International Journal of Technology (2017) 2: 283-291
ISSN 2086-9614

EYE BLINK RATE AS A MEASURE OF MENTAL WORKLOAD IN A DRIVING TASK: CONVERGENT OR DIVERGENT WITH OTHER MEASURES?

Ari Widyanti[1*], Nisha Faradila Sofiani[1], Herman Rahadian Soetisna[1], Khoirul Muslim[1]

[1]*Laboratory for Work System Design and Ergonomics,*
Department of Industrial Engineering, Bandung Institute of Technology (ITB),
Ganesa 10, Bandung 40132, Indonesia

(Received: October 2016 / Revised: December 2016 / Accepted: January 2017)

ABSTRACT

Measuring mental workload requires both subjective and objective measurement. However, as existing objective measures lack applicability due to technical reasons and cost considerations, this study evaluates an easy to use and cost effective method of measuring the sensitivity of the eye blink rate as a potential objective measure of mental workload. Eight participants were instructed to operate a driving simulator in a lab setting and complete a series of driving tasks set at three different levels of difficulty. The completion time and penalty scores were recorded as the performance measures. The eye blink rate data were analyzed as an objective measure, and the NASA Task Load Index (NASA-TLX) was used to assess the participants' mental workload at the end of each task as the subjective measure. Although the completion time, penalties, and NASA-TLX increased as the difficulty level of the tasks increased, the eye blink rate decreased. The implications of these results are discussed.

Keywords: Blink rate; Driving simulation; Mental workload; NASA-TLX

1. INTRODUCTION

Humans have a critical role in work systems alongside other elements such as machines and materials. Therefore, every characteristic, limitation, and capability of humans should be considered when designing and improving a work system. In relation to the limitations and capabilities of humans, workloads have recently gained considerable attention (Johnson & Widyanti, 2011).

According to O'Donnell and Eggemeier (1986), the term workload refers to the portion of the worker's limited capacity that is invested to perform a particular task. This definition indicates that workload is not only dependent on the task itself, but also on the ability and willingness to perform of the human. A worker's workload should be balanced with their ability to achieve the optimal work performance. Thus, workloads that are lower than the worker's ability create boredom and cause a decrease in work performance. Conversely, workloads that exceed the worker's ability can cause fatigue, increase the risk of errors, and lead to a higher chance of a work accident. Workload measures are therefore particularly important in particular for gaining optimal workload.

Workload can be divided into two types: physical and mental. Physical workload, which represents the physical effort that must be invested in completing a task, can be assessed quantitatively using physiological indices, such as heart rate, and biomechanical aspects, such

*Corresponding author's email: widyanti@mail.ti.itb.ac.id, Tel. +62-22-2508124, Fax. +62-22-2508124
Permalink/DOI: https://doi.org/10.14716/ijtech.v8i2.6145

as the tension placed on the human musculoskeletal system. Most definitions of mental workload describe interactions between the task requirement and human capabilities or resources. For example, Rubio et al. (2004) described mental workload as the difference between the resources demanded by the job and the resources available to the worker.

Measures of mental workload have become crucial for developing safe and efficient workplaces of high automated technology. The effect of overload has long been recognized as a factor that leads to injury and accident. However, recent studies have raised concerns about the risk of work *underload*, which may lead to boredom and low motivation to perform the task (Becker et al., 1991; Hancock & Warm, 1989).

In general, mental workload can be measured using three approaches: performance, objective, and subjective. Performance-based measures, which are used to define how well someone performs a task, are based on an underlying assumption that any increase in task difficulty will lead to an increase in demands and will decrease performance, although there is a possibility that performance can be maintained optimally by investing more effort. Objective measurements are measured using a physiological approach such as heart rate, eye blink intervals, or acidity of saliva, and subjective measurements are measured based on the subjective perception of each worker. While objective measurements can detect changes in mental workload without being influenced by the subjectivity factor, they require complex special equipment (such us electrodes and gel for electromyograph/EMG) and special skills (such as frequency based analysis for hart rate variability/HRV measure), and the results of the measurements can be affected by other factors outside of the workload factors, such as environmental factors and conditions preceding the workload measurement (De Waard, 1996). As a measure of mental workload, objective measures are not really objective since many factors can influence the result, including different environments and cultures (see Johnson & Widyanti, 2011; Widyanti et al., 2013a; Widyanti et al., 2013b).

Subjective procedures assume that an increased expense is linked to an increase in perceived effort that can be appropriately assessed and reported by individuals (Hockey, 1986); whereas, physiological indexes assume that mental workload can be measured by the level of physiological activation (e.g., heart rate, oxygen consumption; Rubio et al., 2004). The subjective measures of mental workload are based on workers' perceptions about the workload imposed on them; however, the subjectivity is the main disadvantage of this measure. One of the most widely used subjective measures of mental workload is the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1998), which comprises six dimensions of mental workload (Table 1). This study uses the raw NASA-TLX, because the raw NASA-TLX has been proven to have a sensitivity level that is similar to the traditional NASA-TLX (see Hart, 2006, for a review). The traditional NASA-TLX compose of two steps: pairwise comparison to give weight for each dimension compared to other dimensions and rate of each dimension on 0-100 scale. Whereas the raw NASA-TLX skip the weighting step.

While Western countries have conducted studies to measure mental workloads over approximately 40 years, Indonesia and other Eastern countries have more recently examined the importance of balancing resource demand of task and resource available (Johnson & Widyanti, 2011). Considering the high rate of road accidents due to human error, one of the most frequently conducted measures of mental workload is that placed on car drivers (e.g., Paxion et al., 2014).

As mentioned above, it is possible to measure drivers' mental workloads using three different approaches: performance, objective, and subjective measures. Eye-related measures are relatively well developed objective measures that are often used to assess a driver's mental workload (Marquart et al., 2015). Of these measures, blink rate and blink duration parameters are most commonly used. Blink duration has been shown to decrease as mental workload increases (Kramer, 1990); however, the use of blink rate in a series of driver workload studies has found mixed results, with rates increasing or decreasing depending on the visual demands (Marquart et al., 2015). In an early study, Kramer (1990) stated that more research was required before the eye blink rate could be used as a measure of mental workload due to its mixed result. However, in a real car-driving study, Heger (1998) showed that the eye blink rate decreased as the curves of the road became sharper. Ten years later, Recarte and Nunes (2003) analyzed the blink rate during single- and dual-task (cognitive task plus visual search) conditions and found that the blink rate increased for all cognitive tasks (listening, talking, and calculating) when compared to the control condition. Recent studies have also shown inconsistent results relating blink rate and mental workload (see Marquart et al., 2015 for a review), thus highlighting the need for more research in this area.

Table 1 Dimensions of the NASA-TLX and the descriptions

Dimension	Description
Mental demand (MD)	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, and searching)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical demand (PD)	How much physical activity was required (e.g., pushing, pulling, turning, controlling, and activating)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal demand (TD)	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance (P)	How successful do you think you were in accomplishing the goals of the task set by experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort (EF)	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration level (FR)	How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

Eggemeier (1988) stated that the measures of mental workload should meet the following requirements: sensitivity, which refers to the ability to detect changes in task demands; diagnosticity, which is the ability to identify sources of changes in workload variation and different types of workload (e.g. visual, central); validity, which is the ability to recognize differences in cognitive demands—but not changes in other variables such as physical workload or emotional stress—and to fulfill a classical psychometric concept of validity (face, construct, content, and predictive); intrusiveness, which prevents interference with normal task performance; reliability, which refers to consistency among mental workload measurements in different fields and at different times; implementation, which refers to implementing the measure considering constraints related to time, instruments, and the collection and analysis of data (practical constraints); and acceptability by participants.

This study evaluates the sensitivity of the eye blink rate for measuring participants' mental workload in a driving task. Mental workload was measured at three different levels of difficulty and in a rest condition. Performance measures were assessed based on completion time and error rate. While the blink rate provides an objective measure of mental workload, the NASA-TLX was used to provide a subjective measurement to validate the results.

2. METHODS

Eight participants, four males and four females (mean age = 22.5 years; SD = 1.732 years), participated in this study voluntarily. The participants were required to have one year of driving experience and possess normal vision. Before conducting the experiment, the participants were instructed to have a good and enough sleep (i.e., about 8 hours) the night before the experiment. The physical and mental condition of the participants was quickly overviewed by short interview by experimenter to ensure that the participants were ready for the experiment, indicates by their level of fitness and alertness. Prior to collecting any data, each participant was asked to sign an informed consent form. The experiments were conducted in a driving simulator in a laboratory setting. The driving simulator was used to eliminate variables that cannot usually be controlled, such as the level of difficulty of the activity of driving.

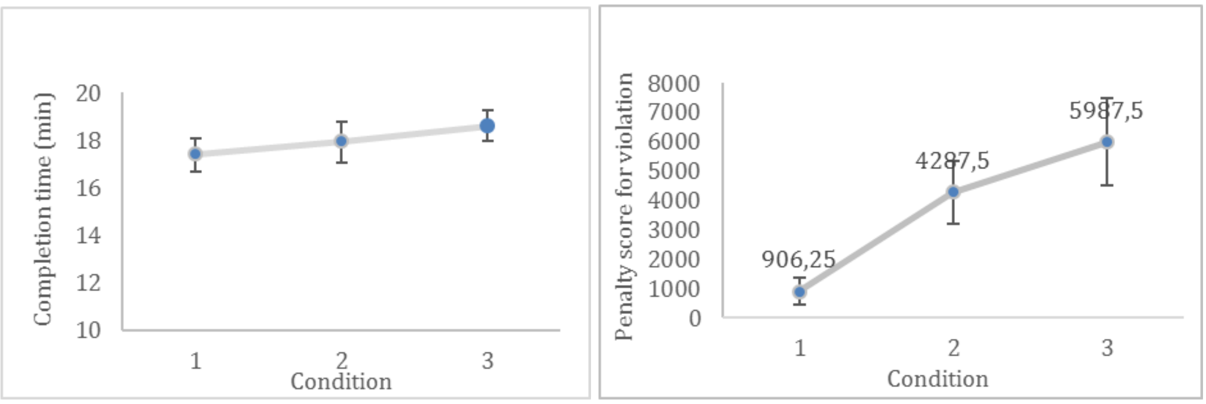
The participants were instructed to complete a series of driving tasks set at three levels of difficulty. Before the experiment was conducted, the participants were given a trial driving task to minimize the learning effect. No time limit was set for this trial session. Then, in the first condition of the experiment—the easiest level—the participants were instructed to drive along a straight road with low traffic density (i.e. number of vehicles per kilometer, 10%). In the second condition, the participants were instructed to drive in a rural area with a few curves and moderate traffic density (50%). In the third condition—the highest level of workload—the participants had to drive in a complex driving situation involving city roads with many curves and high traffic density (80%). A balanced Latin square design was applied to eliminate the influence of task difficulty order on the participants. Each participant was required to follow the same path to ensure they experienced the same workload so that any changes in the participants' mental workload could be monitored.

The driving performance was assessed based on the time needed to complete the task and by the error rate, which was reflected by a penalty for each driving violation during the simulation. The objective measure of the drivers' mental workload, i.e., the blink rate, was assessed during the driving task simulation using Gaze Point apparatus, and the subjective mental workload was assessed at the end of each condition using the NASA-TLX (Hart & Staveland, 1998). To ensure that all participants had the same perceptions, a description of each measured dimension was provided in the NASA-TLX questionnaire. In addition, participants were allowed to ask questions regarding the NASA-TLX questionnaire.

An ANOVA test was conducted to check whether the performance, the NASA-TLX results, and the blink rate results were sensitive to the changes in mental workload. Post-hoc analysis was used to observe the source of differences in mental workload, and MANOVA analysis was performed to determine whether the participants' gender or age could have affected the results of the mental workload measure. All statistical analysis were performed using SPSS, with a significance level of $p < .05$.

3. RESULTS

The findings show that the drivers took more time to complete the task and made more errors as the difficulty level of the task increased (Figure 1). Figure 2 shows that the blink rate for each condition, indicating that drivers blink less frequently as mental workload increases. The subjective measure of mental workload, the NASA-TLX score, also increases as the levels of difficulty increase (Figure 3). Figure 4 shows the score of each dimension of the NASA-TLX for each condition in which all dimensions of the NASA-TLX increased as a function of task difficulty, except for performance.



(a) (b)

Figure 1 Completion time: (a) and error/penalties; (b) for each level of difficulty

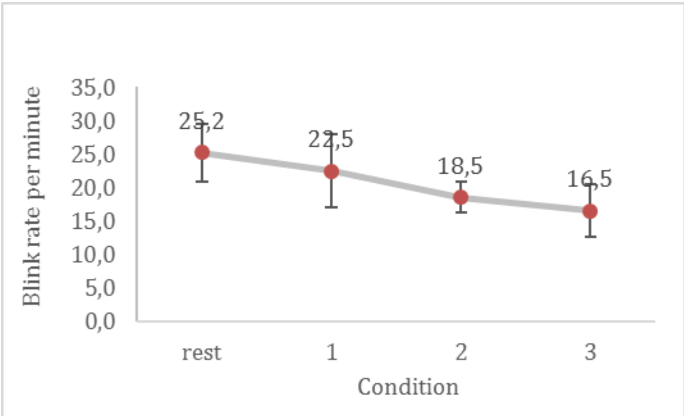


Figure 2 Eye blink rate for each level of difficulty

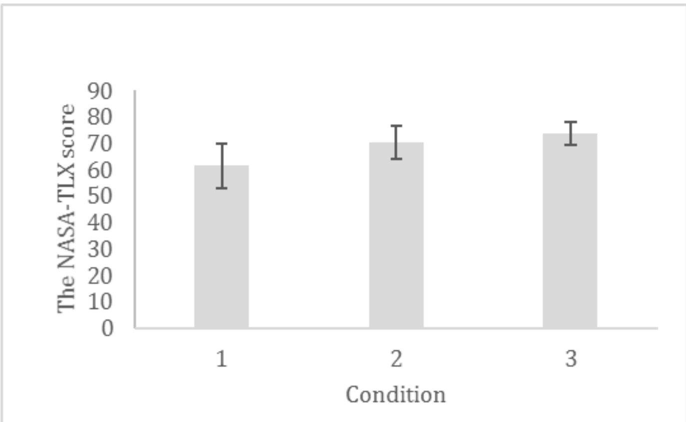


Figure 3 The NASA-TLX score for each level of difficulty

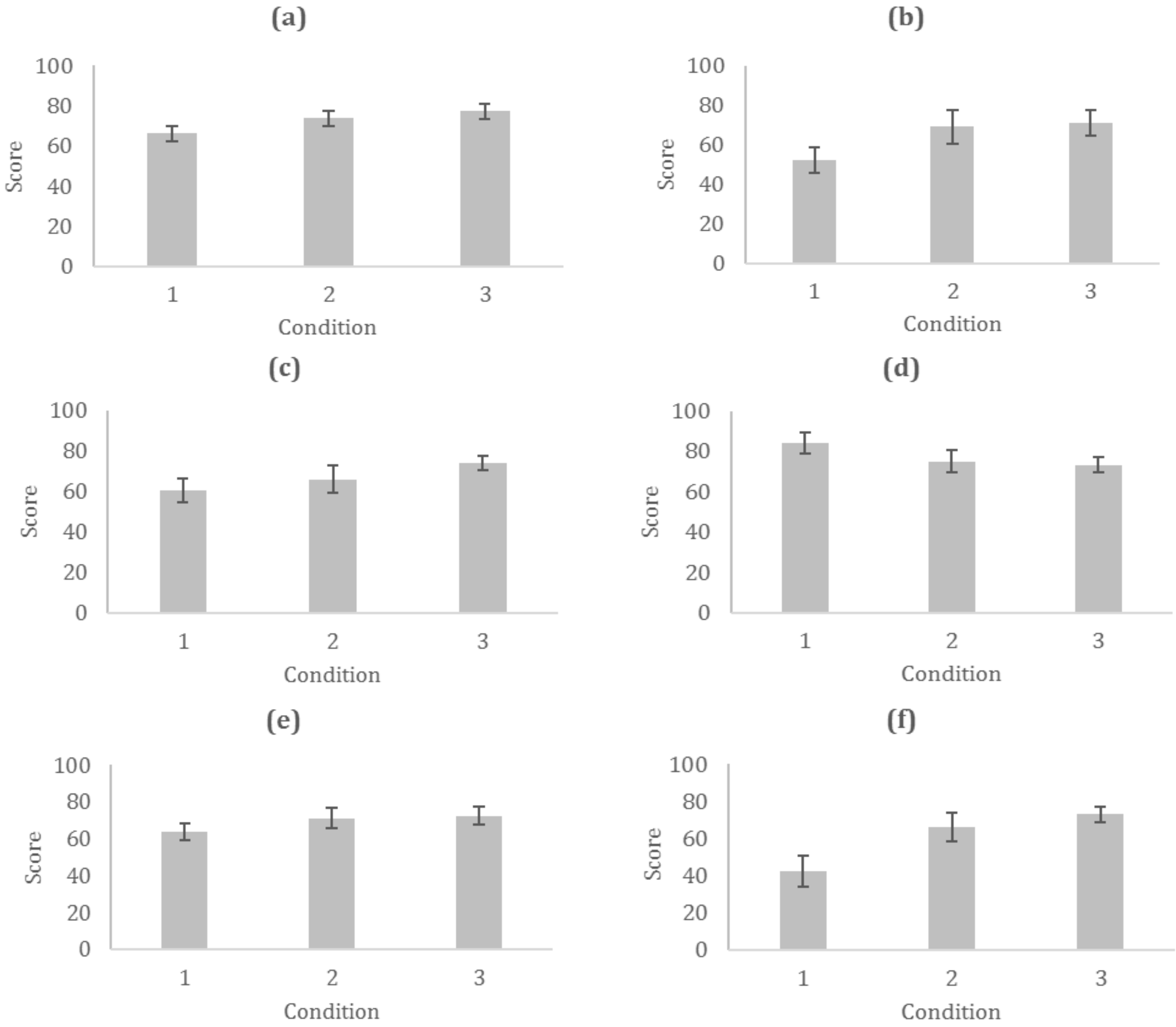


Figure 4 Each dimension of the NASA-TLX score: (a) Mental Demand; (b) Physical Demand; (c) Temporal Demand; (d) Performance; (e) Effort; (f) Frustration Level for each level of difficulty

The ANOVA test showed significant differences in the penalty scores ($F(2,21)=11.308$, $p<.001$), the NASA TLX score ($F(2,21)=6.291$, $p=.007$), and mean blinks per minute ($F(4,35)=4.965$, $p=.003$). Significant differences were also found in all dimensions of the NASA-TLX. The penalty and mean NASA TLX scores both increased as the task difficulty increased; however, the mean blinks per minute decreased as the task difficulty increased. As expected, all dimensions of the NASA-TLX increased with increasing task difficulty, except for the performance dimension.

The post-hoc analysis showed that significant differences existed between the penalty scores of the first and second conditions, and between those of the first and third conditions. The analysis also showed that significant differences existed between the NASA-TLX scores of the first and third conditions, and between those of the first and second conditions. While the post-hoc analysis of the blink rate showed that significant differences existed between the baseline with all conditions, significant differences were also evident between the first and third conditions. A MANOVA test, which was conducted to test whether gender and age had an effect on the results of the blink rate, showed that age ($F=2.478$, $p=.251$) and gender ($F=1.176$, $p=.519$) did not affect the blink rate value.

4. DISCUSSION

This study evaluated the sensitivity of the eye blink rate for measuring mental workload while driving in a simulator. Three levels of task difficulty were used to test the sensitivity of the workload measures. The first workload measure, i.e., performance, was obtained from the completion time and the penalty points awarded for errors by the driving simulator software. The eye blink rate measure was obtained during the driving simulation and was defined as an objective mental workload measure. The NASA TLX scores were obtained during and after the task to measure the subjective mental workload. These three measures were then compared to determine the sensitivity of the eye blink rate for measuring mental workload differences.

The penalty scores, which increased as the task difficulty increased, indicated significant differences between the levels of difficulty. Further, a post-hoc analysis revealed that significant differences existed between the first and second conditions, and between the first and third conditions. However, no significant differences were found between the second and third conditions. The subjective measures found significant differences between the levels of difficulty; however, no significant differences were evident between the second and third conditions. A post-hoc analysis showed that significant differences were observed between the first condition and the second and third conditions. These results mirrored the results for the penalty scores. A balanced Latin square design was applied to eliminate the influence of task order on the results; however, there is no evidence to suggest that the task order caused the non-significant differences between the second and third conditions.

The decrease in blink rate following the increase in task difficulties found in this study can be explained by Recarte and Nunes (2003, p.8), who concluded that “according to blink rate, visual and mental workload produce opposite effects: Blink inhibition for higher visual demand and increased blink rate for higher mental workload.” This theory might support Heger’s (1998) findings because the driver’s task in this experiment might be more visually demanding (scanning the curved road ahead) than mentally demanding (controlling the vehicle), which could explain the decrease in blink rate. This study also found that traffic density and other variables controlled for during the experiment are more visually demanding than mentally demanding for participants. An in-depth analysis of the driving simulation task will thus be conducted.

The sensitivity of mental workload measures needs to be assessed using a number of different methods not only because of its complexity but also because it is influenced by the environment and culture (see Johnson & Widyanti, 2011; Widyanti et al., 2014). In addition, Marquart et al. (2015) stated that future research should focus on combining multiple assessment methods to increase the validity and robustness of assessing mental workloads.

This study has some limitations. The first limitation is the small number of participants in this study. Though a nonparametric test can be used for small samples, a larger number of participants can increase the statistical power and can thus lead to obtaining reliable and rigorous results. The second limitation is that the eye blink rate was used to measure the mental workload only while participants were driving in the simulator. Further studies should thus evaluate the eye blink rate in real driving situations and other real world tasks.

5. CONCLUSSION

This study shows that the eye blink rate, as an objective measure of mental workload, is divergent from performance measures (completion time and error) and subjective measures (the NASA-TLX) of mental workload. The eye blink rate decreases as the level of task difficulty or mental workload increases; thus, the eye blink rate seems more appropriate for measuring visually demanding tasks than mentally demanding tasks.

6. ACKNOWLEDGEMENT

The authors thank the Indonesian Ministry of Research, Technology, and Higher Education (RISTEKDIKTI) for funding this research (Hibah Kompetensi, fiscal year 2016–2017).

7. REFERENCES

Becker, A.B., Warm, J.S., Dember, W.N., Hancock, P.A., 1991. Effects of Feedback on Perceived Workload in Vigilance Performance. *In:* Proceedings of the Human Factors Society, 35, 1991–1994. Santa Monica, CA: Human Factors and Ergonomics Society, pp. 1491–1494

De Waard, D., 1996. *The Measurement of Driver's Mental Workload*. Den Haag: Traffic Research Center, University of Groningen

De Waard, D., Lewis-Evans, B., 2014. Self-Report Scales Alone Cannot Capture Mental Workload. *Cognition Technology & Work*, Volume 16, pp. 303–305

Eggemeier, F.T., 1988. *Properties of Workload Assessment Techniques*, in P. Hancock & N. Meshkati (eds.), Human Mental Workload. North Holland Publisher. Amsterdam, pp. 41–62

Hancock, P.A., Warm, J.S., 1989. A Dynamic Model of Stress and Sustained Attention. *Human Factors*, Volume 31, pp. 519–537

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, in P.A. Hancock, N. Meshkati (eds.), Human Mental Workload. *Elsevier*, New York, pp. 139–183

Hart, S.G., 2006. NASA-Task Load Index (NASA-TLX); 20 Years Later. *In:* Proceedings of the Human Factors and Ergonomics Society 50^{th} Annual Meeting. Santa Monica, CA: Human Factors and Ergonomics Society, pp 904–908

Heger, R., 1998. Driving Behavior and Driver Mental Workload as Criteria of Highway Geometric Design Quality. *Transportation Research Circular*, Washington, DC: Transportation Research Board. pp. 43-1–43-10

Hockey, G.R.J., 1986. *A State Control Theory of Adaptation and Individual Differences in Stress Management*, in G.R.J. Hockey, A.W.K. Gaillard, M.G.H. Coles (eds.), Energetics and Human Information Processing, Martinus Nijhoff Publishers, Dordrecht, The Netherlands, pp. 285–298

Johnson, A., Widyanti, A., 2011. Cultural Influences on the Measurement of Mental Workload. *Ergonomics*, Volume 54(6), pp. 509–518

Kramer, A., 1990. *Physiological Metrics of Mental Workload*. A Review of Recent Progress, in D. Damos (ed.), Multiple-Task Performance. Taylor & Francis, London, pp. 279–328

Marquart, G., Cabrall, C., de Winter, J., 2015. Review of Eye-related Measures of Driver's Mental Workload. *Procedia Manufacturing*, Volume 3, pp. 2854–2861

O'Donnell, R., Eggemeier, F., 1986. *Workload Assessment Methodology*, in K. Boff, L. Kaufman, J. Thomas (eds.), Handbook of Perception and Human Performance. Volume II, Cognitive Process and Performance, John Wiley and Sons, New York, pp. 42-1–42-44

Paxion, J., Galy, E., Berthelon, C., 2014. Mental Workload and Driving. *Frontiers in Psychology*, Volume 5, pp. 1–11

Recarte, M., Nunes, L., 2003. Mental Workload While Driving: Effects on Visual Search, Discrimination, and Decision Making. *Journal of Experimental Psychology: Applied*, Volume 9(2), pp. 119–137

Rubio, S., Diaz, E., Martin, J., Puente, J., 2004. Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology*, Volume 53(1), pp. 61–86

Widyanti, A., de Waard, D., Johnson, A., Mulder, B., 2013a. National Culture Moderates the Influence of Mental Effort on Subjective and Cardiovascular Measure. *Ergonomics*, Volume 56(2), pp. 182–194

Widyanti, A., Johnson, A., de Waard, D., 2013b. Adaptation of the Rating Scale Mental Effort (RSME) for Use in Indonesia. *International Journal of Industrial Ergonomics*, Volume 43(1), pp. 70–76

Widyanti, A., Sunaryo, I., Kumalasari, A.D., 2014. Reducing the Dependency on Rice as Staple Food in Indonesia: A Behavior Intervention Approach. *Journal of International Society for South East Asian Agricultural Sciences*, Volume 20(1), pp. 93–103