



Comparing Eye-Gaze Metrics of Mental Workload in Monitoring Process Plants

Wenyan Huang¹, Yunshu Xu¹, Michael Hildebrandt²,
and Nathan Lau¹✉

¹ Virginia Tech, Blacksburg, USA

{whuang, yunshuxu, Nathan.lau}@vt.edu

² Institute for Energy Technology (IFE), Halden, Norway
michael.hildebrandt@ife.no

Abstract. Continuous, non-invasive workload indicators of operators is an essential component for dynamically assigning appropriate amount of tasks between automation and operators to prevent overload and out-of-the-loop problems in computer-based procedures. This article examines the monitoring task difficulty manipulated by the task type and load, and explores physiological measurements in relation to mental workload. In a within-subject design experiment, forty-five university students performed monitoring tasks in simulated nuclear power plants (NPPs) control room. The performance of monitoring tasks (accuracy), subjective mental workload (NASA Task Load Index), as well as four eye-related physiological indices were measured and analyzed. The results show that as monitoring task difficulty increased, task performance significantly decreased while NASA-TLX, number of fixations, and dwell time significantly increased. Number of fixations and dwell time could be effective, non-invasive continuous indicators of workload for enabling adaptive computer-based procedures.

Keywords: Workload · Eye-related measures · NASA-TLX · Process plants

1 Introduction

Power generation and petrochemical plants rely on procedures extensively [1, 2]. Traditionally, procedures were written on paper, and remain so in many plants. However, there are significant limitations of voluminous paper procedures due to complexity and mental demand with equipment and operations [2–6]. For example, Kontogiannis [3] concluded that paper procedures were inadequate in presenting complex instructions, handling cross-references, tracing suspended or incomplete steps, and monitoring procedural progress. Ockerman and Pritchett [2] also found that paper procedures can be too heavy, delicate, immobile and difficult to follow, preventing operators from executing procedures efficiently.

Computerized procedure systems (CPs) are being developed to resolve the limitations of paper procedures [3, 4, 7–11]. CPs are digital versions of paper procedures that may include additional capabilities to support the operators in executing procedures. These capabilities range from hyperlinks connecting different parts of a

procedure, dynamic displays presenting parameters or controls relevant to procedural steps being executed, automatic checking of preconditions, or automatic execution of control commands [12–14]. CPs can aid process plant operators in reducing operation time and errors while alleviating overall workload. For example, Huang and Hwang [7] showed that average operation time and errors for executing decision and action tasks to deal with alarm signals were significantly reduced with CPs compared to paper procedures.

The benefits of CPs may come with the risk of out-of-the-loop (OOTL) performance problems, the decreased ability of the human operator to intervene or assume manual control when automation fails [10, 15, 16]. Specifically, relieving operators from manually checking pre-conditions and executing control actions to reduce workload may lead them to lose track of procedural steps and misjudge plant state if they haphazardly accept any recommendation of the CPs [12]. Consequently, operators may not abort an inappropriate procedure when the CPs are incorrect, or take inappropriate actions due to wrongly assumed plant state. Taking an inappropriate action might include the control room operator calling field operators to fix equipment that is in an unsafe state because the CPs have changed the equipment setting without operator awareness. For example, when a return-to-normal alarm is reset automatically by CPs, operators may not be aware that such an alarm had sounded, hindering the operator's comprehension and prediction of the plant state [17].

Adaptive automation [18–21] has been proposed as a solution for balancing risk of experiencing OOTL and workload problems. Specifically, real-time assessment of workload can be used to determine appropriate amount of tasks, thereby keeping operators engaged and preventing the OOTL problem [19, 22]. Thus, CPs adaptive to operator workload on monitoring and controlling process plants may reduce the risk of excessive workload and OOTL problem.

As the first step towards developing CPs adaptive to operator workload, this study investigated the use of eye-gaze metrics for assessing operators' workload in monitoring process plants. Specifically, we examined the relationships between eye-gaze measures with respect to a subjective rating scale of workload and task performance. Further, we examined which eye-gaze measures would be most sensitive to manipulation of task difficulty that impacts workload.

1.1 Continuous Indicator of Workload with Eye-Tracking

Eye-tracking can provide nonintrusive, continuous indicators of mental workload experienced by process plant operators, whose tasks involve substantial visual (monitoring) and cognitive processing (diagnosis and self-regulation) [23]. Eye movements are motor responses that are regulated by the cortical and subcortical brain system [24], providing information on the distribution of attention in terms of what stimuli are attended to, for how long, and in what order [25]. Substantial research indicates a correlation between human cognitive workload and eye activity measures, including fixations, saccades and blinks [26–32].

Lin et al. [33] argued that eye fixation and pupil diameter parameters are sensitive indicators to access mental workload. New information is mainly acquired during fixations [24, 34], as suggested by the eye-mind hypothesis postulating that what is

being fixated by the eyes indicates what is being processed in the mind [35]. Only under limited special circumstances can new information be acquired during saccades [36, 37]. Larger number of fixations implies a large magnitude of required information processing and hence higher workload. Longer fixation duration suggests more time spent on interpreting, processing or associating a target with its internalized representation and thus higher workload [33, 38]. Marquart et al. [30] reviewed and concluded that dwell time, the period for a contiguous series of one or more fixations within an area of interest (AOI), can be an indicator of mental workload. Dwell time tends to increase with increasing mental task demands. Pupil diameter usually increases in response to increased difficulty levels of tasks translating to another common indicator of mental workload [6, 26, 39, 40].

1.2 Overview of This Study

Empirical research on eye-tracking for workload assessment in process control appears insufficient for developing adaptive CPs. For this reason, we conducted an experiment involving human participants performing monitoring tasks to provide further empirical evidence on whether eye-gaze measures can be effective, continuous workload indicators.

For monitoring process plants, we hypothesize that eye-gaze measures would be able to reveal the types of monitoring tasks imposing different workload. Also, these eye-gaze measures would reveal difference in task load for the same type of tasks. We further hypothesize that the NASA TLX, a validated subjective workload measure [41, 42], would correlate positively with number of fixations, average fixation duration, dwell time and pupil diameter.

2 Method

2.1 Participants

This experiment recruited 45 Virginia Tech graduate and undergraduate students (age range 18–26, 29 females and 28 males). All participants had normal or corrected-to-normal vision. Participants were compensated \$10/h for about 1.5 h of their time.

2.2 Experimental Apparatus

The experiment was conducted in a quiet room, with a computer workstation presenting an overview display of a nuclear power plant on a 24" LED monitor with 1920×1200 resolution at 60 Hz. Further, the computer workstation collected eye-gaze and heart rate data with the following equipment:

1. SensoMotoric Instruments (SMI) Remote Eye-tracking Device (REDn) recorded eye-gaze data at 60 Hz sampling rate. The REDn sensor was physically attached to the bottom of the monitor and connected to the computer workstation that had the SMI iVIEW software installed for data collection.

2. Shimmer3 ECG Sensors (ECG) recorded electrocardiogram (ECG), the pathway of electrical impulses through the heart muscle, sampling at 1000 Hz. The ECG was wirelessly connected to the computer workstation through Bluetooth. ECG analysis is beyond the scope of this paper.

2.3 Experimental Manipulation

The participant tasks were to identify parameters deemed out-of-range on an overview display of a fictional nuclear power plant (Fig. 1). The contents of the overview display consisted of tanks, pumps, heat exchangers and valves associated with various process parameters such as level (i.e., %), flow rate (i.e., gpm), temperature (i.e., °C), and pressure (i.e., psig & KPph). The locations of these process parameters were the same for all trials but the values of these process parameters were updated for each trial. The Question Box prompted the participants to complete two types of monitoring tasks.

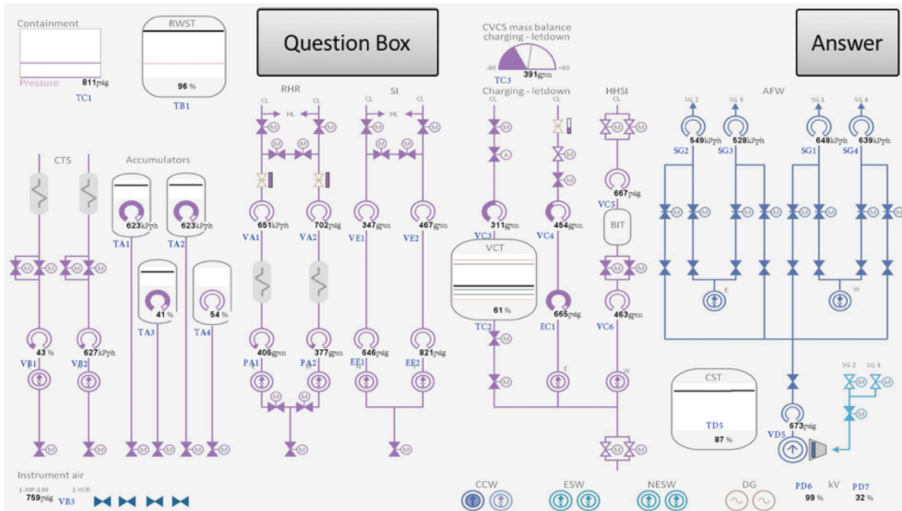


Fig. 1. Nuclear control room monitoring simulation platform presenting the overview display and monitoring tasks.

The two types of monitoring task were target-driven and series-driven verification of process parameters to represent common activities specified by procedures of industrial plants.

Target-driven verification (Fig. 2 left column). Participants were instructed to check specific targets (e.g. TC3, SG1 or VD5) per question or monitoring task. The target-driven task included either one or two targets of parameters to represent low and high task load, respectively.

Series-driven verification (see Fig. 2 right column). Participants were instructed to check all values for a specific type of parameters (e.g. gpm, kPph, psig or %) per

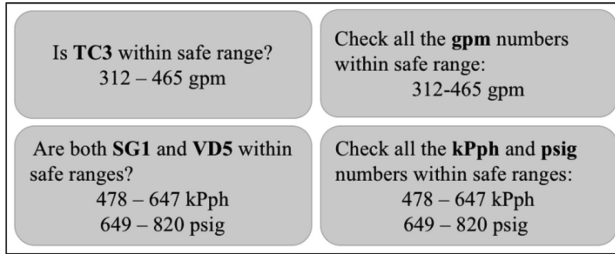


Fig. 2. Examples of question boxes on nuclear control room monitoring simulation, demonstrating manipulation of monitoring task types and task loads.

question. The series-driven task included either one or two series of parameters to represent low and high task load, respectively.

2.4 Procedure

Participants were welcomed with a brief introduction about the study in front of the computer workstation. Then they were asked to give consent and complete a health history questionnaire. The experimenter provided instructions of the control room monitoring task and answered participants' questions.

The participants completed four blocks of control room monitoring tasks for all combination of task type and load conditions: two task types (i.e., target-driven vs series driven) and two task loads (i.e., low vs. high). At the beginning of each block, the participants first completed REDn 9-point eye-gaze calibration. Participants completed a NASA-TLX questionnaire at the end of each four blocks. For each trail of the monitoring task, participants responded by clicking the corresponding out-of-range parameter(s) on the display with a mouse and then clicked the 'Answer' button (see Fig. 1) to submit their responses and proceed to next trial. The experimenter stopped REDn recording at the end of each block.

After completing four blocks of trials, the experimenter helped participants to take off the physiological instruments. Participants were given the opportunity to ask any further questions and \$15 for compensation at departure.

2.5 Experimental Design

The experiment was a 2×2 within-subjects design with two treatments: (1) task type (singular targets or series of targets) and (2) task load (low and high). Four blocks of control room monitoring tasks were assigned in a random order across participants. Each block consisted of 3 min of monitoring tasks. Participants performed tasks at their own pace, leading to different numbers of completed trials in one block.

2.6 Measures

Participants were assessed on three categories of measures: task performance, NASA-TLX, and eye-related measures.

Response Accuracy. The response accuracy was used to assess the task performance. This measure was defined as the percentage of trials for which participant submitted the correct answer by identifying all the out-of-range parameters.

TLX Total Score. The NASA-TLX questionnaire was used to assess the subjective ratings of workload, using a 10-point visual analog scale. This questionnaire is a multidimensional instrument that consists of 6 subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration. The TLX total score was computed by a combination of the six dimensions, resulting in an overall workload scale between 0 and 60.

Eye-Gaze Measures. Number of fixations, fixation duration, dwell time and pupil diameter were used as continuous indicators of workload. Area of interest (AOI) was defined as display area covering the graphic and numerical reading of the parameter(s) that should be monitored in each trial. The AOIs varied between trials depending on the monitoring task type and load. For example, a square was marked as the AOI for the trials with the one-target driven verification task, while eight squares were marked as the AOI for trials with one of the series driven monitoring task. Fixation-based metrics on AOI were extracted to indicate workload. All eye-gaze metrics were computed with SMI BeGaze software. Four metrics were selected for comparison: the total number of fixations on AOIs, average duration of a fixation on AOIs, dwell time (total fixation durations on AOIs), and pupil diameter for fixations on AOIs.

3 Results

The experiment yielded 180 observations (45 participants x 4 experimental blocks), of which twelve were removed due to participants performing the monitoring tasks incorrectly. We further failed to collect NASA TLX for an additional participant. Thus, except for NASA TLX, Pearson-product moment correlation statistics were computed to examine relationships between measures across the 168 observations and two-way analysis of variance (ANOVA) were conducted to examine differences between four experimental conditions. Statistics associated with NASA TLX only contains 167 observations.

Response accuracy was correlated with number of fixations ($r = -0.313$; $p < 0.001$) and dwell time ($r = -0.265$; $p < 0.001$). However, only pupil diameter significantly correlated with NASA TLX ($r = -0.186$; $p < 0.05$). Between eye-gaze measures, dwell time significantly correlated with all three other eye-gaze measures, including number of fixations ($r = 0.877$; $p < 0.001$), fixation duration (0.458; $p < 0.001$), and pupil diameter ($r = 0.173$; $p < 0.05$) (Table 1).

Experimental effects on response accuracy and TLX total score were examined with the nonparametric Kruskal-Wallis rank sum test because their error residuals were not normally distributed.

The nonparametric test results confirmed the hypotheses in revealing that series-driven monitoring tasks significantly hindered response accuracy ($\chi^2(1) = 31.864$, $p < 0.001$, $N = 168$). Further, the nonparametric test also revealed that task load marginally decreased response accuracy ($\chi^2(1) = 2.854$, $p = 0.091$, $N = 168$) and

Table 1. Correlation matrix of the five measures

	Response accuracy	TLX total score	Number of fixations	Avg. fixation duration	Dwell time	Pupil diameter
Response accuracy	1	-0.043	-0.313***	0.034	-0.265***	-0.003
TLX total score	-0.043	1	0.012	-0.152	-0.052	0.186*
Number of fixations	-0.313***	0.012	1	0.056	0.877***	0.068
Avg. fixation duration	0.034	-0.152	0.056	1	0.458***	0.290
Dwell time	-0.265***	-0.052	0.877***	0.458***	1	0.173*
Pupil diameter	-0.003	0.186*	0.068	0.290	0.173*	1

*** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$

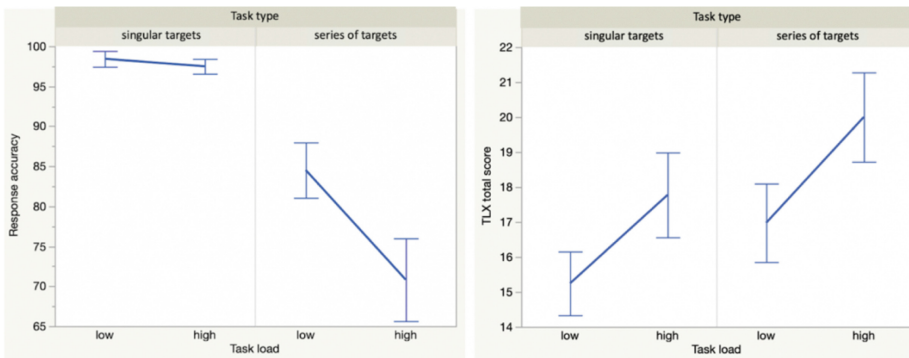


Fig. 3. Mean and standard error plots of response accuracy (left) and subjective workload rating score (right) for each combination of task type and task load.

significantly increased subjective workload ($\chi^2(1) = 4.748$, $p = 0.029$, $N = 167$) (Fig. 3).

All eye-related measures were analyzed in two-way ANOVAs. The main effect of task type was also significant on the number of fixations ($F(1, 159) = 110.634$, $p < 0.001$) and dwell time ($F(1, 159) = 49.117$, $p < 0.001$). Similarly, the main effect of task load was significant on both number of fixations ($F(1, 159) = 31.5963$,

$p < 0.001$) and dwell time ($F(1, 159) = 14.320, p < 0.001$). Furthermore, the interaction effect of task type and load was significant on both number of fixations ($F(1, 159) = 11.997, p < 0.001$) and dwell time ($F(1, 159) = 6.162, p = 0.014$). In other words, increased task load had significantly more impact for performing series-driven than target driven monitoring tasks. However, average duration per fixation and pupil diameter did not reveal any significant effect (Fig. 4).

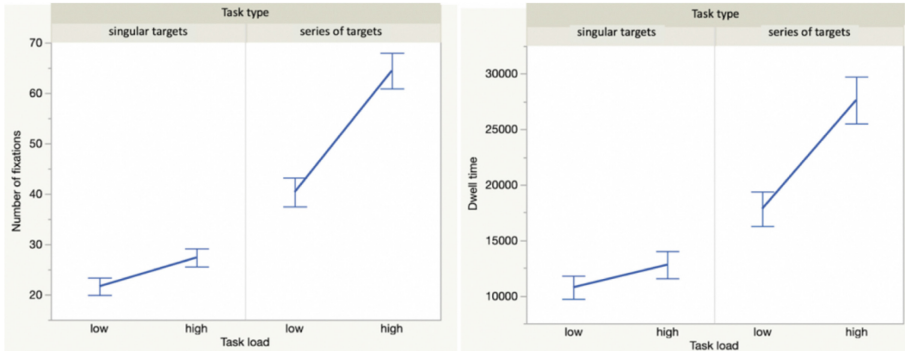


Fig. 4. Mean and standard error plot of the number of fixations (left) and dwell time (right) for each combination of task type and task load.

4 Discussion

The significant main effect of task type and load on response accuracy and NASA TLX confirmed our hypotheses, indicating that the two experimental manipulations were effective at manipulating workload. Thus, we can confidently interpret the eye-gaze metrics with respect to the response accuracy and NASA TLX measures. The number of fixations on AOIs and dwell time on AOIs showed the same main effects as response accuracy and NASA TLX, indicating the sensitivity of these two eye-gaze measures to experimental manipulations. However, these two measures were not sensitive to subjective workload because they did not correlate with NASA TLX. Pupil diameter failed to reveal any significant effects but correlated with NASA TLX. In other words, pupil diameter was sensitive to subjective workload but not to the effect of task type and load. The average fixation duration per AOI did not appear to be a sensitive measure, failing to reveal any significant correlations and experimental effects.

The results of this experiment illustrate how careful consideration is needed in selecting eye-gaze metrics for indicating workload in monitoring process plants. None of the eye-gaze measures showed significance to both correlation with NASA TLX and experimental manipulations (i.e., task type and load), so there is no clear contender of an eye-tracking measure for indicating workload. (Dwell time and number of fixations only showed significant correlation with response accuracy.)

These eye-gaze results must be also be interpreted with respect to the monitoring tasks designed for this experiment. Specifically, there are more targets for the series-driven than target-driven task type, and for the high than low task load. Thus, the

number of fixations may be higher inherently due to the task characteristic of more targets rather than higher mental workload. For this reason, dwell time on AOIs might be a more robust indicator than number of fixations because dwell time is bounded by the allotted time for the block (i.e., 3 min). In the context of this study, the issue on number of targets probably does not present a significant problem for two reasons. First, having more targets is intrinsically linked to the demand of the monitoring tasks, so the results should still be representative for monitoring process plants. Second, dwell time revealed the same experimental effects as NASA TLX, lending empirical support that the experimental manipulations affect dwell time and mental workload similarly.

Another notable result is the weak, positive correlation between dwell time and response accuracy, indicating that eye-gaze behaviors could contribute to task performance. Thus, dwell time might also offer modest and continuous indication of operator engagement with system operations.

The overall empirical results indicated that dwell time could be an effective alternative to NASA TLX as a workload indicator in the context of monitoring parameters prescribed by procedures. Dwell time can be collected in a less invasive manner than NASA TLX while providing continuous indication of workload and engagement. That is, NASA TLX requires interruption of the work tasks whereas the remote eye-tracker can continuously estimate dwell time without any interference. NASA TLX might simply reflect operator initial and final impression of the given task as opposed to their level of cognitive processing as they perform the given task. Once again, the generalization of the study results is limited to task load driven by number of targets.

This research represents the early effort to integrate the concept of adaptive automation into CPs. The results of this experiment highlight the potential of various eye-gaze measures as a continuous indicator of workload to support adaptive features in CPs for control room operators monitoring process plants. Valid and reliable eye-gaze metrics of workload can support continuous, unobtrusive assessment of workload as well as adaptive aiding for display design in the main control room. Future work can examine the use of regression-based machine learning methods on multiple eye-gaze measures to indicate workload while monitoring process plants (see [43]).

References

1. Ludwig, E.E.: *Applied Process Design for Chemical and Petrochemical Plants*, vol. 2. Gulf Professional Publishing, Houston (1997)
2. Ockerman, J., Pritchett, A.: A review and reappraisal of task guidance: aiding workers in procedure following. *Int. J. Cogn. Ergon.* **4**(3), 191–212 (2000)
3. Kontogiannis, T.: Applying information technology to the presentation of emergency operating procedures: implications for usability criteria. *Behav. Inf. Technol.* **18**(4), 261–276 (1999)
4. Niwa, Y., Hollnagel, E., Green, M.: Guidelines for computerized presentation of emergency operating procedures. *Nucl. Eng. Des.* **167**(2), 113–127 (1996)
5. Park, J., Jung, W.: The operators' non-compliance behavior to conduct emergency operating procedures—comparing with the work experience and the complexity of procedural steps. *Reliab. Eng. Syst. Saf.* **82**(2), 115–131 (2003)

6. Xu, S., et al.: An ergonomics study of computerized emergency operating procedures: presentation style, task complexity, and training level. *Reliab. Eng. Syst. Saf.* **93**(10), 1500–1511 (2008)
7. Huang, F.H., Hwang, S.L.: Experimental studies of computerized procedures and team size in nuclear power plant operations. *Nucl. Eng. Des.* **239**(2), 373–380 (2009)
8. Hwang, F.H., Hwang, S.L.: Design and evaluation of computerized operating procedures in nuclear power plants. *Ergonomics* **46**(1–3), 271–284 (2003)
9. Landry, S., Jacko, J.: Improving pilot procedure following using displays of procedure context. *Int. J. Appl. Aviat. Stud.* **6**(1), 47–70 (2006)
10. Lee, S.J., Seong, P.H.: Development of an integrated decision support system to aid cognitive activities of operators. *Nucl. Eng. Technol.* **39**(6), 703 (2007)
11. Lin, C.J., Hsieh, T.L., Yang, C.W., Huang, R.J.: The impact of computer-based procedures on team performance, communication, and situation awareness. *Int. J. Ind. Ergon.* **51**, 21–29 (2016)
12. Naser, J.: Computerized procedures: design and implementation guidance for procedures, associated automation and soft controls, vol. 1015313, Draft Report. EPRI (2007)
13. Yang, C.W., Yang, L.C., Cheng, T.C., Jou, Y.T., Chiou, S.W.: Assessing mental workload and situation awareness in the evaluation of computerized procedures in the main control room. *Nucl. Eng. Des.* **250**, 713–719 (2012)
14. Fink, R.T., Killian, C.D., Hanes, L.F., Naser, J.A.: Guidelines for the design and implementation of computerized procedures. *Nucl. News* **52**(3), 85 (2009)
15. Kaber, D.B., Endsley, M.R.: The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theor. Issues Ergon. Sci.* **5**(2), 113–153 (2004)
16. Lin, C.J., Yenn, T.C., Yang, C.W.: Automation design in advanced control rooms of the modernized nuclear power plants. *Saf. Sci.* **48**(1), 63–71 (2010)
17. Huang, F.H., et al.: Experimental evaluation of human–system interaction on alarm design. *Nucl. Eng. Des.* **237**(3), 308–315 (2007)
18. Byrne, E.A., Parasuraman, R.: Psychophysiology and adaptive automation. *Biol. Psychol.* **42**(3), 249–268 (1996)
19. Kaber, D.B., Riley, J.M.: Adaptive automation of a dynamic control task based on secondary task workload measurement. *Int. J. Cogn. Ergon.* **3**(3), 169–187 (1999)
20. Gevins, A., Smith, M.E.: Neurophysiological measures of cognitive workload during human-computer interaction. *Theor. Issues Ergon. Sci.* **4**(1–2), 113–131 (2003)
21. Haarmann, A., Boucsein, W., Schaefer, F.: Combining electrodermal responses and cardiovascular measures for probing adaptive automation during simulated flight. *Appl. Ergon.* **40**(6), 1026–1040 (2009)
22. Wilson, G.F., Russell, C.A.: Performance enhancement in an uninhabited air vehicle task using psychophysiological determined adaptive aiding. *Hum. Factors: J. Hum. Factors Ergon. Soc.* **49**(6), 1005–1018 (2007)
23. Lau, N., Jamieson, G.A., Skraaning Jr., G.: Situation awareness in process control: a fresh look. In: *Proceedings of the 8th American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation & Control and Human-Machine Interface Technologies (NPIC & HMIT)*, San Diego, CA, USA (2012)
24. Rayner, K.: The 35th Sir Frederick Bartlett Lecture: eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol.* **62**(8), 1457–1506 (2009)
25. Scheiter, K., Van Gog, T.: Using eye tracking in applied research to study and stimulate the processing of information from multi-representational sources. *Appl. Cogn. Psychol.* **23**(9), 1209–1214 (2009)

26. Ahlstrom, U., Friedman-Berg, F.J.: Using eye movement activity as a correlate of cognitive workload. *Int. J. Ind. Ergon.* **36**(7), 623–636 (2006)
27. Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., Babiloni, F.: Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neurosci. Biobehav. Rev.* **44**, 58–75 (2014)
28. Hankins, T.C., Wilson, G.F.: A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. *Aviat. Space Environ. Med.* **69**(4), 360–367 (1998)
29. Kramer, A.K.: Physiological metrics of mental workload: a review of recent progress. In: *Multiple-task Performance*, pp. 279–328 (1991)
30. Marquart, G., Cabrall, C., de Winter, J.: Review of eye-related measures of drivers' mental workload. *Proc. Manuf.* **3**(Suppl. C), 2854–2861 (2015)
31. Ryu, K., Myung, R.: Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *Int. J. Ind. Ergon.* **35**(11), 991–1009 (2005)
32. Wilson, G.F., Russell, C.A.: Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Hum. Factors* **45**(4), 635–644 (2003)
33. Lin, Y., Zhang, W.J., Watson, L.G.: Using eye movement parameters for evaluating human-machine interface frameworks under normal control operation and fault detection situations. *Int. J. Hum. Comput. Stud.* **59**(6), 837–873 (2003)
34. Poole, A., Ball, L.J.: Eye tracking in HCI and usability research. *Encycl. Hum. Comput. Interact.* **1**, 211–219 (2006)
35. Just, M.A., Carpenter, P.A.: Eye fixations and cognitive processes. *Cogn. Psychol.* **8**(4), 441–480 (1976)
36. Matin, E.: Saccadic suppression: a review and an analysis. *Psychol. Bull.* **81**(12), 899 (1974)
37. Campbell, F.W., Wurtz, R.H.: Saccadic omission: why we do not see a grey-out during a saccadic eye movement. *Vis. Res.* **18**(10), 1297–1303 (1978)
38. Loftus, G.R.: Eye fixations and recognition memory for pictures. *Cogn. Psychol.* **3**(4), 525–551 (1972)
39. Kovesdi, C.R., Rice, B.C., Bower, G.R., Spielman, Z.A., Hill, R.A., LeBlanc, K.L.: Measuring human performance in simulated nuclear power plant control rooms using eye tracking. Idaho National Lab. (INL), Idaho Falls, ID (United States), INL/EXT-15-37311, November 2015
40. Palinko, O., Kun, A.L., Shyrovkov, A., Heeman, P.: Estimating cognitive load using remote eye tracking in a driving simulator. In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, pp. 141–144 (2010)
41. Yurko, Y.Y., Scerbo, M.W., Prabhu, A.S., Acker, C.E., Stefanidis, D.: Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. *Simul. Healthc.* **5**(5), 267–271 (2010)
42. Cao, A., Chintamani, K.K., Pandya, A.K., Ellis, R.D.: NASA TLX: software for assessing subjective mental workload. *Behav. Res. Methods* **41**(1), 113–117 (2009)
43. Zhang, X., Mahadevan, S., Lau, N., Weinger, M.B.: Multi-source information fusion to assess control room operator performance. *Reliab. Eng. Syst. Saf.* (2018)