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)