2022-11

Human Performance Test Facility Volume 7

Differences Between Hard Panel and Glass Top Simulator Ergonomics, Workload, and Usability

Date Published: January 2026

Prepared by:
K. Dickerson[1]
H. Watkins[1]
O. Gideon[2]
J. Sollars[1]
R. Boring[2]

[1]U.S. Nuclear Regulatory Commission
[2]Idaho National Laboratory

Kelly Dickerson, NRC Principal Investigator

Research Information Letter
Office of Nuclear Regulatory Research

Disclaimer

Legally binding regulatory requirements are stated only in laws, NRC regulations, licenses, including technical specifications, or orders; not in Research Information Letters (RILs). A RIL is not regulatory guidance, although NRC's regulatory offices may consider the information in a RIL to determine whether any regulatory actions are warranted.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to the Technical Training Center instructors for dedicating their time to serve as participants in the research described in this report. The authors would also like to thank the examiner for their time. The insights shared by this group have been critical to the success of this project. We are not naming the participants to protect their confidentiality as research participants.

ABSTRACT

Increased interest in plant restarts, development efficiencies, and cost considerations have led to an increased interest in using glass top simulators as an alternative to plant referenced simulators for nuclear power plant operator training and licensing. To understand the impacts of this alternative, this report documents qualitative observations and usability, workload, and ergonomic assessments for a direct comparison between hard panel, conventional, and glass top simulators under scenario conditions representative of licensing exam scenarios. We found that usability issues in the hard panel simulator were replicated and exacerbated in the glass top simulator in addition to several unique challenges related to the use of a touchscreen glass top simulator. Workload was higher in terms of effort and frustration while using the glass top simulator for both operators and the examiner participant. Additionally, both operators and the examiner rated their performance lower for the glass top relative to the hard panel simulator. The assessment also revealed that the two simulator types were equivalent in terms of ergonomic challenges. The results are discussed along with suggestions for solutions to the identified issues.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

1 INTRODUCTION

1.1 Regulatory Motivation

Nuclear power plants operating in the U.S. are required to have a high-fidelity simulator that matches the plant main control room for training and testing of reactor operators (10 CFR 55.46). However, there are circumstances when an applicant may need to conduct testing (e.g., validation and verification, operator training, licensing) on a simulator that is not an exact replica of the plant. In some cases, the simulator may incorporate touchscreens or “glass top” simulator bays. The NRC has specific definitions for simulators that qualify as an exact replica of a plant. In 10 CFR 55.4, a plant-referenced simulator (PRS) is defined as a simulator modeling the systems of the reference plant with which the operator interfaces in the control room, including operating consoles, and which permits use of the reference plant’s procedures. 10 CFR 55.46(c) lists several features of PRSs used for administration of operating tests or used to meet experience requirements; these include:

- (c)(2)(i) Utilizes models representing nuclear and thermal-hydraulic characteristics that replicate the most recent core load in the nuclear power reference plant, and
- (c)(2)(ii) Simulator fidelity has been demonstrated so that significant control manipulations are completed without procedural exceptions, simulator performance exceptions, or deviation from the approved training scenario sequence.

A commission approved or non-plant referenced simulator (NPRS) is defined in 55.46(b) and includes these key features:

- Description of system components or the way the simulation facility would be used for the operating tests.
- Description of the performance tests for the simulation facility and their results.
- A description of the procedures for maintaining examination and test integrity consistent with the requirements of § 55.49.

The Commission will approve a simulation facility for administration of operating tests if it finds that the simulation facility and its proposed use are suitable for the conduct of operating tests for the facility licensee's reference plant (10 CFR 55.46(b)(2)) under 10 CFR 55.46(b)(1) or 10 CFR 55.45(b)(3).

1.2 Simulator Fidelity

The NRC and the nuclear power industry have adopted a consensus standard for simulators used in operator licensing exams (ANSI/ANS-3.5-2009, see also RG 1.149, Rev 4). The standard can be applied to plant-referenced simulators or any simulation facility used for meeting the experience requirements in 10 CFR 55.31(a)(5). A real-world example of meeting this standard can be found in the Vogtle Units 3 and 4 Commission Approved Simulator (CAS) report. The Vogtle simulation facility was not plant referenced at the time it was reviewed. However, the Vogtle simulation facility represented the anticipated core model with sufficient fidelity that operators could execute the control manipulations required by 10 CFR 55.31(a)(5) without procedural exceptions, simulator performance exceptions, or deviations from approved training scenario sequences.

Often the NPRS requirements in 10 CFR 55.46(b)(2) are interpreted by applicants as a call to maximize simulator fidelity. Vogtle’s NPRS is an example of efforts to maximize fidelity. However, an alternative approach is emerging. Glass top simulators and their role in licensing, training, and validation have become central to the human factors and operator licensing discussions around plant restarts – applicants are considering using touchscreen-based glass top simulators, modeling the most recent core load prior to decommissioning. Restarting plants simply may not have the ability to build a full fidelity mimic of their control room during the initial operator licensing classes and look to glass tops as a potential solution. Applicants may also be interested in glass top simulators because of cost consideration and the convenience of being able to rapidly iterate and test different HSI designs. However, recent human factors and ergonomics research suggest that glass top simulators as mimics for control rooms may be a unique case where the attempt to maximize fidelity creates a more error prone and difficult environment.

The consequences of introducing new types of human errors and system malfunctions could be significant during operator licensing exams. Future licensed operators are tested on scenarios where at least one critical task must be performed. NRC staff define critical tasks as those used in a simulator scenario to evaluate whether an individual or crew can complete actions that are significant to the safety of the plant and the public (NUREG-1021, ES-3.3(k)). Critical tasks can range from simple, safety-significant tasks (e.g., tripping a reactor coolant pump during a small break LOCA) to tasks that require a higher level of skill and involve multiple crew members. During a licensing exam, failure to properly execute a critical task during a scenario or a job performance measure (JPM) could result in consequences as severe as failing the exam and thus being denied an operator license. The presence of critical tasks is an important distinguishing feature between touchscreen interactions in consumer electronics setting, and a safety critical setting like nuclear process control. The ability for examiners to clearly distinguish between an error in a critical task that is due to the interaction modality (touchscreen vs. hard panel) or a knowledge limitation is a key consideration for the use of touchscreens in a licensing context.

1.2.1 Fidelity Levels and Types Differences

Given the standards, guidelines, and regulations governing the functional properties of either a PRS or NPRS, there are features that must be represented with high fidelity and those that can potentially be represented with lower fidelity. Features that can be represented with lower fidelity will have varying human performance consequences depending on factors like operator experience, workload during the scenario, and the specific type of fidelity that is degraded and to what extent. The consequences of lowering fidelity also depend on the specific simulator use case, for example, the medical simulation literature has demonstrated that inexperienced doctors benefit more from simulations with low physical fidelity because higher fidelity graphics can distract away from learning the underlying cognitive concepts (Offenbacher et al., 2022). There are similar findings from research with air traffic controllers, where new controllers use tabletop airstrips rather than immersive simulations to learn fundamentals without the distraction of the air traffic control technology. There is often benefit for new professionals to have the learning objectives of fundamental skills and technological skills presented separately. While beneficial for new professionals, a very low fidelity simulator (like in Figure 1) would not likely be suitable for reactor operator licensing purposes because it would require changes in the way control manipulations are performed and likely lead to a need for procedure exceptions.

Figure 1. Low Fidelity Example from Air Traffic Control.[1]

Section 11.4.3.3 of NUREG-0711 describes the human system interface (HSI) testbed parameters used for simulator-based evaluations. In the context of section 11.4.3.3, NUREG-0711 refers to integrated system validation (ISV), but simulator-based evaluations are also used for licensing exams and measuring learning outcomes for simulator-based training.

Table 1. Fidelity Types Defined in NUREG-0711.

Fidelity	Definition
Completeness	Testbed should represent the complete integrated system, including HSIs and procedures not specifically required in the test scenarios.
Physical Fidelity	HSIs and procedures should be represented with high physical fidelity to the reference design. Including: presentation of alarms, displays, controls, job aids, procedures, communications equipment, interface management tools, layout, and spatial relationships.
Functional Fidelity	HSIs and procedure functionality should be represented with high fidelity. All HSI functions should be available.
Environmental Fidelity	The testbed's environmental fidelity should be represented with high physical fidelity to the reference design, including the expected levels of lighting, noise, temperature, and humidity.
Data Completeness Fidelity	Information and data provided to personnel should completely represent the plant's systems they monitor and control.
Data Content Fidelity	Data content should be represented with high physical fidelity to the reference design. The presentation of information and controls should rest on an underlying model accurately mirroring the reference plant. The model should provide input to the HSI such that the information accurately matches that which is presented during operations.

[1] Figure sourced from: https://aviationsafetymagazine.com/airmanship/the-atc-academy/).

Data Dynamics Fidelity	Data dynamics should be represented with high fidelity to the reference design. The process model should be able to provide input to the HSI so that information flow and control responses occur accurately and within the correct response time, e.g., information should be sent to personnel with the same delays as occur in the plant.

Fidelity is not the only concern with NPRS, some NPRS may utilize glass top displays with a touchscreen HSI, where the operator's finger or a stylus is used as the input device instead of the conventional keyboard and mouse. Touchscreens have several special ergonomic and performance impacts that could have implications for how operators are tested, the success of training, and the confidence of outcomes from validation events. Unlike simulators in general, the standards covering touchscreens appear to be more limited. ISO 9241 covers touchscreens in Part 410, which describes design criteria and recommendations for a number of physical input devices. Similar to NUREG-0700, ISO 9241 Part 410 provides recommendations for the size, location, and shape of physical and virtual buttons. (Part 410 includes additional button parameters for physical buttons). The standard also describes guidelines for force feedback, actuation force, and button tactile responses along with other feedback modalities to maintain operator situation awareness (SA) during interaction with the HSI.

While NUREG-0700 and NUREG-0711 have information relevant to touchscreens and glass top simulators, there is relatively little research focused on transfer of training, simulator ergonomics, or touchscreen ergonomics that specifically focuses on the nuclear domain, which has special considerations because of the size of the typical control room and the size of nuclear control room control panels. RIL 2022-11, Volume 5 compared touchscreen, analog, and desk top simulators with both operators and students. While the simulator configurations varied, the research demonstrated via physiological measures and performance-based ergonomic information that, while a user may report a preference and even lower workload while using a touchscreen, they experience higher workload generally and reduced accuracy depending on the location of the controls.

1.3 Literature on Touchscreen Ergonomics

Touchscreens ergonomics for glass top control rooms were explored in a modality assessment study by Ulrich, Boring, and Lew (2015). Comparing touchscreens to both track pad and mouse, they found that while operators reported a preference for touchscreen use, operator performance was worse than both the track pad and mouse conditions. RIL 2022-11 (volume 5) reported a similar dissociation between preferences and performance. Students and operators reported lower workload when using a touchscreen compared to a keyboard and mouse, however, physiological measures of workload (electroencephalogram and heart rate variability) demonstrated that their cognitive load was higher when using the touchscreen.

The broader literature on the usability of large screen touchscreens also points to potential challenges. The vertical "bays" in control rooms require the use of unsupported mid-air gestures (the arm has no external support as the operator engages the HSI), which are notorious for inducing fatigue, quickly (Hansberger, 2017). While this gesture type would be present in the control room, hard panel, and glass top simulators, the issue is exacerbated by the repetitive engagement of pointing and pinching gestures. Unsupported mid-air gestures are not the only usability issue. Davis et al. (2014) compared different postural positions for touchscreen, keyboard, and mouse data entry conditions. In the touchscreen conditions, they found participants had the most discomfort in standing positions with low vertical and horizontal screens (Figure 2, panels H and K). This discomfort would be exacerbated if required for a prolonged interaction or repeatedly over the course of a shift. Bodily fatigue can impact motor control and be a potential contributor to gestural errors (e.g., inadequate force, excessive force, missed targets, early release of controls).

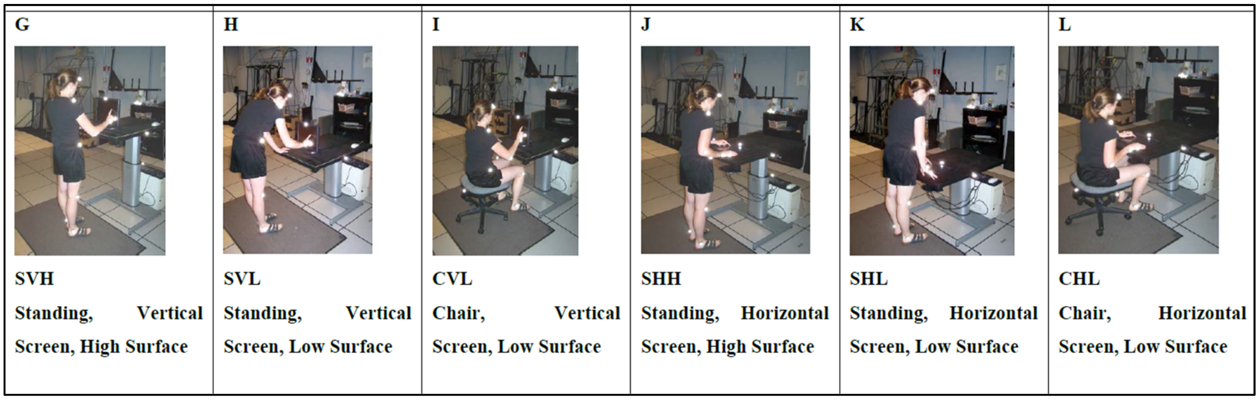


Figure 2. Touchscreen Configurations Tested in Davis et al., (2024).

Additionally, Breuninger (2020) found that the gesture-based interactions and special surface characteristics create a situation where the user is challenged with limited visual and haptic input feedback, mismatch between the size of interactive elements and the digits used for interaction (i.e., occlusion, precision), confusion induced by poorly implemented virtual physics and motion effects, and cognitive interference (e.g., using an upward swipe to access lower parts of a page). Translated into design principles, Breuninger outlined three core issues to manage in a touchscreen interaction design, including:

- Occlusion – finger or other parts of the hand can block items on the screen. The common assumption in Breuninger (2020) and consumer electronic touchscreens in general is that a digit usually approaches from below the device. When this is the case, the finger and hand will occlude most elements directly below the target. Interactive elements must be big enough to be visible and positioned outside of the "occlusion zone". The size and positioning of touchscreen panels in a glass top simulator could also have issues with occlusion if an operator overshoots a control, which then is occluded by not only their hand, but also their arm.
- Feedback – touchscreens lack mechanical buttons and thus do not readily provide feedback about the success of an action. The lack of haptic feedback from the solid surface can be addressed to some degree by adding vibration actuators. Generally, users are relying primarily on visual feedback for assessing the success of their actions. Over reliance on visual cues to make up for the inherent haptic deficiencies of the touchscreen introduces the potential for workload to become unmanageable if the operator must engage in a monitoring or other visually intensive task, while executing manual gestures on the glass top. They would need to shift attention to the virtual button for effective gesture targeting, maintain attention on the target control until feedback is detected and then return to their previous visual task.
- Precision – generally, for experienced users, precision is comparable to mouse and trackball. However, the "fat finger problem" is a consideration. Digits on the hand are larger than the mouse pointer. Interactive elements presented on a touchscreen will have to be larger and further apart to accommodate reduced precision. This could be addressed with a stylus. Other considerations related to precision include:

- High vibration area, where the touchscreen or user’s hand may inadvertently move in the time between targeting a control for action and contact with the touchscreen.
- Accuracy using virtual keyboards: keys must be larger and spaced further apart to prevent typographical errors.
- Parallax errors, where the viewing angle influences the touchscreen user’s ability to align the visual information with their gestural positioning.
- User individual differences. Touchscreens are challenging for those with generalized weaknesses, motor limitations, or a tremor.

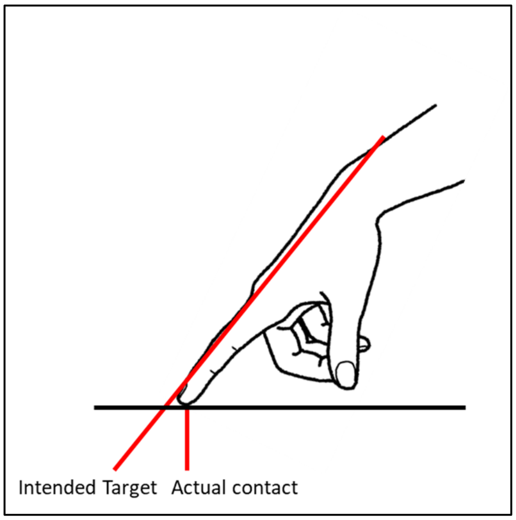


Figure 3. Example of a Parallax Error.

- Posture – with the rare exception of horizontal touchscreens, all touchscreens require the user to hold up the hand or arm for some period of time, during input. There is considerably more shoulder, arm, and hand strain compared to pointer-based input (see also Hansberger, 2017).
- Larger touchscreens – like those intended for use in glass top simulators have more content out of reach of more users (e.g., 5th or 95th percentile people) and have more content requiring prolonged mid-air unsupported gestures, leading to fatigue.
- Complexity – roughly equivalent for selection/activation between point and click, however, touchscreens usually involve many different types of gestures.
- Layered or “secondary functions” are more difficult to integrate for touchscreens.

Orphanides and Nam (2017) provide several other interaction design considerations. They found that users tend to press harder than necessary to activate touchscreens contributing to potential issues with fatigue. They also found that when users move around (e.g., walking, leaning, twisting), error rates tend to be higher, and these errors are not mitigated by familiarity with using gesture-based input. Some consideration for the control room environment, where the user will have been in motion just prior to engaging with the controls may be needed. Orphanides and Nam provided some guidance on the design of touchscreen interfaces to ensure reliable use. Orphanides and Nam’s guidelines are consistent with NUREG-0700, which recommends that multi-touch gestures should have a button-based alternative and that audio, visual, and/or tactile feedback should be incorporated into the interaction design. NUREG-0700 also suggests additional visual cues on icons assist with visual targeting and provides specific suggested guidelines for virtual QWERTY keyboard.

2 RESEARCH QUESTIONS

Previous research, the broader literature on touchscreen usability, and the recent interest in using glass top simulators highlight the need to gain a better understanding of the impact of glass top simulators in a licensing, training, and validation context.

2.1 Fidelity and Licensing

- How will differences between a glass top and analog (hard panel) simulator impact training effectiveness?
- Are there differences between a glass top and a hard panel simulator that will make conducting licensing examinations more difficult?
- Are there clear mitigation strategies for addressing limitations associated with glass top simulators?

2.2 Ergonomic, Usability, and Workload Differences

- Will glass top simulators put more strain on the operator's shoulders, arms, and wrists?
- Are there operator actions or bodily positions unique to the glass top simulator, such that it would impact examiner confidence or ability to conduct an exam?
- What types of workload are most impacted by changing from a hard panel to a glass top simulator?
- Do operators rate the overall usability of the glass top simulator lower than the hard panel simulator?

The present report (Volume 7) will attempt to address these questions based on qualitative observations and survey-based assessments. A future RIL in this series (Volume 8) will expand on the insights reported here by adding simulator-based performance data and an analysis of eye movement patterns and their potential differences between the two simulator types. Volume 8 will also address questions about action efficacy and efficiency differences between the hard panel and glass top simulators.

3 METHODS

The studies included reactor operators performing relevant identical scenarios on glass top and hard panel training simulators for the same plant (see procedure section for details). This report includes survey data, qualitative observations, and an ergonomic assessment. As such, only method-related information relevant to those techniques is reported here. A follow-up report will be published that contains the performance and eye tracking data that were collected. The scenarios and overall procedures are the same as all measures were collected during the same session.

3.1 Participants

Six NRC Reactor Technology Instructors (RTIs) formed two 3-person crews with one person acting as the Control Room Supervisor (CRS) and two acting as reactor operators (ROs). Each person performed the same role across both simulator types and all scenarios. The study also included observations from one qualified NRC examiner.

3.2 Scenarios

The study included 2 scenarios. Table 2 provides a summary of those activities. Appendix A provides full scenario descriptions. This report includes data that was collected during, before, or after the scenarios. Data collected from JPMs will be included in the follow up report (Volume 8). Appendix A provides full scenario descriptions. The scenarios used were based on previously used exams from Callaway and South Texas Project, the simulators used for this study were configured to represent the Turkey Point control room. Callaway, South Texas Project, and Turkey Point are all similar sized 4-loop Westinghouse plants, therefore using previous exams from those sites as the starting point for study scenario development, is a reasonable approach. The procedures listed in Table 2 are those used in the NRC's Technical Training Center (TTC) "Trojan" hard panel and glass top simulator.

Table 2. Study Scenarios

Scenario	Description	Source	TTC Procedures
1. Tube leak and rupture, reactor fails to trip	50% power mode 1. Shutdown bank D – dropped rod. A component cooling water (CCW) pump trip with other pump failing to start. A tube leak on steam generator (SG) D, grows into a tube rupture. The reactor fails to trip (manual rod control). Operator will insert negative reactivity (boration) and depressurize the reactor cooling system (RCS) to the E-3 SI criteria and isolate the rupture.	Callaway 2022-1 Scenario 2	E-0 E-3 FR-S.1 GOI 5 ONI 2-4 ONI 3-12
SG level channel fails low, pressurizer PORV leak, SB	SG level median selection circuit fails. Pressurizer power-operated relief valve (PORV) develops leak in addition to a small break loss of	South Texas Project 2022 Scenario 1	E-0 GOI-5 ONI 2-6 ONI 8-1

LOCA, CCW pump failure	coolant accident (SB LOCA) on a cold leg (loop does not matter). Steam Generator Full-Power Test (SGFPT) trips overspeed, causing automatic runback. The other SGFPT fails leading to manual reactor trip.	ONI 36

3.3 Simulator Set Up

Both the hard panel Trojan simulator and glass top were configured as 4 loop pressurized water reactors (PWRs) set up to conduct scenarios with a normal crew complement. This included a CRS, an operator at the controls (OATC) and the balance of plant operator (BOP). Both simulators used the same initial conditions and were controlled by an instructor station running the same software. The glass top simulator had identical controls to the hard panel simulator. It additionally had the capabilities to review electronic copies of procedures from any of the panels; however, this was not used during the scenarios. The glass top simulator was 75% of the size of the hard panel simulator. The impact of the scaled reduction was that the location of controls were slightly different between the hard panel and glass top simulators.

3.4 Ergonomic Assessment

The Rapid Upper Limb Assessment (RULA, see appendix B) was used to perform an ergonomic assessment for both the hard panel and glass top simulators. The RULA was used to assess physical fatigue and overall body positioning. To perform the ergonomic assessment the experimenter asked the participant to reach out and touch a single control item on the back panel and another on the lower panel (approximately at waist level). The participant was asked to reach out several times (between 3-5 per panel) so that the researcher could observe their reaching action and score the gesture using the RULA's rubric. The upper panel was not assessed for upper body or postural ergonomics as those panels were primarily for displaying information and not interaction. Participants were also asked to estimate the force required to engage the controls. The RULA measures force on a scale of 0 (0 kg, no resistance) to 3 (10 kg or more of static or repeated load). The maximum total RULA score is 7 (see Table 3).

Table 3. RULA Score Interpretation

Level	Range	Interpretation
1	1-2	Posture is acceptable.
2	3-4	Ergonomics of the system or device need further evaluation.
3	5-6	Ergonomics of the system are poor and changes should be considered.
4	7	Ergonomics are very poor. Immediate changes are recommended.

3.5 Surveys

3.5.1 System Usability Scale (SUS)

The SUS (Brooke, 1996) is a 10-item questionnaire for measuring the perceived usability of a system (see Table 4). It is widely used across many industries as a quick standardized method and provides an overall assessment on a scale from 0 to 100. The long history and widespread use of the SUS led to the establishment of published industry-specific benchmarks. The average range for domains within industrial process control (e.g., chemical processing, oil and gas, manufacturing) is 60-75, with an industry average around 68. The SUS is also recognized as a method suitable for research with small samples.

Table 4. System Usability Scale Questions.

Reverse Coded?	SUS Question
No	I think that I would like to use this system frequently.
Yes	I found the system unnecessarily complex.
No	I thought the system was easy to use.
Yes	I think that I would need the support of a technical person to be able to use this system.
No	I found the various functions in this system were well integrated.
Yes	I thought there was too much inconsistency in this system.
No	I would imagine that most people would learn to use this system very quickly.
Yes	I found the system very cumbersome to use.
No	I felt very confident using the system.
Yes	I needed to learn a lot of things before I could get going with this system.

Note. "Reverse coded" means that a higher score indicates that the system performed worse for that item. Questions that are not reverse coded map higher scores to better performing systems. Using both conventional and reverse coding reduces participant response bias.

3.5.2 National Aeronautics and Space Administration Task Load Index (NASA-TLX)

NASA-TLX (Hart & Staveland, 1988) has six dimensions including mental, physical, temporal demand, performance, effort, and frustration level. For the mental, physical, temporal demand, effort, and frustration subscales, higher scores indicated more workload (or more workload-associated frustration). The performance subscale is interpreted differently. When performance scores are high, it reflects the participant's view that they performed better on the task. Together these dimensions provide an overall and domain estimate of an individual's subjective workload experience.

3.5.3 Demographic Information

Participants were asked to optionally disclose their age, years of experience as an RO, SRO, shift supervisor, and instructor. They were also asked about years of experience with both PWR and BWR technology. The ergonomic assessment combined with a small sample introduces the possibility of a prior injury history biasing the results. Participants were asked to disclose (via yes or no question) if they have previously had any surgical procedures or medical conditions that would limit their range of motion in their shoulder, arm, hands, fingers, and their height.

3.6 Procedure

For practical reasons both crews started in the hard panel simulator[2]. The procedures were identical in the hard panel and glass top simulators. Before starting the first scenario, all participants completed the demographic questionnaire. The first crew moved to the position in

[2] While it was not possible to counterbalance the order of using the glass top and hard panel, the high level of experience by the operators minimizes strong practice effects.

the room associated with their roles (e.g., operator at the controls (OATC), balance of plant (BOP), or CRS who read the procedures). The crew that was not participating in the scenario was not allowed to remain in the room during the scenario run to avoid additional uncontrolled exposure to the study materials. Participants not involved in the scenario were given three options for how to wait for their turn in the simulator. They could (1) participate in the second experimental task unrelated to the scenarios (thermal process dispatch task; see Gideon, Dickerson, Ulrich, & Lew, 2025), (2) return to their offices, or (3) wait in the study prep room. Each scenario started with the simulator set to the appropriate initial condition and placed in a freeze state. Crews did not have knowledge prior to completing the hard panel scenarios what the malfunctions would be. Once the crew was in position, the person at the instructor's station would provide a shift turn over crew brief. Upon stating "end brief" the simulator would be placed in a run state, and the scenario would be active. Each scenario was designed to take approximately 45 minutes.

At the end of each scenario, each crew member completed the NASA-TLX and SUS surveys. The RULA was conducted either before or after a crew had finished a scenario, as it required the simulator to be in freeze. Each crew completed two different scenarios on the hard panel simulator, followed by repeating the same scenarios on the glass top simulator the following day[3].[4]

During the scenarios, each member of the study team was assigned a specific operator to monitor. Each study team member shadowed the same crew member for the duration of the study. The purpose of the shadowing was to (1) capture qualitative aspects of performance, (2) generate a timestamp to "anchor" the simulator logs that will be used in the performance-based analysis, and (3) quickly capture actions that were not expected based on the scenario design. In some cases, it was possible for operators to begin attempting to diagnose a malfunction using a different than expected procedure, whereby the shadows could capture notes about what actions were taken that were not already documented.

[3] The simulators and TTC instructors that supported this research had limited availability. All data collection activities had to be completed within a five-day period.

[4] There were not enough crews to counterbalance effectively. The crews had slightly less familiarity with the glass top simulator so it was decided that the experienced gained from seeing the scenarios during the hard panel simulator conditions would address the potential limitations of the crews entering the scenario cold in the less familiar simulator.

4 RESULTS AND DISCUSSION

4.1 There are Limitations to Both the Hard Panel and Glass Top HSIs

Based on qualitative observations, generally all the operators were successful navigating both control rooms, and understanding the information presented through both digital and analog HSIs. However, there were challenges in both the hard panel and glass top simulator that impacted the efficiency of operator performance. Despite some challenges being present in the hard panel control room, the primary focus of this discussion will be on the glass top simulator.

4.1.1 Hard Panel Simulator

There were two general issues with the hard panel (which were replicated on the glass top because the control room is an exact match). First, operators expressed uncertainty about the status of valves with longer stroke times. The binary green/red status information light did not illuminate until the valve was fully open or fully closed. The result was operators questioning if they had applied sufficient force on the control to start the mechanical opening process. Additionally, analog controls for things like turbine level controls lacked tactile feedback (i.e., stops) for both the rate and extent of the change to guide the operator during operation. The operators had to rely on dedicating their visual attention to the specific indication associated with those controls. It was observed that when a malfunction required manual control, it created an attentional tunneling effect for the BOP operator.

Interestingly, because the dial's interaction requirements were demanding for operators on the hard panel simulator, the glass top simulator was developed with a virtual button placed next to each of those controls (see NUREG-0700 for similar design recommendation). The operators could turn the knob or could click the button to incrementally control steam generator level. This button also made a subtle clicking sound when depressed, which was easily overtaken by noise associated with control room activities.

4.1.2 Glass Top Simulator

The glass top simulator had a couple of usability issues that were specific to use of glass tops and have implications for ISV and examination. The two issues of note observed in the current study are virtual buttons "sticking" and glare.

4.1.2.1 Virtual Button Issues

There were a handful of instances where a virtual button got stuck and this sticking was not detected by the operator. The steam generator level control button did not change color when depressed or give any other visual indication it had been depressed and reset (or not). For example, if the operator wanted to increase the steam generator level by one unit by clicking on the virtual button, the simulator acted as if the operator was continuing to press the button, even after the operator's finger was removed from the virtual button. This caused a runaway level control action that was not intended and not immediately detected by the operator.

This button sticking issue could have been caused by several factors. Typically, software implemented on a touchscreen will contain bounding regions around active elements so that an operator action within a particular region is mapped to a specific identified control. It is possible the sticking was caused by the operator clicking at the edge of one of these software regions.

Another possible cause is that depending on the spacing of the electrodes in the electrode matrix layer, the operator's action could have active regions mapped to more than one control or that indicated a pinch rather than pointing action and confused the system or briefly changed the electrostatic field in an unusual way, causing the button to remain in the depressed position.

Spacing between electrodes in a projected capacitive (PCAP) touchscreen will vary depending on the size of the screen and specific implementation. A simulator used for validation or licensing purposes may opt for greater electrode density (if available for the required screen dimensions) than a training or research simulator, where gestural accuracy is less critical. A purely software cause is that events like clicking a virtual button are triggered by touching an area of the screen. Stopping touching that area of the screen should trigger the corresponding unclick function in the software. It is possible that the simulator software for the glass top had not undergone sufficient quality control to ensure all button press functions had a corresponding button depress function. This type of touchscreen glitch can be insidious – In the case observed, the glitch was noticed because it caused the plant to trip. It is unknown if this issue was present in portions of other panels that contained similar virtual buttons and the potential changes in plant parameters were minor enough to go unnoticed.

4.1.2.2 Glare

NUREG-0700 section 11.3.1.3-1 states that "glare should not interfere with the readability of the displays, labels, or indicators on the display device". The rationale for the glare guidelines is that reflections from ambient light sources can be disruptive, distracting, and potentially debilitating.

Glare was a non-issue in the hard panel control room. Every panel was equally visible from any position in the room. There was a substantial amount of visible glare produced on the touchscreen panels. There was glare regardless of lighting conditions. With the room lights on, the specular glare from the overhead lights was prominent (see Figure 4, left side). There was also some reflective glare from the back panel onto the lower panel. Operators were observed using one hand to block the glare coming from the back panel so they could access controls on the lower panel with greater certainty. When the room lights were on, the glare issues were minimal for the examiner and easily addressed by re-positioning relative to the operator and the screen.

When the room lights were off, the reflective glare improved somewhat for the operators; however, it completely obstructed the examiner's view of the controls and the operator's actions on those controls (see Figure 4, right side). The reflection of controls from the back panel onto lower panels made it nearly impossible to see what control was accessed. In the context of licensing exams, there would need to be some consideration of how to manage visibility issues for examiners, particularly for cases where an operator makes an error or is delayed in executing an action. It would be very difficult to know by visual observation alone (and the simulator log may be ambiguous) if an error was due to poor visibility or limited operator knowledge. Examiners would need to rely on asking follow-up questions to make that determination.

Importantly, as recommended in NUREG-0700 Section 11.3.1.3-5, the touchscreen panels were all equipped with an anti-glare coating (TouchPro, [projected capacitive touchscreen] PCAP). However, this coating was not sufficient to address the glare. Taken together, the glare challenges for the glass top simulator were significant, and licensees will need to specifically consider the lighting conditions in the room where the glass top simulator will be used to avoid this issue. This may mean that the control room and simulated control room may need to have very different lighting conditions to make the controls equally visible.

Some color schemes may also be more susceptible to glare, and faithfully replicating the colors of the hard panels may contribute to glare. The dark mode settings found in many software applications is an example of a glass top accommodation used to minimize eye strain and reduce glare. However, any deviation in the representation of the glass top displays and the hard panels means a difference between the simulator and the actual plant panels, resulting in a decrease in the fidelity of the simulator.

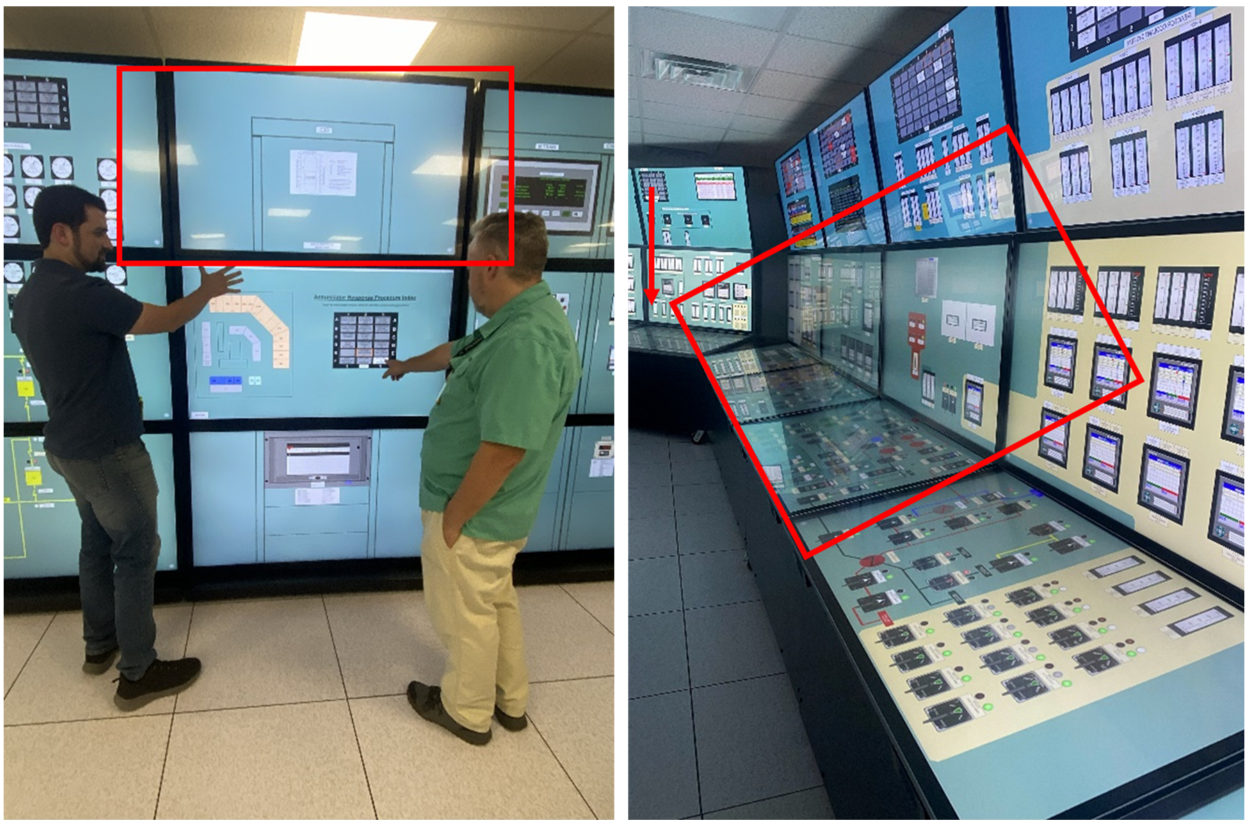


Figure 4. Glare Examples Under Different Lighting Conditions.

4.2 Survey Results

Given the small sample size, the surveys were analyzed using paired samples t-tests targeting comparisons of the two simulator types. There was *a priori* expectation that there would be differences between the hard panel and glass top simulators, specifically and that operators would experience more workload and find the glass top simulator less usable.

4.2.1 System Usability Survey

All six participants contributed data to the SUS. The SUS ratings were collected after both scenarios were completed on a simulator type (hard panel, glass top). The literature on the benchmarking and evaluation of the SUS reports that an average SUS score is 68. A poor or marginally usable score is 51, and the 90th percentile score is 80 (Lee, Lee, Ha, & Seong, 2016; Fernandes, Bloch, Braseth, & McDonald, 2023).

There was a significant difference between the hard panel (*M* = 76.88, *SE* = 5.62) and glass top (*M* = 41.25, *SE* = 4.41) simulators in terms of overall SUS score ($t(5)$ 3.63, p = .008, Figure 4). Importantly there was only a minor numerical difference in scores between the crews (around 1.0 for the hard panel and .5 for the glass top). These results suggest that operators found the glass top simulator to be more difficult to use and less useful than the hard panel simulator. Additionally, the hard panel simulator rating was well above the SUS benchmark average, and the glass top was in the marginal usability range. It is important to note that while all operators were fully proficient on both the location and operation of the glass top and hard panel controls, they had more experience using the hard panel simulator, so it is possible that the higher usability ratings for the hard panel are based in part on a bias favoring the hard panel simulator.

While bias is always a possibility, particularly with smaller sample sizes, recent research from NASA (Shelat, Homer, Jarasinski, & Marquez, 2025) found that there is a significant positive correlation between SUS ratings and perceived novelty, specifically, participants will rate a system as more usable on the SUS if that system is more novel. The NASA results suggest that, in this case, the additional time with the hard panel simulator would be unlikely to produce the SUS differences observed in hard panel vs. glass top comparison.

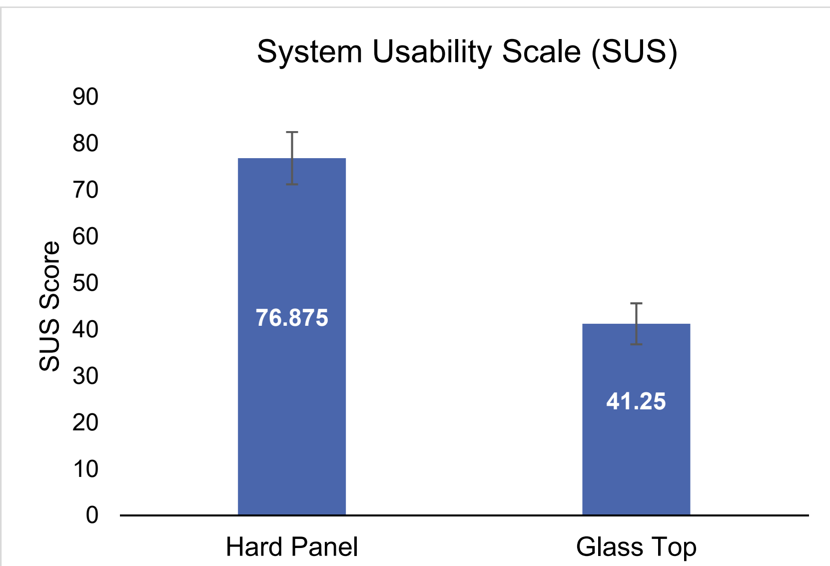


Figure 5. Average System Useability Scores (N = 6) for Hard Panel and Glass Top Simulators.

A review of the individual SUS questions revealed that the glass top simulator was rated lower than the hard panel simulator on all items except 1 – “I would imagine that most people would learn to use this system very quickly.” On a scale from 1 to 5 operators rated the hard panel simulator as 2.20 and the glass top as 2.5, suggesting that the operators felt that it would be slightly easier to learn how to use the glass top simulator than the hard panel simulator. See Table 3 for the 10 SUS questions.

4.2.2 NASA-TLX: Operators

Both operators and the examiner filled out the NASA-TLX after each scenario for both the hard panel and the glass top simulator.

4.2.2.1 Scenario 1 Steam Generator Tube Leak, Rupture, and Reactor Failed to Trip

Paired samples t-tests comparing the hard panel and glass top simulator were conducted (see Figure 6). There was a statistically significant difference between the hard panel and glass top simulator types on the performance ($t(4)$ 1.81 p = .072[5]), effort ($t(4)$ -2.75, p = .026), and frustration (t(4) -7.09, p =.001) subscales. No other comparisons were significant.

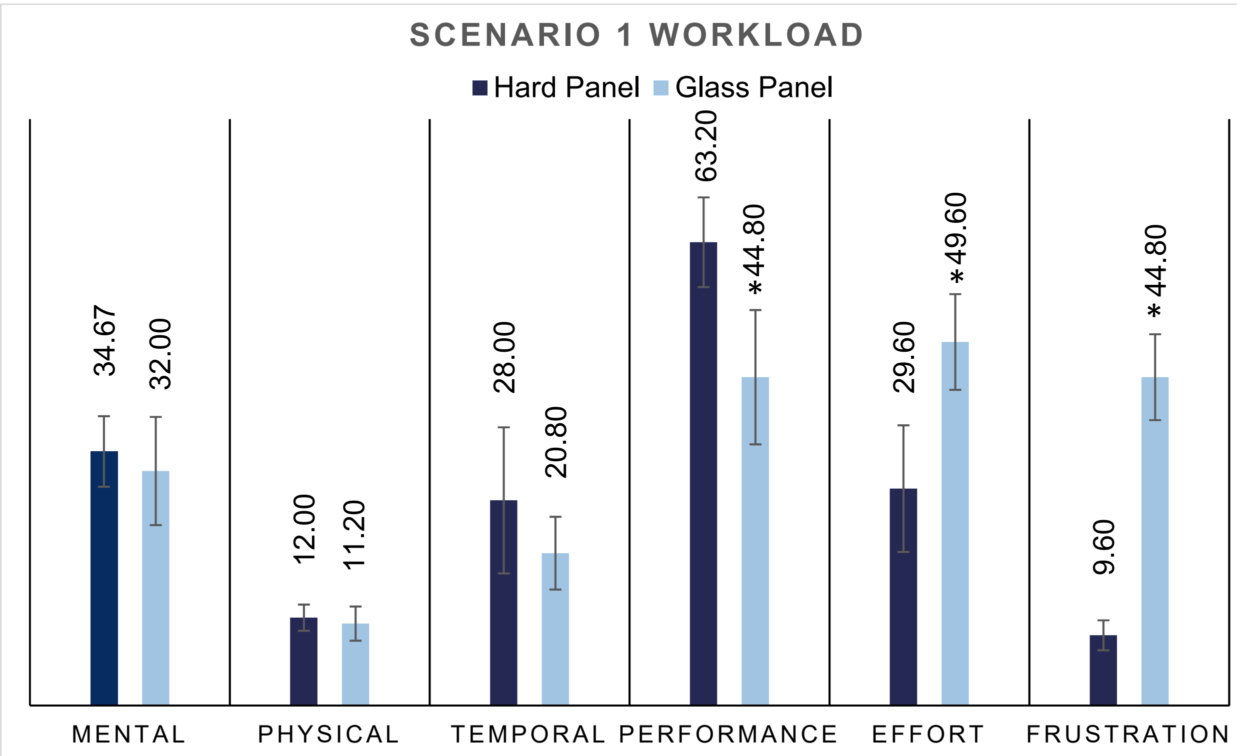


Figure 6. NASA-TLX Ratings for Operators - Scenario 1.

4.2.2.2 Scenario 2 – Steam Generator Level Channel Fails Low, PORV Leak, LOCA.

Paired samples t-tests comparing the hard panel and glass top simulator were conducted (see Figure 7). The results for scenario 2 were consistent with scenario 1. There was a statistically significant difference between the hard panel and glass top simulator types on the effort ($t(2)$ –2.302, p = .074), and frustration (t(2) -7.00, p =.010) subscales.[6] No other comparisons were significant.

5 Trending towards significance on a one tailed test.
6 Some operators opted to skip some NASA-TLX workload classification ratings. These missing data caused the change in degrees of freedom for subscale tests in Scenario 2.

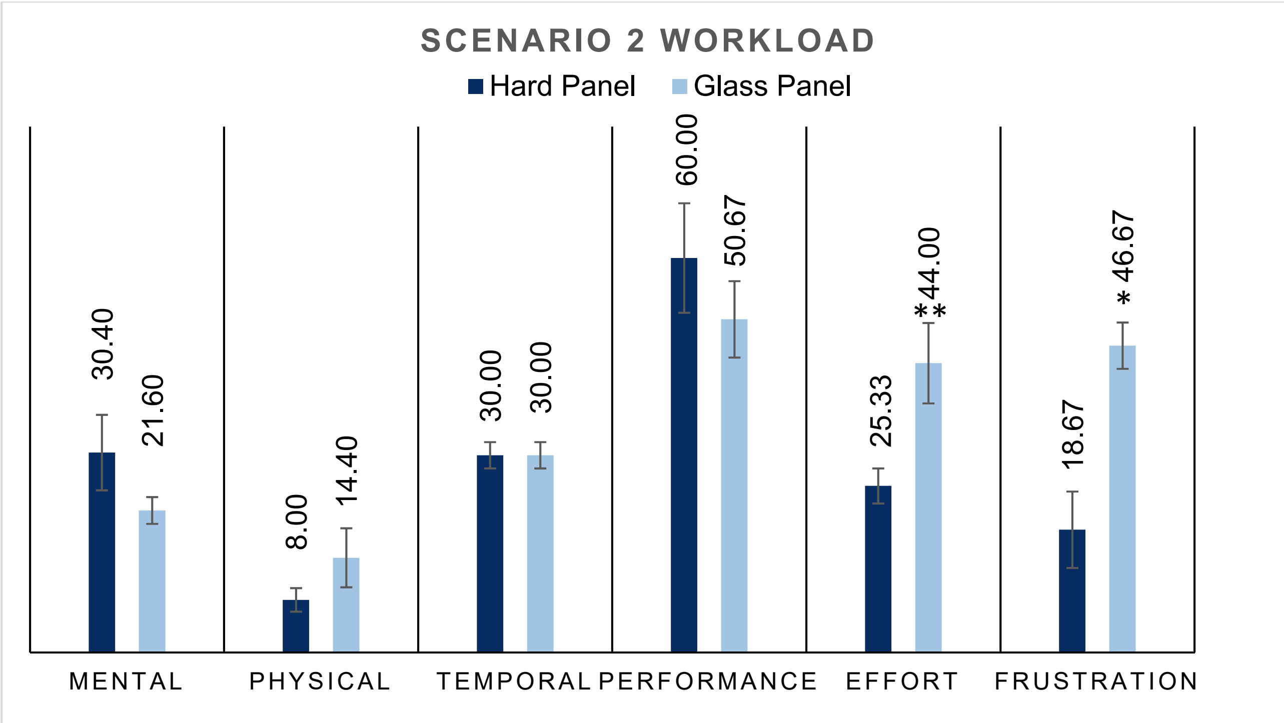


Figure 7. NASA-TLX Ratings for Operators - Scenario 2.

4.2.3 NASA-TLX: Examiner

The examiner completed 3 out of 4 NASA-TLX survey opportunities (see Figure 8 for results). Workload was generally perceived as very low during the hard panel scenario. The examiner also felt their performance was better when observing an operator during a hard panel scenario. Their workload when observing operators during glass top scenarios was approximately double their hard panel scenario workload across all subscales. The examiner indicated that they had to stay much closer to the operator to complete their observation and that they had persistent difficulty seeing the operators' actions and judging if the correct control was used. This was both because the operators were generally standing closer to the panel for a greater proportion of the time in the scenario and differences in height between the examiner and the operators. In this case the examiner's height was 15th percentile for women and she was tasked with observing operators who on average were in the 65th height percentile for men. The examiner standing closer to the operators during the glass top scenarios, the glare, and the height differential made it possible for the operators to occlude the examiner's view of the controls with their body and limit the examiner's available options for repositioning themselves to improve visibility.

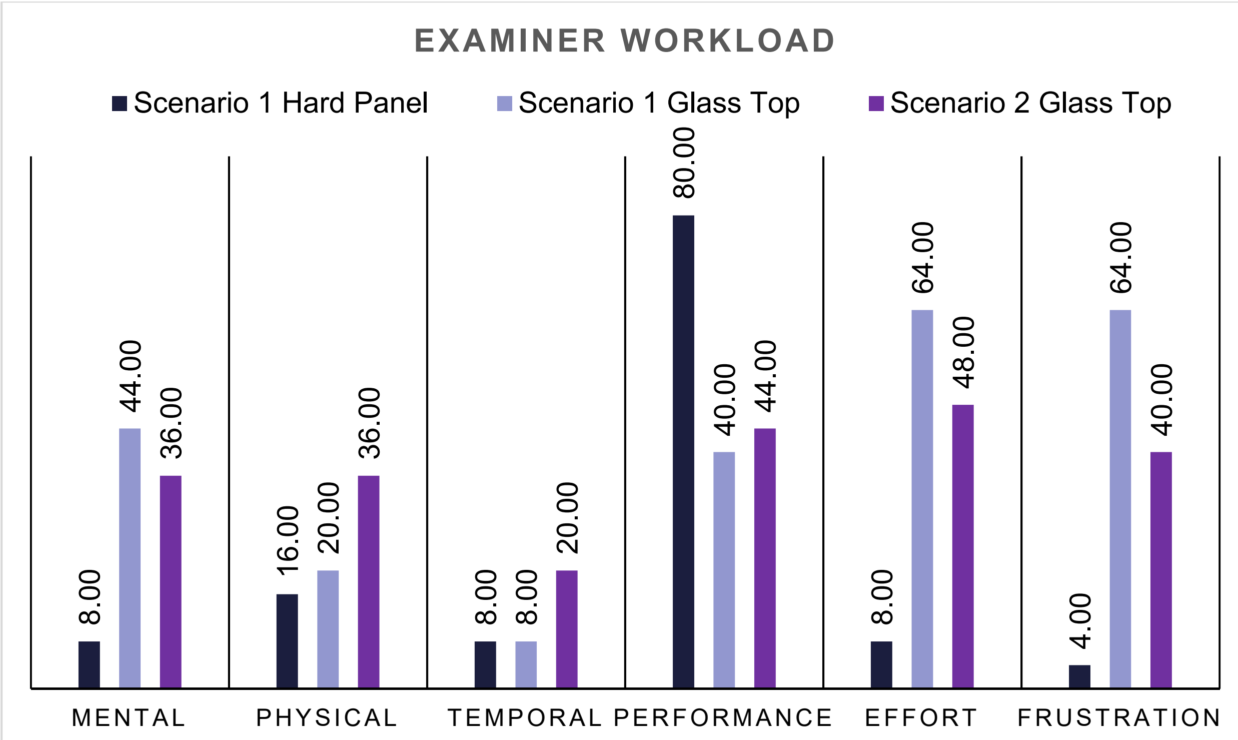


Figure 8. Examiner Workload Results.

4.2.4 RULA

The RULA was used to assess the ergonomics for both the hard panel and glass top simulators. RULA data was collected for operators executing control gestures on both the lower panel (see Figure 9) and the back panel (see Figure 10). Counter to the predictions about the glass top simulator, operators rated the lower panel as more comfortable to use than the hard panel simulator, with the exception of the lower body postural score and the overall neck and trunk positioning. It is important to note that any score over 5 on the RULA suggests the potential for overuse and strain injuries and excessive fatigue. The hard panel simulator is representative of plants designed in the 1980s prior to widespread use of formalized ergonomic assessment. It is not unexpected that the hard panel simulator would have some ergonomic issues. However, the glass top was 75% scale of the hard panel simulator. This size reduction should have also reduced the lower body, neck, and trunk strain.

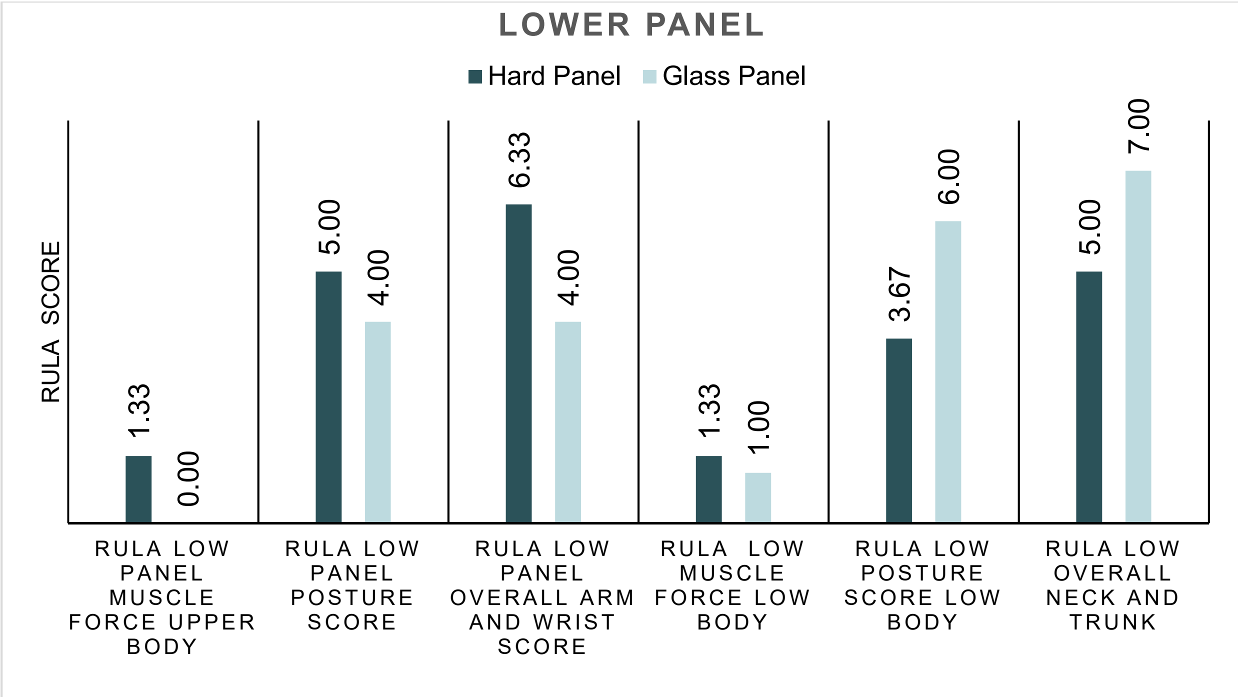


Figure 9. RULA Subscales for the Lower Panel.

As seen in Figure10, the glass top performed better when comparing the back panel ergonomics. Overall, both simulators received high RULA scores, suggesting that the panel type (hard, glass) may have been less important than the position of the controls in determining general discomfort during use, and the potential risk for excessive physical fatigue and overuse and strain injuries. Importantly, the majority of controls necessary to control the plant are on the lower panel, and the back panel is primarily used to display information.

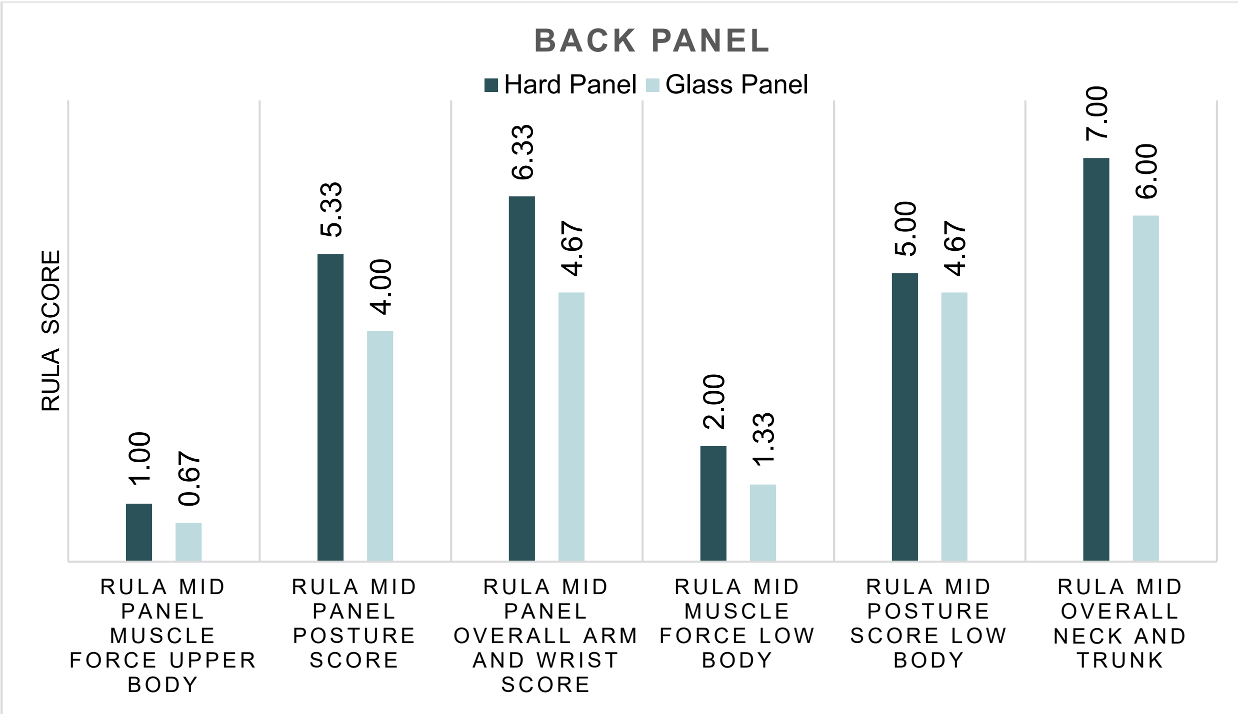


Figure 10. RULA Subscales for the Back Panel.

Operator height was collected to aid in interpretation of the RULA scores; however, due to small sample size, a meaningful analysis of a potential connection between operator height and use discomfort was not possible. Based on the RULA scores for both the lower and back panels and the representativeness of these two simulators, a formalized ergonomic assessment using NUREG-0700 guidelines should be conducted to ensure any operator errors observed are not due to an ergonomic issue like fatigue or awkward body positioning influencing their ability to effectively control the plant using touchscreen-based controls.

4.3 Additional Observations

4.3.1 Glitches and Unplanned Freezing

Simulators will malfunction, regardless of panel type. During the study, which was conducted across four full days of testing, the hard panel simulator functioned as expected and did not require any unplanned restarts or resets. The glass top simulator experienced several malfunctions with the simulator graphics. Specifically, the graphics would freeze despite the scenario continuing, leaving the operators blind to the ongoing plant conditions. These graphics malfunctions were identified in a timely manner, but it was not immediate. This type of issue has serious implications for examination, training, and operations.

During an exam, when a simulator's graphics freeze in an uncontrolled and unplanned manner, it introduces an opportunity to question the exam validity. In situations where an applicant is unhappy with their performance on an exam, and a graphics malfunction occurs, the applicant could question the examiner's decision and the overall validity of the exam process. Infrequent graphics malfunctions would be an annoyance in a training context; however, frequent malfunctions, particularly those that are not identified right away, can interfere with training quality and learning of key temporal relationships between plant parameters indicated through the HSI and plant state. It is unlikely that an operating plant would implement a 100% glass top control room; however, any glass top panel would be vulnerable to graphics malfunctions. Depending on the operational conditions and the systems impacted, a prolonged, unidentified graphics malfunction could have impacts on operator SA.

Additional issues related to virtual button sticking on the glass top simulator were discussed in Section 4.1.2.1. Such malfunctions are assumed to negatively impact the SUS scores.

4.3.2 Implementation and Engineering Processes

The development of a control room is a highly complex engineering project. This is true of a simulated digital/glass top control room and extends to both the functional and graphical elements of the system. During the study, in addition to periodic graphics malfunctions, there were also typos in labels of key systems. For example, the four steam generators, which should have been uniquely labeled, had two steam generators using the same label. As a recently developed training simulator, it is not unexpected that a few minor errors would be identified during a lengthy data collection exercise. If a glass top simulator is used in an exam or validation context by an applicant, it would be important (as with a hard panel simulator) to document processes for data and configuration management and understand the validation and verification processes and outcomes for the simulator during its development. NUREG/CR-6734 (Volumes 1 and 2) describes digital systems software requirements guidelines and common software failures. Despite being published in 2001, these guidelines and illustrative faults highlight the need for systematic software engineering practices with testing and maintenance processes that easily and proactively identify errors and system limitations (see also NASA-STD-8739.8B and NPR 7150.2).

5 SUMMARY CONCLUSIONS AND NEXT STEPS

5.1 Answers to Research Questions

5.1.1 How will differences between a glass top simulator and an analog simulator impact training effectiveness?

Based on the SUS and workload results, there is potential for the glass top simulator to be more of a distraction than a training aid. While the group of operators participating in the study had less experience with the glass top than the hard panel simulation, they had sufficient time to become fluent in the gestural requirements, the amount of force necessary to engage controls, and the positions of those controls. The functional layout of the glass tops was identical to the hard panels, meaning there should be minimal new learning required for the layout of indicators and controls across the boards. Despite their experience, operators were observed having to make multiple attempts to engage some controls and experienced inadvertent activation of controls, breaking the link between action and outcome that would be crucial to learning and understanding system state. Additionally, the gestures required to engage glass top controls are sufficiently different from the hard panel controls. If training occurred on a glass top for operators to use hard panels in an actual plant control room, there is likely the need for supplemental training on analog controls. Operators would need to learn specific actions, such as holding a switch in an open position for several minutes to open a valve or turning a J-Handle to an intermediate rather than terminal positions. However, in the scenarios tested, there were relatively few of these types of gestures (despite selecting scenarios that included a full range of control gestures). Supplemental training with an analog “action board” to learn plant-specific control gestures is a feasible path forward for plants using NPRSs for training and licensing purposes while a PRS is under construction.

5.1.2 Are there differences between a hard panel and glass top simulator that will make conducting licensing examinations more difficult?

The need for specific lighting and the greater potential for simulator-software glitches both have the potential to add uncertainty and complexity to the licensing exam process. Glare was an issue despite anti-glare coating being applied to the glass top panels. The glare added uncertainty and workload to the examiner’s decision making. It also created conditions where they felt they needed to remain at the panel with the operator for the duration of the scenario. The glare was also an issue for operators; they were observed using a hand to shield an area of the controls to make them more easily visible. During the study, there were no noticeable software glitches in the hard panel simulator. The glass top simulator experienced several serious glitches. Instances of virtual button sticking caused an unintended plant trip. There were several instances where graphics failed to update, giving widespread false indications of the plant’s current state. Glitches in general can call testing integrity into question and add unnecessary delays to an often tight testing schedule. It could also create a need for examiners to prepare additional backup scenarios so that they can accommodate a greater number of simulator issues. It is important to note that a malfunction like frozen graphics would not necessarily show up in a simulator’s output log. If the graphics do not freeze across all panels, it could go unnoticed for long enough that an operator acts on inaccurate information. This is because in the log the simulator’s behavior would look normal, but it would look like an operator error to the examiner. If this error occurs on a critical action this could potentially invalidate any critical tasks that occurred during the freeze. If the operator and examiner are unaware of the frozen graphics and no follow up is offered to reassess the operator’s knowledge, a relatively minor simulator performance issue could have major consequences. In general, it is reasonable to expect that when conducting exams using a glass top simulator, examiners should prepare additional backup scenarios, build extra time into their schedule, and potentially go as far as including an exam observer on their team whose primary focus is to monitor for simulator-related issues. The touchscreens used in glass top simulators are a much newer technology compared to what is used in hard panel control rooms. Similarly, the software and hardware infrastructure driving the glass top panels is also a relatively recent innovation. Technology will improve, but given current capabilities additional preparation for glass top exams is likely needed.

5.1.3 Are there clear mitigation strategies for addressing limitations associated with glass top simulators?

The previous section discussed some examiner-centric glass top simulator use risk mitigation strategies. Plants can mitigate the risk of hardware and software glitches and failures in their simulators by adopting the same high reliability engineering practices applied to other technology through the control room and broader plant (see NUREG/CR-6734, Volumes 1 and 2 for examples and guidelines related to high reliability software engineering; see also NPR 71.50.2 and NASA-STD-8739.8B). Plants may also consider a full simulator walkdown exercise prior to exams to ensure all controls, labels, and graphics are presented and operate as expected.

5.1.4 Will glass top simulators put more strain on the operator's shoulders, arms, and wrists?

There is no indication from the observations or the RULA scores that the glass top simulator is worse than the hard panel simulator from an ergonomic perspective. The two simulator types are on average approximately equal in terms of poor ergonomics scores. This is not unexpected for the hard panel simulator and more broadly analog control rooms in operating plants. If a glass top is used as a high-fidelity non-plant referenced simulator for ISV and licensing purposes during plant modernization or a restart of a legacy plant, the same poor ergonomics from the analog control room would be replicated on the glass top, just rendered in two dimensions. The likelihood of input errors and poor ergonomics overall suggest that if the gestures to engage plant controls are different in the glass top simulator, it may be better from an ergonomic and accuracy perspective to use a keyboard and mouse or simplify the gestures rather than try to maximize fidelity. For example, having all controls represented with high fidelity but using simple actions like tapping to engage all controls would alleviate this issue.

5.1.5 Are there operator actions or bodily positions unique to the glass top simulator that would impact examiner confidence or ability to conduct the exam?

It was observed generally that all operators had to stand closer to the glass panels compared to the hard panels. The change from 3-dimensional to 2-dimensional controls also impacted the possible viewing distance for the examiner during the scenarios. These changes in operator positioning changed examiner positioning. More research with a larger group of operators that represent a greater height range is needed to evaluate the extent to which changes in examiner positioning would be (1) necessary, (2) an adequate solution, and (3) impact an examiner's confidence or ability to conduct an exam.

5.1.6 What types of workload are most impacted by changing from a hard panel to a glass top simulator?

For both operators and the examiner, the performance, frustration, and effort subscales were the most impacted by the change from the hard panel to the glass top simulator. Operators and the examiner rated their performance higher during the hard panel simulator scenarios than the glass panel scenarios. Frustration and effort were higher when completing the scenarios in the glass top than hard panel simulator. Taken together with the general observations, it is possible that the differences in workload for operators are due to situational factors related to poor usability of the glass top simulator. For the examiner, it is possible that the workload effects were driven by novelty. The data collection reported here would have been that examiner's first attempt to observe a glass top exam.

5.1.7 Do operators rate the overall usability of the glass top simulator lower than the hard panel simulator?

The SUS scores for the hard panel were surprisingly good. While the hard panel simulator does have all the human factors quirks of an older control room (e.g., red means go/open for some controls), there have also been small modifications, such as color-coding areas of panels based on system type and digital trend indicators. These updates, coupled with the operators' years of experience using the hard panels, are likely contributors to the particularly strong SUS scores for the hard panel simulator. The SUS scores for the glass top simulator were poor, which was somewhat expected. Attempting to maintain gestural fidelity on a glass top simulator exacerbates usability challenges already present in the hard panel simulator. Additionally, there was a minor loss of operational efficiency in the glass top simulator due to the need to make multiple attempts to complete an action and visibility issues.

5.2 Next Steps

This study also collected eye tracking and performance data (observer and simulator logs). These data are planned to be published in a follow-up report. The results from this direct comparison of two full scope simulators with expert operators paired with the RIL 2022-11 Volume 5 report on touchscreen ergonomics highlight the need for additional systematic and controlled research on the impact of specific types of fidelity on both operator and examiner performance. Additional research focused on gathering a higher volume of complex gestures with a glass top is also needed to determine the broader impact of this limitation of glass top simulators.

The current study was highly realistic, which reduced the available participant pool leading to a smaller than ideal sample size. A study with both the complexity and realism of the current study is limited by potential confounding factors, including individual differences in size, technology fluency and preferences. Simulator-related factors, such as the scaled size reduction of the glass top simulator, poor lighting conditions, and software glitches were also factors driving differences in perceived usability and workload between the two simulators. While these unintended differences between the simulators yielded useful insights, follow up research with the necessary controls and narrower scope (i.e., examination, training, or ISV) would enable further development of general recommendations about the use of glass top simulators for different types of regulatory activities.

6 REFERENCES

ANSI/ANS-3.5-2009, *Nuclear Power Plant Simulators for Use in Operator Training and Examination. American Nuclear Society,* 2009

Breuninger, J. (2020). *Suitability of touch gestures and virtual physics in touch screen user interfaces for critical tasks* (Doctoral dissertation, Technische Universität München).

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Dickerson, K., & Green, N, Human Performance Test Facility Volume 5 Touchscreen Ergonomic Considerations, *RIL-2022-11*, Washington, D.C., 2025. ADAMS Accession No: ML25226A197

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139-183). North-Holland.

Hecht, M., Hecht, H., & Shaffer, R. (2001). Digital Systems Software Requirements Guidelines. (NUREG/CR-6734). Washington, D.C.: U.S. Nuclear Regulatory Commission. ADAMS Accession No: ML012330160

Hecht, M., Hecht, H., & Shaffer, R. (2001). Digital Systems Software Requirements Guidelines, Failure Descriptions. (NUREG/CR-6734). Washington, D.C.: U.S. Nuclear Regulatory Commission. ADAMS Accession No: ML012330184

ISO 9241-410:2008 Ergonomics of human-system interaction-Part 410: Design criteria for physical input devices, ISO, Geneva, 2008.

Davis, K. G., Hammer, M. J., Kotowski, S. E., & Bhattacharya, A. (2014). An ergonomic comparison of data entry work using a keyboard vs. touch screen input device while standing and sitting. *J Ergonomics S*, *4*(2).

Fernandes, A., Bloch, M., Braseth, A. O., & McDonald, R. (2023). Perceived usefulness and usability of overview displays in nuclear control rooms. In *Proceeding of the 33rd European Safety and Reliability Conference. Research Publishing Services, Southampton* (pp. 804-811).

Gruse, M., Logan, R., Flerlage, E., Kilby, E., Cupples, J., Hale, K., & Baumgartner, N. (2017). A physical demands analysis to establish critical physical tasks for US Air Force explosive ordnance disposal and battlefield airmen occupations. *Journal of Science and Medicine in Sport*, *20*, S156.

Hansberger, J. T., Peng, C., Mathis, S. L., Areyur Shanthakumar, V., Meacham, S. C., Cao, L., & Blakely, V. R. (2017, May). Dispelling the gorilla arm syndrome: the viability of prolonged gesture interactions. In *International conference on virtual, augmented and mixed reality* (pp. 505-520). Cham: Springer International Publishing.

Lee, S. M., Lee, H. C., Ha, J. S., & Seong, P. H. (2016). Development of digital device based work verification system for cooperation between main control room operators and field workers in nuclear power plants. *Nuclear Engineering and Design*, *307*, 1-9.

NASA. Software Engineering Requirements. NPR 7150.2B. November 19, 2014.

NRC (2010). RG 1.149. NUCLEAR POWER PLANT SIMULATION FACILITIES FOR USE IN OPERATOR TRAINING. LICENSE EXAMINATIONS, AND APPLICANT EXPERIENCE REQUIREMENTS. ADAMS Accession No: ML100491773

O'Hara, J.M., Higgins, J.C., Fleger, S.A., & Pieringer, P.A. (2012). Human Factors Engineering Program Review Model (NUREG-0711, Revision 3). Washington, D.C.: U.S. Nuclear Regulatory Commission. ADAMS Accession No: ML12324A013

O'Hara, J. M. & Fleger, S. (2019). Human-System Interface Design Review Guidelines. (NUREG-0700). Washington, D.C.: U.S. Nuclear Regulatory Commission. ADAMS Accession No: ML19339H052

Offenbacher, J., Kim, J. G., Louie, K., Patel, S., Genes, N., Smith, S. W., Nikolla, D., Carlson, J., Gulati, R., Sinha, S., Sagalowsky, S., Boatright, D., & Glimcher, P. (2025). Clinical decision making during supervised endotracheal intubations in academic emergency medicine. *The American Journal of Emergency Medicine*.

Operator Licensing, 10 CFR 55.31(a)(5) (2015). https://www.ecfr.gov/current/title-10/chapter-I/part-55/subpart-D/section-55.31

Operator Licensing, 10 CFR 55.45 (b) (2015). https://www.ecfr.gov/current/title-10/chapter-I/part-55/subpart-E/section-55.45

Operator Licensing, 10 CFR 55.46(c) (2015). https://www.ecfr.gov/current/title-10/chapter-I/part-55/subpart-E/section-55.46

Operator Licensing, 10 CFR 55.46 (2015). https://www.ecfr.gov/current/title-10/chapter-I/part-55/subpart-E/section-55.46

Operator Licensing, 10 CFR 55.49 (2015). https://www.ecfr.gov/current/title-10/chapter-I/part-55/subpart-E/section-55.46

Orphanides, A. K., & Nam, C. S. (2017). Touchscreen interfaces in context: A systematic review of research into touchscreens across settings, populations, and implementations. *Applied ergonomics*, *61*, 116-143.

Scheetz, M., Buchanan, T., Nist, L., Seymour, J. (2023). NUREG-1021, Revision 12, Effectiveness Review Interim Report September 2023 (NUREG-1021, Revision 12). Washington, D.C.: U.S. Nuclear Regulatory Commission. ADAMS Accession No: ML23304A006

Shelat, S., Homer, K. E., Karasinski, J. A., & Marquez, J. J. (2025, June). Multidimensional usability assessment in spaceflight analog missions. In *SpaceCHI 4.0*.

"Software Assurance and Software Safety Standard, NASA-STD-8739.8B," 2022. URL https://standards.nasa.gov/standard/NASA/NASA-STD-87398, last accessed 11 December 2025.

Ulrich, T., Boring, R., & Lew, R. (2018, June). Extrapolating nuclear process control microworld simulation performance data from novices to experts-a preliminary analysis. *In International Conference on Applied Human Factors and Ergonomics* (pp. 283-291). Cham: Springer International Publishing.

.

APPENDIX A

Scenario 1
Major – Steam Generator Tube Rupture (SGTR)
Set up: Use IC-164

Time	Event No.	Malfunction	Notes	Event Description
0	1	None		Continue to raise power from 50% to 100%. (Allow 5-10% power rise)
10	2	Malf ROD1C: Rod Drop M14		Shutdown Bank Rod M14 will drop due to loss of power to its stationary coil.
20	3	Malf RCS8D-Rupture of SG Tube set to .12		50 GPM Tube Leak from the D SG
30	4	Malf RCS8D-Rupture of SG Tube set to 1.0		Increase to 400 GPM leak on the D SG
	5	EFF 2A and ESF 2B	Part of IC-164 setup.	Reactor fails to trip, FRP S-1 actions required
On request	6	Manual valve will fail to close. Reported by field operator.		MS-027 for SG D will not close. Requires RNO action to trip the TDAFW pump.

Expected Operator Action Summary:

1. The crew will use the guidance of GOI-5 Power Operation and Load Changes Above 20% Power to raise power from 50% to 55-60%. Preparations may be initiated to start the second MFW Pump.
2. M14 Shutdown Bank Rod Will drop into the core. The crew observe the Rod on the bottom on the Individual Rod Position indication and K12 E1 ROD BOTTOM alarm in alarm on panel C13. The crew will take actions to stabilize the plant per **ONI 2-4 Control Rod and Rod Position Indication section 3.3**. Rod Control will be placed in Manual using the ROD BANK SELECT SW on C02. The Main Turbine may be ramped to control Tavg on program using load limit set adjust on panel C05.
3. A 50 gpm tube leak will start, and the crew will respond to the Rad Monitor alarm K01 B1 Process Monitor Alert and B3,C3,D3 S/G Tube Leak alarms. The crew will be directed to enter **ONI 3-12 SG Tube Leak** by the alarm response procedure to mitigate the transient. PZR level will be stabilized using charging controls on panel C12 and leak rate calculated.
4. The leak will get larger causing the crew to attempt to manually trip the reactor and actuate a Safety Injection using the Reactor Trip switches on panels C02 and C12 and per **ONI-3.12 guidance**. This is also a transition to the **E-0 Reactor Trip or Safety injection procedure**.

5. The reactor will not trip, and E-0 will be exited and immediate action for **FR-S.1 Response to Nuclear Power Generation/ATWS** will be performed. The crew will open the orange handle B01 and B04 Bus Feeder breakers on panel C11 to deenergize the Rod Drive MGs. Opening these breakers will succeed in inserting all control rods. FR-S.1 will transition the crew back to E-0.
6. The crew will diagnose the SGTR using Radiation Monitor readings and transition out of E-0 and into **E-3 SGTR procedure**. During isolation of the SG with a tube rupture IAW step 3 of E-3, the crew will identify MS-027 can't be closed locally and will secure the TDAFW Pump IAW **step 3.c. RNO of E-3**. using switch P-102A West AFW Pump on panel 02. The Operator must take the pump to pull-to-lock or it will auto start back to run.
7. The Scenario will end when the initial RCS cooldown and depressurization is complete.

Note: During the scenario for the ramp or to control reactivity, the operators may use the Makeup System controls on Panel C12 to adjust RCS Boron Concentration. Additionally, an operator may make load adjustments on the Main Turbine using the Load Limit Adjust on panel C05.

1. The crew will use the guidance of GOI-5 Power Operation and Load Changes Above 20% Power to raise power from 50% to 55-60%. Preparations may be initiated to start the second MFW Pump

Time	Position	Required Operator Actions	Notes	Correct?
	OATC/ BOP	5.1 Load Escalation 5.1.1 IF the load escalation is stopped AND power is stabilized at any point during the performance of this procedure:	These are steps they will continue to use while raising load.	
		a. REFER to Subsection 5.2 of this procedure for steady-state operating guidelines.		
		b. ADJUST the load limit pot to limit load at the intermediate power level. **(BOP)**		
		5.1.2 Throughout the power escalation: b. MAINTAIN the AFD within the target band (TS 3.2.1){Within limits of TS 3.2.3} using a combination of the following: **(OATC)**	These are steps they will continue to use while raising load.	
		I) Control rod insertion-withdrawal in AUTOMATIC or MANUAL. **(OATC)**		

		2) RCS boration-dilution [Operating Instruction (OI) 3-7]. **(OATC)** b. Using Attachment I for guidance, MAINTAIN the load limit pot setting ≤100 MWe above the load demand.		
	BOP	5.1.3 IF the power goal is for ≥70% power, VERIFY the following are being performed:		
	BOP	b. PREPARATIONS are being made (OI 8-I) to place the second feedwater/condensate train in service prior to exceeding 70% power: I) If not already running, START the second condensate pump. 2) WARM-UP of the second MFW pump (OI 8-1). 3) AFTER the second MFW pump has been reset, RESET the turbine runback circuit by turning the turbine load limiter to <5.0, then return it to the desired setpoint. RECORD turbine runback circuit reset in the CO's Log for future verification (Subsection 5.1.12 of this procedure).	BOP, if they get to it. It is more important to see the rise to 55-60% power.	
		5.1.4 -5.1.9 steps are complete.	The scenario starts at 50% with these steps complete.	

Scenario 2
Small Break Loss of Coolant Accident (SLOCA)

Time	Event No.	Malfunction	Notes	Event Description
0	1	None		The crew will reduce power to 90% for Turbine valve testing. (5%-10%)
10	2	SGN 20C- 0		Median Level circuit for the C SG will fail low.
30	3	PZR PZR5A Auto Controller 30%		Pressurizer Pressure Controller failure in Auto. Fail to 30%
35	4	MFW-MFW1A-Trip		Feed Pump trip-will result in a Runback
45	5	MFW MFW1B-Trip		Second feed pump trip/ requires Rx trip
45	6	PZR PZR2A-5-10%		Leaking PORV after trip. May need to adjust 5%-10% to prevent causing an SI before the next event.
45	7	RCS RCS5B-1500		SBLOCA on Loop B cold leg 1500 gpm. Insert while in ESP-0.1 (critical task).
45	8	Malf pump PMP_SIS002 Auto start fail	Part of IC setup	B SI pump fails to auto start (critical task).

1. The crew will use the guidance of GOI-5 Power Operation and Load Changes Above 20% Power to reduce power 5-10%.

Time	Position	Required Operator Actions	Notes	Correct?
		5.3.2 INITIATE load reduction as follows:		
	OATC	a. NOTIFY the dispatcher. b. INITIATE RCS boration as required.		
	BOP	c. VERIFY the power system stabilizer is removed from service.		
		d. INITIATE turbine load reduction using either the Load Decrease push button or the Load Limiter.		

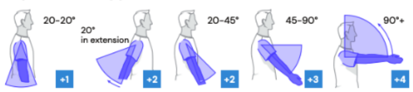
APPENDIX B

B.1 RULA Worksheet and Example Scores

TuMeke Ergonomics

A. Arm and Wrist Analysis

Step 1 : Locate Upper Arm Position



Step 1a: Adjust...
If shoulder is raised: +1
If upper arm is abducted : +1
If arm is supported or person is leaning : -1

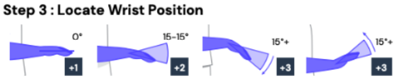
Upper Arm Score

Step 2 : Locate Lower Arm Position

Lower Arm Score

Step 2a : Adjust...
If either arm is working across midline or out to side of body: +1

Step 3 : Locate Wrist Position



Step 3a : Adjust...
If wrist is bent from midline : +1

Wrist Score

Step 4 : Wrist Twist
If wrist is twisted in mid-range : +1
If wrist is at or near end of range: +2

Wrist Twist Score

Step 5 : Look-up Posture Score in Table A
Using values from steps 1-4 above, locate score in Table A

Posture Score A
+

Step 6 : Add Muscle Use Score
If posture mainly static (i.e. held > 1 minute),
Or if actions repeated occurs 4x per minute : +1

Muscle Use Score
+

Step 7 : Add Force/Load Score
If load < 4.4 lbs (intermittent) : +0
If load 4.4 to 22 lbs (intermittent) : +1
If load 4.4 to 22 lbs (static or repeated) : +2
If more than 22 lbs or repeated or shocks : +3

Force/Load Score
=

Step 8: Find Row in Table C

Wrist & Arm Score

SCORES

Table A		Wrist Posture Score 1		2		3		4	
Upper Arm	Lower Arm	Wrist Twist 1	Wrist Twist 2	Wrist Twist 1	Wrist Twist 2	Wrist Twist 1	Wrist Twist 2	Wrist Twist 1	Wrist Twist 2
1	1	1	2	2	2	2	3	3	3
	2	2	2	2	2	3	3	3	3
	3	2	3	3	3	3	3	4	4
2	1	2	3	3	3	3	4	4	4
	2	3	3	3	3	3	4	4	4
	3	3	4	4	4	4	4	5	5
3	1	3	3	4	4	4	4	5	5
	2	3	4	4	4	4	4	5	5
	3	4	4	4	4	4	5	5	5
4	1	4	4	4	4	4	5	5	5
	2	4	4	4	4	4	5	5	5
	3	4	4	4	5	5	5	6	6
5	1	5	5	5	5	5	6	6	7
	2	5	6	6	6	6	7	7	7
	3	6	6	6	7	7	7	7	8
6	1	7	7	7	7	7	8	8	9
	2	8	8	8	8	8	9	9	9
	3	9	9	9	9	9	9	9	9

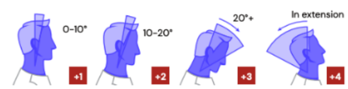
Table C		Neck, trunk and leg score 1	2	3	4	5	6	7+
Wrist and Arm Score	1	1	2	3	3	4	5	5
	2	2	2	3	4	4	5	5
	3	3	3	3	4	4	5	6
	4	3	3	3	4	5	6	6
	5	4	4	4	5	6	7	7
	6	4	4	5	6	6	7	7
	7	5	5	6	6	7	7	7
	8+	5	5	6	7	7	7	7

RULA Scoring (Final score from Table C)	
1 or 2	acceptable posture
3 or 4	further investigation, change may be needed
5 or 6	further investigation, change soon
7	investigate and implement change

RULA Score

B. Neck, Trunk and Leg Analysis

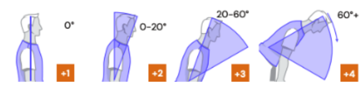
Step 9 : Locate Neck Position



Step 9a: Adjust...
If neck is twisted: +1
If neck is side bending: +1

Neck Score

Step 10 : Locate Trunk Position



Step 10a : Adjust...
If trunk is twisted: +1
If trunk is side bending: +1

Trunk Score

Step 11 : Legs
If legs and feet are supported: +1
If not: +2

Leg Score

Table B	Trunk Posture Score 1		2		3		4		5		6	
Neck Posture Score	Legs 1	Legs 2	Legs 1	Legs 2	Legs 1	Legs 2	Legs 1	Legs 2	Legs 1	Legs 2	Legs 1	Legs 2
1	1	3	2	3	3	4	5	5	6	6	7	7
2	2	3	2	3	4	5	5	5	6	7	7	7
3	3	3	3	4	4	5	5	6	6	7	7	7
4	5	5	5	6	6	7	7	7	7	7	8	8
5	7	7	7	7	7	8	8	8	8	8	8	8
6	8	8	8	8	8	8	8	9	9	9	9	9

Step 12 : Look-up Posture Score in Table B
Using values from steps 9-11 above, locate score in Table B

Posture Score B
+

Step 13 : Add Muscle Use Score
If posture mainly static (i.e.hold > 1 minute),
Or if actions repeated occurs 4x per minute: +1

Muscle Use Score
+

Step 14 : Add Force/Load Score
If load < 4.4 lbs (intermittent) : +0
If load 4.4 to 22 lbs (intermittent) : +1
If load 4.4 to 22 lbs (static or repeated) : +2
If more than 22 lbs or repeated or shocks : +3

Force/Load Score
=

Step 15 : Find Column in Table C
Add values from steps 12-14 to obtain
Neck, Trunk and Leg Score. Find Column in Table C

Neck, Trunk & Leg Score

Automate this asssessment with Computer Vision! Find us at www.tumeke.io

based on RULA: a survey method for the investigation of work-related upper limb disorders, McAtamney & Corlett, Applied Ergonomics 1993, 24(2), 91-99