator respond dynamically[3] to operator input did not allow for sufficient control for the novice population. Therefore, we determined it necessary to remove the physics forgoing the dynamic simulation environment for a controlled experimental environment able to be systematically presented to participants allowing for statistical analysis of their performance. However, the order in which certain steps occurred within each task type, as well as the timing and incremental changes in temperature and pressure were maintained in accordance with the would-be physics of a dynamic environment experienced by real operators.

3.4 Training Participants

Participants were trained so that they could be proficient at performing the tasks successfully and support assertions of a cognitively simplified, yet appropriately similar task environment. Training consisted of three phases using a scaffolding approach. Participants were required to pass a proficiency test for each phase with a score of 80 % or greater. They were tested on their abilities in three areas: communication, navigation, and task performance. Participants were allowed a maximum of two attempts to pass each phase of training and only completed a second attempt of a training phase if they did not achieve an 80 % or greater on their first attempt. In addition, if participants did not receive a score of 80 % or greater on the second attempt of any of the three phases, the researcher classified them as ineligible to participate in the study, and they were dismissed.

[3]Dynamic response of simulator refers to the resulting change to the state (i.e., the physics) of the simulator based on operator input.

3.5 Use of Confederates

The use of confederates is another aspect of the experimental design that supports the creation of an ecologically valid environment [8]. Participants served in the role of RO1 while confederates served as RO2. Confederates were extensively trained on the experimental tasks and proper interactions with the participants. The confederates were paired with experimenters who served in the role of SRO for the duration of the data collection. Crew composition in NPP MCRs is often stable across shifts, therefore, that consistency was adhered by fixed partnering across data collection sessions. Using a confederate model allowed experimenters to emulate the “team” dynamic experienced by real NPP operators, but maintain control over the experience of the participant.

4 Conclusions

Nuclear specific human performance data collection efforts large enough for quantitative analysis is not widely practiced. The staff at the U.S. Nuclear Regulatory Commission determined it necessary to develop its own such research program with the hope that others might follow suit. Our focus was to develop a methodology to gather meaningful data from novices using a simplified operating environment to inform us about the highly complex operational environment of the NPP MCR.

Only one participant was dismissed from the experiment due to failure to reach proficiency on the progressive training module, providing evidence that university students were able to become proficient in performing realistic (rule-based and skill-based) operator tasks in the simplified controlled environment.

Using this research design strategy to develop a baseline, we anticipate being able to identify measures of workload best suited for particular tasks or combination of tasks, the levels of workload associated with tasks, and the kind of workload induced (e.g. physical, cognitive) by tasks. Further, we expect that our method will improve data collection techniques for use with the operator population, such that lab results may be further validated.

References

1. Lackey, S., Reinerman-Jones, L.E., Salcedo, J.: Equal but different: 5 research strategies for improving conclusions drawn from novice populations. Proceedings of the 2014 MODSIM World Conference, Hampton, VA (April, 2014)
2. Reinerman-Jones, L.E., Guznov, S., Mercado, J., D’Agostino, A.: Developing methodology for experimentation using a nuclear power plant simulator. Foundations of Augmented Cognition **8027**, 181–188 (2013). doi:10.1007/978-3-642-39454-6_19

3. Whaley, A.M., Xing, J., Boring, R.L., Hendrickson, S.T.L., Joe, J.C., Le Blanc, K.L.: Building a psychological foundation for human reliability analysis, NUREG-2114. U.S. Nuclear Regulatory Commission, Washington, DC (2015)
4. O'Hara, J.M., Higgins, J.C., Brown, W.S.: Human factors considerations with respect to emerging technology in nuclear power plants, NUREG/CR-6947 U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research (2008)
5. O'Hara, J., Higgins, J., Fleger, S., Barnes, V.: Human-system interfaces for automatic systems (No. BNL–93950-2010-CP). Brookhaven National Laboratory (BNL) (2010)
6. O'Hara, J.M., Higgins, J.C., Pieringer, P.A., Fleger, S.A.: Human factors engineering program review model, NUREG-0711, Rev. 3. Washington, DC (2012)
7. Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. **63**(2), 81 (1956)
8. Leis, R., Reinerman-Jones, L.: Methodological implications of confederate use for experimentation in safety-critical domains. Proceedings of the International Meeting of Applied Human Factors and Ergonomics, Las Vegas, NV (2015)

Using the Human Systems Simulation Laboratory at Idaho National Laboratory for Safety Focused Research

Jeffrey C. Joe and Ronald L. Boring

Abstract Under the United States (U.S.) Department of Energy (DOE) Light Water Reactor Sustainability (LWRS) program, researchers at Idaho National Laboratory (INL) have been using the Human Systems Simulation Laboratory (HSSL) to conduct critical safety focused Human Factors research and development (R&D) for the nuclear industry. The LWRS program has the overall objective to develop the scientific basis to extend existing nuclear power plant (NPP) operating life beyond the current 60-year licensing period and to ensure their long-term reliability, productivity, safety, and security. One focus area for LWRS is the NPP main control room (MCR), because many of the instrumentation and control (I&C) system technologies installed in the MCR, while highly reliable and safe, are now difficult to replace and are therefore limiting the operating life of the NPP. This paper describes how INL researchers use the HSSL to conduct Human Factors R&D on modernizing or upgrading these I&C systems in a step-wise manner, and how the HSSL has addressed a significant gap in the process for upgrading systems and technologies that are built to last, and therefore require careful integration of analog and new advanced digital technologies.

Keyword Human factors · Nuclear power plant · Control room modernization · Instrumentation and control systems upgrades

J.C. Joe (✉) · R.L. Boring
Idaho National Laboratory, Idaho Falls, ID, USA
e-mail: Jeffrey.Joe@inl.gov

R.L. Boring
e-mail: Ronald.Boring@inl.gov

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_17

1 The Need for Nuclear Power Plant Control Room Modernization

In 2014, nuclear power provided approximately 20 % of all the electricity generated in the United States (U.S.) [1], and did so safely and reliably (i.e., non-intermittently). Low carbon replacement technologies for electrical generation, including renewable energy and new nuclear power plants (NPPs), have not materialized as quickly as some expected. In 2016, the Bill and Melinda Gates Foundation highlighted this concern [2], reiterating research showing, as seen in Fig. 1, that transitioning from one energy source to others has historically taken decades [3].

Without suitable electrical generation replacement technologies in place, it becomes even more important to ensure that the current fleet of NPPs continues to generate electricity safely and reliably. The U.S. Department of Energy's (DOE) Light Water Reactor Sustainability (LWRS) program, operated in close collaboration with industry research and development (R&D) activities, provides the scientific basis for licensing and managing the long-term, safe, and economical operation of commercial NPPs. In short, the LWRS program focuses on research that contributes to the national policy objectives of energy and environmental security.

One of the principal LWRS R&D focus areas is the Advanced Instrumentation, Control, and Information Systems Technologies pathway [4]. Two interrelated goals of this pathway are: (1) to ensure that legacy analog instrumentation and control (I&C) systems are not an obstacle to the continued operation of commercial NPPs, and (2) to implement digital I&C technologies that facilitate broad innovation and improve the NPP operating business model. Idaho National Laboratory (INL) researchers [5] have pointed out that empirically rigorous Human Factors R&D that improves I&C design, implementation, and operator performance is an

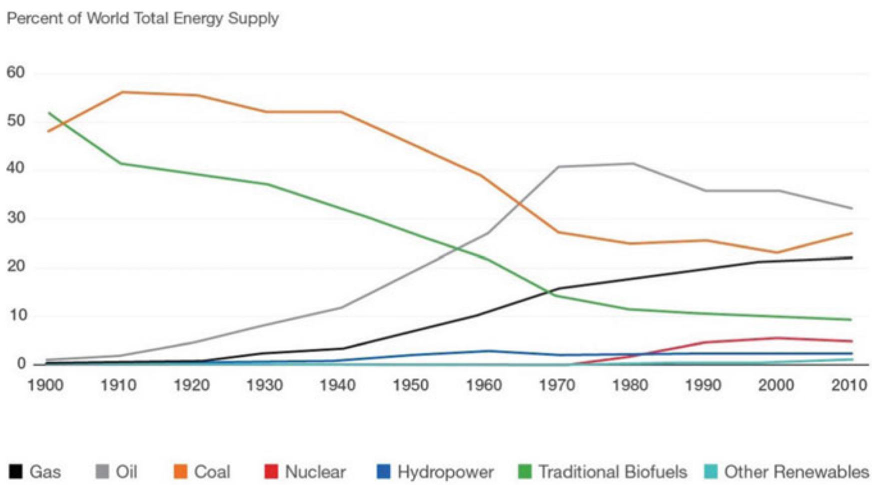

Fig. 1 Energy transitions take decades

essential link in the value chain that is at the core of improving NPP operating business models. Thus, INL is conducting this LWRS sponsored R&D to develop the requisite scientific knowledge on advanced I&C technologies that are needed to support the safe, reliable, and cost-competitive production of electricity from NPPs. As mentioned in [6], this often involves developing new capabilities to optimize process control and implementing them cost effectively in existing NPPs. It also requires developing and substantiating optimal approaches to achieve sustainability of I&C systems throughout the period of extended operation, as there are challenges with integrating new digital technologies with existing I&C systems, and the obsolescence time frame for digital technologies is much shorter than it is for analog technologies, especially analog systems that are certified for use to control safety functions in NPPs. To meet these requirements, R&D must be conducted on new methods for visualization, integration, and information use to enhance operator situation awareness in order to achieve safer, more reliable electricity generation through the installation of new or enhanced I&C systems.

2 The Need for a Research Simulator for Control Room Modernization

Every commercial NPP has a full-scale, full-scope, high fidelity NPP simulator on site that they use to train and qualify main control room (MCR) operators. Yet, INL has built the Human Systems Simulation Laboratory (HSSL) and installed a full-scope, full-scale, reconfigurable NPP simulator to support the LWRS R&D activities described above [7]. The HSSL simulator is an essential tool INL uses to accomplish this research, but given that NPPs already have training simulators on site begs the question of why another NPP MCR simulator is needed for this R&D. There are a number of inter-related answers:

- *Training simulators at NPPs are booked to capacity to support operator training* [8]. The training simulator at each NPP is a valuable and highly utilized resource. The simulator is an essential tool that NPPs use to maintain the operator's qualifications to operate the plant. NPP MCR operators also go through training on a regular basis. Anecdotally, operators at one U.S. commercial NPP are on-shift for 4 weeks, go to training on the 5th week, and then have the 6th week off. Furthermore, because licensed NPPs always need a crew of operators in the MCR (whether at full-power or in refueling), this means crews of operators are always cycling between being on shift, in training, or off. To keep up with this demanding schedule, the training simulator is also in near constant use.
- *Training simulators must maintain an identical configuration to the MCR, and modifying them introduces some risk.* The training simulator is, for all practical purposes, an exact replica of the MCR, and needs to maintain a layout and functionality that is identical to the MCR. Testing new I&C technologies in the training simulator would change its configuration. Furthermore, cutting,

grinding, and welding the steel of the simulator's control boards risks damaging adjacent devices and under-board cabling. Wire bundles would likely need to be separated, introducing the possibility of damaging signal cables to devices that simply need to be moved to make space on the boards for the upgrades [9].

- *NPPs are complex systems.* Given the complexity of commercial NPPs, it is useful to have a test bed, such as an R&D simulator, to evaluate and thoroughly test new I&C technologies before they are installed in the MCR and put into operation. The new technology requires testing to ensure that it functions properly (e.g., as expected), that safety is not compromised with its installation, and that any unintended consequences with its installation (e.g., unanticipated adverse interactions) are investigated to the fullest extent possible. Additionally, NPPs are commercial ventures that work to minimize the time the plant is down (e.g., for maintenance and refueling) and maximize the time it is generating electricity. Having a full-scope, full-scale, easily reconfigurable R&D simulator allows utilities to perform thorough integrated system testing without increasing down time for the actual NPP.
- *Regulatory environment for nuclear industry.* The U.S. Nuclear Regulatory Commission (NRC) closely regulates the nuclear industry, and changes to MCR are examined carefully in terms of whether they might require a license amendment because they significantly increase the risk of known accident scenarios, introduce new accident scenarios, and/or generally reduce safety margins at the plant [10]. For example, functionality gained through new I&C technologies (e.g., automation) may be perceived to affect safety margins, requiring a license amendment with the NRC. Yet, in many cases, the basic research to demonstrate how the new I&C technology affects operator performance, system performance, and safety margins (presumably in a net positive manner) is not readily available to all utilities in the industry. Therefore, an R&D simulator that can perform fundamental Human Factors R&D meets an important need for the nuclear industry.
- *NPPs have long expected service lives.* One attribute of modern technology, in the broadest sense of this term, is that it lives on a broad continuum in terms of its expected service life. As Fig. 2 shows, some technologies are disposable after one use. Others are designed to last for days, weeks, months, years, or even decades. The NRC originally licensed NPPs to operate for 40 years, but many have applied for 20-year license extensions that will allow them to continue to operate. With this expected service life for NPPs, they must be designed and built to last. The built to last design philosophy, as a consequence, dictates the strategies and methods that must be employed if they are to undergo any modernization efforts. For example, challenges associated with merging original analog technology with new digital technology are a problem unique to technologies that are built to last. For disposable technologies, it is apparently more profitable to produce a new version than it is to try to maintain backwards compatibility. As such, for NPPs, it is useful to have an R&D MCR simulator that has the fidelity to evaluate these issues, and others that arise from their long service lives.

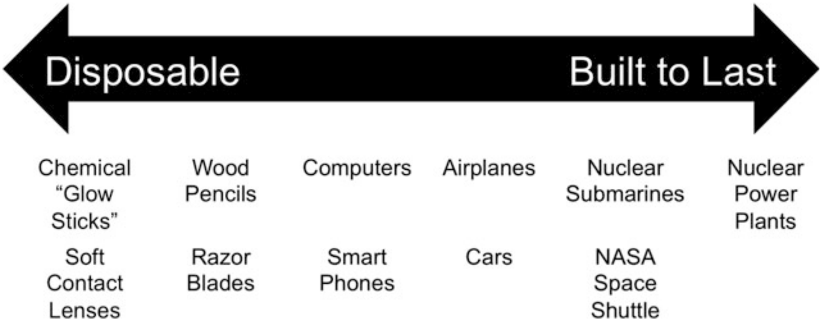

Fig. 2 A continuum of the expected service life of technologies

3 The Human System Simulation Laboratory at Idaho National Laboratory

3.1 General Characteristics of the Human System Simulation Laboratory

Given the inter-related reasons listed in Sect. 2, the INL has built the HSSL to conduct this LWRS sponsored safety focused R&D. Figure 3 depicts the HSSL and the reconfigurable, full-scope, full-scale NPP simulator (see [11] for a detailed description). The HSSL simulator is reconfigurable both in terms of the physical configuration of its constituent 15 bays or kiosks, and in terms of the NPP simulations it can run. That is, the HSSL simulator is designed to support the different physical layouts of MCRs, and numerous models of currently operating Pressurized and Boiling Water Reactors (i.e., NPPs). The HSSL simulator is also capable of

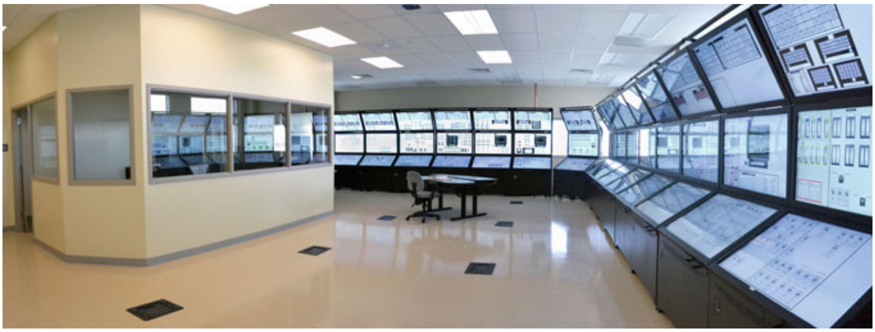

Fig. 3 The human systems simulation laboratory at Idaho national laboratory

supporting small modular reactor simulations, and potentially other advanced control rooms for next generation NPP designs. Full scope means that the simulator encompasses all of the critical functions found in a NPP MCR. It is a high fidelity simulator that is able to simulate both normal conditions and a wide range of abnormal plant conditions. Full scale means that the simulator is capable of faithfully reproducing the physical layout of the displays and controls for many different MCRs. The 15 bays, each containing 3 large screen monitors, are capable of displaying the front panels of many different MCRs.

The 45 large screen displays of the HSSL simulator display virtual representations of both analog and digital indicators and controls. This is an obvious departure from realistically representing the physical ergonomics of a NPP MCR (analog controls in particular), but it was necessary to do this for a number of reasons. First, each NPP MCR is unique in terms of the layout of the displays and controls, even at many multi-unit stations. The ability to quickly represent the different layouts of displays and controls at different NPPs necessitates the simulator presenting them virtually. Second, because the HSSL is an R&D simulator and the researchers want to rapidly prototype new digital I&C solutions, they need to have the flexibility to change displays and controls quickly, which is easily achieved through virtually representing them. A physical reconfiguration of the MCR boards each time the simulator runs a different NPP model, or when the researchers alter a digital I&C solution, in terms of its physical location on the control boards, functionality, look and feel, etc. would be labor intensive and not cost-effective relative to doing these activities in a virtual environment.

4 Research Approach

Generally speaking, the goal of these R&D activities is to evaluate the effects of new digital I&C technologies on human and overall system performance to ensure that performance with the new digital I&C technologies is at least as good as, if not better than, performance with the existing I&C system [12]. This goal is achieved by using the Guideline for Operational Nuclear Usability and Knowledge Elicitation (GONUKE) framework [13] as the general research approach. GONUKE is a general methodology derived from standard Human Factors usability testing and pedagogical evaluation [14], and as Fig. 4 shows, is comprised of four R&D activities (heuristic evaluation, usability testing, design verification, and integrated system validation), which are a function of the evaluation phase (formative vs. summative) and the evaluation type (expert review vs. user testing). Integrated system validation is a nuclear domain specific term, but refers to running ‘operator-in-the-loop’ studies whereby operators run through normal operating scenarios and critical abnormal scenarios using both the existing analog and new digital I&C technologies to assess human and overall system performance.

Evaluation Type	Evaluation Phase: Formative	Evaluation Phase: Summative
Expert Review (Verification)	Heuristic Evaluation	Design Verification
User Testing (Validation)	Usability Testing	Integrated System Validation

Fig. 4 Simplified GONUKE usability matrix

Additionally, as it is the case for most other Human Factors research (e.g., smart phone design), it is critical for this research approach to factor in how the human system interface design of new digital I&C technologies is affected by parameters such as:

- Desired information density
- Monitor/screen size
- The number of monitors/screens to be used
- The type(s) of input device(s) that will be used
- The underlying navigation philosophy (e.g., navigation structure and capabilities)

and how these factors subsequently affect human cognition and behavior, overall system performance, and the economic competitiveness of the NPP relative to other electrical generation sources.

Given the research goals and Human Factors Engineering design parameters listed above, the R&D approach INL researchers use is a blend of the GONUKE framework, standard Human Factors measurement constructs, tools, and methods, and approaches that are specific to the nuclear domain. For example, for the GONUKE R&D activities involving user testing, INL researchers rely on a standard set of Human Factors measurement constructs to assess performance, such as task success, task time, efficiency, satisfaction, errors, and learn-ability [15, 16], but also make use of analytical assessment techniques derived from Human Reliability Analysis. With respect to measurement tools, INL researchers use both standard tools, including: mobile eye trackers, physiological measures, scenario 'freeze probes', simulator logs, audio-video recordings, behavioral observations, interviews, and surveys, as well as specialized versions of these tools [17]. For GONUKE activities involving expert review, INL researchers rely on standard Human Factors Engineering analytical methods [18], but use nuclear domain specific standards and guidelines [19].

5 Conclusion

Since 2012, researchers at INL have been using simulation to conduct safety-focused research under the U.S. DOE LWRS program to develop the scientific basis to extend the operating life of existing NPPs. One focus area for LWRS is the NPP MCR, because many of the I&C system technologies installed in the MCR, while highly reliable and safe, are now difficult to replace and are therefore limiting the operating life of the NPP. INL researchers have been using the HSSL simulator to evaluate new I&C technologies, and get a head start on training operators to the new technologies, before the MCR, or even the training simulator at the plant, is modified. The HSSL is currently the only opportunity for many U.S. utilities to work with new I&C systems at full scale to test how it will integrate with their existing plant I&C systems. With the HSSL simulator, the preliminary design of new I&C technologies can be modified based on what is learned to further improve plant safety and efficiency prior to implementation, which is a significant advantage and cost-savings opportunity for any NPP engaged in MCR modernization.

Acknowledgments INL is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517. This work of authorship was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights. The United States Government retains, and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof. The INL issued document number for this paper is: INL/CON-16-38104. The lead author also thanks Dr. Allison C. Joe for her help in preparing this article.

References

1. World Nuclear Association: Nuclear Power in the USA, http://www.world-nuclear.org/information-library/country-profiles/countries-t-z/usa-nuclear-power.aspx
2. Gates, B., Gates, M.: Annual Letter 2016, https://www.gatesnotes.com/2016-Annual-Letter
3. Smil, V.: Energy transitions: history, requirements, prospects. LLC. Santa Barbara, CA, ABC-CLIO (2010)
4. Hallbert, B., Persensky, J., Smidts, C., Aldemir, T., Naser, J.: Report from the light water reactor sustainability workshop on advanced instrumentation, Information, and control systems and human-system interface technologies. Technical Report INL/EXT-09-16631, Idaho National Laboratory, Idaho Falls, ID (2009)
5. Joe, J.C., Thomas, K., Boring, R.: Establishing the value chain for the inclusion of new human factored technologies in U.S. commercial nuclear power plants. Procedia Manufact. **3**, 1312–1318 (2015)

6. Boring, R.L., Agarwal, V., Joe, J.C., Persensky, J.: Digital full-scope mockup of a conventional nuclear power plant control room, Phase 1: installation of a utility simulator at the Idaho national laboratory. Technical Report INL/EXT-12-26367, Idaho National Laboratory, Idaho Falls, ID (2012)
7. Boring, R.L.: Overview of a reconfigurable simulator for main control room upgrades in nuclear power plants. In: Proceedings of the human factors and ergonomics society 56th annual meeting, pp. 2050–2054. Sage Publications, Thousand Oaks, CA (2012)
8. Boring, R.L.: The use of simulators in human factors studies within the nuclear industry. In: Skjerve, A.B., Bye, A. (eds.) Simulator-Based Human Factors Studies Across 25 Years. Springer-Verlag, London, UK (2011)
9. Hugo, J., Boring, R., Hanes, L., Thomas, K.: A reference plan for control room modernization: planning and analysis phase. Technical Report INL/EXT-13-30109, Idaho National Laboratory, Idaho Falls, ID (2013)
10. Claude, E., Fink, R., Paul, K.: Guideline on licensing digital upgrades EPRI TR-102348 revision 1 NEI 01-01: a revision of EPRI TR-102348 to reflect changes to the 10 CFR 50.59 rule. Technical Report EPRI TR-1002833, Palo Alto, CA (2002)
11. Boring, R., Agarwal, V., Fitzgerald, K., Hugo, J., Hallbert, B.: Digital full-scope simulation of a conventional nuclear power plant control room, Phase 2: installation of a reconfigurable simulator to support nuclear plant sustainability. Technical Report INL/EXT-13-28432, Idaho National Laboratory, Idaho Falls, ID (2013)
12. Boring, R., Agarwal, V.: Beyond integrated system validations: use of a control room training simulator for proof-of-concept interface development. 8th International Topical Meeting on Nuclear Power Plant Instrumentation. Control, and Human-Machine Interface Technologies (NPIC & HMIT), pp. 1389–1399. American Nuclear Society, La Grange Park, IL (2012)
13. Boring, R.L., Ulrich, T., Joe, J.C., Lew, R.: Guideline for operational nuclear usability and knowledge elicitation (GONUKE). Procedia Manufact. **3**, 1327–1334 (2015)
14. Scriven, M.: The methodology of evaluation. In: Stake, R.E. (ed.) Curriculum Evaluation. Rand McNally, Chicago, IL (1967)
15. Boring, R.L., Joe, J.C., Ulrich, T., Lew, R.: Early-stage design and evaluation for nuclear power plant control room upgrades. In: Proceedings of the 2014 International Annual Meeting of the Human Factors and Ergonomics Society, pp. 1909–1913. Sage Publications, Thousand Oaks, CA (2014)
16. Boring, R., Joe, J.C., Ulrich, T., Lew, R.: Light water reactor sustainability program operator performance metrics for control room modernization: a practical guide for early design evaluation. Technical Report INL/EXT-14-31511 (Rev. 1), Idaho National Laboratory, Idaho Falls, ID (2015)
17. Skråning, G.: The operator performance assessment system (OPAS). Technical Report HWR-538, OECD Halden Reactor Project, Halden, Norway (1998)
18. Vicente, K.: Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. Lawrence Erlbaum Associates, Mahwah, NJ (1999)
19. O'Hara, J., Brown, W., Lewis, P., Persensky, J.: Human-system interface design review guidelines. NUREG-0700, Rev. 2, U.S. Nuclear Regulatory Commission, Washington, DC (2002)

Erratum to: Consumer Electric Energy Management Strategies and Preferences in Emergency Demand Response: Results from a Survey

Huiyang Li, Haya Salah and Ziang Zhang

Erratum to: “Consumer Electric Energy Management Strategies and Preferences in Emergency Demand Response: Results from a Survey” in: S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_8

The book was inadvertently published with an incorrect last name for the chapter author as Salah1 instead of Salah. The erratum book and the chapter has been updated with the changes.

The updated original online version for this chapter can be found at
DOI 10.1007/978-3-319-41950-3_8

H. Li (✉) · H. Salah · Z. Zhang
State University of New York at Binghamton, Binghamton, NY, USA
e-mail: hli@binghamton.edu

H. Salah
e-mail: hsalah1@binghamton.edu

Z. Zhang
e-mail: zhangzia@binghamton.edu

S.M. Cetiner et al. (eds.), *Advances in Human Factors in Energy: Oil, Gas, Nuclear and Electric Power Industries*, Advances in Intelligent Systems and Computing 495, DOI 10.1007/978-3-319-41950-3_18