

Advantages and Disadvantages of Physiological Assessment for Next Generation Control Room Design

Joint 8th IEEE HFPP / 13th HPRCT

Tuan Q. Tran
Ronald L. Boring
Donald D. Dudenhoeffer
Bruce P. Hallbert
M. David Keller
Tessa M. Anderson

August 2007

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint should not be cited or reproduced without permission of the author. This document 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, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights. The views expressed in this paper are not necessarily those of the United States Government or the sponsoring agency.

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



Advantages and Disadvantages of Physiological Assessment For Next Generation Control Room Design

Tuan Q. Tran⁽¹⁾, Ronald L. Boring⁽¹⁾, Donald D. Dudenhoeffer⁽¹⁾, Bruce P. Hallbert⁽¹⁾, M.David Keller⁽²⁾, and Tessa M. Anderson⁽³⁾

⁽¹⁾Human Factors and I&C Department, Idaho National Laboratory, Idaho Falls, ID, USA.

⁽²⁾Department of Psychology, New Mexico State University, Las Cruces, NM, USA.

⁽³⁾Department of Psychology, Idaho State University, Pocatello, ID, USA

Email: {tuan.tran, ronald.boring, donald.dudenhoeffer, bruce.hallbert@inl.gov; mdavidk@nmsu.edu; obortess@isu.edu}

Abstract - We propose using non-obtrusive physiological assessment (e.g., eye tracking,) to assess human information processing errors (e.g., loss of vigilance) and limitations (e.g., workload) for advanced energy systems early in the design process. This physiological approach for assessing risk will circumvent many limitations of current risk methodologies such as subjective rating (e.g., rater's biases) and performance modeling (e.g., risk assessment is scripted and is based upon the individual modeler's judgment). Key uses will be to evaluate (early in the design process) novel control room equipment and configurations as well as newly developed automated systems that will inevitably place a high information load on operators. The physiological risk assessment tool will allow better precision in pinpointing problematic design issues and will provide a "real-time" assessment of risk. Furthermore, this physiological approach would extend the state-of-the-art of human reliability methods from a "static" measure to more "dynamic." This paper will discuss a broad range of the current popular online performance gauges as well as its advantages and disadvantages for use in next generation control room.

I. INTRODUCTION

Next generation advanced energy systems (e.g., nuclear power plants, air traffic control, system surveillance rooms) have been conceptualized to be more computerized, intelligent, integrated, and automated [1]. The operator's role in such advanced systems will be more supervisory control with emphasis on monitoring (i.e., checking the system status), detecting (i.e., recognizing that something is not operating normally in the system), and problem-solving (i.e., resolve the detected anomaly). Similar to control room settings, many military technical systems will be heavily integrated and automated. For example, the U.S. Army conception of their Future Force Warrior (FFW) as an integrated "system of systems" in which soldiers will be outfitted with advanced communication, global-positioning, and tactical systems for unsurpassed "survivability and responsiveness" [2].

These technological advances whether in the control room or wearable system suits will inevitably present new information processing demands on the human operator leading to a high-level of workload. High workload is known to induce stress as well as fatigue, which can severely diminish one's ability to perceive,

recognize, and respond to emergency or unanticipated events and, thus, can place both the operator and system at risk. Traditional approaches in human factors to alleviate heavy workload have focused primarily in using subjective workload assessment or performance modeling approaches early in the design process to identify performance bottlenecks. Design mitigations are then implemented to improve human performance [3].

In this paper, we argue that traditional approaches (i.e., subjective workload assessment and performance modeling) are limited due to their subjective vulnerability (e.g., confirmation biases). Instead, this paper proposes the use of physiological assessment (e.g., EEG, ECG) to assess individual's workload early in the design process. Moreover, we will discuss how physiological assessment can enhance current state-of-the-art human reliability analysis (HRA) in understanding and quantifying human performance shaping factors (PSFs) as well as examining the role of dependency.

II. PHYSIOLOGICAL MEASURES TO ASSESS WORKLOAD

Physiological measures have become an increasingly popular approach in assessing workload of newly developed system design. Physiological measurements are not constrained to several limitations that are inherent in self-report questionnaires and performance modeling. In subjective measures, participants quantify their experience of workload [4] either during or immediately after completing a task. Researchers tend to be suspicious of subjective data given the inherent biases (e.g., social desirability) that comes with self-evaluation. Furthermore, subjective measures are limited in their ability to capture "real-time" workload change without provoking a task-interruption nuisance onto the subject. In addition, workload may not be "cognitively penetrable" [5] meaning the person who is self-assessing his or her workload may not have access to the cognitive factors affecting performance. Because of this, subjective measures rely heavily on a subject's memory to remember the task experience therefore reducing accuracy and precision in identifying system design problems.

Performance modeling is also subjective and limited in assessing workload. Typically, in performance modeling, a modeler would first perform a task analysis decomposing the necessary functions needed to successfully complete the task. For each function of the task, the modeler determines (based upon the modeler's understanding of the task) and assigns the appropriate level of workload demand. Thus, over or under decomposing the task can severely distort the validity of the model. Furthermore, the scripted-nature (i.e., task analysis) of performance modeling does not allow an inspection of "emergent" or unexpected behaviors between the human and the system; thus, this approach is very limited in detecting novel design problems.

In the physiological measurement approach, subjects are monitored by an array of physiological sensors, some requiring contact with the subject's body through electrolyte sensors (e.g., electroencephalography, EEG; electrocardiography, ECG) while others are standoff sensors (e.g., eye tracking device embedded in the physical system). Physiological measures permit a more objective workload assessment and can provide "real-time" evaluation, thus allowing the system designer to quickly and accurately identify usability problems as they occur [6]. Moreover, it fosters the assessment of "emergent" behaviors allowing designers to detect novel design issues.

Concerns with many popular physiological measures are its obtrusive nature (e.g., interfering or restricting natural body movement), lack of comfort, inability to be worn for a long duration of time, and interference of bodily fluid (e.g., sweat) with the devices [7]. Thus, we have selectively identified performance gauges in the literature that can be easily built onto or within the control system. Such physiological measures can assess human performance without interfering or restricting the operator's natural movements. Below is a brief list of performance gauges we have identified as being suitable for use in control room.

A. Eye Tracking: Eye-tracking research has become a popular tool in many disciplines in generating understanding of cognitive behaviors such as what information people use, where they look for the information (i.e., scanning patterns), when they use the information, and how much time is needed to process various pieces of information [8]. Furthermore, recent research advancements using eye tracking have shown it to be a good measure of workload by assessing different eye-movement behaviors (e.g., eye-blinking frequency, eye-closure speed, pupil dilation, and gaze duration) [Ikehara]. Depending on the system used, an eye tracker can easily be embedded onto a computer monitor or display with no physical contact to participant, thus not limiting natural behavior.

B. Electroencephalograms (EEGs): EEGs has been demonstrated in the literature to correlate well with vigilance and workload [9]. EEGs measure neurological activities by placing electro-skin-sensors across various positions on the skull (currently the skin-sensors are embedded in a wearable headset). Recent efforts to improve usability issues such as comfort and non-restraint have lead to embedding wireless electro-skin-sensors in glasses or goggles. Participants simply wear the glasses similar to reading glasses [8] and thus do not interfere with an operator's primary area of action.

C. Body Postures: The body posture is hypothesized to be sensitive to changes in workload and is based on the hypothesis that changes in automatic behavior (i.e., postural adjustment) correlates with changes in cognitive activity. For example, a decline in postural stability is correlated with an increase in mental activity [10]. Body posture can be measured by a 16 x 16 pressure sensor array covering the seat and back cushion. Changes in pressure are detected by examining the standard deviations across the 256 sensors over time. Because the sensors are built right into the chair the process of collecting body posture data is unobtrusive.

D. Galvanic skin response (GSR): GSR measures the autonomic nervous system (ANS) activity to gauge a person's arousal level. The relationship between ANS and arousal has been long known in the field of neuropsychology. GSR electrodes can be applied under the person's toes, thus, body movement is not restricted.

E. Inter-Heart Beat Interval (IHBI): IHBI measures the time between successive heartbeats in milliseconds. The IHBI gauges a person's arousal (e.g., short IHBI) and fatigue (e.g., long IHBI) levels. The IHBI uses three electrodes, one placed on the person's right side (below the collarbone), one on the left side (below the left breast), and one on the person's right abdomen. Although IHBI seems to have the most contact with users' bodies, electrodes are placed away from their primary area of actions.

F. Image-Based Technology: Similarly, image-based technology has also been used to assess individuals' cognitive and affective states. Image-based technology rests on the theoretical foundation that a person's states can be exhibited "in certain visual behaviors that are easily observable from changes in facial features such as the eyes, head, and face" [11]. For instance, yawning, sluggishness in facial expression, sagging posture, and body- and face-movement can be used collectively to assess a person's cognitive and affective states. Applications of such technology have been validated as evidence of its increasing use in the field of human-

computer interactions in exploring the speed and ease of interface usability [7] and in the driving industry in exploring a driver’s state of vigilance and fatigue [12]. Image-based technology (as mentioned with the eye tracker) can also be implemented in the system itself and does not inhibit the physical actions of the operator.

III. USING PHYSIOLOGICAL ASSESSMENT TO BETTER EXTEND HRA “STATE-OF-THE-ART”

PSFs and dependency play a critical role in quantifying risk in the majority of HRA methods. Operational experience has noted that environmental contextual demands shape risky behaviors. Thus, identifying the contextual demands that make humans more likely to make errors in their performance is critical in predicting the likelihood of an error. HRA methods traditionally represent these environmental contextual demands as PSFs. Dependency is the extent to which success or failure of a previous action impacts the probability of success or failure on a subsequent action(s). Dependency may be present when knowledge or performance or lack of knowledge from a previous task influences the success or failure of the current task. For example, dependency may be present when failure on Task A causes a crew member extreme anxiety with a resultant increase in the probability of failure above the nominal case on Task B.

In brief, HRA typically begins with the system analyst to discuss the Probability Risk Analysis (PRA) model, in which potential performance issues are identified and defined that may impose significant risk to the system safety. In the error identification stage, the analyst identifies possible existing errors as well as its classification (commission, omission, slip, lapse, mistake, etc). Screening for significant potential errors also occurs at this stage. Errors that have been deemed high risk are then modeled in the error representation stage. The *error representation* process incorporates the identified significant high risk errors into fault trees and event trees to identify various forms of human interactions with system hardware. This allows an inspection of the sequence of possible events (or steps) and its consequences (e.g., core damage) as well as dependencies across different human action sequences and recoveries. Representation also allows insight to co-existing PSFs that can affect human performance across various action sequence stages as well as PSF interactions. The error quantification step then quantifies the identified human errors into error probabilities by taking the nominal Human Error Probability (HEP) for the task, modifying it to reflect PSF influence to obtain basic HEP, and ultimately, conditional HEP when dependencies are considered. The final HEPs are then submitted within PRA to determine the plant’s risk.

Thus, understanding the dynamics of PSF is an important prerequisite in estimating and reducing risk. A shortcoming of current HRA methods are their inability to assess PSF over time (i.e., dynamic). Instead, they typically assess PSFs within a time-slice or event-base (i.e., static) [13]. Because of this, HRA methods may have great difficulty in distinguishing and characterizing PSF’s effect. For instance, the duration of a PSF’s effect is unclear as well as the propensity of the PSF to change. Reference [13] refers to the former as *PSF latency* and the latter as *PSF momentum*. Similarly, reference [13] distinguished the difference between a “static” PSF and a “dynamic” PSF, where the former remains constant across events and the latter naturally changes across events. Finally, there is a high degree of uncertainty how PSFs affect other PSFs even though there is a high acceptance in the HRA community that PSFs have a high correlation among each other. Table 1 summarizes PSF characteristics in a dynamic HRA approach.

TABLE 1. PSF CHARACTERIZATION

PSF Characteristic	Descriptions	Example
Latency	time duration of PSF effect	when an operator experiences high stress, how long will that stress linger until the operator recuperates and returns to a state of homeostasis
Momentum	propensity of PSF to change	stress progressively increases as it approaches the closing window of available time to perform an action
Static	PSF that remains constant across the events or tasks in a scenario	poor fitness for duty due to lack of sleep the night before has a constant effect throughout an event
Dynamic	PSFs evolve naturally across events or tasks in a scenario	during an unusually long work shift, in which fatigue—a negative contributor to fitness for duty—would be expected to set in
Intiator	a PSF effect initiates another PSF effect	low time available window initiate high stress and workload

Furthermore, the current event-based approach by most HRA methods may prevent a clear understanding of dependency (i.e., successful performance of Task B is dependent on performance on Task A). Because of

this event-driven approach, HRA limited understanding of PSFs and dependency [13]. A failure to account for PSFs and dependency can lead to inadequate calculations, inflating or deflating the final HEPs. Because of this, it has proposed that the HRA community consider a “dynamic HRA” approach in which the analyst observes the evolution of PSFs (and dependency) and accounts for their consequences to the outcome of events [13].

The proposed use of physiological assessment can help overcome the shortcomings of HRA quantification by using physiological assessment that can provide cumulative PSF data that are captured dynamically and in “real-time.” Thus, physiological assessment can monitor how a PSF progress changes across time and multiple events as well as examine the dependencies across tasks. Moreover, analysts can time-stamp the physiological assessment results with that of operator’s behavioral performance allowing a finer precision in identifying high risk occurrences during task performance. This ability to pinpoint performance bottlenecks can provide analysts with important information about the challenges of developing next generation control rooms such as updating the crew staffing model, defining shared resources between the automated system and human operators, and evaluating newly designed interfaces of such complex system.

Current limitations of physiological assessment are:

- Precision – overall, the precision of physiological assessment in assessing workload is improving. However, it is still not highly accurate. Accuracy estimates are at approximately 70-80% [14]. Thus, we propose that physiological assessment should be used jointly with traditional behavioral measures (e.g., workload questionnaire).
- Practical Application – physiological assessment technology has not reached the point in which physiological assessment can be used in practical application under actual operational condition due to long-term usability and comfort (e.g., wearing physiological sensors for a complete shift).
- Data Analysis – Because data can be collected in real time, the amount of data collected from the proposed physiological assessment tools might be overwhelming if collected over a long period of time (complete shift). Researchers will have to evaluate what data is important and learn how to evaluate it. Nevertheless, because the potential benefits are encouraging we propose that physiological assessment tools be considered especially for use during the design process of novel interfaces or system configurations.

IV. CONCLUSION

Using non-obtrusive physiological approaches for assessing physiological demands on a user of a newly developed system can overcome some of the limitations of current risk methodologies for the design of next generation control rooms. For example, next generation control rooms will be more sophisticated and complex, presenting many challenges to system designers such as updating crew staffing model, defining shared resources between the automated system and human operators, and evaluating newly design interfaces of such complex system. The proposed use of physiological assessments can assist in providing useful data for system designers who face these upcoming challenges. Using physiological assessment tools will provide more accurate feedback in determining problematic design issues and will provide a “real-time” assessment of risk. Furthermore, this physiological approach would extend the state-of-the-art of human reliability methods from a “static” measure to more “dynamic.” Finally, the use of physiological assessment allows the designers to simulate and assess the risk of different human-system configurations.

V. REFERENCES

- [1] Abbott, K., Corker, K., Figarol, S., Hoekstra, J., Pritchett, A., & Sheridan, T. Panel on human-automation interaction in the next generation air transportation system. *Proceedings of the International Conference on Human-Computer Interaction in Aeronautics*, 2006, pp. 246-247.
- [2] U.S. ARMY. Future force warrior. Available online at: <http://www.natick.army.mil/soldier/wsit/>
- [3] Pavel, M., Wang, G., & Li, K. Augmented cognition: Allocation of attention. Proceedings of the 36th Hawaii International Conference on System Sciences, 2002.
- [4] Tsang, P.S., & Vidulich, M.A. Mental workload and situation awareness. In G. Salvendy (Eds.). *Handbook of Human Factors and Ergonomics*, 3rd Ed. New Jersey: John Wiley & Sons, Inc., 2006.
- [5] Fodor, Jerry A. *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Mass.: MIT Press, 1983.
- [6] Pool, A., & Ball, L.J. Eye tracking in human-computer interaction and usability research: Current status and future prospects. Available at: <http://www.alexpoole.info/academic/Poole&Ball%20EyeTracking.pdf>
- [7] Ikehara, C. S., & Crosby, M. E. Assessing cognitive load with physiological sensors. Proceedings of the 38th Hawaii International Conference on System Sciences, 2005.
- [8] Salvucci, D. Mapping eye movements to cognitive processes. Department of Computer Science, Carnegie Mellon University. An unpublished dissertation.
- [9] Matthews, R., McDonald, N.J., Anumula, H., & Trejo, L. Novel hybrid sensors for nonobtrusive recording of human biopotentials. Proceedings of the 2nd Annual Augmented Cognition, San Francisco, CA, 2006.
- [10] St. John, M., Kobus, D.A., Morrison, J.G. DARPA augmented cognition technical integration experiment (TIE), Technical Report 1905, 2003.

- [11] Ji, Q., Zhu, Z., & Lan, P. Real-time nonintrusive monitoring and prediction of driver fatigue. *IEEE Transportation on Vehicular Technology*, 53(4), 1052-1068, 2004.
- [12] Ji, Q., & Yang, X. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, 8, 357-377, 2002.
- [13] Boring, R.L. Modeling human reliability using MIDAS. *Proceedings of the American Nuclear Society's Fifth International Topical Meeting on Nuclear Plant Instrumentation, Controls, and Human Machine Interface Technology*, Albuquerque, NM, 2006.
- [14] St. John, M., Risser, M.R., & Kobus, D.A. Toward a Usable Closed-Loop Attention Management System: Predicting Vigilance from Minimal Contact Head, Eye, and EEG Measures. *Proceedings of the 2nd Annual Augmented Cognition*, San Francisco, CA, 2006.