Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira,
Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria

doi: 10.3850/ESREL2026061419_esrel26-p26101-cd

Exploring Head-Tracking for use in Overview Display study in a Nuclear Control Room Setting

Hoa Thi Nguyen

Department of Applied Data Science, Institute for Energy Technology, Norway. E-mail: hoathi.nguyen@ife.no

Jens-Patrick Langstrand

Department of Applied Data Science, Institute for Energy Technology, Norway.
E-mail: jens.patrick.langstrand@ife.no

Human performance in control room environments has traditionally been assessed using expert observations, task-based performance measures, and operator self-reports. Recent advances in sensing technologies enable the collection of biometric data that can complement these established methods. However, a key challenge lies in bridging the gap between data capture and meaningful interpretation: while computer vision and sensor technologies provide new means for measuring operator behavior, valid interpretation requires domain knowledge from human factors and process experts.

This paper presents a study investigating the use of head-orientation tracking as a non-intrusive, biometric measure for evaluating operator interaction with digital displays in a nuclear control room. Certified operators participated in simulated scenarios at the Halden Man–Machine Laboratory, including an unexpected loss of a large overview display. Head-orientation data, captured through computer vision and mapped into a digital twin of the control room, was analyzed alongside operator self-evaluations and process expert assessments.

The results indicate that head-orientation tracking provides objective and scalable indicators of display usage and visual engagement, aligning more closely with expert evaluations than with self-reports in several cases. The biometric data also revealed behavioral patterns not captured by traditional methods, such as continued visual attention toward the overview display after its removal. These findings demonstrate the value of biometric-driven analysis as a complementary tool for human performance evaluation and highlight the importance of interdisciplinary collaboration in developing robust interpretation frameworks for biometric data in safety-critical environments.

Keywords: headtracking, human assessment, overview display, nuclear control room.

1. Introduction

Nuclear control room operation is safety-critical and places high demands on operators' ability to monitor, interpret, and communicate system information. As digital control room systems become increasingly prevalent, operators must interact with layered displays and navigate across multiple screens, which can increase cognitive and interaction demands compared to analogue interfaces.

To support monitoring and team communication in digital environments, overview displays (ODs) have been introduced as a complementary information source. ODs are typically large, centrally mounted displays designed to present key plant parameters and system states at a glance Härefors (2008), supporting situation awareness and shared understanding among operators Fernandes et al. (2020). While several studies report positive effects of overview displays—particularly in reducing information search and supporting efficient monitoring under certain conditions, including studies conducted at HAMMLAB Fernandes et al. (2023); Braseth and Fernandes (2024); Fernandes et al. (2024)—their effectiveness remains dependent on display design, task context, and operator strategies.

Usability evaluations of overview displays have traditionally relied on expert observations, performance measures, and subjective operator assessments. Although well established, these methods provide limited insight into how operators physically attend to and interact with displays during task execution. In this study, head-orientation tracking is investigated as a non-intrusive metric to complement traditional usability evaluation methods, with a focus on analyzing operator interaction with digital displays. By combining head-orientation data with established human factors methods, the study explores how objective, behavior-based measurements can complement traditional performance evaluations. The potential benefits and limitations of head-orientation tracking are addressed through the following research questions:

(i) How can head-orientation data be collected in a meaningful and reliable manner for evaluating display usage in a control room context?
(ii) How can head-orientation data be interpreted in conjunction with traditional usability evaluation methods?
(iii) What limitations exist when using head-orientation tracking to capture and interpret operator behavior?

To answer these questions, an overview display user test was conducted at HAMMLAB, where certified operators responded to a sudden and unexpected loss of the overview display. The results indicate a positive relationship between usability ratings and perceived usefulness, supported by operator feedback from post-scenario debrief interviews. Head-orientation data demonstrates potential for revealing operators' visual engagement and interaction patterns in a multi-display control room environment.

2. Methodology

2.1. *Procedure*

The experiment was conducted over three sessions during the summer of 2024, with a different operating crew in each session. Each crew consisted of four roles: Reactor Operator (RO), Balance of Plant Operator (BOP), Shift Supervisor (SS), and Shift Technical Advisor (STA). Prior to the experiments, the crews received approximately six hours of training to familiarize themselves with the digital displays and the overview display.

In each session, the crew completed four simulated scenarios using a Generic Pressurized Water Reactor (GPWR) simulator, each lasting 25–30 minutes, with a 10-minute break after the second scenario. A crew debrief interview was conducted at the end of each day. All scenarios began with a standard turnover and visible displays, followed by triggered events—primarily reactor trips—requiring operators to assess the situation and follow established procedures. Table 1 summarizes the scenario runs and indicates when the overview display was hidden.

Table 1. Overview of scheduled runs. "All" indicates the overview display was visible throughout the run; "OD loss" indicates temporary removal.

Run	Scenario	Crew	Display
201	4	1	All
202	1	1	All
203	3	1	OD loss
204	2	1	OD loss
206	3	2	OD loss
207	1	2	All
208	2	2	All
209	4	2	OD loss
211	2	3	OD loss
212	3	3	All
213	1	3	OD loss
214	4	3	All

The Synopticon system Langstrand et al. (2018) was used to collect head-position data via webcams mounted at each operator workstation, enabling tracking of which displays were viewed, when, and by whom. Further details of the setup are provided in Section 2.2.

After each scenario, participants completed a brief self-evaluation questionnaire (2) assessing perceived performance. In addition, one or two Process Experts (PEs) observed each run and evaluated crew performance using a 7-point scale based on the criteria listed in Table 3. These evaluations were informed by observations of head orientation and crew communication patterns, with delayed responses used as indicators of higher disruption.

All participating crews were informed about the study procedures and provided written consent prior to participation. All data were collected and stored anonymously.

2.2. *Head Tracking System*

Synopticon is a sensor fusion platform that integrates biometric sensing with digital twin technology to capture and annotate human behavior in real time Langstrand et al. (2018). The platform provides the following capabilities:

(i) Virtual twin construction of the control room: Synopticon integrates the OptiTrack position tracking system to enable virtual twin creation. A custom modeling tool with attached position-tracking markers was developed to record the positions of three screen corners (bottom-left, top-left, and top-right), which are sufficient to reconstruct each display within the control room layout. Digital representations of the control room were created in Unreal Engine 5 (Figure 2). OptiTrack provided millimeter-level tracking accuracy. Three wide-angle OptiTrack cameras, mounted on a tripod, were systematically repositioned throughout HAMMLAB to capture the positions and orientations of the operators' workstations and the overview display (Figure 1). The resulting 3D model was generated once and used consistently throughout the study.

(ii) Head-orientation monitoring: Each of the four operator workstations was equipped with a webcam mounted on the central screen. Head orientation was tracked using OpenTrack Halik (Halik), which provides XYZ position data and roll, pitch, and yaw orientation of the operator's head relative to the camera. This data was streamed into Synopticon's virtual control room twin, where head movements were mapped onto the 3D model to automatically identify which screens the operator was viewing (Figure 2). This head-orientation extraction method was selected for its non-intrusive nature, ensuring minimal interference with operator tasks. A single webcam was sufficient to reliably capture the primary and adjacent displays at each workstation, although tracking accuracy decreased for screens located farther from the camera.

3. Results Headtracking analysis

Self-evaluations were completed individually, while process experts assessed overall crew performance. During data collection, event timings were recorded in Synopticon to relate the data to different scenario stages and to distinguish between periods when the overview display (OD) was visible or removed. The total time operators spent viewing the overview display at each stage was aggregated and exported using Synopticon (Figure 3, Figure 4, Figure 5).

3.1. *First crew*

Discrepancies were observed between the First Crew's self-evaluations (4) and the recorded OD viewing time from Synopticon (Figure 3), except for the STA, whose reported usage closely matched the measured viewing time, with low scores in Runs 203 and 204. In contrast, the RO reported higher usage in Run 201 despite spending over 40 seconds viewing the OD, while in Run 202 a higher reported usage was given despite less recorded viewing time. Other crew members also reported high usage in Runs 203 and 204, even though the OD was hidden shortly after the

Fig. 1. Optitrack cameras setup for 3D modelling

Table 2. Sefl-evaluation questions for operators. The evaluation scale is from 1 to 7

Self-evaluation Questions	Scale 1	Scale 7
How often have you used the overview display in this scenario	Not at all	Very often
To what extend did the loss of the overview display disrupt the crew's work	Not at all	Significantly

Table 3. Process Experts' evaluations criterias. The evaluation scale is from 1 to 7

Process Experts' Questions	Scale 1	Scale 7
To what extend did the crew rely on the overview display in this scenario?	Not at all	Very often
To what extend did the loss of the overview display disrupt the crew's work	Not at all	Significantly

Fig. 2. 3D model of HAMMLAB with head-tracked crew in Synopticon. Operator roles clockwise from bottom left are: SS, RO, BoP, STA

first trigger and recorded viewing times were low. Notably, the RO continued glancing at the OD after it was hidden. Overall, self-reported disruption among the First Crew remained high; although they felt more prepared for the loss of OD, they still reacted negatively during the second occurrence, as reflected in elevated disruption scores.

The Process Experts' (PE) evaluations generally aligned more closely with the actual time the crew spent viewing the OD than the crew's self-evaluations. Usage for Run 203 was rated as high despite the OD being hidden, as the crew continued to glance at its location. From Run 203 to Run 204, both usage and disruption ratings decreased in the PE evaluations, consistent with the crew's reports of adapting to the expectation that the OD would eventually be removed. While high disruption was noted when the OD was first hidden in Run 203, the crew appeared to adapt quickly, as reflected in the lower disruption score for Run 204.

In this crew, the STA also had access to the OD on a right-side workstation monitor due to limited visibility of the OD. This was not accounted for in the current analysis, and the recorded OD viewing time is therefore likely underestimated. Synopticon data revealed patterns not fully captured by the PE evaluations, including a marked decline in OD usage during the final two runs.

Table 4. First crew self-evaluation and process expert-evaluation. "U." is short for usage and "D." is short for disruption.

Role	Run 201		Run 202		Run 203		Run 204	
	U.	D.	U.	D.	U.	D.	U.	D.
Crew self-evaluation								
RO	4	-	6	-	5	6	5	6
SS	5	-	6	-	6	6	6	6
STA	6	-	5	-	1	7	2	6
BoP	7	-	4	-	7	6	7	7
Avg	5.5	-	5.25	-	4.75	6.25	5	6.25
Std	1.12	-	0.83	-	2.28	0.43	1.88	0.43
Process expert-evaluations								
PE1	5	-	4	-	4	6	4	3
PE2	4	-	3	-	6	6	5	3

3.2. *Second Crew*

During the Second Crew's scenarios, technical issues with the webcam prevented head-orientation tracking for the BOP role (Figure 4). As observed for the First Crew, self-reported OD usage did not align with the OD viewing time recorded by Synopticon. The RO reported maximum usage across all runs, although recorded OD viewing time in Runs 206 and 207 was substantially lower than in

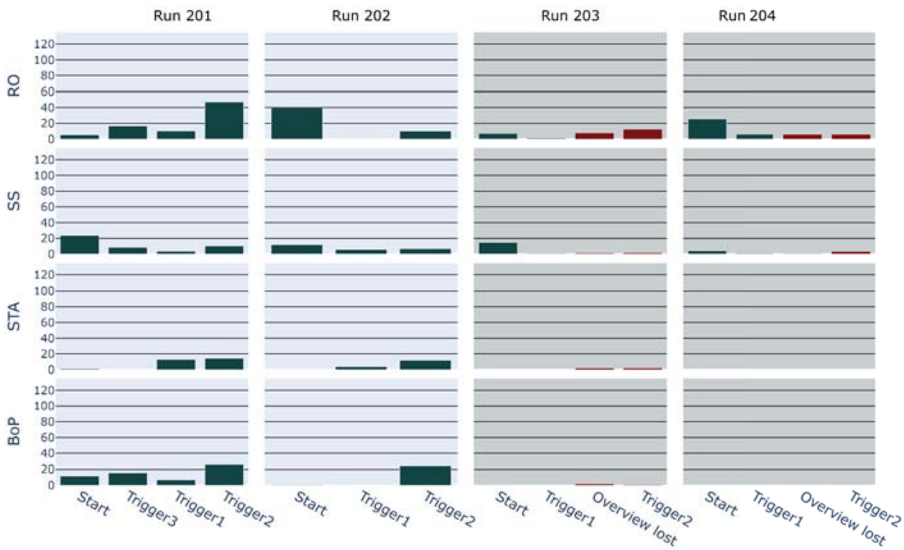

Fig. 3. The total time(s) the RO, SS, STA, and BoP of the First crew spent looking at the Overview Display

Runs 208 and 209. Self-reported disruption was slightly lower during the second instance of OD removal (Run 209) than during the first (Run 206), despite increased OD viewing time in Run 209.

In contrast, the PE evaluations more closely reflected the recorded OD usage, with the lowest usage rating in Run 206 and increasing scores in Runs 207 and 208. The crew spent less time viewing the OD prior to its first removal in Run 206 than in the second removal in Run 209, which may explain the higher disruption rating assigned by the PE. This pattern was not reflected in the crew's self-evaluations.

Table 5. Second crew self-evaluation and process expert-evaluation. "U." is short for usage and "D." is short for disruption.

Role	Run 206		Run 207		Run 208		Run 209	
	U.	D.	U.	D.	U.	D.	U.	D.
Crew self-evaluation								
RO	7	5	Missing	-	7	-	7	5
SS	6	6	6	-	7	-	6	5
STA	5	6	5	-	6	-	6	5
BoP	4	3	3	-	4	-	5	3
Avg	5.5	5	4.6	-	6	-	6	4.5
Std	1.12	1.22	1.25	-	1.22	-	0.71	0.87
Process expert-evaluations								
PE1	3	3	4	-	6	-	5	5

3.3. *Third Crew*

During data collection for the Third Crew, the BOP frequently adopted a relaxed posture, often leaning back in the chair, which placed them outside the tracking range and resulted in very low recorded OD viewing time in Synopticon. Overall, the Third Crew's self-evaluations (Table 5) aligned more closely with the recorded viewing times (Figure 5), with higher reported usage corresponding to longer observed durations. This crew was also the only one to report lower OD usage scores in scenarios where the OD was hidden. For the RO and STA, slightly higher disruption scores were reported during the second OD removal (Run 213), which corresponded with increased glancing toward the OD after it was removed.

Two process experts evaluated the Third Crew, but their assessments were inconsistent. PE1 reported the highest OD usage for Run 211, while Synopticon data aligned more closely with PE2's evaluation and the crew's self-assessment. In Run 213, substantial glancing toward the OD after its removal by the RO and STA was reflected in PE2's disruption rating, whereas PE1 reported a much lower score. This discrepancy may be due to fatigue leading to an accidental reversal of the

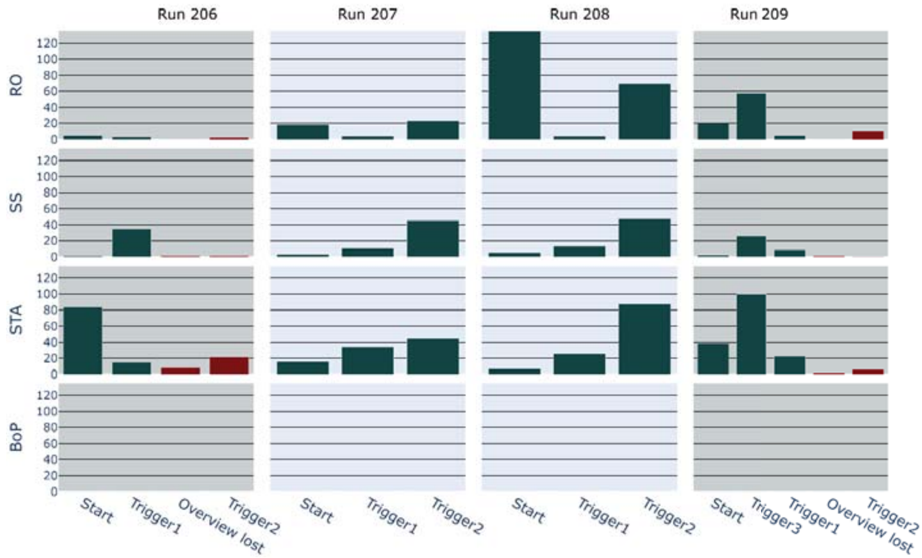

Fig. 4. The total time(s) the RO, SS, STA, and BoP of the Second crew spent looking at the Overview Display

rating scale by PE1.

Additionally, the SS in this crew had access to the OD on a third workstation screen in addition to the large display. Viewing time from this screen was not included in the present analysis but will be considered in future work.

Table 6. Third crew self-evaluation and process expert-evaluation. "U." is short for usage and "D." is short for disruption.

Role	Run 211		Run 212		Run 213		Run 214	
	U.	D.	U.	D.	U.	D.	U.	D.
Crew self-evaluation								
RO	2	5	6	-	3	6	6	-
SS	3	5	6	-	3	6	6	-
STA	2	4	3	-	2	5	2	-
BoP	5	5	4	-	3	3	4	-
Avg	3	4.75	4.75	-	2.75	5	4.5	-
Std	1.22	0.43	1.30	-	0.43	1.22	1.66	-
Process expert-evaluations								
PE1	6	2	5	-	5	1	5	-
PE2	4	3	5	-	5	6	5	-

4. Discussions

4.1. *Regarding the research questions*

Head-orientation data alone lacks sufficient context to yield meaningful insights and must therefore be interpreted in relation to the operational environment. Synopticon's virtual twin capabilities enable head-orientation data to be projected into a digital twin of the control room, allowing automatic annotation of gaze targets by identifying which process screen intersects with the projected head-orientation vector. This enables real-time statistical analysis of operator interaction patterns with control room displays.

To support meaningful interpretation, head-orientation data must be linked to contextual information. In this study, scenario events were recorded to allow interpretation of head-orientation data across different stages of each scenario. A basic statistical analysis of aggregated display usage was conducted, with a particular focus on overview display (OD) usage. The results revealed patterns comparable to both crew self-evaluations and process expert assessments, while also capturing additional behaviors—such as continued glancing toward the OD after it was hidden—highlighting the potential of biometric-driven analysis as a complementary, objective source of information alongside traditional evaluation methods.

A key limitation of head-orientation data is its reliance on human interpretation, which requires

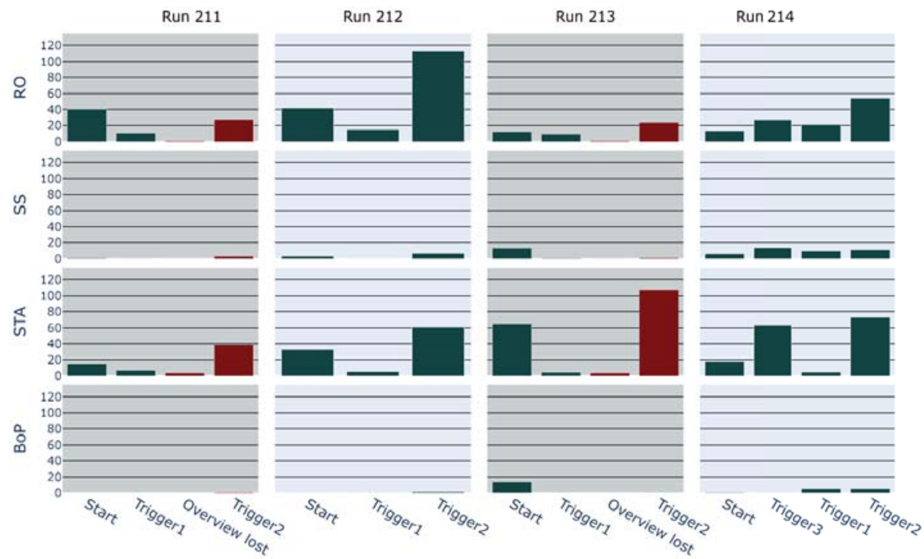

Fig. 5. The total time(s) the RO, SS, STA, and BoP of the Third crew spent looking at the Overview Display

specialized expertise from process experts and subjective self-evaluation. Addressing this limitation requires the development of a conceptual framework for interpreting display usage patterns. Further research is needed to collect data across a wider range of conditions and at a larger scale. Moreover, data analysis must be conducted collaboratively with process experts and human factors specialists to ensure valid interpretation. Such coordinated efforts between human factors experts, process experts, and data scientists are essential to bridge the gap between human interpretation and automated analysis, and to enable more effective extraction of insights from head-orientation data and other biometric signals.

4.2. *Technical limitations*

The use of a single webcam per operator workstation limited the tracking range, resulting in reduced tracking accuracy when operators attended to displays located farther from the camera. In some cases, operators moved completely outside the tracking range, preventing reliable data capture. In addition, the virtual twin of the environment was static, meaning changes such as adjustments to workstation height or screen positioning were not reflected in the model, reducing spatial accuracy.

One approach to mitigating erroneous data is to incorporate confidence scores into the tracking output, allowing low-confidence samples to be treated differently during analysis. Tracking coverage could also be improved by installing multiple cameras at each workstation and selecting the camera with the highest confidence for analysis.

Finally, recent advances in computer vision and deep learning offer alternative approaches for collecting head-orientation and display-attention data through direct visual analysis. Exploring such methods could eliminate the need for explicit environment modeling, thereby reducing setup complexity and study overhead.

5. Conclusions and Future Works

This study explored the potential of biometric and data-driven methods for analyzing human performance by applying two complementary approaches. The first relied on traditional evaluation methods, including expert assessments, operator self-evaluations, and direct observation. The second employed a biometric approach, using head-orientation tracking combined with a digital twin of the control room to quantify the time operators spent viewing the overview display (OD). The biometric approach was intended to complement traditional methods, which also served as a benchmark for interpreting the biometric findings.

Head-tracking data were analyzed in relation to both crew self-evaluations and process expert (PE) assessments. The results revealed discrepancies between self-assessments and expert evaluations, particularly with respect to disruption scores. Overall, the PE assessments showed closer alignment with the biometric data. These differences suggest that crew members and experts may interpret concepts such as “usage” and “disruption” differently.

While PEs appeared to base their evaluations largely on observable visual engagement with the OD, crew members may have focused more on the perceived usefulness of the information obtained from it. This variation in interpretation led to inconsistencies across roles and crews that were not always reflected in the recorded viewing time. In this context, biometric data collected through Synopticon provides an objective empirical basis for understanding operator behavior, helping to reduce subjectivity and support more consistent evaluations across individuals. A limitation of the traditional expert evaluations was that ratings were assigned at the crew level, potentially obscuring role-specific differences.

In addition, head-tracking data revealed behavioral patterns not captured by either expert assessments or self-evaluations. Notably, operators continued to glance toward the OD even after it was removed, suggesting a habitual or subconscious reliance on its presence that may not be fully reflected in subjective reports. These findings highlight the value of biometric and data-driven methods as a complement to traditional evaluation techniques, offering more objective, continuous, and scalable insights into human–system interaction in complex and safety-critical environments.

Overall, the results indicate that biometric data can provide objective and scalable support for human performance evaluation. With further refinement of metrics and analysis methods in collaboration with human factors specialists, such approaches have the potential to become valuable tools for human factors research and human performance assessment.

Acknowledgement

This research was supported by the OECD Nuclear Energy Agency (NEA) Halden Human–Technology–Organization (HTO) Project. We thank the Programme Review Group of the Halden HTO Project for their constructive comments and review.

References

Braseth, A. O. and A. Fernandes (2024). Overview of Displays for nuclear control rooms: a good practices study. In *Human Factors in Software and Systems Engineering*, Volume 140. AHFE Open Acces.

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

Fernandes, A., A. O. Braseth, and R. McDonald (2024). Measuring nuclear control room operators’ performance in the context of overview displays. In *Human Factors in Software and Systems Engineering*, Volume 140. AHFE Open Acces.

Fernandes, A., A. O. Braseth, R. McDonald, and M. Eitrheim (2020). Exploring the effects of large screen overview displays in a nuclear control room setting. In *Engineering Psychology and Cognitive Ergonomics. Cognition and Design*, Cham, pp. 18–29. Springer International Publishing.

Halik, S. Opentrack: Open source head tracking software. Version accessed 2024.

Härefors, E. (2008). *Use of Large Screen Displays in Nuclear Control Room*.

Langstrand, J.-P., H. T. Nguyen, and M. Hildebrandt (2018, September). Synopticon: Sensor Fusion for Real-Time Gaze Detection and Analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 62*(1), 311–315. Publisher: SAGE Publications Inc.