

# Measuring cognitive load in a simulation game

B. Larraga-García<sup>1</sup>, V. Ruiz Bejerano<sup>1</sup>, Á. Gutiérrez<sup>1</sup>

<sup>1</sup> ETSI de Telecomunicación, Universidad Politécnica de Madrid, Madrid. España.  
{blanca.larraga, veronica.ruiz, a.gutierrez}@upm.es

## Abstract

*Cognitive Load (CL) measurement has gained interest across various domains as performance and efficiency can be affected if CL is not properly managed. The human cognitive structure has some limitations which are important to consider when accomplishing a task. Additionally, the study on how a primary task is impacted when a secondary task is accomplished has been studied. This affects to the working load memory and, therefore, to the CL. This work uses a tool, a multitasking game, that allows to infer cognitive load. This game considers a primary task and different types of secondary tasks: visual, memory and auditory. An experimental test has been performed using this game and objective, subjective and performance information has been gathered. The objective information was obtained thorough different sensors, the subjective information through the NASA-TLX and Paas questionnaires and the performance information through a score value provided by the game. An impact on performance on the primary task was perceived when secondary tasks were activated. Nevertheless, this impact is different depending on the type of secondary task accomplished. The memory task seems to impact the most on the performance on the primary task. Additionally, the subjective information did not correlate with any other variable gathered along the whole experimental test. With respect to the objective information, the eye measurements correlate with different physiological variables, showing that they provide reliable information with respect to measuring CL. Finally, interesting results on how the different difficulty levels impact performance is obtained.*

## 1. Introduction

Interest in measuring Cognitive Load (CL) has increased not only in the learning environment but also in many other domains such as surgery, driving, manufacturing or emergency services [1]. This is due to the fact that, an appropriate management of the CL may support performance and efficiency [2]. The CL Theory (CLT) began with the aim to solve problems on students' learning in the late 1970s. Since then, a lot of research has been made on this field providing information about how the human cognitive structure works. Considering this cognitive structure, the new information is managed by the working memory, which has a limitations in time and capacity. Nevertheless, with time, some information goes to the long-term memory in which such capacity and time limitations do not exist. Therefore, to avoid unnecessary information while performing a task will avoid to overload the working memory capacity. This involves also to limit distractions and to deliver the information in an appropriate manner so that the working memory is not affected. As a consequence, the CL will not increase. However, if this is not properly man-

aged, problems may arise. When mental task demand increases, people become unable to maintain adequate engagement in the task performed which leads to an increase in errors that, in some areas, might have an enormous impact. Nevertheless, it is not easy to measure the CL and different approaches have been taken in the literature. These approaches rely on either subjective, objective or performance information. The most frequent approaches for assessing CL have been subjective measurements [3, 4]. These measurements gather the information from users through specific questionnaires. Nevertheless, new approaches have been developed to integrate objective information, such as physiological data from the body [5, 6]. In these studies the impact on cardiovascular, temperature, respiration and electrodermal measures are presented. They state that these variables are the most popular in CL assessment. Additionally, the eyes provide valuable information according to [7], which states that the pupil diameter increases when CL increases. This is because the autonomous nervous system is affected when the CL increases and therefore, the pupil diameter changes accordingly. Moreover, other eye measures such as fixations and blinks are also of interest [8]. As the CL changes the number of fixations as well as the duration is affected; similarly, the number of blinks and their duration change. These measures have been proposed as the most reliable physiological information [2]. Finally performance measures, which consider how a task is performed, are also considered. This is specially studied when different tasks are performed simultaneously, as the human cognitive structure is composed of multiple interdependent modules which have a great potential but, they are limited to execute one task at a certain time, which is a human cognitive structure limitation. These studies focus on measuring how performance on a task is affected when another one is performed at the same time. This phenomenon is known as multitasking, and its influence on CL has also been extensively examined in the literature[9]. Considering all this, there are studies that find some correlations between the information provided by this three sources of data (subjective, objective or performance) whereas other do not; nevertheless, all of them agree that further research is needed. The aim of this study is to analyze the impact that a multitasking game that combines a primary and a secondary task has on CL by analyzing objective, subjective and performance information. Furthermore, the investigation will encompass an examination of how various types of secondary tasks influence performance on the primary task.

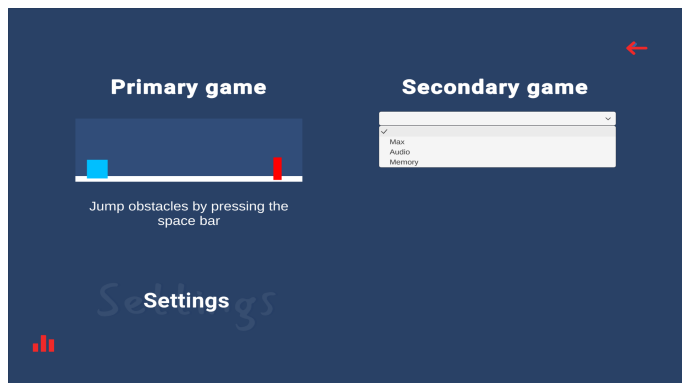
## 2. Materials

### 2.1. Participants

Seven healthy participants volunteered to participate in the study: 3 women and 4 men, with an average age of 28 years ( $SD \pm 6.8$ ). The total time of the study lasted 60 minutes. All the participants had normal or corrected-to-normal vision and were right-handed.

### 2.2. Software used

The software used within this study has been developed in Unity 3D [10]. The game presents a multitasking environment in which different types of tasks are implemented to be executed simultaneously. There is a primary task, which consists on a blue square jumping a red obstacle as shown in Figure 1.



**Figure 1:** Multitasking game in which the primary task and the different secondary tasks are shown.

A secondary task is implemented taking into account visual, auditory and memory inputs.

- **Visual:** the secondary task that deals with visual inputs presents a list of numbers and the participant needs to find the maximum one. This list of numbers changes considering if the answer provided is correct or not. If the answer is correct the range of number presented increase and if it is not correct, the range of numbers remains the same.
- **Memory:** the secondary task that uses the memory is based on the n-back task. A white square appears with a small blue square inside. This blue square could be placed in any of the four corners. The participant must remember the position of the small square of the first figure, for the 0-back task. Therefore, a sequence of figures appears and each of them must be compared with the first one to identify if they match or not. For the 1-back task, each figure must be compared with the previous one within the sequence of figures. And finally, for the 2-back task, each figure must be compared with the one that appeared two positions before.
- **Auditory:** the secondary task that deals with auditory input consists on listening to a list of figures and colours with the goal to identify the one that is repeated. The list will increase its length if the

answer is correct and it will remain with the same length if the answer is not correct.

Additionally, both primary and secondary tasks, have different difficulty levels: low, medium and high. This means that for the primary and secondary tasks that involves visual and auditory inputs, the speed in which the bar moves or in which the sequence of numbers or figures are displayed/listened changes. For the secondary task that involves memory, the difficulty changes from low (0-back task), to medium (1-back task) and to high (2-back task).

### 2.3. Hardware used

Several hardware devices are used to measure the vital signs and the eye measurements. The Zephyr BioModule inserted into the Zephyr strap gathers the respiration rate [11]. The Empatica wristband is used to gather the superficial temperature and the galvanic skins response [12]. The Fitbit Versa 2 gathers the heart rate [13]. Finally, the Pupil Core eye tracking platform shown in Figure 2 processes the information with respect to the pupil size, the blink and the fixations duration [14]. The Pupil Core eye tracker has three cameras: the world camera that records data at 120 Hz and two eye cameras that gathers all the eye-related information at 200 Hz. For this pilot study, the pupil size in millimeters as well as the duration of the blinks and the fixations, in seconds, are analyzed.



**Figure 2:** Pupil core eye tracker[14].

Finally, the subjective information is also gathered by using the NASA-TLX questionnaire [15] and the Paas scale [16]. The NASA-TLX questionnaire measures the workload experienced when a task is performed and takes into account six dimensions: mental, physical, temporal, effort, performance and frustration. It is currently the most used one. On the other hand, the Paas scale focuses only on the mental effort experienced. Even though it considers only one aspect, according to [17] this scale is appropriate to detect Cognitive Load changes.

## 3. Methods

The study accomplished within this work consisted of a set of games in which the different complexities of both, primary and secondary tasks were involved. Each participant performed a session composed of 18 trials of 0.65 minutes each of them. Every trial comprised either just the primary task or a primary and a secondary task at the same time. These trials were grouped in three blocks with 6 trials each of them. After each block, a

5-minutes break was made. The combination of trials were randomized and in Table 1, one example of a block of trials is shown.

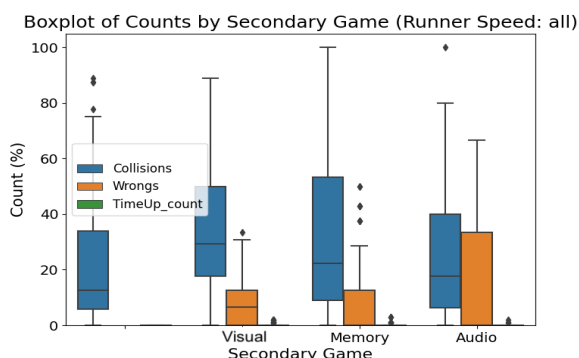
Trial	Task	Task	Task	Task
1	P1	P1 & V1	P1 & V2	P1 & V3
2	P1	P1 & M1	P1 & M3	P1 & M2
3	P1	P1 & A3	P1 & A1	P1 & A2
4	P3	P3 & M2	P3 & M1	P3 & M3
5	P3	P3 & A1	P3 & A2	P3 & A3
6	P3	P3 & V3	P3 & V1	P3 & V2

**Table 1.** A set of trials in which P corresponds to the primary task, V to the visual, M to the memory and A to the auditory secondary task. Additionally, the numbers 1, 2, 3 refer to the difficulty levels low, medium and high respectively.

Once a block was finished and before the 5-minutes break, all the participants fulfilled a NASA-TLX questionnaire which measures workload and a Paas scale to measure mental effort experienced.

#### 4. Results and discussion

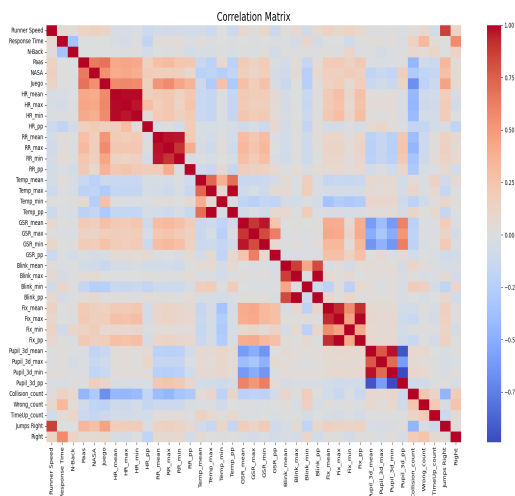
All the physiological information from the participants were gathered and analyzed. First of all, an analysis on the performance on primary and secondary tasks was performed. As shown in Figure 3, it is perceived that the primary task is affected when a secondary task is performed at the same time. In blue, the collisions which refer to the performance on the primary tasks increases when a secondary task is included. In fact, the secondary task that affects the most on the primary task performance is the visual task, followed by the memory task and finally by the auditory one. Nevertheless, when having a look at the performance of the secondary tasks, wrongs in Figure 3, it is seen that the auditory game seems to be the one that performed the worst. Finally, the Time-Up variable which counts the percentage of times that the answer is not provided within the given time is also shown. Nevertheless, its impact is almost imperceptible.



**Figure 3:** Primary and secondary task performance considering the percentage of collisions, errors and time-up events.

Then, a deep analysis on all the relationship between the physiological variables, the subjective information

provided by the questionnaires and the performance on the game is carried out. To accomplish this analysis, the values per physiological variable considered are: the maximum, the minimum, the average and the peak-to-peak values; and Figure 4 shows the results obtained. The subjective information is the one provided by the NASA-TLX and Paas questionnaires, and the performance information is provided by the score value obtained once a block of trials is finished. First of all, the questionnaire responses do not present a correlation higher than 0.5 with any other variable. This coincides with previous works, that highlight that the response to these questionnaires may vary a lot between individuals [4, 18] and that, additionally, the cognitive load varies along a task and these questionnaires are fulfilled once the task has finished [19]. Therefore, they do not estimate cognitive load in real time.



**Figure 4:** Correlations between all the objective and subjective information gathered from the whole session for all the participants.

The performance on the game correlates specially with heart rate and respiration rate increasing the score obtained on the game when the heart and respiration rates increase. This could be related with the motivation and excitement on the game [20]. Then, an analysis on the physiological variables is performed as they are able to detect changes along the game providing real-time information to assess cognitive load. Between these variables, some correlations are obtained. First of all, the change in the dimension of the pupil (pupil peak-to-peak variable) correlates with the galvanic skin response value as, according to [2] the pupil variation is an indicator of cognitive load as well as the galvanic skin response [21]. Additionally, the fixations duration correlates also with the heart rate and the galvanic skin response. Usually, when the cognitive load increases the fixation duration increases as the attention is focused in order to try to solve a task [22]. According to [5, 23] eye measures provide reliable information that can be measured easily by using a single device whereas to obtain similar

information from other physiological variables may require the use of more devices. On top of that, the right and wrong responses do not show any correlation with the objective, subjective and performance information. What it is found relevant is that when the response time increases there are more right answers. This makes sense as a higher response time means that there are more seconds given to the user to provide an answer. Therefore, if this time is higher, there is more time to provide an answer and the number of correct responses increases. What is interesting to see is that, the number of times that the obstacle is avoided increases with the speed of the obstacle. This was perceived during the sessions as to avoid the obstacle at lower speeds seems to be more complex as more attention is needed. With higher speeds, the jumps are more automatically performed.

## 5. Conclusions

This work has successfully measured objective, subjective and performance information when a multitasking activity is accomplished. Additionally, how a secondary task impacts on the performance of a primary task has been studied. Moreover, relationships between all these variables have been extracted. Of course, this is a preliminary work with some limitations such as the number of participants that should be increased to confirm the preliminary findings obtained within this work.

## Acknowledgements

This research has been supported by the European Commission within the context of the project RESCUER, funded under EU H2020 Grant Agreement 101021836. Additionally, it has been partially supported by the Universidad Politécnic de Madrid under the call First Projects 2023 and under the Técnicas de Control y Análisis de Variables Cinemáticas Aplicadas a la Extracción de Patrones de Movimiento project (RP2109560107).

## References

- [1] A.T. Duchowski, C. Biele, A. Niedzielska, K. Krejz, and Kiefer P. The index of Pupillary Activity. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1(1), 2018.
- [2] R. Cong, K. Tago, and J. Qun. Measurement and verification of cognitive load in multimedia presentation using an Eye Tracker. *Multimedia Tools and Applications*, 81(7), 2022.
- [3] G. Cierniak, K. Scheiter, and P. Gerjets. Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, 25(1), 2009.
- [4] J. Leppink, F. Paas, C. P. van der Vleuten, T. van Gog, and J. J. van Merriënboer. Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(1), 2013.
- [5] F. N. Biondi, B. Saberi, F. Graf, J. Cort, P. Pillai, and B. Balasingam. Distracted worker: Using pupil size and blink rate to detect cognitive load during manufacturing tasks. *Applied Ergonomics*, 106(1), 2023.
- [6] O. Anmarkrud, A. Andresen, and I. Braten. Cognitive load and working memory in multimedia learning: Conceptual and measurement issues. *Educational Psychologist*, 54, 2019.
- [7] S. Chen, J. Epps, and F. Chen. An investigation of pupil-based cognitive load measurement with low cost infrared webcam under light reflex interference. *Proceedings of 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1(1), 2013.
- [8] J. Hyönä, Radach R., and Deubel H. *The mind's eye cognitive and Applied Aspects of Eye Movement Research*. North-Holland, 2003.
- [9] Ö. Örun and Y. Akbulut. Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention. *Computers in Human Behavior*, 92:216–229, 2019.
- [10] Unity 3D. <https://unity.com/es>, 2023.
- [11] Medtronic. <https://www.zephyranywhere.com/system/components>, 2023.
- [12] Empatica. <https://www.empatica.com/research/e4/>, 2023.
- [13] Fitbit. <https://www.fitbit.com/global/es/products/smartwatches/versa>, 2023.
- [14] Pupil Labs. <https://pupil-labs.com/products/core/>, 2023.
- [15] NASA:B. Gore. <https://humansystems.arc.nasa.gov/groups/TLX/>, 2023.
- [16] F. Paas. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84:429–434, 1992.
- [17] P. Ayres. Something old, something new from cognitive load theory. *Computers in Human Behavior*, 113:106503, 2020.
- [18] A. Skulmowski and D.G. Rey. Measuring Cognitive Load in Embodied Learning Settings. *Frontiers in Psychology*, 8(2), 2017.
- [19] M. Ahmad, I. Keller, D. Robb, and Katrin Lohan. A framework to estimate cognitive load using physiological data. *Personal and Ubiquitous Computing*, 27(1), 2020.
- [20] G. Chanel, C. Rebetez, M. Betrancourt, and T. Pun. Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. *Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era*, 7(1), 2008.
- [21] N. Nourbakhsh, F. Chen, Y. Wang, and RA. Calvo. Detecting users' cognitive load by galvanic skin response with affective interference. *ACM Transactions on Interactive Intelligent Systems*, 7(3), 2017.
- [22] M. Holmqvist, K. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer. *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press, 2011.
- [23] S. Chen and J. Epps. Automatic classification of eye activity for cognitive load measurement with emotion interference. *Computer Methods and Programs in Biomedicine*, 110(2), 2013.