Predicting Driver's Work Performance in Driving Simulator Based on Physiological Indices

Cong Chi Tran[1,2(✉)], Shengyuan Yan[1], Jean Luc Habiyaremye[1], and Yingying Wei[3]

[1] Harbin Engineering University, Harbin 150001, China
[2] Vietnam National University of Forestry, Hanoi 10000, Vietnam
trancongchi_bk@yahoo.com, yanshengyuan@hrbeu.edu.cn,
habijealuc@yahoo.fr
[3] East University of Heilongjiang, Harbin 150086, China
weiyingying2007@126.com

Abstract. Developing an early warning model based on mental workload (MWL) to predict the driver's performance is critical and helpful, especially for new or less experienced drivers. This study aims to investigate the correlation between human's MWL and work performance and develop the predictive model in the driving task using driving simulator. The performance measure (number of errors), subjective rating (NASA Task Load Index) as well as six physiological indices were assessed and measured. Additionally, the group method of data handling (GMDH) was used to establish the work performance model. The results indicate that different complexity levels of driving task have a significant effect on the driver's performance, and the predictive performance model integrates different physiological measures shows the validity of the proposed model is well with $R^2 = 0.781$. The proposed model is expected to provide a reference value of their work performance by giving physiological indices. Based on this model, the driving lesson plans will be proposed to sustain the appropriate MWL as well as improve work performance.

Keywords: Driving simulator · Work performance · Predictive model

1 Introduction

Reducing road accident is an important issue. Contributing factors to crashes are commonly classified as human, vehicle or roadway and environmental [1]. Driving is often heavy mental workload (MWL) tasks, because in order to prevent accidents, drivers of must continually acquire and process much information from their eyes, ears, and other sensory organs. The information includes the movements of other vehicles and pedestrians, road signs and traffic signals, and various situations and changes in the road environment. These incidents require a lot of driver's attention. Human errors such as misperception, information processing errors, and slow decision making are frequently identified as major reasons can cause the accidents [2]. Therefore, improving driver's MWL could be helpful in improving driver performance and reducing the number of accidents.

For most drivers, both excessive and low MWL could degrade their performance, and furthermore, may affect the safety of the driver and others. Because of when the situation is low-demanding (e.g., in long and boring roads), or conversely when the situation is high demanding (e.g., in the city with much information to process), drivers are overloaded with an increase of workload leading to performance impairments [3, 4]. Only with an appropriate level of MWL, the drivers can perform the right tasks. Therefore, for the purpose of driver's safety, developing an early warning model based on MWL to predict the driver's performance is critical and helpful, especially for new drivers or little experience in driver training.

MWL refers to the portion of operator information processing capacity or resources that is actually required to meet system demands [5]. The MWL is induced not only by cognitive demands of the tasks but also by other factors, such as stress, fatigue and the level of motivation [6, 7]. In many studies of driving task, the MWL was measured by subjective measures, such as NASA task load index (NASA-TLX) [8–10]. However, a major limitation of subjective measures is that they can only assess the overall experience of the workload of driving but cannot reflect changes in workload during the execution of the task. Also, rating scale results also can be affected by characteristics of respondents, like biases, response sets, errors and protest attitudes [11, 12]. Thus, the continuous and objective measures (e.g. physiological signal) to assess the MWL in addition to evaluating the overall workload in driving tasks is necessary [13].

Recently, many driving simulators can measure performance accurately and efficiently, and they are more and more used in driving education tasks. It is commonly accepted that the use of driving simulators presents some advantages over the traditional methods of drive learning because their virtual nature, the risk of damage due to incompetent driving is null [14]. In addition, simulators make it possible to study hazard anticipation and perception by exposing drivers to dangerous driving tasks, which is an ethically challenging endeavor in real vehicles [15], and also offers an opportunity to learn from mistakes in a forgiving environment [16, 17]. In this study, we conducted an experiment to simulate the car driving tasks to assess the relation between work performance, subjective rating, and physiological indices for new drivers. According to these relationships, the study developed a predictive model by using the group method of data handling (GMDH) to integrate all physiological indices into a synthesized index. The physiological indices used in this study were the eye activities (pupil dilation, blink rate, blink duration, fixation duration) and cardiac activities (heart rate, heart rate variability). The performance of the task was measured by the number of errors, and the subjective rating was rated by the NASA-TLX questionnaire.

2 Methodology

2.1 Participants

Twenty-six male engineering students voluntary, age 19.2 ± 1.1 years (mean $\pm$ SD) participated in the experiment. They have very little (less than two months) or no driving experience. They have normal eyesight (normal or corrected to normal vision in both eyes) and good health. For ensuring the objectivity of experimental electrocardiography

(ECG) data, all participants were asked to refrain from caffeine, alcohol, tobacco, and drug six hours before the experiment. All participants completed and signed an informed consent form approved by the university and were compensated with extra credit in extracurricular activities in their course.

2.2 Apparatus

A driving simulator (Keteng steering wheel and City car driving software version 1.4.1) was used in this study. The city car driving is a car simulator, designed to help users feel the car driving in a city or a country in different conditions. Special stress in the City car driving simulator has been laid on the variety of road situations and realistic car driving.

IView X head mounted eye-tracking device (SensoMotoric Instruments) was used to record participants' dominant eye movements. Software configuration has the video recording and the BeGaze version 3.0 eye movement data analysis, sampling rate 50/60 Hz (optional 200 Hz), tracking resolution, pupil/Corneal reflection <0.1° (typical) and gaze position accuracy <0.5°–1.0° (typical). ANSWatch TS0411 was used to measure the heart rate (HR) and HRV (heart rate variability) data.

2.3 Work Performance and Mental Workload Measures

Various MWL measurements have been proposed, and these measurements could be divided into three categories: performance measure, physiological measures and subjective ratings [18]. Performance measures can be classified into many categories such as accuracy, task time, worst-case performance, etc. [19]. In this study, the number of errors of driving task was calculated because of some reasons: (1) driving errors to involve risky behaviors that we need to understand to prevent accidents and fatalities [20]. In addition, many studies had shown that the number of errors has a sensitive to differences in the visual environment [21, 22]. (2) in the City car driving software, all driving errors include such as didn't follow the speed limit, driving on the red light, no turn signal when changing the lane, accident and so forth are displayed when driving and counted after finish the task.

Subjective ratings are designed to collect the opinions from the operators about the MWL they experience using rating scales. With the low cost and the ease of administration, as well as adaptability, have been found highly useful in driving tasks [20, 23]. In this study, subjective ratings NASA-TLX [24] was used to evaluate the driver's MWL because of there are many studies successfully applied to measure MWL in the driving [8, 9]. NASA-TLX is a multi-dimensional rating scale using six dimensions of workload to provide diagnostic information about the nature and relative contribution of each dimension in influencing overall operator workload. Six dimensions to assess MWL including mental demand (MD), physical demand (PD), temporal demand (TD), own performance (OP), effort (EF) and frustration (FR).

Physiological measures can be further divided into central nervous system measures and peripheral nervous system measures [25]. These methods do not require the user to generate overt responses, they allow a direct and continuous measurement of the current

workload level, and they have high temporal sensitivity and can thus detect short periods of elevated workload [26]. Although central nervous system measures (i.e. electrocardiogram) has high reliability in measurement of driver's MWL [13], the applicability of these measures is limited due to the expensive instruments so it was not suitable to the conditions of this experiment. Therefore, the central nervous system measures were not used in this study.

Eye activity is a technique that captures eye behavior in response to a visual stimulus, and this technique has become a widely used method to analyze human behavior [27]. Eye response components that have been used as MWL measures include pupil dilation, blink rate, blink duration and fixations. Human pupil dilation may be used as a measure of the psychological load because it is related to the amount of cognitive control, attention, and cognitive processing required for a given task [28]. It also has been previously shown to correlate with the cognitive workload, whereby increased frequency of dilation is associated with increased degree of difficulty of a task [29]. In the driving study, pupil dilation was able to reflect the load required by tasks [30], and it would measure the average arousal underlying the cognitive tasks [31]. The blink of the eye, the rapid closing, and reopening of the eyelid is believed to be an indicator of both fatigue and workload. It is well known that eye blink rate is a good indicator of fatigue. Blink rate has been investigated in a series of driver and workload studies with mixed results attributable to the distinction between mental and visual workload [31]. They suggested that blink rate is affected by both MWL and visual demand, which act in opposition to each other, the former leading to blink rate increase, the latter to blink rate decrease. Besides blink rate, blink duration has been shown to be affected by visual task demand. Blink duration has been shown to decrease with increases in MWL. The studies mentioned in Kramer's review all found shorter blink durations for increasing task demands (both mental and visual) [32]. Some studies show that blink duration is a sensitive and reliable indicator of driver visual workload [8, 33]. Eye fixation duration is also extensively used measures and is believed to increase with increasing mental task demands [34]. Recently, fixation duration and the number of fixations have also been investigated in a series of studies about driver hazard perception and they found that increased fixation durations during hazardous moments, indicating increased MWL [20].

The heart rate (HR) and heart rate variability (HRV) potentially offer objective, continuous, and nonintrusive measures of human operator's MWL [26]. Numerous studies show that HR reflects the interaction of low MWL and fatigue during driving [35, 36]. In addition to basic HR, there has also been growing interest in various measures of HRV. Spectral analysis of HRV enables investigators to decompose HRV into components associated with different biological mechanisms, such as the sympathetic/parasympathetic ratio or the low frequency power/high frequency power (LF/HF) ratio, the mean inter-beat (RR), the standard deviation of normal RR intervals (SDNN), etc. The SDNN reflects the level of sympathetic activity about parasympathetic activity and has been found to increase with an increase in the level of MWL [13, 25].

2.4   Experimental Task

There were three levels of task complexity in this experiment such as high, medium and low. Special stress in the City car driving simulator has been laid on the variety of road situations and realistic car driving. The condition setting of task shown in Table 1.

Table 1.   Experiment task setting

Condition setting	Task complexity		
	Low	Medium	High
Type and gears	Manual, 6	Manual, 6	Manual, 6
Steering wheel	Left	Left	Left
Vehicular traffic density	10%	20%	40%
Traffic behavior	Quiet traffic	Usual traffic	Fast-moving traffic
Pedestrian traffic density	20%	20%	40%
Frequency of assignments	medium	medium	medium
Point limit	10	10	10
Sudden change of lane by a car	Rare	Often	Very often
Sudden stop of a car in the traffic	Rare	Often	Very often
Driving out of a car on oncoming lane	Often	Often	Very often
Time and weather	Daytime/clear	Daytime/clear	Nighttime/clear

2.5   Group Method of Data Handling

Actually, there are many methods used to develop the predictive model such as GMDH, Neural Networks, Logistic regression, Naive Bayes, etc. This study used the GMDH method [37] to establish a prediction model of work performance. This is a widely used neural network methodology which requires no assumptions of the relationship between predictors and responses [38]. The GMDH algorithm has been widely used in various fields, e.g. nuclear power plants [25], Stirling engine design [39], education [40]. This study investigated the relationship between seven physiological indices and work performance on different levels of task complexity.

2.6   Procedure

All participants received about two hours of training. During the training, they were taught how to use the eye tracking equipment, complete the NASA-TLX questionnaire and driving simulator. After that, each participant was received about 30 min to practice by himself on the driving simulator. This practice served the purpose of familiarizing subjects with the simulator and the general feel of the pedals and steering. The practice step would end until the participant was sure that he understood all procedures. The experiment was conducted on the next day.

Before the experiment, the participant took a 20 min rest, and then wore the measurement apparatus and proceeded with system adjustment. The initial physiological indices were acquired as a baseline before the experiment. During experiment, the physiological indices were collected during each phase (level of task complexity), and the NASA-TLX questionnaire was conducted after each phase to evaluate the subjective MWL of different levels of task complexity. Each phase lasted for about 20 min and had 5 min break after each phase. The limitation of driving speed limits in this study was required less 45 km/h.

The scenario included a normal driving environment in the city (2 km of city roads with some stop signs or crossing lights). Each participant was made to test the three level of the task in a randomized order (Fig. 1). They were asked to follow speed limits and to comply with traffic laws throughout the course of the experiment. Three level of workload with high, medium and low of task complexity in this experiment shown in Table 1.

Fig. 1. Driving task in the experiment: (A) Low task; (B) Medium task; (C) High task

3 Results

3.1 Sensitivity with the Workload Level

At alpha level of .05, a MANOVA results showed that there are a statistically significant difference in task levels, $F(16, 136) = 3.52$, $p < .0005$; Wilk's $\Lambda = .50$, partial $\eta^2 = .293$ with the high observed power of 99.1%. Descriptive statistics was presented in Table 2. There were significant differences in almost methods between workload levels in this driving task, however; no significant difference was found in pupil dilation ($p = .574$) and fixation duration ($p = .143$). The number of errors in performance measure showed that the high task has significantly higher error than the low task by almost 23.3% (Tukey HSD $p = .036$). However, there was no significant difference between the high task and medium task (Tukey HSD $p = .261$), and medium task and low task (Tukey HSD $p = .561$).

Table 2. Sensitivity with the workload level

Method	Task complexity (mean ± SD)			F statistic	Partial	p-value
	Low	Medium	High			
Number of errors (times)	4.1 ± 1.9	4.5 ± 1.8	5.3 ± 1.6	3.241	.080	.045*
NASA-TLX	37.7 ± 11.8	40.0 ± 11.49	53.1 ± 12.8	12.310	.247	.000**
Pupil dilation (px)	50.6 ± 9.2	49.4 ± 8.8	51.1 ± 9.1	1.299	.033	.574
Blink rate (times/min)	46.9 ± 6.1	45.2 ± 5.9	40.5 ± 5.1	8.689	.188	.000**
Blink duration (ms)	134.8 ± 18.9	145.5 ± 21.4	151.4 ± 23.2	4.085	.098	.021*
Fixation duration (s)	46.1 ± 23.7	47.3 ± 17.3	55.6 ± 13.1	1.998	.051	.143
HR (times/min)	77.3 ± 8.1	81.1 ± 8.2	86.1 ± 7.5	8.122	.178	.001*
SDNN (ms)	47.2 ± 17.0	50.3 ± 13.6	57.3 ± 11.7	3.400	.083	.039*

* $p \leq .05$, ** $p \leq .001$

3.2   Correlation Between the Number of Errors and Other Methods

The analysis of correlation was used to examine the relationship between the number of errors and other methods as shown in Table 3. It indicated that the number of errors and the NASA-TLX was positively correlated with each other. The correlation coefficient of $r = 0.563$ was found to be statistically significant at $p < 0.01$ (two-tailed). Mean of NASA-TLX score and the number of errors of each participant shown in Fig. 2.

Table 3. Correlation between the number of errors and other methods

Number of errors	1							
NASA-TLX	.563**	1						
Pupil dilation	.539**	.508**	1					
Blink rate	−.129	−.287*	.037	1				
Blink duration	.449**	.498**	.577**	−.061	1			
Fixation duration	.402**	.330**	.377**	−.055	.408**	1		
HR	.392**	.357**	.384**	−.250*	.511**	.249*	1	
SDNN	.150	.200	.259*	−.081	.189	.084	.280*	1

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

The statistic also showed that most physiological measures in this study correlate significantly with the number of errors indicate that physiological measures may assess the work performance by participants in the driving complex task.
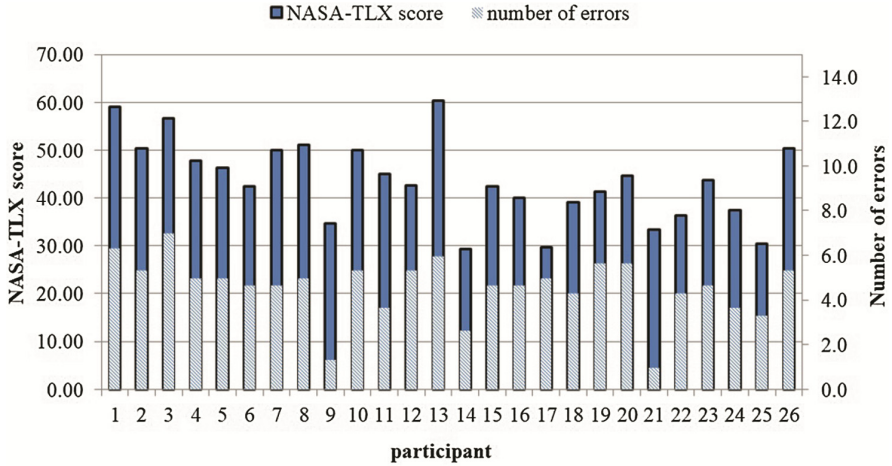
Fig. 2. Mean of the number of errors and NASA-TLX score of each participant in driving task

3.3   Predicting the Number of Errors by Integrating Physiological Measures

Six physiological indices, including pupil dilation (X_1), blink rate (X_2), blink duration (X_3), fixation duration (X_4), HR (X_5) and SDNN (X_6) into a synthesized index and to establish a model of work performance, this study used the GMDH method and the predictive modeling software DTREG version 10.6. The ratio of training and testing in this study was selected as 80%:20% to fit in with the available experimental sample size of 26. Each input variable (X_i) was normalized to a range of 0 and 1 before the training and testing process begins. The network was trained by using a random training data set, and the training data was also never used in the test data.

The results indicated that physiological indices of X_1, X_2, and X_5 were the best significant predictor factors in the performance by the subject. The model is expressed by Eq. (1) with the mean square error was 1.03, and R^2 of the model was 78.1%.

$$
\begin{aligned}
Y = {} & 4.816 + 1.152X_5 + 0.588X_2 + 0.233X_1 - 0.477X_5^2 + 0.091X_2^2 + 0.095X_1^2 - 0.433X_5X_2 \\
& - 0.290X_5X_1 - 0.163X_2X_1 + 1.125X_5X_2X_1 - 0.451X_5^3 - 0.276X_2^3 + 0.027X_1^3 - 0.467X_5X_2^2 \\
& - 0.309X_5X_1^2 - 0.844X_2X_5^2 + 0.010X_2X_1^2 + 1.079X_1X_5^2 + 0.217X_1X_2^2
\end{aligned} \tag{1}
$$

In the validation data, the result showed that the mean target value for predicted values is 4.62 while mean target value for input data is 4.5 (97.4%). Therefore, this model was suitable to estimate the performance of different MWL based on physiological measures in driving tasks.

4   Discussion

The number of errors was calculated as performance measures for the driving tasks in this study. The evaluation result showed that increasing task complexity makes increase the number of errors. This result is consistent with numerous studies which had found

that the human's performance was affected when the MWL was low [41]. On the other hand, the NASA-TLX scores showed a significant correlation with the different levels of MWL. For most of the subjects, the highest NASA-TLX score occurred in the high task complexity phase whereas the lowest score happened in the low task complexity phase. This result indicated that these tasks used in this experiment could distinguish the different levels of MWL.

Eye response measures are useful to reflect temporal distribution workload levels in driving task. However, no significant difference was found in pupil dilation and fixation duration. This result indicated that the pupil dilation in this experiment might not represent an increased processing need but rather reflects an increased attention and arousal caused by errors. This finding is consistent with Bradshaw's study in which he found that the pupil size change was not linked to the task complexity, but instead to the level of arousal of participants in problem-solving tasks [42]. Fixation duration index is extensively used measures and is believed to increase with increasing mental task demands [34], and Goldberg and Kotval [43] also found a negative correlation between fixation time and performance. Although the overall significance in fixation duration between different task levels was not found, there was a significant difference between the high task and low task. This result could be explained that the difference between the task levels (low-medium-high) is small.

Cardiac responses such as HR, HRV were used, and these responses seem more sensitive to the accumulative workload than eye response measures do. The experimental result indicated that mean of participants' HR and HRV components increased when the task complexity increased. These findings were consistent with previous studies [13, 44]. The participants in driving task needed to continuously exert mental effort to keep alert, and fatigue may have reduced the participants' attention. O'Hanlon [45] found that the initial decrease was changed into a gradual increase in HRV in long-time continuous driving and Tripathi, Mukundan and Mathew [46] also found that HRV increased in high-demand vigilance tasks that also require continuous exertion. Another plausible reason is the interaction influence of respiration on HR and HRV. A cognitive load promotes oxygen demand by cells and leads to the production of more cardiac output by increasing HR [47]. During the execution of tasks, participants breathed deep and long, which will increase.

Finally, this study used GMDH method to construct a model to predict the driver's work performance on different workload levels. Although the statistic in table showed that blink rate and HRV measure no correlates with the number of errors significantly at the level of .05, the predictive model that integrates different physiological measures explains 78.1% of the number of errors. With this model, it could provide a reliable reference tool to predict the work performance of drivers.

4.1 Limitations

Some limitations of this study should be mentioned. First, the experiment has used a small sample of a student population to evaluate and predict model; the small sample size reduced the statistical power. These students also do not represent the characteristics of the people who want to learn the driving car. In addition, in the simulation condition,

the participants often have psychological comfort because they must suffer the consequences of their mistakes when the operation fails or does not fulfill the requirements of the task. This causes for lack of significant differences among the outcomes and assessment results had limits of reliability. Finally, this result has been not shown the causal relationship between the physiological measures and the error rate but show a correlation between them under certain situations.

5   Conclusions

This paper reports the correlation of human's MWL and work performance in the driving task using driving simulator based on NASA-TLX and six physiological indices. The results show that different complexity levels of the driving task have a significant effect on the new driver's performance. In six physiological indices were used, three indices of pupil dilation, blink duration, and HR were the significant predictor factors, and the validity of this model was very well with $R^2 = 0.78$. Therefore, this model can be used to predict the new driver's work performance and maybe apply for actual. Although the model development process is still in an early phase, it can be used to predict the value of a new driver or little experience driving people on practice phase procedure.

References

1. Knipling, R.R.: Evidence and dimensions of commercial driver differential crash risk. In: Driving Assessment 2005: 3rd International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, pp. 2–8 (2005)
2. Todoskoff, A., Dillies, M.-A., Popieul, J.-C., Angue, J.-C.: Characterization of driver's behavior: a joint study in a car driving simulator and in real driving situations. In: Driving Simulation Conference, pp. 295–308 (1999)
3. Meister, D.: Behavioral Foundations of System Development. Wiley, New York (1976)
4. Waard, D.D., Jessurun, M., Steyvers, F.J., Reggatt, P.T., Brookhuis, K.A.: Effect of road layout and road environment on driving performance, drivers' physiology and road appreciation. Ergonomics **38**, 1395–1407 (1995)
5. Eggemeier, F.T., Wilson, G.F.: Performance-based and subjective assessment of workload in multi-task environments. In: Damos, D.L. (ed.) Multiple-task Performance, 217–278. Taylor & Francis, London (1991)
6. Sheridan, T., Stassen, H.: Definitions, models and measures of human workload. In: Moray, N. (ed.) Mental Workload. NATOCS, vol. 8, pp. 219–233. Springer, Boston (1979). https://doi.org/10.1007/978-1-4757-0884-4_12
7. Xie, B., Salvendy, G.: Review and reappraisal of modelling and predicting mental workload in single-and multi-task environments. Work Stress **14**, 74–99 (2000)
8. Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., Montanari, R.: Driver workload and eye blink duration. Transp. Res. Part F Traffic Psychol. Behav. **14**, 199–208 (2011)

9. 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**, 991–1009 (2005)
10. Yan, S., Tran, C.C., Wei, Y., Habiyaremye, J.L.: Driver's mental workload prediction model based on physiological indices. Int. J. Occup. Safety Ergon. 1–9 (2017)
11. Dyer, R.F., Matthews, J.J., Wright, C.E., Yudowitch, K.L.: Questionnaire construction manual. DTIC Document (1976)
12. Johnson, A., Widyanti, A.: Cultural influences on the measurement of subjective mental workload. Ergonomics **54**, 509–518 (2011)
13. Heine, T., Lenis, G., Reichensperger, P., Beran, T., Doessel, O., Deml, B.: Electrocardiographic features for the measurement of drivers' mental workload. Appl. Ergon. **61**, 31–43 (2017)
14. Hoeschen, A., Vervey, W., Bekiaris, E., Knoll, C., Widlroiter, H., Ward, D., Uneken, E., Gregersen, N., Falkmer, T., Schelin, H.: Inventory of drivers training needs and major gaps in the relevant training procedures. In: European Commission GROWTH Programme (2001)
15. Underwood, G., Crundall, D., Chapman, P.: Driving simulator validation with hazard perception. Transp. Res. Part F Traffic Psychol. Behav. **14**, 435–446 (2011)
16. Allen, R.W., Park, G.D., Cook, M.L., Fiorentino, D.: The effect of driving simulator fidelity on training effectiveness. DSC 2007 North America (2007)
17. Flach, J.M., Dekker, S., Jan Stappers, P.: Playing twenty questions with nature (the surprise version): reflections on the dynamics of experience. Theor. Issues Ergon. Sci. **9**, 125–154 (2008)
18. Tsang, P.S., Vidulich, M.A.: Mental Workload and Situation Awareness. Wiley, New York (2006)
19. Gawron, V.J.: Human Performance, Workload, and Situational Awareness Measures Handbook. CRC Press, London (2008)
20. Di Stasi, L.L., Álvarez-Valbuena, V., Cañas, J.J., Maldonado, A., Catena, A., Antolí, A., Candido, A.: Risk behaviour and mental workload: multimodal assessment techniques applied to motorbike riding simulation. Transp. Res. Part F Traffic Psychol. Behav. **12**, 361–370 (2009)
21. Baker, C.A., Morris, D.F., Steedman, W.C.: Target recognition on complex displays. Hum. Factors J. Hum. Factors Ergon. Soc. **2**, 51–60 (1960)
22. Downing, J.V., Sanders, M.S.: The effects of panel arrangement and locus of attention on performance. Hum. Factors **29**, 551–562 (1987)
23. Milleville-Pennel, I., Charron, C.: Do mental workload and presence experienced when driving a real car predispose drivers to simulator sickness? An exploratory study. Accid. Anal. Prev. **74**, 192–202 (2015)
24. Hart, S., Staveland, L.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv. Psychol. **52**, 139–183 (1988)
25. Gao, Q., Wang, Y., Song, F., Li, Z., Dong, X.: Mental workload measurement for emergency operating procedures in digital nuclear power plants. Ergonomics **56**, 1070–1085 (2013)
26. Wilson, G.F., Eggemeier, F.T.: Psychophysiological assessment of workload in multi-task environments. In: Damos, D.L. (ed.) Multiple-task Performance, 329–360. Taylor & Francis, London (1991)
27. Yan, S., Tran, C.C., Chen, Y., Tan, K., Habiyaremye, J.L.: Effect of user interface layout on the operators' mental workload in emergency operating procedures in nuclear power plants. Nucl. Eng. Des. **322**, 266–276 (2017)

28. Wierda, S.M., van Rijn, H., Taatgen, N.A., Martens, S.: Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. Proc. Nat. Acad. Sci. **109**, 8456–8460 (2012)
29. Hampson, R., Opris, I., Deadwyler, S.: Neural correlates of fast pupil dilation in nonhuman primates: relation to behavioral performance and cognitive workload. Behav. Brain Res. **212**, 1–11 (2010)
30. Recarte, M.A., Nunes, L.M.: Mental workload while driving: effects on visual search, discrimination, and decision making. J. Exp. Psychol. Appl. **9**, 119 (2003)
31. Recarte, M.Á., Pérez, E., Conchillo, Á., Nunes, L.M.: Mental workload and visual impairment: differences between pupil, blink, and subjective rating. Span J. Psychol. **11**, 374–385 (2008)
32. Kramer, A.F.: Physiological metrics of mental workload: a review of recent progress. In: Damos, D.L. (ed.) Multiple-task Performance, 279–328. Taylor & Francis, London (1991)
33. Häkkänen, H., Summala, H., Partinen, M., Tiihonen, M., Silvo, J.: Blink duration as an indicator of driver sleepiness in professional bus drivers. Sleep **22**, 798–802 (1999)
34. Recarte, M.A., Nunes, L.M.: Effects of verbal and spatial-imagery tasks on eye fixations while driving. J. Exp. Psychol. Appl. **6**, 31 (2000)
35. Myllylä, T., Korhonen, V., Vihriälä, E., Sorvoja, H., Hiltunen, T., Tervonen, O., Kiviniemi, V.: Human heart pulse wave responses measured simultaneously at several sensor placements by two MR-compatible fibre optic methods. J. Sens. **2012**, 1–8 (2012)
36. Solovey, E.T., Zec, M., Garcia Perez, E.A., Reimer, B., Mehler, B.: Classifying driver workload using physiological and driving performance data: two field studies. In: Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, pp. 4057–4066. ACM, New York (2014)
37. Ivakhnenko, A.G.: The group method of data handling-a rival of the method of stochastic approximation. Sov. Autom. Control **13**, 43–55 (1968)
38. Stern, H.S.: Neural networks in applied statistics. Technometrics **38**, 205–214 (1996)
39. Ahmadi, M.H., Ahmadi, M.-A., Mehrpooya, M., Rosen, M.A.: Using GMDH neural networks to model the power and torque of a stirling engine. Sustainability **7**, 2243–2255 (2015)
40. Náplava, P., Šnorek, M.: Modelling of students quality by means of GMDH algorithms. Syst. Anal. Model Simul. **43**, 1415–1426 (2003)
41. Wilson, G.F., Russell, C.A.: Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. Hum. Factors **45**, 635–644 (2003)
42. Bradshaw, J.: Pupil size and problem solving. Q. J. Exp. Psychol. **20**, 116–122 (1968)
43. Goldberg, J.H., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. **24**, 631–645 (1999)
44. Brookhuis, K.A., de Waard, D.: Monitoring drivers' mental workload in driving simulators using physiological measures. Accid. Anal. Prev. **42**, 898–903 (2010)
45. O'Hanlon, J.F.: Heart Rate Variability: A New Index of Driver Alertness/Fatigue. Automotive Engineering Congress; 0148–7191 (1972)
46. Tripathi, K., Mukundan, C., Mathew, T.L.: Attentional modulation of heart rate variability (HRV) during execution of PC based cognitive tasks. Ind. J. Aerosp. Med. **47**, 1–10 (2003)
47. Brookhuis, K.A., de Waard, D.: Assessment of Drivers' Workload: Performance and Subjective and Physiological Indexes. Stress, Workload and Fatigue. Lawrence Erlbaum Associates Inc., Mahwah (2001)