Review of Cognitive Ergonomic Measurement Tools

B Fista1, H A Azis1, T Aprilya1, S Saidatul1, M K Sinaga1, J Pratama1, F A Syalfinal1, Steven1 and S Amalia1

1Department of Industrial Engineering, Bandung Institute of Technology (ITB)
Ganesa 10 Bandung, Indonesia

sheila.amalia@students.itb.ac.id

Abstract. In measurement and evaluation for a cognitive performance carried out on various tasks can use objective and subjective measurement tools. This study aims to review research on measuring instruments and provide its potential to be used in research related to cognitive ergonomics. The method that is used in this study is a review of article literature on studies that are subjective and objective measurement tools. The results of this study indicate that in each study usually does not only use one measuring instrument, to validate measurements, but another measurement tool is also used. And in research usually uses subjective and objective measurement tools for the same task.

1. Introduction
Cognitive ergonomics, defined by the International Ergonomics Association is concerned with mental processes, such as perception, memory, reasoning, and motor response, as they affect interactions among humans and other elements of a system. Cognitive ergonomics studies cognition in work and operational settings, in order to optimize human well-being and system performance. It is a subset of the larger field of human factors and ergonomics. The relevant topics include mental workload, decision-making, skilled performance, human-computer interaction, human reliability, work stress, and training as these may relate to human-system design.

Mental workload is the portion of the operator’s limited mental capacities actually required to perform a particular task [1]. There is a body of literature describe the measurement of mental workloads that are performance measures, objective measure (based on physiological indices), and subjective measures (based on the perceived load of human). Performance measures were conducted based on reaction time and error. Psychophysiological indices include heart rate variability /HRV, Galvanic Skin Response/GSR, eye blink rate and such advances of brain activity measures for example electroencephalogram (EEG). Subjective measures include the NASA-TLX, RSME, and SWAT.

Subjective measures, which require that individuals reflect on their state and report the degree of effort expended, can be expected to be sensitive in a wider range of circumstances than performance measures, as operators, in general, can indicate well how much effort the performance of a task requires [2]. Psychophysiological indices such as pupil diameter and heart rate measures have the advantage that they can be taken continuously and relatively unobtrusively [3]. Among the psychophysiological measures of mental workload, heart rate and HRV are frequently used [4].

Subjective measurements are easy to use and usually have a high degree of acceptance by the operator. However, there are some disadvantages to these measures. First, if the assessment is conducted after the work has been completed, there is a possibility that the worker may fail to remember all the workload he or she experienced while performing the task. The results of
measurements with this method could also produce a biased result because the worker does not provide the real answer [5]. In addition, subjective mental workload measures have been shown to be culturally dependent [6].

Objective measurement has the advantage of being carried out continuously and is sensitive in detecting the changes in mental workload [7]. The disadvantages of this measurement include the need for special measurement equipment and the influence of other factors such as physical and environmental on the measurement results [8].

This study aims to review research on measuring instruments commonly used in research related to cognitive ergonomics based on mental workload assessment. This literature review aims to provide a potential study of cognitive measurements that have been developed. There are subjective measurement tool and objective measurement tool.

2. Method
A literature review is conducted using the search engine of Google Scholar, including Science Direct, Taylor and Francis, SAGE, Plos One, IEEE, etc. The keywords used are a combination of the words 'objective measurement tools for cognitive' and 'subjective measurement tools for cognitive'. For 'objective measurement tools', the keyword used is 'heart rate variability, ‘eye tracker, ‘salivary’ ‘galvanic skin response’ and ‘electroencepalogram'. For 'subjective measurement tools', the keywords used are ‘NASA-TLX’, ‘SWAT’, and ‘RSME’. Only publications written in English are used which are published in journals from 2000-2018 (18 years). Later, the result of the google scholar is screening based on the Scimago journal rank, and only quartile 1, 2, and 3 (Q1, Q2, and Q3) is selected due to consideration of high qualified of a journal.

3. Result
The review result can be seen in Table 1 – 2 where the amount of final papers is 23 papers. For these 23 papers, an abstract review was applied in which the abstract must meet these following criteria:

- The method used must be an experimental design.
- The dependent variable or variable discussed is in the form of measurement tools for the cognitive task.

From the table, it can be concluded that there are several measurement tools for cognitive task evaluation, with subjective and objective tools.

4. Discussion

4.1. Objective Measurement
The eye tracker is a device for measuring eye positions and eye movement. Eye tracker basically functions as a USB 3.0 camera device that sends and receives data via USB 3.0 to a server running on a computer. The binocular gaze and pupil data can be accessed through the API using a variety of programming languages including Java, C++, and C#. It can be used to determine a lot of data such as the pupil size [9][10], blink rate [9][11][12], fixation [8][9][12], and spread of search[14]. From the result, the eye tracker can be used to driving simulator [8][9][11][14], measuring working memory performance [12], and measuring attention performance [10][12].
Table 1. Search term using ‘objective measurement tools’

| Source | Cognitive Task | Measuring Instrument | Parameter | Object | Country | Conclusion |
|--------|----------------|----------------------|-----------|--------|---------|------------|
| [5] Psychomotor vigilance test (PVT) and working memory task | Heart rate variability (HRV), NASA-TLX | Root mean square of the successive differences (RMSDD), R-R intervals, high-frequency (HF), mental demand (MD) | 24 males of undergraduate students | Spain | HRV is highly sensitive to the overall demands of sustained attention over and above the influence of other cognitive processes. |
| [12] Still | The salivary transcription device, a testing-strip, and an optical analyzer | salivary amylase activity (kU/l), Cortisol (pmol/ml) | 11 Healthy Young adult and 10 Healthy Young adults (for psychological Stressor) | Japan | SA is an easy method to monitor salivary amylase activity used to distinguish eustress and distress, a salivary transcription device was fabricated to control the enzymatic reaction time |
| [13] Attention | Salivettes, a finger-pulse-plethysmograph, two Ag/AgCl surface electrodes | cortisol (nmol/l), Heart Rate(BPM), salivary amylase Activity (U/ml), Skin conductance level (SCL) | 87 Male participants | | Social evaluative stress can be successfully induced in a virtual environment resulting in stress responses on several physiological measures associated with the HPA axis and the SAM system |
| [14] Driving simulator and n-back task | Eye Tracker, Button attached to the index finger for measuring DRT | Pupil size, Blink rate, Eye movement, Fixation rate, and DRT | 22 participants | Slovenia | When assuming that head movements are negligible when drivers look at the forward roadway, |
| [8] | Eye tracker, and dongle that connected on the OBD port of the car | Fixation frequencies, fixation duration, visual scan | 30 participants | Belgium | | |
| Reference | Methodology | Details |
|-----------|-------------|---------|
| [7]       | Memory      | Eye Tracker, CRT monitor, and 432 Photographs. Fixation density, fixation duration, memorability, and RMSD of all fixation during the trial from the initial fixation cross. 77 undergraduate students from Canada. |
| [19]      | Memory      | Galvanic Skin Response (GSR). Accuracy, reaction time, NASA-TLX, HRV LF band, and mean GSR. 195 Indonesian university students. |
| [20]      | Memory      | NASA-TLX, HRV, GSR. Accuracy, reaction time, NASA-TLX, HRV MF band. 72 Indonesian students. |
| [21]      | Memory      | HRV, NASA-TLX. Accuracy, reaction time, NASA-TLX, HRV MF band. 195 university students from Indonesia. |
| [22]      | Driving simulator | Eye tracker, NASA-TLX. Completion time, penalty score, Eye blink rate, Ratings of NASA-TLX. 8 participants from Indonesia. |

Eye fixations are more spread over the visual field during hands-free phoning. Eye movements during both study and test are related to overall memory performance, which includes both correctly recognizing a previously studied scene and correctly rejecting a scene that was not studied.

For the conceptual design of the development of real-time and adaptive mental workload measures based on GSR, the steps are filtering and classifying. Like the HRV, GSR showed to be sensitive in distinguishing rest and task condition significantly but not sensitive in distinguishing different levels of mental workload. The sensitivity of the subjective measures of mental workload has been shown to depend on culture. Culture influences more how invested mental effort is reported than how it is experienced psychophysiological ly. Although the completion time, penalties, and NASA-TLX increased as the difficulty level of the tasks increased,
| Reference | Task Type | Measurement | Subjects | Location |
|-----------|-----------|-------------|----------|----------|
| [26]      | Flying simulator | Heart rate (HR), Heart Rate Variability (HRV) | 26 participants | Finland |
| [27]      | Divided attention task | Electroencephalogram (EEG), eye tracker | EEG workload, pupil diameter | 47 ATC students | USA |
| [28]      | Sustain attention task | Eye tracker | Eyeblink (frequency and duration) | 19 civilian and active duty military personnel | USA |
| [29]      | Driving simulator | fNIRS, eye tracer NASA-TLX, BIOAPC equipment (skin conductance, heart rate, and respiration rate) | NASA-TLX Mental Demand, nSCRs per minute, mean skin conductance, hemoglobin concentration, the spread of search | 30 participants | UK |

The eye blink rate decreased.

HR and HRV differentiate varying task demands in situations where variations in performance are insignificant.

The EEG workload metric and pupil diameter were sensitive to workload manipulations but did not differentiate experience groups.

Commercially available EEG cognitive state metrics may be a viable tool for enhancing ATC training.

Eyeblink information may be an indicator of arousal levels. Using an eye-tracker to detect changes in eye blinks in an operational environment would allow preventative measures to be implemented.

Increases in subjective ratings of mental workload caused by changes in road type were accompanied by increases in skin conductance, acceleration signatures and horizontal spread of search. Such changes were also associated with increases in the concentration of oxygenated hemoglobin in the prefrontal cortex.
The driving simulator, working memory task

Electrocardiogram (ECG), Electroencephalogram (EEG), NASA-TLX

fro-theta power, parietal alpha power (par-alpha), subjective load, HR, HRV

26 participants Germany

Electroencephalogram (EEG)

theta power, alpha power

8 participants California

**Table 2.** Search term using ‘subjective measurement tools’

| Source | Cognitive Task | Measuring Instrument | Parameter | Object | Country | Conclusion |
|--------|----------------|----------------------|-----------|--------|---------|------------|
| [6]    | Driving simulator | Subjective Workload Assessment Technique (SWAT) | Driving lap times and errors, RVCB time and accuracy scores, SWAT scores | 12 licensed drivers | United States (US) | Cognitive workload increased as a result of sharing a dual task. |
| [3]    | Sternverg’s Memory Searching task, Tracking task. | NASA TLX, SWAT, Workload Profile (WP) | Subject performance, mental workload. | 36 students | United Kingdom (UK) | Workload Profile is recommended for comparing between mental workload of 2 or more tasks with different objective levels of difficulty. NASA-TLX is recommended of a particular individual task. Mental task demands reduce situation awareness, and that not only type-of-task, but also time-on-task, should be considered in Human Factors research of automated driving. Particularly when the driver is distracted or fulfilling other tasks during automated driving mode, the additional haptic information encoding can be lifesaving. |
| [16]   | Driving an Automated Vehicles | NASA Task Load Index | Mental task Demand | 19 male and 14 female | UK | More automated tasks do not imply performance improvement. It is also not |
| [17]   | Driving | NASA Task Load Index | Mental Demand | | German | |
| [18]   | Operator tasks on automated | RSME | Mental Load | 6 students | Netherland | |
true that more task allocated to human operators, a higher overall mental load perceived. The adapted RSME with a reduced number of descriptive labels appears to be used differently than the original RSME. Because a broader range of values is used, the adapted instrument may be more sensitive to changes in demand than the original one.

Galvanic skin response (GSR) is a psychophysiological phenomenon exhibited by skin containing sweat glands. It consists of a decrease in skin resistance or a decrease in impedance or a change in potential in response to an attention-getting or alerting stimulus [15]. GSR sensor measures sweat gland activity using silver chloride electrodes. The device can be attached to the wrist and ankle, or anywhere on the body with sweat glands. Many parameters can be found, such as skin conductance level [14][15] and mean GSR[16]. From the result, the GSR can be used to driving simulator [14], measuring attention performance [15], and measuring memory performance [16].

Salivary Amylase (SA) is a method that can quantify α-amylase activity in saliva (salivary amylase activity (SMA)) the psychological stress and to distinguish comfortable and uncomfortable physical states, that have been investigating and establishment [18], it can be used to determined what happen in human mind/ people, and then to make decisions maker about worked, whether he or she allowed to work or not. From the result, the salivary amylase can be used to distinguish eustress and distress [18] and measuring attention performance [16].

Heart Rate Variability (HRV) is a variation in the time interval between heartbeats. HRV usually analyzed based on frequency domain analysis with three frequency bands where each band is associated with the mechanism of the human body involving high frequency (HF) band (0.15 to 0.40 Hz), low frequency (LF) band (0.07 to 0.14 Hz), and very low frequency (VLF) band (0.02 to 0.06 Hz). From the result, HRV can be used to measuring attention [19] and memory performance [11][17][19] and flying simulator [20]. Beside HRV, Electroencephalogram (ECG), that reflects heart activity, has been introduced for driver workload assessment [21].

Electroencephalogram (EEG) signal is a representation of the brain's electrical activity recorded from electrodes placed on the scalp. The EEG spectral components, for example, theta (4–8 Hz) and alpha (8–12 Hz), are used to determine activity levels during different cognitive activities. From the result, [8] indicate that increased workload leads to increased frontal theta (fro-theta) activity and decreased parietal alpha (par-alpha) activity. Besides that, the EEG can be used to measuring attention performance [10], measuring memory performance [21][22], and using in driving simulator [21].

4.2. Subjective Measurement

The NASA-TLX method was developed by Sandra G. Hart from NASA Ames Research Center and Lowell E Staveland from San Jose State University in 1981[22]. This method in the form of a questionnaire was developed based on the emergence of needs of subjective measurements that are
easier but more sensitive to the measurement of workload. From the result, the NASA-TLX can be used to measuring memory performance [2][4] and driving simulator [4][23][24].

Rating Scale Mental Effort (RSME) is developed to evaluate workload. The uses are simple, quick, cheap, and requires no special equipment. From the result, the RSME can be used to measuring memory performance [4][6].

Subjective Workload Assessment Technique (SWAT) is a multidimensional scaling method designed to assess time load, mental effort load and stress load [25]. Time load reflects the amount of spare time available in planning, executing, and monitoring task. Mental effort load reflects how much conscious mental effort and planning are required to perform a task. Psychological stress loads measures the amounts of risk, confusion, frustration, and anxiety associated with task performance. From the result, the SWAT can be used to measuring memory performance.

5. Conclusion
The result of this study potentially will provide a significant contribution to the mapping of cognitive measurement research. There are two types of cognitive ergonomics measurement tools. Objective measurements consist of GSR, SA, eye tracker, HRV, and EEG, whereas subjective measurements consist of SWAT, NASA-TLX, and RSME. The tools can be combined as needed to measure cognitive load. These tools can be used for measuring cognitive performance. In Indonesia, cognitive measurement tool was developed. For objective measurement, there are GSR, HRV, and eye tracker that used to measuring memory performance and driving simulator. For subjective measurement, there are NASA-TLX and RSME that used to measuring memory performance.

Reference
[1] O'Donnell RD and Eggemeier FT 1986 Handbook of perception and human performance England 2 1-49
[2] Rubio S, Diaz E, Martin J and Puente J 2004 Applied Psychology: An International Review 53 61-86.
[3] Brookhuis KA, de Waard D 2001 Stress, Workload and Fatigue 321-33
[4] Widyanti A, Sofiani N F, Soetisna HR and Muslim K 2017 IJ Tech 8 283-291
[5] Tsang P, Wilson GF 1997 Mental workload in Salvendy G. editor. Handbook of Human Factors and Ergonomics (New York)
[6] Johnson A and Widyanti A 2011 Ergonomics. 54 509-518
[7] De Waard D 1996 The Measurement of Drivers’s Mental Workload (Den Haag: Traffic Research Center)
[8] Bridger RS 1995 Introduction to Ergonomics (Singapore: McGraw-Hill Book Co)
[9] Cegovnik T, Stojmenova K, Jakus G and Sodnik J 2018 Appl. Ergonomics 68 1-11
[10] Bernhardt K A, Poltavskia D, Petrosa T, Ferraroa F R, Jorgensonb T et al 2019 Appl. Ergonomic 77 83-91
[11] Widyanti A, Johnson A and Waard D 2014 Int. J.of Industrial Ergonomics 43 70-6
[12] McIntire L K, McKinley R A, Goodyear C, McIntire J P 2014 Appl. Ergonomic 45 354-62
[13] Damiano C, Walther DB 2019 Cognition 184 119-29
[14] Foy HJ and Chapman P 2018 Appl. Ergonomics 73 90-9.
[15] Nater UM, Rohleder N, Gaab J, Berger S and Jud A 2004 Int. J. of Psycho-Physiology. 55 333-42
[16] Zimmer P, Buttlar B, Halbeisen G, Walther E and Domes G 2019 Psychoneuroendocrinolog 101 186-92
[17] Widyanti A, Hanna, Muslim K and Sutalaksana Z 2017 Work 56 111-7
[18] Klaus K, Doerr JM, Strahler J, Skoluda N, Linnemann A and Nater UM 2019 Psychoneuroendocrinolog 101 80-6
[19] Luque CA, Perales JC, Cardenas D and Sanabria D 2016 Biological Psychology 113 83-90
[20] Mansikka H, Virtanen K, Harris D and Simola P 2019 Appl. Ergonomics 56 213-19
[21] Lei S and Roetting M 2011 *Human Factors: The Journal of the Human Factors and Ergonomics Society* 53 168-79
[22] Gevins A, Smith ME, Leong H, McEvoy L, Whitfield S, Du R and Rush G 1998 *The Journal of the Human Factors and Ergonomics Society* 40 79–91
[23] Hancock P and Meskhati N 1988 *Human Mental Workload* (North Holland: Elsevier Science)
[24] Heikkoop D, Winter J, Arem B and Stanton N 2018 *Transportation Research Part F* 58 193-209
[25] Schwallk M, Kalogerakis N, and Maier T 2015 6th Int. Conf. on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences 2466-73
[26] Reid B and Nygren T 1988 Advances in Psychology 52 185-218
[27] Waugh J, Glumm M, Kilduff P, Tauson R, Smyth C and Pillalamarri R 2000 *Proc. of the IEA 2000/HFES 2000 Congress* 44 276-79
[28] Desmet C and Diependaele K 2019 *Transportation Research Part F: Traffic Psychology and Behaviour* 60 549–59
[29] Toker D, Conati C, Steichen B and Carenini G 2013 *Proc. of the SIGCHI Conference on Human Factors in Computing Systems – CHI* 295-304
[30] Sudheesh N N and Joseph KP 2000 *ITBM-RBM* 21 158-63
[31] Wei Z G, Macwan A P, Wieringa P A 1995 *IFAC Man-Machine System*
[32] Widyanti A 2017 *IOP Conf. Series: Materials Science and Engineering* 277
[33] Widyanti A, de Waard D, Johnson A and Mulder B 2013 *Ergonomics* 56 82-194
[34] Ion A, Ilesiau D, Aldhafri S, Rana N, Ratanadilok K, Widyanti A and Nedelcea C 2017 *J. of Personality Assessment*. 99 25-34