

# **A Guide for Selecting Appropriate Human Factors Methods and Measures in Control Room Modernization Efforts in Nuclear Power Plants**

Kovesdi, Casey R., Joe, Jeffrey C., Boring  
, Ronald Laurids PhD

July 2018



The INL is a U.S. Department of Energy National Laboratory  
operated by Battelle Energy Alliance

# **A Guide for Selecting Appropriate Human Factors Methods and Measures in Control Room Modernization Efforts in Nuclear Power Plants**

**Kovesdi, Casey R., Joe, Jeffrey C., Boring , Ronald Laurids PhD**

**July 2018**

**Idaho National Laboratory  
Idaho Falls, Idaho 83415**

**<http://www.inl.gov>**

**Prepared for the  
U.S. Department of Energy  
Under DOE Idaho Operations Office  
Contract DE-AC07-05ID14517**

# A Guide for Selecting Appropriate Human Factors Methods and Measures in Control Room Modernization Efforts in Nuclear Power Plants

Casey Kovesdi<sup>1,\*</sup>, Jeffrey Joe<sup>1</sup>, and Ronald Boring<sup>1</sup>

<sup>1</sup> Idaho National Laboratory, Idaho Falls, Idaho, USA  
{Casey.Kovesdi, Jeffrey.Joe, Ronald.Boring}@inl.gov

**Abstract.** Many of the U.S. nuclear power plants are approaching the end of their 60-year licensing period. The U.S. Department of Energy Light Water Reactor Sustainability Program is conducting targeted research to extend the lives and ensure long-term reliability, productivity, safety, and security of these plants through targeted research, such as integrating advanced digital instrumentation and control technologies in the main control room. There are many challenges to this, one being the integration of human factors engineering in the design and evaluation of these upgrades. This paper builds upon recent efforts in developing utility-specific guidance for integrating human factors engineering in the control room modernization process by providing commonly used data collection methods that are applicable at various phases of the upgrade process. Advantages and disadvantages of each method are provided for consideration of an optimal human factors evaluation plan to be used throughout the lifespan of the upgrade process.

**Keywords:** Control Room Modernization · Human-systems Interface Evaluation · Human Factors Engineering · Nuclear Power Plants

## 1 Introduction

Nuclear power annually accounts for approximately 20 percent of the electrical generation in the United States (U.S.), and has done so over the past couple decades. However, many of these U.S. nuclear power plants (NPPs) are now approaching the end of their 60-year licensing period. In order to extend the lives of these NPPs beyond their existing licensing period and to ensure their long-term reliability, productivity, safety, and security, one area that is being explored under the U.S. Department of Energy (DOE) Light Water Reactor Sustainability (LWRS) Program concerns main control room (MCR) modernization. Specifically, the integration of advanced digital instrumentation and control (I&C) technologies is being researched and implemented to ensure safety and enhance productivity and situation awareness of the plant.

These upgrades can vary in scope and level of effort; nonetheless, the modernization process often entails a stepwise approach where individual systems of the plant

---

\* Corresponding author: Tel.: +1-208-526-2336 · Fax: +1-208-526-0528  
*E-mail address:* Casey.Kovesdi@inl.gov

are upgraded and scheduled around planned outages. The result of such differences in the upgrade process requires grading the level of the human factors engineering (HFE) design and evaluation effort [1]. For utilities, this graded approach is no easy feat, often requiring HFE expertise on hand to sift through the required regulatory documentation and to select the appropriate HFE methods and measures to apply.

Recent work to support the industry has focused on developing a user-centered design process for utilities to use as part of ensuring that HFE is properly integrated into their planned upgrade process. Namely, this process is described as the Guideline for Operational Nuclear Usability and Knowledge Elicitation (GONUKE) [2]. While GONUKE provides a user-centered design process to follow, no specific HFE methods or measures are explicitly suggested to utilities. Hence, there is opportunity to elaborate on the specific HFE methods and measures that are commonly used within each of the phases and types of evaluation described in GONUKE.

The next section describes the existing HFE documentation to support control room modernization: the U.S. Nuclear Regulatory Commission (NRC) HFE Review Model, the Electric Power Research Institute (EPRI) HFE Guidance for Control Room and Digital Human-System Interface (HSI) Design and Modification, the Institute of Electrical and Electronics Engineers (IEEE) Guide for the Evaluation of Human-System Performance in Nuclear Power Generating Stations, and the GONUKE process. This paper concludes with providing and exemplifying commonly used HFE data collection methods and measures used for HFE activities specific to control room modernization.

## **2 Existing HFE Guidance for Control Room Modernization**

### **2.1 NUREG-0711: Human Factors Engineering Program Review Model**

The U.S. NRC NUREG-0711 Rev. 3, *Human Factors Engineering Program Review Model*, supports NRC staff in HFE reviews for new NPP builds, as well as for upgrades of existing NPPs [3]. NUREG-0711 essentially serves as a detailed resource for the NRC staff to verify that state-of-the-art HFE principles have been incorporated throughout each phase of development. These phases are described as *Planning and Analysis*, *Design*, *Verification and Validation (V&V)*, and *Implementation and Operation*. There are twelve review elements arranged within each of the four phases.

A comprehensive description of these elements is beyond the scope of this paper; the reader should refer to NUREG-0711 for further details. However, it is important to note that each element is interlinked. For example, a modification is assumed to at least in part be driven by lessons learned from operating experience of past events, which drives what functions should and should not be automated. Human actions are then expected to be analyzed in detail through analytical approaches such as task analysis to identify important human actions that have implications to plant safety. This information is then used to drive the design of the HSI(s), procedures, and training. The integration of the HSI(s), procedures, and training are then evaluated during V&V. The validated design is then implemented into the plant, where at this point, human performance is continuously done.

NUREG-0711 emphasizes performance measurement in the selection of appropriate HFE methods and measures, especially in integrated system validation (ISV) during V&V. NUREG-0711 discusses a need for a hierarchical set of performance measures including aspects of plant performance, personnel task performance, situation awareness, cognitive workload, and anthropometric/physiological factors. *Plant performance* concerns measures of performance for various plant functions, systems, or components. *Task performance* concerns measures of time, accuracy, amount accomplished, subjective reports, and behavioral categorization by observers. *Situation awareness* concerns the degree to which personnel's perception of plant parameters and understanding of the plant's condition correspond to its actual condition at any given time and influences of predicting future states. *Workload* comprises the physical, cognitive, and other demands that tasks place on the plant personnel. Finally, *anthropometric/physiological factors* concern the visibility of displays, accessibility of control devices, and ease of manipulating control devices.

## **2.2 EPRI Human Factors Guidance for Control Room and Digital HSI Design and Modification**

EPRI 3002004310, *Human Factors Guidance for Control Room and Digital Human-System Interface Design and Modification*, provides comprehensive guidance for utilities, their suppliers, and contractors to integrate HFE into the overall design efforts [5]. One important piece to EPRI 3002004310 is the emphasis on tailoring and grading the HFE approach based on the scope of the modernization effort. For instance, an HFE program should be appropriately based on the impact the modification may have to the safety of the plant. Modifications with direct impact to plant safety should take greater priority where greater rigor is applied to the design and evaluation process compared to lower-risk modifications. For selecting appropriate HFE methods and measures, it may not be necessary to use costlier data collection for a finer grained analysis when the proposed modification is less impactful on plant safety. However, modifications that impact multiple systems or have direct safety consequences should use more rigorous data collection methods that build higher confidence that no new human error modes were introduced and that state-of-the-art HFE principles are reflected in the design.

## **2.3 IEEE Guide for the Evaluation of Human-System Performance in Nuclear Power Generating Stations**

The IEEE *Guide for the Evaluation of Human-System Performance in Nuclear Power Generating Stations*, IEEE Std 845-1999(R2011) outlines specific evaluation characteristics that should be considered when evaluating human-system performance related to systems, equipment, and facilities in NPPs [6]. Application of IEEE Std 845-1999(R2011) in NPP MCR modernization activities can range from gathering informal design input such as user opinions to tightly controlled experimental techniques used for answering specific design questions (i.e., hypotheses). IEEE Std 845-1999(R2011) emphasizes the use of a diverse set of measures to the extent that it is cost justifiable, as a single measure may not provide sufficient validity. Lastly, a distinction is made between subjective and objective measures. *Subjective* measures are

comprised from data collected from judgement and opinions of users or experts. *Objective* measures are comprised from data collected from observable human behavior. IEEE Std 845-1999(R2011) describes how objective measures are less apt at being biased from opinion like subjective measures. However, objective measures can be difficult to collect from unobservable human behavior like cognitive processes.

## **2.4 The GONUKE Process for Control Room Modernization**

The intent of the GONUKE process is to provide direct support for utilities in NPP MCR modernization activities that ensures HFE is adequately and consistently integrated throughout the upgrade process [2]. Hence, GONUKE fits within the review model of NUREG-0711 by covering key HFE activities for each of the four phases covered. However, one important distinction with GONUKE is that there is greater emphasis on HFE involvement in the earlier phases to ensure success at later phases.

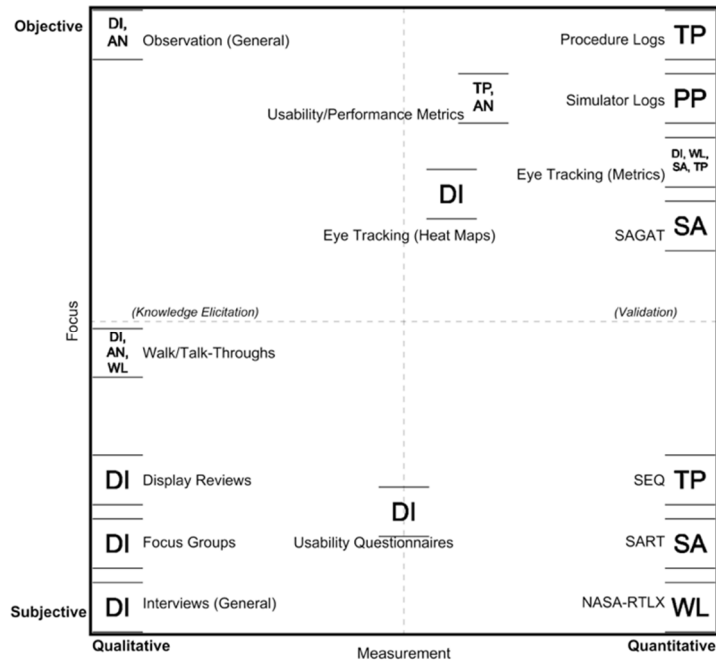
GONUKE describes three types of evaluation: expert review (i.e., herein described as verification), user testing (i.e., herein described as validation), and knowledge elicitation. The latter two involve data collection methods with plant personnel while the former involves human factors subject matter experts using human factors standards like NUREG-0700 [4] in analyses.

There are four fundamental HFE evaluation phases as described by GONUKE, which correspond to the phases of NUREG-0711 [3]. The first phase, Planning and Analysis, includes baseline evaluation (i.e., for validation) and cognitive walkthrough (i.e., for knowledge elicitation) as data collection activities. Baseline evaluation entails evaluating the as-built system so that it can serve as a benchmark for the new system. Cognitive walkthrough entails collecting plant personnel input to capture needs and expectations of the new system. Next, Design (i.e., described as *formative* evaluation) is done through usability testing and feedback about the HSI design to support refining the HSI design through this collection of early feedback; identification of acceptance criteria can also be considered. These formative activities during Design are most successful when done iteratively [7]. Third, a *summative* evaluation (i.e., V&V) is done through ISV and feedback about operator performance after the design process is complete to confirm the usability and performance of a design. Finally, operator training and operator experience reviews are collected with the built design as part of Implementation and Operation for continuous monitoring of human-system performance.

## **3 Common HFE Data Collection Methods and Measures**

Over the past several years, human factors researchers from the LWRS Program have conducted numerous workshops with various utilities. This work used a range of HFE data collection methods for collecting design input, as well as evaluating human-system performance through operator-in-the-loop studies. These methods were applied in various phases described in NUREG-0711 [3]. There was a range of testing environments, spanning from less controlled platforms such as informal discussions and field visits to controlled studies using full-scope testbeds. A collection of com-

mon HFE methods previously used from this past work is presented in Fig. 1, which is an adaptation of Rohrer's *Landscape of User Research Methods* [8].



**Fig. 1.** Commonly used HFE methods and measures for control room modernization in NPP.

The y-axis denotes the focus of evaluation, being the degree to which the method is subjective or objective. As discussed in IEEE Std 845-1999(R2011), objectivity is the degree to which the method comprises data from observed human behavior whereas subjectivity is the degree to which the method comprises data from judgement and opinions of users or experts [6]. IEEE Std 845-1999(R2011) suggests using a diverse set of these measures to strengthen validity of findings.

On the x-axis, the methods lie on the degree of measurement, denoted as being qualitative or quantitative. *Qualitative* measurement provides descriptive information regarding the qualities of a topic [9]. For example, qualitative data may take the form of the description and categorization of various higher-level themes extracted from open-ended comments. This information is important for identifying and correcting potential HSI design issues [10]. Contrarily, *quantitative* measurement provides information regarding the quantity of a topic [9]. Quantitative analyses deal with both discrete and continuous data to make numerical measurement for both descriptive and comparative purposes. One may reasonably assume that quantitative measures are most important in later-staged efforts like ISV; however, quantitative measures still can be useful earlier in the development cycle [7]. To that end, qualitative measures can carry important insights in later phases such as collecting feedback about operator performance.

A third dimension of Fig. 1 pertains to the human performance dimension(s) that can be collected from the HFE method. These dimensions include: *DI* – Design Input, *PP* – Plant Performance, *TP* – Task Performance, *SA* – Situation Awareness, *WL* – Workload, and *AN* – Anthropometric/Physiological Factors.

The latter five correspond to the hierarchical set of performance measures described in NUREG-0711 [3]. *Design input* is defined as the thoughts, feelings, and experiences expressed by the users that can be used to inform the design of the HSI. For example, comments and suggestions made by plant personnel to enhance the usability of the HSI would classify as design input. The methods presented in Fig. 1 are discussed in detail in the subsequent sub-sections, provided with a short description and a list of their advantages and disadvantages. Their applications to GONUKE [2] and to NUREG-0711 [3] are summarized in Fig. 2.

	Planning and Analysis	Design	Verification and Validation	Implementation and Operation
<b>Verification</b>	Verification incorporates use of various human factors standards and guidelines as opposed to data collection methods from plant personnel.			
<b>Validation</b>	<u>Baseline Evaluation</u> <ul style="list-style-type: none"> <li>• Eye Tracking (Heat Maps)</li> <li>• Usability Questionnaires</li> <li>• Usability/performance Metrics</li> <li>• Procedure Logs</li> <li>• Simulator Logs</li> <li>• Eye Tracking (Metrics)</li> <li>• SAGAT</li> <li>• SEQ</li> <li>• SART</li> <li>• NASA-RTLX</li> </ul>	<u>Usability Testing</u> <ul style="list-style-type: none"> <li>• Eye Tracking (Heat Maps)</li> <li>• Usability Questionnaires</li> <li>• Usability/performance Metrics</li> <li>• Procedure Logs</li> <li>• Simulator Logs</li> <li>• Eye Tracking (Metrics)</li> <li>• SAGAT</li> <li>• SEQ</li> <li>• SART</li> <li>• NASA-RTLX</li> </ul>	<u>Integrated System Validation (i.e., ISV)</u> <ul style="list-style-type: none"> <li>• Usability/performance Metrics</li> <li>• Procedure Logs</li> <li>• Simulator Logs</li> <li>• Eye Tracking (Metrics)</li> <li>• SAGAT</li> <li>• SEQ</li> <li>• SART</li> <li>• NASA-RTLX</li> </ul>	<u>Operator Training</u> <ul style="list-style-type: none"> <li>• Procedure Logs</li> <li>• Simulator Logs</li> <li>• SART</li> <li>• NASA-RTLX</li> </ul>
<b>Knowledge Elicitation</b>	<u>Cognitive Walkthrough</u> <ul style="list-style-type: none"> <li>• Observation (General)</li> <li>• Walk/talk-throughs</li> <li>• Interviews (General)</li> <li>• Usability Questionnaires</li> </ul>	<u>Operator Feedback on Design</u> <ul style="list-style-type: none"> <li>• Walk/talk-throughs</li> <li>• Focus Groups</li> <li>• Display Reviews</li> <li>• Interviews (General)</li> <li>• Usability Questionnaires</li> </ul>	<u>Operator Feedback on Performance</u> <ul style="list-style-type: none"> <li>• Focus Groups</li> <li>• Interviews (General)</li> </ul>	<u>Operator Experience Reviews</u> <ul style="list-style-type: none"> <li>• Observation (General)</li> <li>• Focus Groups</li> <li>• Interviews (General)</li> </ul>

Fig. 2. HFE methods and measures mapped to GONUKE (rows) and NUREG-0711 (columns).

### 3.1 Description of Common HFE Data Collection Methods and Measures

**Observation (General).** Observation is a general HFE technique used to collect data about the physical and verbal aspects of a task or scenario [11]. It can be used to inform design or to inform task analysis. In this context, observation here is described as passive (i.e., without intervening or interfering with what operators are doing). Various tools of ranging complexity can support in observation such as audio/video recordings, spreadsheets, as well as pen and paper [12]. *Advantages:* Provides rich contextual information, minimally intrusive, does not necessarily require specialized



equipment. *Disadvantages:* Cannot formally evaluate performance, does not reveal plant personnel's thoughts and rationale for making decision observed or why an observed error occurred.

**Walk/talk-throughs.** A walkthrough is an HFE technique where an expert user 'walks through' or demonstrates a set of tasks or scenario to describe a task, highlight potential issues with the system, or highlight important actions that may be influenced by the upgrade. A talkthrough is a verbal demonstration of a walkthrough [5]. *Advantages:* Provides accurate descriptions of a task, observers can query of particular topics in 'real time,' collects behavioral and attitudinal aspects of human interactions. *Disadvantages:* Provides only descriptive data (i.e., cannot be used for validation), requires access to expert users such as plant personnel, requires degree of effort synthesizing data collected from verbal reports and observations made.

**Display Reviews.** In a display review, static HSI displays are systematically evaluated with expert users (e.g., plant personnel) to uncover potential design issues [5]. This activity can be completed remotely, at the plant, or at a simulator. Likewise, this activity can be done one-on-one or as a group (i.e., see Interviews versus Focus Groups). *Advantages:* Provides readily actionable design recommendations, incorporates identified issues and suggestions directly from actual users of the system, very flexible to administer. *Disadvantages:* Identified issues and suggestions are not directly tied to observable human-system performance data, issues and suggestions may be biased by past experience with an existing system.

**Focus Groups.** A focus group is a general HFE technique used to collect qualitative (i.e., attitudinal) data about a specific topic. This activity is done with a group of users (ideally around 5 individuals), where verbal notes are collected from semi-structured questions administered by a facilitator [5]. Data can be collected in digital (e.g., spreadsheet) or pen and paper format. Raw notes are ultimately synthesized to make meaningful insights to the research question at hand. *Advantages:* Data collection is more efficient than with traditional interviews. *Disadvantages:* Collects only attitudinal data (not behavioral), can be strongly susceptible to response bias between individuals compared to interviews.

**Interviews (General).** Like focus groups, interviewing is a general HFE technique used to collect qualitative (i.e., attitudinal) data about a specific topic. However, these activities are done one-on-one, where verbal notes are recorded from the interviewee [11]. *Advantages:* Less prone to response bias compared to focus groups, can provide rich data when done with observation. *Disadvantages:* Cannot formally evaluate performance, data collection is more labor intensive than focus groups.

**Usability/performance Metrics.** Usability/performance metrics, as described in ISO 9241-11 [14], incorporate measurement to the extent of effectiveness, efficiency, and satisfaction with which a specified user achieves specified goals in a particular envi-

ronment. Typical measures include task success, error frequency, completion time, and usability ratings. It should be noted that there are different formal integrated platforms to which usability/performance metrics are collected. These platforms include the Operator Performance Assessment System (OPAS) [15] and the Supervisory Control and Resilience Evaluation (SCORE) framework [16]. These methodologies incorporate both structured evaluation of operator actions as well as expert evaluation. The reader should refer to the provided references for an in-depth discussion of these methodologies. *Advantages:* Can formally evaluate objective performance, low cost, does not require specialized equipment, data collection and post-processing is not labor intensive. *Disadvantages:* Sensitivity of measures are controversial for HFE studies in NPP MCRs, methods like OPAS and SCORE require considerable upfront effort identifying important human actions.

**Eye Tracking (Heat Maps).** A heat map is a visualization technique that uses different colors to show the amount of fixations or dwell durations over an area (i.e., such as a control board) [13]. These visualizations illustrate where users looked and for how long. *Advantages:* Illustrates through objective data where users looked and for how long. *Disadvantages:* Cannot formally evaluate performance, requires specialized eye tracking equipment, data collection and post-processing is labor intensive, does not explain why users looked at a particular place, does not provide information about visual scanning patterns or sequence of eye movements.

**Usability Questionnaires.** A usability questionnaire provides a set of structured questions that address different design elements of usability through allowing users to either rate or comment about each design element. Information collected from the questionnaire can be used to support identifying potential design issues and areas for improvement with the HSI. Questions can be created from content in NUREG-0700 [4], EPRI 3002004310 [5], or other relevant resources (e.g., operating experience). *Advantages:* Provides both qualitative and quantitative data, information collected can be adaptable depending on project needs, can be used to compare different designs, easy to develop and administer, low cost. *Disadvantages:* Provides only subjective data from user opinions, custom questions need to be pilot tested to ensure they address the usability topic at hand, no available normative benchmark scores to compare and validate to since questions are customized each time.

**Procedure Logs.** Procedure logs are a form of observation done in a controlled setting/lab, where the observer collects step-by-step actions from plant personnel. Each step is then time-stamped. For design and evaluation, the data collected here can be used to identify correct path as well as collect completion times per step. *Advantages:* Provides objective performance data of task performance, low cost, does not require specialized equipment. *Disadvantages:* Requires formal acceptance criteria for evaluation or an alternative design to compare to, data collection can be labor intensive.

**Simulator Logs.** Simulator logs capture the values of key parameters through the lifetime of a scenario. The data captured here is fully applicable to observing plant

performance. *Advantages:* Provides objective data of plant performance. *Disadvantages:* Requires a testbed that is capable of recording and readily exporting key process values.

**Eye Tracking (Metrics).** Eye tracking is a general methodology used in HFE to capture measures of visual attention, mental workload, and situation awareness [17]. As such, physiological data can be collected to make inferences about a certain HSI design. *Advantages:* Captures a rich set of performance measures which are continuously tracked throughout the course of a scenario. *Disadvantages:* Requires specialized eye tracking equipment, data collection and post-processing is labor intensive, requires additional level of inference when associating measures to HFE constructs (e.g., workload), not all testing environments or individuals are capable of using the equipment (e.g., issues range from lighting considerations to individual differences [17]).

**Situation Awareness Global Assessment Technique (SAGAT).** SAGAT is a freeze-probe/recall HFE method used to assess the three levels of situation awareness. SAGAT provides an objective means to evaluate situation awareness where participants' responses to key queries are compared to what actually happened during the scenario. Questions are crafted to address Level 1, Level 2, and Level 3 situation awareness [18]. *Advantages:* Provides objective data of situation awareness. *Disadvantages:* Freeze-probes used can be highly intrusive, requires testing environment capability to freeze and blank out indication status, requires considerable upfront effort developing meaningful queries to administer during each probe.

**Single Ease Question (SEQ).** The SEQ is a standardized single post-scenario usability question that asks users to rate (i.e., typically using a seven-point scale; 1=very difficult; 7=very easy) their overall ease of completing a task [7]. The SEQ can be administered digitally, via pen and paper, or verbally after completion of a task. *Advantages:* Very easy to administer, low cost, provides quantitative data of perceived task difficulty. *Disadvantages:* Not diagnostic regarding identification of usability issues when used alone, may not necessarily correlate with observed task performance.

**Situation Awareness Rating Technique (SART).** The SART is a post-scenario rating method used to derive a measure of perceived situation awareness [11]. The questionnaire comprises 10 dimensions (i.e., questions) using a seven-point rating scale (1=low; 7=high). A composite situation awareness score is derived from SART where a greater value denotes greater situation awareness. A simpler version (i.e., the 3D-SART) is also available, comprising of only 3 questions. *Advantages:* Easy to administer, low cost, not intrusive. *Disadvantages:* Provides subjective (perceived) data of situation awareness, may be more representative of workload than situation awareness.

**National Aeronautics and Space Administration Raw Task Load Index (NASA-RTLX).** The NASA-RTLX is a post-scenario rating method to assess workload, comprising six different dimensions [11]. Each dimension (i.e., question) typically uses a 20-point scale (1=low; 20=high) where higher values denote greater workload. The NASA-RTLX is a shortened version of the NASA-TLX where the set of 15 pairwise comparison is omitted. Workload can be evaluated by each dimension and holistically from aggregating the individual scales. *Advantages:* Easy to administer, low cost, not intrusive, there is a rich database of benchmark values across various industries to compare to [19]. *Disadvantages:* Provides a subjective (perceived) assessment of workload, responses can be confounded with task performance.

### 3.2 Application of HFE Data Collection Methods and Measures

Researchers from the LWRS Program partnered with a large utility as part of integrating HFE into the control room upgrades for several non-safety related systems. The upgrade significantly affected I&C behind the board. Moreover, there was also a number of indications and controls from control boards of the MCR that were affected, including several new digital HSIs and single loop interface module (SLIM) replacements for manual-auto stations. In this partnership, LWRS human factors researchers were able to successfully integrate common HFE methods within the utility's major engineering project milestones such as their software-in-the-loop (SWIL) testing, factor acceptance testing (FAT), and licensing operator requalification training (LORT). For instance, researchers iteratively incorporated lightweight methods like usability questionnaires, interviews, and display reviews to collect design feedback and knowledge elicitation with licensed operators between planned activities for each milestone.

Data collected from each milestone was used to inform the HSI design to which updated versions were used in a full-scale HFE workshop to evaluate the upgrades in an operational context. This workshop used formal objective methods such as eye tracking metrics and usability/performance metrics, as well as subjective methods such as usability questionnaires, SEQ, SART, and NASA-RTLX. Data collected was evaluated against industry accepted criteria or plant expected criteria for usability testing. Results from the study showed significant improvements to the design of the HSIs. No safety-critical issues were identified.

## 4 Conclusions

This paper summarizes commonly used HFE methods and measures for MCR modernization in NPPs. As addressed in the paper, there are certain advantages and disadvantages of each method that should be accounted for when planning HFE integration into control room modernization efforts. A recommended approach is to use a diverse set of methods to collect human factors data across all quadrants described in Fig. 1. By including a balanced and diverse set of these methods and measures, this paper offers a complete, but not exhaustive, set of methods that can be used for design input and validation of human-system performance as described in NUREG-0711 [3].

Nonetheless, future work should investigate alternative HFE methods that may serve useful.

**Acknowledgments.** INL is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517. This work of authorship was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights. The United States Government retains, and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof. The INL issued document number for this paper is: INL/CON-17-44009.

## References

1. Joe, J.C., Boring, R.L., & Persensky, J. J. (2012, in press). Commercial Utility Perspectives on Nuclear Power Plant Control Room Modernization. In Proceedings of the 2012 American Nuclear Society Nuclear Power Instrumentation and Control and Human Machine Interface Technology Conference, pp. 2039-2046.
2. Boring, R. L., Ulrich, T. A., Joe, J. C., & Lew, R. T. (2015). Guideline for Operational Nuclear Usability and Knowledge Elicitation (GONUKE). *Procedia Manufacturing*, 3, 1327–1334.
3. U.S. Nuclear Regulatory Commission. (2012). Human Factors Engineering Program Review Model, NUREG-0711, Rev. 3.
4. U.S. Nuclear Regulatory Commission. (2002). Human-System Interface Design Review Guidelines, NUREG-0700, Rev. 2.
5. Electric Power Research Institute (2015). Human Factors Guidance for Control Room and Digital Human-System Interface Design and Modification: Guidelines for Planning, Specification, Design, Licensing, Implementation, Training, Operation, and Maintenance for Operating Plants and New Builds. EPRI, Palo Alto, CA. 3002004310.
6. Institute of Electrical and Electronics Engineers (1999). IEEE Guide to the Evaluation of Human-System Performance in Nuclear Power Generating Stations (IEEE Std. 845-1999). New York: Institute of Electrical and Electronics Engineers.
7. Sauro, J., & Lewis, J. R. (2016). Quantifying the user experience: Practical statistics for user research. Morgan Kaufmann.
8. Rohrer, C. (2014). When to use which user experience research methods. Jakob Nielsen's Alertbox. Retrieved from <https://www.nngroup.com/articles/which-ux-research-methods/> on January 24, 2018.
9. Kuniavsky, M. (2003). Observing the user experience: a practitioner's guide to user research. Morgan kaufmann.
10. Kovesdi, C.R. & Joe, J.C. (2017). A Novel Tool for Improving the Data Collection Process during Control Room Modernization Human-System Interface Testing and Evaluation Ac-

- tivities. In 10<sup>th</sup> International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human Machine Interface Technologies pp. 1261-1271.
11. Stanton, N., Salmon, P. M., & Rafferty, L. A. (2013). Human factors methods: a practical guide for engineering and design. Ashgate Publishing, Ltd.
  12. Kirwan, B., & Ainsworth, L. K. (Eds.). (1992). A guide to task analysis: the task analysis working group. CRC press.
  13. Bergstrom, J. R., & Schall, A. (Eds.). (2014). Eye tracking in user experience design. Elsevier.
  14. ISO/IEC. "9241-11 Ergonomic Requirements for Office Work with Visual Display Terminals (VDT)s- Part II Guidance on Usability," ISO/IEC 9241-11,1998 (E).
  15. Skraaning, G., 2004, Experimental Control Versus Realism: Methodological Solutions for Simulator Studies in Complex Operating Environments, HWR-361, Halden Reactor Project, Halden, Norway.
  16. Braarud, P. Ø., Eitrheim, M. H. R., & Fernandes, A. (2015). SCORE – An Integrated Performance Measure for Control Room Validation. In 9<sup>th</sup> International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human Machine Interface Technologies pp. 2217-2228.
  17. Kovesdi, C., Rice, B., Bower, G., Spielman, Z, Hill, R., and LeBlanc, K. (2015). Measuring Human Performance in Simulated Nuclear Power Plant Control Rooms Using Eye Tracking, INL/EXT-15- 37311. Revision 0.
  18. Endsley, M. R. (2016). Designing for situation awareness: An approach to user-centered design. CRC press.
  19. Grier, R. A. (2015, September). How high is high? A meta-analysis of NASA-TLX global workload scores. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 59, No. 1, pp. 1727-1731). Sage CA: Los Angeles, CA: SAGE Publications.