



Application of Eye Tracking for Measurement and Evaluation in Human Factors Studies in Control Room Modernization

March 2018

Changing the World's Energy Future

Casey R Kovesdi, Spielman, Z.A, Rice, B.C., Le Blanc, K.L.



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

DISCLAIMER

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

Application of Eye Tracking for Measurement and Evaluation in Human Factors Studies in Control Room Modernization

Casey R Kovesdi, Spielman, Z.A, Rice, B.C., Le Blanc, K.L.

March 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**

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

Casey Kovesdi^a, Zachary Spielman^a, Katya LeBlanc^a, Brandon Rice^a

^aIdaho National Laboratory, Human Factors, Controls, and Statistics Department

PO Box 1625, Mail Stop, 3818 Idaho Falls, Idaho 83415, USA

*Email: Casey.Kovesdi@inl.gov

Number of pages: 21

Number of tables: 7

Number of figures: 4

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

I. 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 a 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 describing the effects of 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. Technical considerations and the unique challenges with using eye tracking in full-scaled simulations are addressed. 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.

II. APPLICATION OF EYE TRACKING IN CONTROL ROOM MODERNIZATION

II.A. 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]. In 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.

Decades of research support the case for including eye tracking as a method to provide 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 the 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 evaluate 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.

II.B. 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.



Fig. 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 result in a suboptimal placement of the eye tracking when worn, which can negatively affect data accuracy. 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

II.C. Applicable Eye Tracking Measures for Control Room Modernization

Eye tracking can be described as a 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 interest (AOIs), within the context of a particular task [15]. For example, an HMI display that enables efficient 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 of 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

The following section shares experiences with using eye tracking in a CRM design workshop using a full-scale and full-scope simulator to gain qualitative insights into its feasibility, sensitivity, and usability when used in this context. To assess feasibility, the researchers determined how well eye movements mapped to relevant AOIs. In a full-scale simulator, operators are moving around the room, which can lead to variation in the viewing distances of important information. Viewing distance can greatly impact eye tracking accuracy. Another aspect of feasibility is how well the equipment functions; therefore, technical difficulties specific to the equipment were documented. For sensitivity, preliminary eye tracking measures were collected (i.e., refer to Table 1) to qualitatively examine sensitivity of these selected measures evaluate their ability to characterize aspects of human performance in future studies. Finally, operator acceptance of eye tracking was assessed to explore the equipment’s usability when used in a full-scope simulator study and 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. A brief summary of the testing environment and eye tracking equipment used in this effort is discussed next.

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

III.A. Summary of the Testing Environment and Eye Tracking Equipment

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 houses a full-scale and full-scope configuration of the main control room

(refer to Fig. 1 as a reference). Two three-person crews of licensed operators participated in this study. Each crew completed realistic scenarios that exercised different plant systems (e.g., Turbine Control System and Chemical and Volume Control System) using a conventional control board configuration compared to a hybrid control board configuration (i.e., with and without overview displays). The HSSL enabled these operators to interact naturally within a virtualized environment of their main control room.

The eye tracking glasses (refer to Fig. 1) were used, which 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 AOIs were broadly classified as specific regions of the control board that corresponded to each monitor on the glasstop simulator bay. Fig. 2 illustrates these generalized AOIs used. Differences in visual search could be compared across the different board configurations. Most of the controls were arranged on the 'Bottom' AOIs. The hybrid control boards with overview displays utilized the 'Top' AOIs. The new HMI displays were presented on the 'Middle' AOIs.

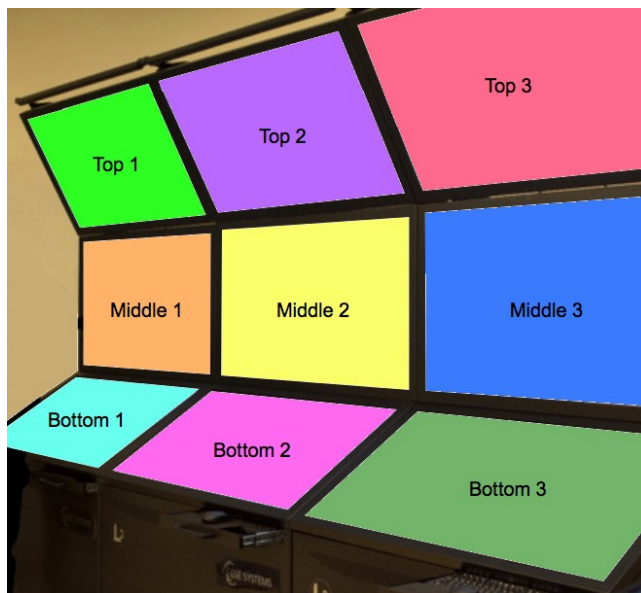


Fig. 2. Generalized AOIs created to compare visual search.

During each scenario, the operator that was anticipated to interact the most with the relevant control boards wore the eye tracking glasses. A three-point calibration was completed before each eye tracking

trial to ensure optimal accuracy. Lighting conditions were controlled for as all scenarios, and operators were instructed to interact as they normally would for each of the eye tracking scenarios.

III.B. Feasibility of Eye Tracking

There was a total of three different scenarios that were analyzed in preliminary evaluation. Two of these three scenarios were from a single reactor operator (RO1) interacting with the existing control room configuration compared to the hybrid control room incorporating the new HMI digital technology including an overview display. The third usable scenario came from a second reactor operator (RO2) using the new HMI digital technology including an overview display.

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.

Regarding how well the equipment functioned, there were instances where the data was not successfully recorded due to the micro-USB-to-USB connector from the eye tracking glasses to the mobile recording device becoming loose, resulting in 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.

III.C. Sensitivity of Measures

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. Fewer fixations and saccades, shorter scan duration and length, and a lower matrix density all indicate an HMI that supports efficient visual search.

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. Fig. 3 provides a heat map to visually represent these dwell durations; darker red colored regions indicate maximum dwell duration (i.e., 2000 ms) while blue and lighter green colored regions indicate minimum dwell duration (i.e., 200 ms – 800 ms). 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%

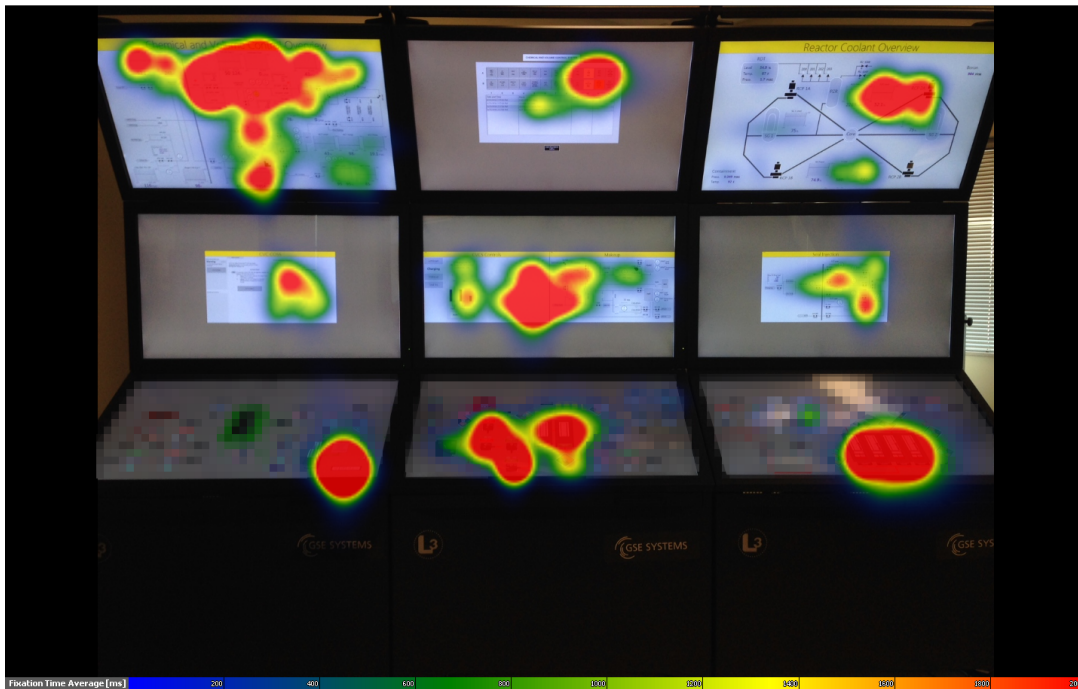


Fig. 3. Heat map to illustrate dwell durations (ms) over selected AOIs.

Workload: Fixation duration (seconds), pupil diameter (millimeters), blink frequency (count), blink rate (count per second), and blink duration (seconds) are presented in Table 5 with a comparison to the operators' NASA-TLX responses (1= Low Workload; 100 = High Workload). 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. 4 illustrates the effect of applying this post-processing technique to the raw pupil data of a single operator in scenario condition. Greater fixation duration, greater pupil diameter, fewer blinks, lesser blink rate, lesser blink duration, and greater NASA-TLX responses may indicate increased workload.

TABLE V.
Observed measures of workload

Control Board	Fixation Duration	Pupil Diameter	Blink Frequency	Blink Rate	Blink Duration	NASA-TLX
RO1 Conv.	0.38	2.98	18	2.82	0.29	26.67
RO1 New HMI	0.23	2.85	38	6.87	0.34	25
RO2 New HMI	0.29	2.09	8	6.37	0.75	28.33

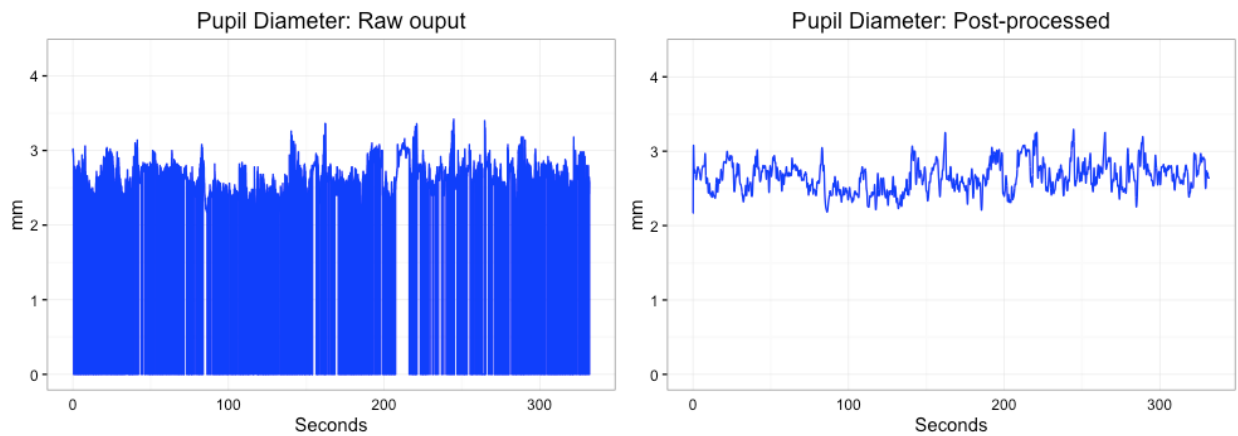


Fig. 4. Raw versus post-processed pupil diameter: A comparison across on scenario.

This preliminary evaluation was not intended to provide generalizable results, the measures collected were noticeably different between conditions. The different eye tracking measures were consistent with one another in terms of the direction and what each measure indicated for the various interface conditions. For example, a qualitative examination of the results from RO1's visual search efficiency between the different HMI conditions shows all measures of visual search efficiency were lower in one condition (i.e., New HMI) compared to the other condition (i.e., Conventional). This is consistent with the workload results.

Another important point is that a within-subjects design should always be used when using eye tracking to compare different candidate technologies to ensure potential measurement bias from the use of eye tracking in any one condition is experimentally controlled. For example, when looking at pupil diameter, RO2's diameter is noticeably smaller than RO1's. Without a comparison between conditions, it is unclear if this difference is due to individual differences or from the control board configuration under study.

III.D. Usability of Equipment

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. A total of 5 operators answered the questionnaire. Each operator was provided the opportunity to respond to the questionnaire after each eye tracking scenario. Tables 6 and 7 present the findings concerning the usability of the eye tracking equipment.

TABLE VI.

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

Operator	Not comfortable	Moderately comfortable	Comfortable
RO1	0	1	2
RO2	0	1	0
RO3	1	2	0
RO4	0	1	0
RO5	0	0	0

Overall (%)	13%	63%	25%
-------------	-----	-----	-----

TABLE VII.

Tabulation of responses: Questions 2 through 4

Operator	Question 2: Were the eye tracking glasses distracting?		Question 3: Did the eye tracking glasses affect your performance?		Question 4: Would you be willing to wear the eye trackers for future studies?	
	Yes	No	Yes	No	Yes	No
RO1	0	3	0	3	3	0
RO2	0	1	0	1	1	0
RO3	2	1	0	3	3	0
RO4	0	1	0	1	1	0
RO5	0	1	0	1	1	0
Overall (%)	22%	78%	0%	100%	100%	0%

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 indicated 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. A follow-up interview with RO3 identified

the primary contributor for these negative responses to be associated with the placement of the eye tracker's heat sync being placed in his pocket, as opposed to its designated external pack, which caused the device to overheat during these selected scenarios, making it uncomfortable to wear the device.

These preliminary results from the questionnaire suggest that use of eye tracking is an acceptable method for operators to use when participating in full-scale simulator studies. Including eye tracking should not negatively impact the operator's perceived level of performance. Though, care should be given with managing the basic logistics of fitting the equipment to the operator. For instance, the misplacement of the eye tracker's heat sync caused reported discomfort from one operator during the study. Having a formal process to ensure that all equipment is properly fitted and calibrated should reduce risk of these logistical problems. 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.

IV. CONCLUSIONS

This paper presented preliminary findings of the overall feasibility, sensitivity, and usability (i.e., operator acceptance) with including eye tracking into a full-scale simulator study for CRM. Despite there being limitations in the overall experimental design, which limited analysis to qualitative examination of the results. 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.

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/JOU-16-404000.

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] N. STANTON, P. M. SALMON, AND L. A. RAFFERTY, *Human Factors Methods: A Practical Guide for Engineering and Design*. Ashgate Publishing, Ltd. (2013).
- [4] A. D. ANDRE, AND C. D. WICKENS, "When Users Want What's Not Best for Them," *Ergonomics in design*, **3**, 4, 10-14 (1995).
- [5] K. LEBLANC, R. BORING, J. JOE, B. HALLBERT, AND K. THOMAS, "A Research Framework for Demonstrating Benefits of Advanced Control Room Technologies," INL/EXT-14-33901 Revision 1, Idaho National Laboratory (2014).
- [6] J. R. BERGSTROM AND A. SCHALL, *Eye Tracking in User Experience Design*. Elsevier (2014).
- [7] P. M. FITTS, R. E. JONES, AND J. L. MILTON, "Eye Movements of Aircraft Pilots during Instrument Landing Approaches," *Aeronautical Engineering Review*, **9**, 2, 24-29 (1950).
- [8] R. J. JACOB AND K. S. KARN, "Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises," *Mind*, **2**, 3, 4 (2003).
- [9] R. PARASURAMAN AND M. RIZZO, *Neuroergonomics: The brain at work*. Oxford University Press (2008).

- [10] P. BHAVSAR, B. SRINIVASAN, AND R. SRINIVASAN, "Pupillometry Based Real-Time Monitoring of Operators Cognitive Workload to Prevent Human Error during Abnormal Situations," *Industrial & Engineering Chemistry Research*, **55**, 12, 3372-3382 (2015).
- [11] J. S. HA, Y. J. BYON, J. BAEK, AND P. H. SEONG, "Method for Inference of Operators' Thoughts from Eye Movement Data in Nuclear Power Plants," *Nuclear Engineering and Technology*, **48**, 1, 129-143 (2016).
- [12] C. KOVESDI, B. RICE, G. BOWER, Z. SPIELMAN, R. HILL, R., AND K. LEBLANC, "Measuring Human Performance in Simulated Nuclear Power Plant Control Rooms Using Eye Tracking," INL/EXT-15-37311. Revision 0, Idaho National Laboratory (2015).
- [13] C. D. WICKENS, J. G. HOLLANDS, S. BANBURY, AND R. PARASURAMAN, *Engineering Psychology & Human Performance*. Psychology Press (2015).
- [14] J. H. GOLDBERG AND A. M. WICHANSKY, *Eye Tracking in Usability Evaluation: A Practitioner's Guide*, IN R. RADACH, J. HYÖNÄ, AND H. DEUBEL (Eds), *The Mind's Eyes Cognitive and Applied Aspects of Eye Movements*. Amsterdam: Elsevier Science, 493-516 (2002).
- [15] A. POOLE AND L. J. BALL, "Eye Tracking in HCI and Usability Research," *Encyclopedia of Human Computer Interaction*, **1**, 211-219 (2006).
- [16] K. F. VAN ORDEN, W. LIMBERT, S. MAKEIG, AND T. P. JUNG, "Eye Activity Correlates of Workload during a Visuospatial Memory Task," *Human Factors*, **43**, 1, 111-121 (2001).

- [17] Q. GAO, Y. WANG, F. SONG, Z. LI, AND X. DONG, "Mental Workload Measurement for Emergency Operating Procedures in Digital Nuclear Power Plants," *Ergonomics*, **56**, 7, 1070-1085 (2013).
- [18] B. NOAH, J. H. KIM, L. ROTHROCK, AND A. THARANATHAN, "Evaluating Alternative Visualization Techniques for Overview Displays in Process Control," *IIE Transactions on Occupational Ergonomics and Human Factors*, **2**, 3-4, 152-168 (2014).
- [19] N. A. KLOOSTERMAN, T. MEINDERTSMA, A. M. LOON, V. A. LAMME, Y. S. BONNEH, AND T. H. DONNER, "Pupil Size Tracks Perceptual Content and Surprise," *European Journal of Neuroscience*, **41**, 8, 1068-1078 (2015).