APPLICATION OF EYE TRACKING FOR MEASUREMENT AND EVALUATION IN HUMAN FACTORS STUDIES IN CONTROL ROOM MODERNIZATION

Kovesdi, C., Spielman, Z., LeBlanc, K., Rice, B.
Idaho National Laboratory
PO Box 1625, Mail Stop, 3818 Idaho Falls, Idaho 83415, USA
Casey.Kovesdi@inl.gov; Zachary.Spielman@inl.gov; Katya.LeBlanc@inl.gov; Brandon.Rice@inl.gov

ABSTRACT

An important element of human factors engineering (HFE) pertains to measurement and evaluation (M&E). The role of HFE-M&E should be integrated throughout the entire control room modernization (CRM) process and be used for human-system performance evaluation and diagnostic purposes with resolving potential human engineering deficiencies (HEDs) and other human machine interface (HMI) design issues. NUREG-0711 describes how HFE in CRM should employ a hierarchical set of measures, particularly during integrated system validation (ISV), including plant performance, personnel task performance, situation awareness, cognitive workload, and anthropometric/ physiological factors. Historically, subjective measures have been primarily used since they are easier to collect and do not require specialized equipment. However, there are pitfalls with relying solely on subjective measures in M&E such that negatively impact reliability, sensitivity, and objectivity. As part of comprehensively capturing a diverse set of measures that strengthen findings and inferences made of the benefits from emerging technologies like advanced displays, this paper discusses the value of using eye tracking as an objective method that can be used in M&E. A brief description of eye tracking technology and relevant eye tracking measures is provided. Additionally, technical considerations and the unique challenges with using eye tracking in full-scaled simulations are addressed. Finally, this paper shares preliminary findings regarding the use of a wearable eye tracking system in a full-scale simulator study. These findings should help guide future full-scale simulator studies using eye tracking as a methodology to evaluate human-system performance.

Key Words: eye tracking, human factors engineering, control room modernization

1 INTRODUCTION

An important element of human factors engineering (HFE) pertains to measurement and evaluation (M&E). The role of HFE-M&E should be integrated throughout the entire control room modernization (CRM) process and be used for human-system performance evaluation and diagnostic purposes with resolving potential human engineering deficiencies (HEDs) and other human machine interface (HMI) design issues. NUREG-0711 [1] describes how HFE in CRM should employ a hierarchical set of measures, particularly during integrated system validation (ISV), including plant performance, personnel task performance, situation awareness, cognitive workload, and anthropometric/physiological factors. Further, NUREG-0711 describes expectations of the characteristics these measures should comprise: construct validity, reliability, sensitivity, unobtrusiveness, and objectivity.

Construct validity is described as the extent of accuracy to which a measure represents the aspect of performance it is intended to measure. *Reliability* is the degree of repeatability, or the extent of being able to obtain the same results if a measure is used exactly the same way under the exact same experimental conditions. *Sensitivity* is the degree of appropriateness, regarding a measure's scale and frequency of being collected, that it is capable of assessing aspects of performance. For example, some performance measures may only be sensitive under tightly controlled conditions, but fail to be sensitive when used in complex simulations such as full-scale scenarios. *Unobtrusiveness* is the extent to which a measure does not alter the psychological or physical processes that are under investigation. Finally, *objectivity* is the quality that a measure is based on easily observable phenomena.

These characteristics are also highlighted in the Institute of Electrical and Electronics Engineers, Inc. (IEEE) *Guide for the Evaluation of Human-System Performance in Nuclear Power Generation Stations* [2], which provides a basis for using a diverse set of measures, both subjective and objective, to strengthen findings made during M&E. These measures should support the comprehensive evaluation of human-system performance by considering all influencing factors such as environmental conditions, organizational design, training, as well as the physiological, perceptual, and cognitive processes of relevant plant personnel. Historically, subjective measures have been considered 'more practical measures' since they are easier to collect and do not require specialized equipment [3]. However, there are pitfalls with relying solely on subjective measures in M&E, especially when using only these measures for evaluation of complex safety-critical systems such as a nuclear power plant. For instance, subjective measures (e.g., operator preference) can be vulnerable to individual bias and perspective [2]. Paper and pencil methods such as NASA-TLX, while acknowledged as a valid measure of workload, can be susceptible to confounding with task success and cannot continuously track workload throughout task [3]. Similarly, user preference data does not always reliably correlate with human performance [4]. Such disparities support the need to also include objective measures into M&E as part of a diverse set of measures for CRM.

As part of comprehensively capturing a diverse set of measures that strengthen findings and inferences made of the benefits from emerging technologies like advanced displays, this paper discusses the value of using eye tracking as an objective method that can be used in M&E [5]. A brief description of eye tracking technology and relevant eye tracking measures is provided. Additionally, technical considerations and the unique challenges with using eye tracking in full-scaled simulations are addressed. Finally, this paper shares preliminary findings regarding the use of a wearable eye tracking system in a full-scale simulator study. These findings should help guide future full-scale simulator studies using eye tracking as a methodology to evaluate human-system performance.

2 APPLICATION OF EYE TRACKING IN CONTROL ROOM MODERNIZATION

2.1 The Value of Eye Tracking in Control Room Modernization

The advent of eye tracking research can be traced from early work in the late 19th century. These eye tracking systems have been described as 'medieval torture devices', which required participants to place coverings with sticks attached to their eyes in order to trace eye movements [6]. Moving forward into the mid 20th century, eye tracking equipment progressed to being film-based and was used in some of the earliest HFE efforts of informing cockpit designs [7]. While eye tracking has been very useful in answering fundamental human factors research questions, the legacy eye tracking equipment was intrusive, often requiring the participant to be constrained within a head restraint or bite bar [6]. Further, the eye tracking software for data processing and analysis was onerous to use and required extensive knowledge and training. These constraints left eye tracking inaccessible to applied work such as with the evaluation of human-system performance in CRM.

Fortunately, today's eye tracking equipment is not constrained by the technical limitations of its predecessors. Modern eye trackers generally use infrared technology to accurately track pupil coordinates relative to the corneal reflection. These eye trackers are capable of rapidly calibrating to participants and have been designed so that participants have greater freedom of movement, allowing them to behave as they normally would in the built environment [6]. The software for data processing and analysis has also automated many of the tedious tasks required of the user previously such as with defining areas of interest and creating meaningful measures. Collectively, these hardware and software improvements have positioned this technology as a viable, unobtrusive option to include as part of a diverse set of measures for CRM.

There is also a large body of literature from the decades of research, supporting the case for including eye tracking as a method to providing valid, reliable, and sensitive measurement of human factors constructs such as workload, situation awareness, and visual attention in control room settings. For instance, Jacob and Karn [8] provided a detailed table of 21 eye tracking usability studies, which lists common measures that can be used in human-system performance (i.e., usability) M&E. Moreover, eye tracking in HFE has been used in a variety of domains such as automotive, healthcare, aerospace, and more recently nuclear power [9]. Recent research in nuclear power domain included eye tracking in desktop simulator tasks to continuously track operator workload, as part of evaluating different HMI technologies [e.g., 10, 11]. Collectively, these studies have provided a wealth of knowledge regarding use of various eye tracking measures within the area of CRM. This prior research further supports the application of eye tracking in full-scaled simulation environments, enabling operators to interact with candidate CRM technologies more naturally with minimal mobility restrictions. The use of eye tracking to human-system performance M&E may be applicable to later-staged formative studies and ISV using full-scale testbeds, as well as during focused tasks to test very specific research questions [12]. The next section describes the different types of eye tracking systems that can be used in each of these CRM efforts.

2.2 Types of Modern Eye Tracking Systems

Modern eye trackers come in different forms, and can be categorized as being *remote* or *wearable*. Remote eye trackers are the least obtrusive since they do not require the participant to wear any device. These devices are generally integrated to a dedicated stimuli display placed below the monitor. Because the visual space is already integrated, data processing can be significantly more efficient and less complex than with the wearable systems. For example, some wearable systems require mapping each and every eye movement from a video file to a reference scene. The time required to complete this mapping task can be equally as long as the data collection trial itself. A remote system can automatically map eye movements onto a reference scene to begin data analysis. An obvious disadvantage, however, of remote eye tracking equipment in M&E for CRM lies in its limitation to collect data in larger spaces such as with a full-scale control room. Remote systems maybe a useful tool when designing focused experiments that aim at evaluating a very specific HMI or plant process from a single monitor.

Wearable eye trackers enable freedom to move naturally throughout the experimental environment (e.g., a full-scale control room). These systems are worn by the participant and fit like traditional glasses. Fig. 1 shows one type of wearable eye tracking system that was used during a full-scale simulator study.

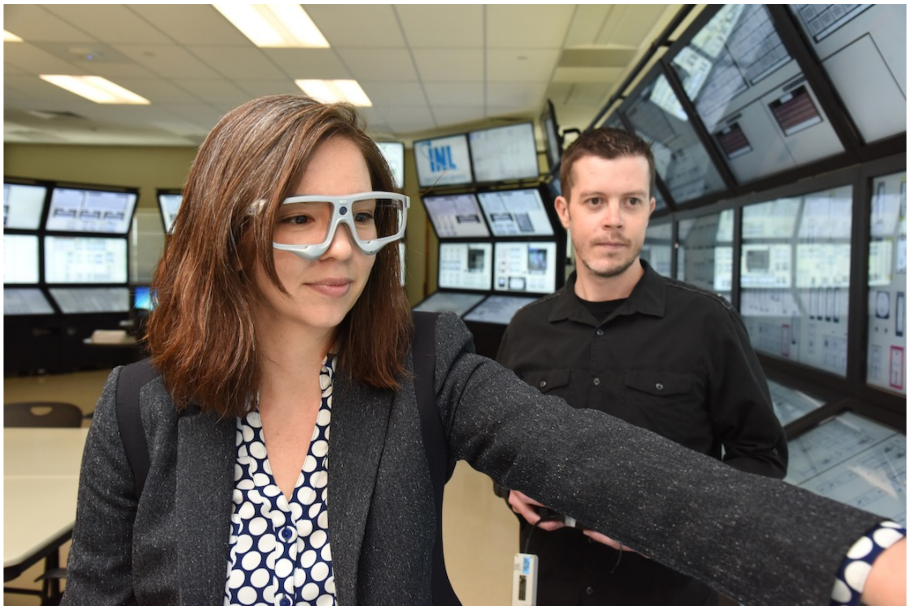

Figure 1. Illustration of a wearable eye tracking system used during a full-scale simulator study.

In addition to the added time required for data processing, there are other challenges that are unique to wearable systems. For one, an individual's facial structure may for a suboptimal placement of the eye tracking to collect accurate data. For example if the eye tracking glasses are positioned too high or low relative to the eyes, the cameras used to record eye movements will be unable to obtain an accurate recording. Similarly, the frame of the eye tracking glasses can reduce peripheral vision; potentially limiting the visibility of information relative to how one normally would be able to see. Finally, certain measures that can be collected from remote systems may not be available if using a wearable system (e.g., saccade amplitude or velocity). The next section discusses the identified eye tracking measures applicable to CRM. To close, it should be noted that the experimental design might mitigate potential biases resulting from the technical constraints of the wearable eye tracking system. Providing a comparative evaluation between technologies (e.g., benchmarking) is one way that can help reduce technology-centric effects on human performance.

2.3 Applicable Eye Tracking Measures for Control Room Modernization

Eye tracking can be described as method used to objectively capture when, where, and for how long someone is looking at something in space [6]. The application of eye tracking is vast, and an in-depth discussion of its use in other scientific disciplines is outside the scope of this paper. However within the realm of HFE, eye tracking enables objective measurement of visual search behavior as it applies to HMI design. These results can then be used to evaluate and improve HMI design through descriptive (e.g., observing the proportion of visual search spent on a display) and inferential analyses (e.g., comparing workload between two alternative HMI designs).

Eye tracking captures two fundamental characteristics of eye movements: the *fixation* and the *saccade*. The fixation can be described as a pause in eye movement over a specific region of the visual field [6]. The duration of a fixation is suggested to be indicative of the degree of cognitive effort required to process information; this is consistent with the eye-mind hypothesis [8]. This mechanism of information processing that takes place during a fixation is achieved by neuronal activation at the fovea [9], which is inferred to be traceable to where one's attention is being directed [8]. The fovea is a small region of the retina that is responsible for the perception of fine visual details. The spatial coverage of a fixation is limited to no more than 4 visual degrees [13]. As such, eye movements are required to refocus the fovea to different regions of space in order to maintain a comprehensive understanding of the environment. An additional measure from eye tracking is fixation frequency. Fixation frequency can be indicative of either (1) the degree to which an HMI supports efficient visual search [14] or (2) the level of importance of certain information, described as areas of interests (AOIs), within the context of a particular task [15]. For example, an HMI display that enables efficiency visual search will be designed logically to the operator based upon an operational context. Fewer overall fixations for a task would be expected if the design supports efficient visual search. Similarly, important AOIs (e.g., key indications) that are most relevant to the task would be expected to yield more fixations relative to ancillary AOIs.

The saccade is a type of eye movement where one's point of gaze is rapidly shifted to a new location of the visual field [9]. Information processing is suppressed during the saccade [13]. Nonetheless, useful insight into visual search patterns can be gained from observing saccadic behavior. Generally, the amplitude or length of a saccade can be indicative workload where the saccade shortens when workload increases [16]. Frequency of saccades is analogous to the number of fixations and can be useful in determining search efficiency [8, 14].

Together, fixations and saccades provide an aggregate understanding of one's visual scan pattern. The total length and duration of one's scanning under a task can be used to measure visual search efficiency where a short scan path length and duration would suggest greater efficiency. Likewise, comparing the transitions of scanning AOIs to a baseline (i.e., scan transition matrix density) can be valuable to understanding optimal display layouts where a HMI display that requires more revisiting of the same AOI would be inferred as less efficient. Lastly, eye tracking can capture other physiological measures such as pupil diameter and blinks that are both correlated with workload and have been used in control room research [17, 18]. Workload has shown to be positively related to pupil diameter and negatively correlated to blink measures [18]. These measures can provide continuous tracking of workload, as opposed to pen and paper methods like NASA-TLX. The reader is referred to INL/EXT-15-37311 [12] for an in-depth discussion of these measures; however, Table 1 of this paper outlines the relation of select eye tracking measures to their corresponding relevant to human-system performance M&E in CRM.

Table I. Summary of eye tracking measures as they relate to human-system performance characteristics in control room modernization

Human-System Performance Characteristic	Measure	Relation
Visual Search Efficiency	Frequency of Fixations & Saccades	Negative
	Scan Path Duration & Length	Negative
	Scan Transition Matrix Density	Negative
Visual Attention	Fixations per Area of Interest	Positive
	Dwell Duration	Positive
Workload	Fixation Duration	Positive
	Pupil Diameter	Positive
	Frequency of Blinks	Negative
	Blink Rate	Negative
	Blink Duration	Negative

These measures were collected during a CRM design workshop to support three key objectives. The first objective was test the feasibility of including eye tracking in a full-scale control room simulator study. For example, the accuracy of the eye tracker was explored to determine how well eye movements mapped to relevant AOIs. In a full-scale simulator, the viewing distances of important information can vary, which can greatly impact eye tracking accuracy. Observed technical difficulties specific to the equipment were documented as well. The second objective was to provide preliminary findings about the sensitivity and validity of the selected eye tracking measures to determine their value for future studies. Finally, operator acceptance of eye tracking was collected to understand if there were any negative impacts to the validity of the scenarios run or unwillingness to participate in future studies using eye tracking.

3 EVALUATION OF THE VALUE USING EYE TRACKING IN FULL-SCALE SIMULATOR STUDIES

3.1 Summary of the Simulator Study Protocol

A human-system performance study was conducted to evaluate and inform the design of select HMI upgrades for a U.S. commercial light water reactor nuclear power plant. This evaluation was part of a smaller tasking effort to test the use and feasibility of eye tracking methodology in a full-scale simulator study. This study took place in the Idaho National Laboratory (INL) Human Systems Simulation Laboratory (HSSL), which housed a full-scale and full-scope configuration of the main control room (refer to Fig. 1 as a reference). Two crews of three-person licensed operators participated in this study, each on separate occasions. The HSSL enabled these operators to interact naturally within a virtualized environment of their main control room. Operators completed different scenarios that exercised use of the new HMI upgrades when compared to their existing control room. The new HMI upgrades incorporated new digital CRM technologies such as overview displays, computer based procedures, and an advanced prognostic diagnosis system known as the Computer Operator Support System (COSS). These new technologies were designed using state-of-the-art HFE design principles (e.g., Information Rich Design) to enhance the detection, diagnosis, and response of changes in plant conditions.

There were separate scenarios for each of the primary plant systems involved in this CRM effort. These systems entailed the Turbine Control System (TCS) and Chemical and Volume Control System (CVCS). For the TCS, one set of scenarios involved and introduced vibration abnormality to one of the turbine bearings while operators interacted with their existing boards, a hybrid board with task-based HMI displays without an overview display, and a hybrid board with task-based HMI displays with an overview display. For the CVCS, the scenario involved a letdown isolation event produced by an instrumentation and control malfunction. Operators interacted with their existing boards, a new a hybrid board with task-based and overview HMI displays, and a new hybrid board with the COSS implementation.

Eye tracking data was collected separately from the subjective measures to ensure visual search patterns made by the operators were representative to actual operation. The eye tracking glasses were used for data collection in the HSSL (refer to Fig. 1). This equipment comprised of two eye facing cameras and an outward-facing camera to record the study environment. The eye tracking glasses recorded at 60 Hz with an accuracy of 0.5 visual degrees. The glasses were connected to a mobile smart device by a micro-USB-to-USB connector. The mobile device utilized proprietary eye tracking software needed to record the eye tracking data onto a micro-SD card. Data captured from the micro-SD card were processed using the eye tracking vendor’s proprietary software, as well as with custom software created by INL for additional processing of select measures (e.g., pupil diameter).

The AOIs were broadly classified as specific regions of the control board, which corresponded to each glasstop monitor of the HSSL for either the TCS or CVCS system. Fig. 2 illustrates the generalized AOIs used for this preliminary study. Differences in visual search could be compared across the different conditions (i.e., conventional board layout versus hybrid control board with COSS) between these AOIs. The conventional layout presented a mimic representation of the CVCS sub-systems, dispersed across all of the AOIs. Most of the controls were arranged on the 'Bottom' AOIs. The hybrid control board with COSS utilized the 'Top' AOIs for overview display. The COSS system was presented on the 'Middle 1' AOI, while task-based I&C presented from a new HMI was presented on the 'Middle 2' AOI.

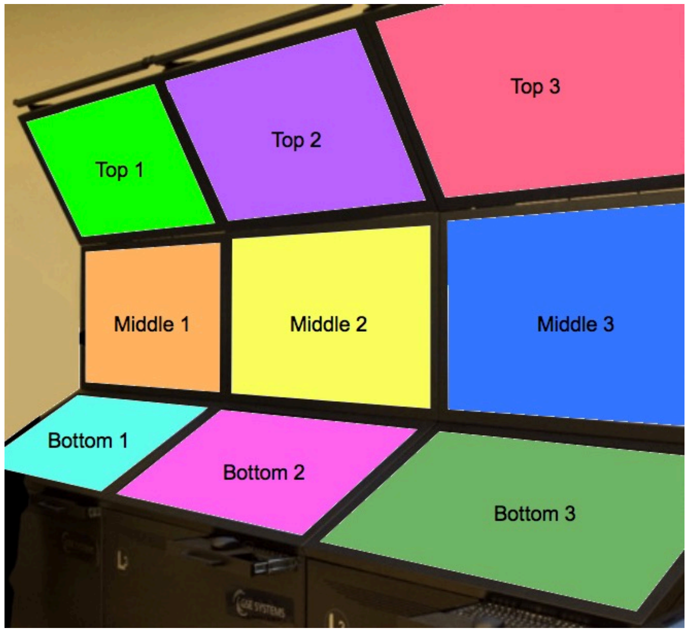

Figure 2. Generalized areas of interests created to compare visual search.

For simplicity, only one operator at a time wore the eye tracking glasses under select scenarios. This operator was chosen based on the level of anticipated interaction with the new HMI upgrade. A three-point calibration was completed before each eye tracking trial to ensure optimal accuracy. Lighting conditions were controlled for as all scenarios were completed in the HSSL. Operators were asked to interact as they normally would for each of the eye tracking scenarios.

3.2 Results

3.2.1 Feasibility of using eye tracking in a full-scale simulator study

A subset of scenarios that collected eye tracking data was determined usable for subsequent data analysis. Part of the reason for the lack of usable data was due to the inclusion of subjective data collection using a verbal protocol. These trials asked operators to verbalize their thoughts and opinions of the HMI interface as part of identifying design issues with the HMI upgrades. While these results had direct benefit to the primary objective for this design workshop, the validity of the eye tracking data was jeopardized during these trials as visual search patterns reflected eye movements both relevant and irrelevant to the scenario. Additionally, there were instances where the data was not successfully recorded. The cause for the lost data during this study was determined to be device-specific, rather than a methodological limitation of the use of eye tracking in a full-scale simulator study. That is, the micro-USB-to-USB connector from the eye tracking glasses to the mobile recording device had become loose in scenarios with lost data. To note, the models of eye tracking glasses used have recently been updated with new hardware that no longer uses the micro-USB-to-USB connector.

There were a total of three scenarios that were analyzed in this preliminary evaluation. Two of these three scenarios were from a single reactor operator interacting with the existing control room configuration compared to the hybrid control room incorporating the new HMI digital technology (i.e., with COSS) for the CVCS. The third usable scenario came from a reactor operator who completed the

CVCS scenario under the COSS condition. From this data, there were no observed issues with the accuracy of the eye tracking glasses. The average dwell percentage spent outside relevant regions of the control board was 0.2%, with a maximum percentage of 0.6% percent. These findings suggest that eye tracking can be included in full-scale simulator studies. Wearable eye tracking technology can capture accurate recordings of visual search patterns when AOIs were defined by each generalized simulator bay display. Nonetheless, investigation of how accuracy may be impacted when more detailed AOIs are defined should be pursued. Future studies that include eye tracking to collect objective measures must also consider separating simulator trials with verbal protocol from eye tracking in order to ensure valid interpretation of visual search.

3.2.2 Preliminary eye tracking measures

This section presents the eye tracking measures collected during the CVCS scenarios of this preliminary study. It must be emphasized that due to the lack of experimental control and limited sample size, these findings are not intended to show human-system performance characteristics that are representative of the new digital technologies. Rather, these findings simply demonstrate how such measures may be used in future carefully planned simulator studies that support CRM.

Measures of Visual Search Efficiency: Fixation frequency (count), saccade frequency (count), scan path duration (seconds), scan path length (total pixels), and scan transition matrix density (percent AOI coverage) are presented in Table 2. A HMI display that support efficiency visual search is indicative of requiring fewer fixations and saccades, shorter scan duration and length, and a lower matrix density.

Table II. Observed measures of visual search efficiency

Control Board	Fixation Frequency	Saccade Frequency	Scan Path Duration	Scan Path Length	Matrix Density	Scenario Time
RO1 Conv.	378	294	21.28	18648.38	0.29	169.31
RO1 New HMI	901	806	61.27	45723.00	0.41	332.11
RO2 New HMI	515	471	42.48	27851.54	0.31	169.67

Visual Attention: Fixations per AOI (count) is presented in Table 3 while dwell duration (percent) is presented in Table 4. These measures capture the proportion of visual attention devoted to a certain AOI, and can be used to determine what information or region of the control room captured operators' attention most, as well as what areas were captured least.

Table III. Observed measures of visual attention: Fixations per AOI

Control Board	Bottom			Middle			Top		
	1	2	3	1	2	3	1	2	3
RO1 Conv.	12	83	99	2	1	8	1	7	13
RO1 New HMI	69	81	58	36	159	24	358	43	72
RO2 New HMI	2	4	2	90	200	15	177	5	14

Table IV. Observed measures of visual attention: Dwell duration

Control Board	Bottom			Middle			Top		
	1	2	3	1	2	3	1	2	3
RO1 Conv.	2.0%	20.7%	22.9%	1.0%	0.3%	29.2%	0.5%	3.3%	7.9%
RO1 New HMI	8.6%	8.0%	10.1%	2.3%	11.1%	7.4%	29.7%	3.4%	6.3%
RO2 New HMI	0.4%	0.5%	0.2%	12.9%	29.2%	2.6%	27.4%	0.5%	1.7%

Workload: Fixation duration (seconds), pupil diameter (millimeters), blink frequency (count), blink rate (count per second), and blink duration (seconds) are presented in Table 5. Increased workload is suggested to be indicative of greater fixation duration, greater pupil diameter, greater blink frequency, lesser blink rate, and lesser blink duration.

Table V. Observed measures of workload

Control Board	Fixation Duration	Pupil Diameter	Blink Frequency	Blink Rate	Blink Duration
RO1 Conv.	0.38	2.98	18	2.82	0.29
RO1 New HMI	0.23	2.85	38	6.87	0.34
RO2 New HMI	0.29	2.09	8	6.37	0.75

Tables 2 through 5 present potential eye tracking measures that can be captured using wearable eye tracking equipment. It should be noted that additional pupil diameter processing was required after exporting the raw data from the eye tracking proprietary analysis software in order to remove artifacts such as blinks. Thus, additional processing entailed (1) removing all instances where pupil diameter was recorded at 0 millimeters, (2) applying a linear correction to the removed instances, and (3) smoothing the data using a low-pass third order Butterworth filter with a cutoff of 4 Hz to capture cognitive processes [19]. Fig. 3 illustrates the effect of applying this post-processing technique to the raw pupil data of a single operator in scenario condition.

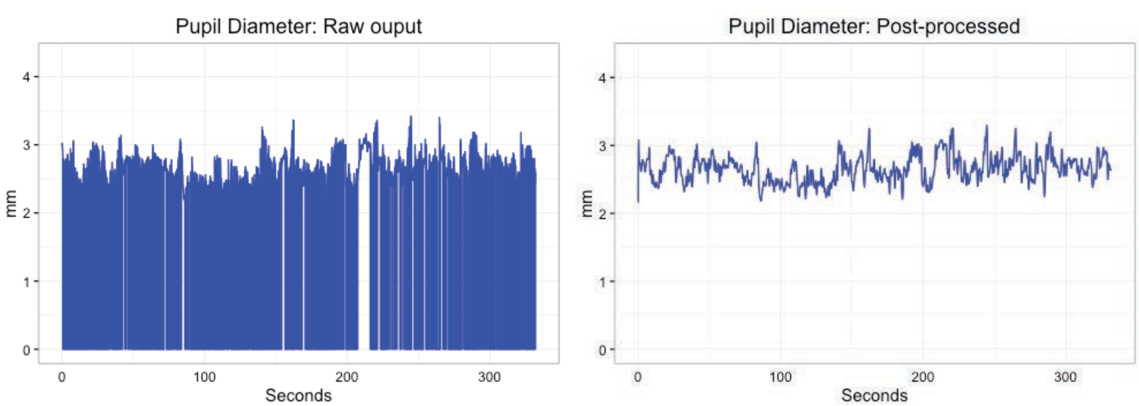

Figure 3. Raw versus post-process pupil diameter: A comparison across one scenario.

While the results from this preliminary evaluation are not statistically powered to allow for generalizable results between the conventional CVCS board versus a hybrid CVCS with COSS, the measures collected were noticeably different between conditions. Interestingly, the directionality of these selected measures concurred with what would be expected when using a diverse set of eye tracking measures in evaluating various human-system performance characteristics. For example when qualitatively looking at RO1's visual search efficiency between conditions, all measures of visual search efficiency were lower in one condition (i.e., New HMI) compared to the other condition (i.e., Conventional). Similar directionality can be qualitatively noticed with workload.

3.2.3 Operator acceptance of eye tracking

Operator acceptance was collected from the following multiple-choice questions administered as a questionnaire at the end of each simulator scenario that included eye tracking. These questions were:

- *Question 1:* How comfortable were the eye tracking glasses (and associated equipment)? *Responses:* Not comfortable, moderately comfortable, comfortable, N/A.
- *Question 2:* Were the eye tracking glasses distracting? *Responses:* Yes, No, N/A
- *Question 3:* Did the eye tracking glasses affect your performance? *Responses:* Yes, No, N/A
- *Question 4:* Would you be willing to wear the eye trackers for future studies? *Responses:* Yes, No, N/A

A total of 5 operators answered the questionnaire. Each operator was provided the opportunity to respond to the questionnaire during for 4 different scenarios (i.e., all conditions for the CVCS and HMI only condition for the TCS). Table 6 presents the responses for Question 1. Table 7 presents the responses for Questions 2 through 4.

Table VI. Tabulation of responses: How comfortable were the eye tracking glasses (and associated equipment)?

Ops.	Not comfortable	Moderately comfortable	Comfortable	N/A
RO1	0	1	2	1
RO2	0	1	0	3
RO3	1	2	0	1
RO4	0	1	0	3
RO5	0	0	0	4
Overall (%)	5%	25%	10%	60%

Table VII. Tabulation of responses: Questions 2 through 4

Ops.	Question 2			Question 3			Question 4		
	Yes	No	N/A	Yes	No	N/A	Yes	No	N/A
RO1	0	3	1	0	3	1	3	0	1
RO2	0	1	3	0	1	3	1	0	3
RO3	2	1	1	0	3	1	3	0	1
RO4	0	1	3	0	1	3	1	0	3
RO5	0	1	3	0	1	3	1	0	3
Overall (%)	10%	35%	55%	0%	45%	55%	45%	0%	55%

There were a large proportion of responses marked as ‘n/a’ across all of the questions. One reason for the ‘n/a’ responses was due to the limited time given at the end of each scenario to administer this specific questionnaire. Of the available responses, qualitative inspection suggests that operators were generally accepting of including eye tracking in the full-scale simulator study. Most coded responses from operators for Question 1 was that wearing the equipment was ‘moderately comfortable.’ Operators generally responded that wearing the eye tracking equipment was not distracting. No operators responded that the eye tracking equipment impacted their performance during the scenarios. All operators were willing to wear the eye tracking equipment in future studies. There was one operator (i.e., RO3) who responded that the eye tracker was ‘not comfortable’ during one scenario, but also responded that the eye tracker was ‘moderately comfortable’ during two of the scenarios. This same operator indicated that the eye tracking equipment was distracting (i.e., responded ‘Yes’ to Question 2) during two of the scenarios.

Collectively, these preliminary results suggest that use of eye tracking is an acceptable method for operators in full-scale simulator studies; further, including eye tracking should not negatively impact to operator’s perceived level of performance. A within-subjects design should always be used when comparing different candidate technologies to ensure potential measurement bias from the use of eye tracking on any one condition is experimentally controlled. Likewise, future full-scale simulator studies must consider how potential eye tracking hardware limitations might negatively impact operator performance or become a distraction. These considerations are particularly important for ISV.

4 CONCLUSIONS

This paper presented preliminary findings of the overall feasibility, data characteristics, and operator acceptance with including eye tracking into a full-scale simulator study for CRM. Despite there being limitations in the overall experimental design and statistical power, these preliminary results offer some insight into the potential value of including eye tracking as part of a diverse set of measures that support objective human-system performance evaluation in full-scale simulator studies. Moving forward, future work will consider including eye tracking in a formally designed simulator study specific to using eye tracking. For instance, developing more focused tasks that capture important points in time that tax operators’ workload or situation awareness can improve the sensitivity and diagnosticity of these measures. Moreover, having access to a greater sample of operator will improve that statistical conclusion validity made with these objective measures.

5 ACKNOWLEDGMENTS

INL is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517. This work of authorship was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights. The United States Government retains, and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof. The INL issued document number for this paper is: INL/CON-16-404000.

6 REFERENCES

1. U.S. Nuclear Regulatory Commission, *Human Factors Engineering Program Review Model, NUREG-0711, Rev. 3,* U.S. Nuclear Regulatory Commission, Washington, DC (2012).
2. 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.
3. Stanton, N., Salmon, P. M., & Rafferty, L. A. (2013). *Human factors methods: a practical guide for engineering and design*. Ashgate Publishing, Ltd.
4. Andre, A. D., & Wickens, C. D. (1995). When users want what's not best for them. *Ergonomics in design*, *3*(4), 10-14.
5. LeBlanc, K., Boring, R., Joe, J., Hallbert, B., & Thomas, K. (2014). *A Research Framework for Demonstrating Benefits of Advanced Control Room Technologies*, INL/EXT-14-33901 Revision 1.
6. Bergstrom, J. R., & Schall, A. (Eds.). (2014). *Eye tracking in user experience design*. Elsevier.
7. Fitts, P.M., Jones, R.E., & Milton, J.L. (1950). Eye movements of aircraft pilots during instrument landing approaches. *Aeronautical Engineering Review, 9*(2), 24-29.
8. Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind, 2*(3), 4.
9. Parasuraman, R., & Rizzo, M. (Eds.). (2008). *Neuroergonomics: The brain at work.* Oxford University Press.
10. Bhavsar, P., Srinivasan, B., & Srinivasan, R. (2015). Pupillometry based Real-time Monitoring of Operator’s Cognitive Workload to Prevent Human Error during Abnormal Situations. *Industrial & Engineering Chemistry Research*, *55*(12), 3372-3382.
11. Ha, J. S., Byon, Y. J., Baek, J., & Seong, P. H. (2016). Method for Inference of Operators' Thoughts from Eye Movement Data in Nuclear Power Plants. *Nuclear Engineering and Technology*, *48*(1), 129-143.
12. 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.
13. Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance*. Psychology Press.
14. Goldberg, J. H., & Wichansky, A. M. (2002). Eye tracking in usability evaluation: A practitioner’s guide. In R. Radach, J. Hyönä, & H. Deubel. (Eds), *The Mind's Eyes Cognitive and Applied Aspects of Eye Movements*. Amsterdam: Elsevier Science, 493-516.
15. Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. *Encyclopedia of human computer interaction*, *1*, 211-219.
16. Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T. P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 43*(1), 111-121.
17. Gao, Q., Wang, Y., Song, F., Li, Z., & Dong, X. (2013). Mental workload measurement for emergency operating procedures in digital nuclear power plants. *Ergonomics*, *56*(7), 1070-1085.

18. Noah, B., Kim, J.H., Rothrock, L., & Tharanathan, A. (2014). Evaluating Alternate Visualization Techniques for Overview Displays in Process Control. *IIE Transactions on Occupational Ergonomics and Human Factors, 2*(3–4), 152–168.

19. Kloosterman, N. A., Meindertsma, T., Loon, A. M., Lamme, V. A., Bonneh, Y. S., & Donner, T. H. (2015). Pupil size tracks perceptual content and surprise. *European Journal of Neuroscience*, *41*(8), 1068-1078.