XLIV Jornadas de Automática 2023

XLIV Jornadas de Automática 2023 – Zaragoza
https://jautomatica.es

Cognitive Load measurement by using a concurrent multitasking simulation game

Larraga-García, Blanca.[a,*], Ruiz, Verónica.[a], Gutiérrez, Álvaro.[a]

[a] *Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad Politécnica de Madrid, 28040 Madrid, Spain*

To cite this article: Larraga-García, B., Ruiz, V., Gutiérrez,Á. 2023. Cognitive load measurement by using a concurrent multitasking simulation game.
XLIV Jornadas de Automática, 59-64. https://doi.org/10.17979/spudc.9788497498609.059

Resumen

La Teoría de la Carga Cognitiva establece que el proceso de aprendizaje o el desempeño de una tarea se ven afectados cuando hay una sobrecarga en la memoria de trabajo. Además, realizar múltiples tareas de manera simultánea también tienen un impacto en est sobrecarga. Este trabajo presenta una herramienta, un juego de multitareas, que permite inferir carga cognitiva. Se ha llevado a cabo una prueba experimental utilizando el juego y recogiendo información tanto subjetiva como objetiva de todos los participantes. Adicionalmente se analizó el impacto de aumentar la complejidad de la tarea primaria. Todas las variables fisiológicas se vieron afectadas por este cambio en la tarea primaria, especialmente la temperatura corporal y los movimientos de los ojos. Además, la información subjetiva proporcionada con el cuestionario NASA-TLX y la escala Paas también aumentaron con la complejidad de la tarea primaria. Finalmente, se obtuvo una importante correlación entre las dos medidas subjetivas.

Palabras clave: Carga cognitiva, Multitarea, Tarea primaria, Datos subjetivos, Datos objetivos

Abstract

The Cognitive Load Theory states that the learning process or the performance of a task is affected when there is an overload on the working memory. Additionally, concurrent multitasking has also an impact on this overload. This work presents a tool, a multitasking game that allows to infer cognitive load. An experimental test has been performed using the game and gathering subjective and objective information. Within this test, an analysis of the impact of increasing the complexity of the primary task has been performed. All the physiological variables were affected, especially the body temperature and the eye movements. Moreover, the subjective information provided with the NASA-TLX questionnaire and the Paas scale also increased with the complexity of the primary task. Additionally, a strong correlation between the two subjective measures was obtained.

Keywords: Cognitive load, Multitasking, Primary task, Subjective information, Objective information

[*] Autor para correspondencia: blanca.larraga@upm.es

1. Introduction

Cognitive Load Theory (CLT) emerged in the late 1980s and it is a theory that states that the best learning occurs when the learning conditions are in accordance with the human cognitive architecture (Cooper, 1990). This architecture considers cognitive structures such as the working, the long-term memory and how they are organized to handle information. The CLT sets that new information is first managed by the working memory which has a limitation in capacity and time; then, it goes to the long-term memory in which such limitations are not present (Mutlu-Bayraktar et al., 2019). Therefore, to guarantee a better learning, the information must consider the limitations of the working memory. This includes to avoid inadequate information, to deliver the information in an appropriate manner and to limit distractions during the learning process. Additionally, the better the working memory processes novel information, the faster it will be transferred to the long-term memory (Sweller et al., 2019). Consequently, to avoid an overload of the working memory will support the learning process, transferring the learnt information to the long-term memory which gathers the information for years or even a lifetime (Camina and Güell, 2017). Cognitive Load (CL) is the concept used to define the overload of the working memory (van Merriënboer and Kirschner, 2018). CL increases when the cognitive system perceives unnecessary information. There are three categories of CL on the working memory: intrinsic, extraneous and germane (Paas and Sweller, 2014). Intrinsic CL refers to the complexity of the task to be learnt; the more complex the task is, the higher intrinsic CL is perceived. The extraneous CL considers the adequacy of the learning material used; the content used to deliver a material is key and it should be carefully considered to support learning. Finally, the germane CL arises when mental structures are being developed; these structures will support the transfer from the working memory to the long-term memory. The capacity for the germane CL is the one left from the intrinsic and extraneous loads; therefore, paying attention to how to deliver a content considering its complexity is key to support the learning process on the working memory. Additionally, CL is affected by personal factors such as prior knowledge, emotional state, age, or experience (Hudlicka, 2003). Maintaining an appropriate CL is critical not only for learning purposes but also for workplace operations. If mental requests increase while performing a specific task, the primary task, the CL increases and the performance is affected by that load (Cohen et al., 2017). Multitasking and its impact on CL has been studied in the literature (Örun and Akbulut, 2019). Multitasking is defined as performing multiple activities. Nevertheless, there are some discussion on how to define multitasking and some classifications have been proposed (Colom et al., 2010), (Hambrick et al., 2009). These classifications differentiate between concurrent multitasking and sequential multitasking. Whereas in the concurrent multitasking the different tasks are performed at the very same time, in the sequential multitasking, different tasks are performed but they are performed sequentially. According to (Anderson, 2007) 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. Consequently, concurrent multitasking increases CL as stated in (Örun and Akbulut, 2019). With respect to CL measurement, there are several methods that have been developed but it is still a challenge to asses it. The most common methods to measure CL have been through subjective measures (Cierniak et al., 2009), (Leppink et al., 2013) whereas new approaches have been developed which incorporate objective information such as the physiological information from the body (Biondi et al., 2023), (Anmarkrud et al., 2019). Currently, there are studies that combine both, subjective and objective information to find the relationship between them (Minkley et al., 2021), (Makransky et al., 2019). There are studies that find some correlations whereas other do not; nevertheless, all of them agree that further research is needed.

The aim of this study is to develop a concurrent multitasking game that combines a primary and a secondary tasks to be able to measure CL by analyzing objective and subjective information. Additionally, the impact of different complexities of the primary task will be studied.

2. Methods

2.1. Simulation game design

A simulation game has been implemented in Unity 3D (3D, 2023). The game has been designed considering a multitasking exercise in order to measure how CL is affected when different tasks are performed simultaneously. Therefore, a primary task was designed to jump to avoid an obstacle and a secondary task that consists on selecting the maximum number from a set of different numbers (see Figure 1).

Figure 1: Screen in which the primary and secondary game are presented with a short explanation.

The primary game has been programmed to introduce different speeds for the obstacle. By doing this, the primary task can be customized in order to deal with an easier or more complex primary task. To avoid the obstacle, the space bar of the keyboard must be used. With respect to the secondary task, a set of numbers is shown with the objective to select the maximum number by using the mouse. The first set of numbers shown is between 0 and 9; if the number selected is correct, the upper bound is increased a tenth being the next set of numbers between 0 and 19 and so on. If the selected number is not correct, the set of numbers does not change. Additionally there is a time limit within an answer has to be provided, if not, the highest number of the set is highlighted in red as shown in Figure

2 and a new set of numbers is shown, which contains the same range of numbers than in the previous set.

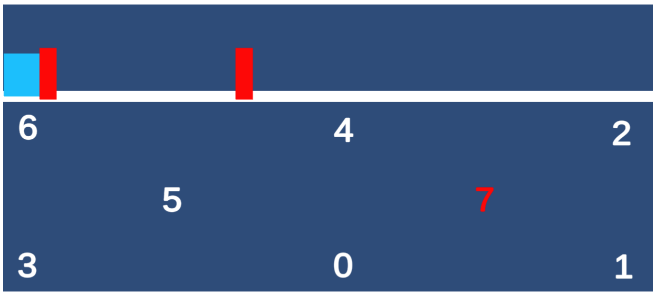


Figure 2: Example in which the number is highlighted in red as no response was provided by the user.

2.2. *Participants*

Ten healthy volunteers participated in the study: 7 men and 3 women, with an average age of 29 years (SD ±7.4). All participants were right-handed and had normal or corrected-to-normal vision. The test execution lasted 90 minutes.

2.3. *Equipment used*

In (Ayres, 2020) evidence is presented with respect to the best physiological information to detect changes in CL: cardiovascular measurements, respiratory, eye activity, electrodermal and temperature measurements. Additionally, each physiological measure could be more or less sensitive to CL (Minkley et al., 2021). The cardiovascular measure used is the heart rate (HR), which is sensitive to changes in the mental workload increasing when the workload increases (Solhjoo et al., 2019).The respiratory measure used is the breathing rate (BR) as it is also directly related to CL (Hidalgo-Muñoz et al., 2019).The eye gaze movement is also an important aspect when measuring CL as the movements of the eyes provide useful information in addition to the previous measurements when CL is analyzed (Chen and Epps, 2013). More frequent saccades imply a higher CL (Biondi et al., 2023). Finally, the body temperature and the electrodermal activity have also an important correlation with CL (Kimura et al., 2020) as the temperature increases when the CL increases. Therefore, the physiological variables gathered for this study are: heart rate (HR), breathing rate (BR), galvanic skin response, body temperature and eye gaze movement. To acquire this information the following equipment has been used:

- Chest band: the tool to measure heart rate and breathing rate is Zephyr BioModule (Medtronic, 2023) inserted into the the Zephyr strap. The range for the heart rate is between 0 and 240 beats per minute and for the breathing rate, between 0 and 120 breaths per minute. The transmission frequency for each value is 1 Hz.

- Wrist band: the tool to measure body temperature and galvanic skin response is the Empatica(Empatica, 2023) wristband. The electrodermal sensor measures the skin conductance in microSiemens, transmitting the data four times per second. Body temperature is measured, in Celsius, using an infrared thermometer. The data are also transmitted four times per second. Additionally, a Fitbit Versa 2 (Fitbit, 2023) is included as a second wristband to gather the heart rate measurements due to problems with the heart rate gathered with the Zephyr strap.

- Hololens: the tool to gather the eye gaze information is the Microsoft Hololens (Microsoft, 2023) and all the tools needed for the development and deployment of this tool. The Hololens gather information about the movement of both eyes, from the origin to the direction of the gaze. When using the Hololens, they must be calibrated for each participant. Additionally, the Hololens store all the data to avoid sending all the information continuously.

The Zephyr BioModule, the Empatica and the Fitbit shown in Figure 3 send the data to a mobile phone. Then all that data together with the data from the Hololens are sent to an embedded computer following MQTT and HTTP protocols. Additionally, a web interface to control the data acquisition is created as well as a MongoDB database. A dockerized system has been created with four docker containers: the docker-phone, the docker-holo, the docker-flask for the web interface and the docker-mongo (Swani and Tyagi, 2017).

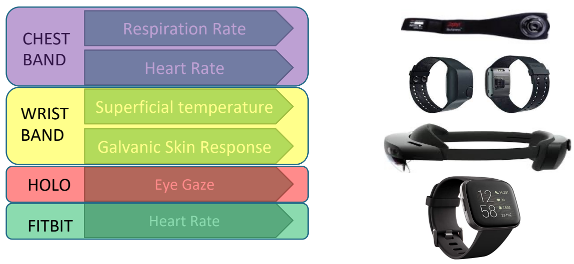


Figure 3: Equipment used to measure the physiological variables. They are presented in order: Zephyr strap, Empatica wristband , Hololens and Fitbit Versa 2.

The subjective information gathered is through two questionnaires: the NASA-TLX questionnaire (Hart and Staveland, 1988),(Hart, 2006) and the Paas scale (Paas, 1992). The NASA-TLX questionnaire is a multidimensional tool to assess workload and consists of two different parts: the first one consists of rating from 0 to 100 six different categories (Mental Demand, Physical Demand, Temporal Demand, Performance, Effor and Frustation); the second one consists of selecting, between two of the categories previously mentioned, one of them for all the possible combinations. The NASA-TLX test has been selected as it is the most used nowadays (Gore, 2023). The Paas scale allows participants to rate their mental load after performing a specific task. It has been used considering its simplicity and that, even though it is a uni-dimensional scale, it is appropriate to detect changes in CL (Ayres, 2006), (Gimino, 2002).

2.4. Experimental Procedure

Each participant performed one session with 18 blocks and each block contained four trials as shown in Table 1. Each trial lasted 0.75 minutes. After each six blocks, a NASA-TLX and Paas questionnaire was fulfilled and then, a 5-minutes break was made. The primary task was tested for three different speeds: A is the lowest, B is the intermediate and C is the highest one. The secondary task was tested also for three different speeds. That is the speed for the set of numbers change: 1 is the lowest speed, 2 is the intermediate speed and 3 is the highest one. A calibration screen is used prior to the execution of the primary and secondary tasks. The aim of this calibration screen is to gather the eye gaze movement of each participant to filter the results and make them comparable amongst participants. Once the calibration is performed, the session starts providing a first screen with some explanations as shown in Figure 1. Then, when the start button is selected, the trials start. The first trial of each block starts always with the primary task and then, the rest of the trials were made randomized across participants (Table 1 is shown as an example). The total time of the whole session was 90 minutes.

Table 1: A set of trials performed by a random user

Block	Trials			
1	A	A+1	A+2	A+3
2	A	A+3	A+1	A+2
3	A	A+3	A+1	A+2
4	A	A+2	A+3	A+1
5	A	A+2	A+3	A+1
6	A	A+1	A+2	A+3
5 minutes break				
Block	Trials			
1	B	B+1	B+2	B+3
2	B	B+3	B+1	B+2
3	B	B+3	B+1	B+2
4	B	B+2	B+3	B+1
5	B	B+2	B+3	B+1
6	B	B+1	B+2	B+3
5 minutes break				
Block	Trials			
1	C	C+1	C+2	C+3
2	C	C+3	C+1	C+2
3	C	C+3	C+1	C+2
4	C	C+2	C+3	C+1
5	C	C+2	C+3	C+1
6	C	C+1	C+2	C+3

3. Results

With respect to the subjective information, the results obtained from both questionnaires are quite similar, showing a correlation value of 0.66 which means that there is a strong and positive correlation between these two subjective methods. In Figure 4, NASA-TLX score increases according to the complexity of the primary task. Nevertheless, the dispersion on the answers is quite diverse.

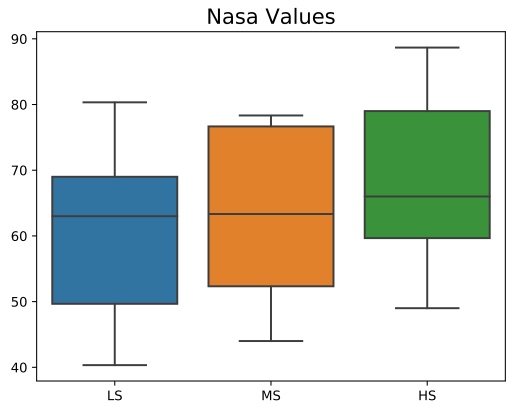


Figure 4: NASA-TLX score obtained from the participants when the primary task speed is the lowest (LS), intermediate (MS) and the highest speed (HS).

With respect to the objective information, Figure 5 shows a sample of the data obtained by the chest band along time is shown, both the heart rate and the respiratory rate. It is possible to perceive the problems mentioned with respect to the heart rate of this device. Due to this fact, the Fitbit gathered the heart rate information as previously explained.

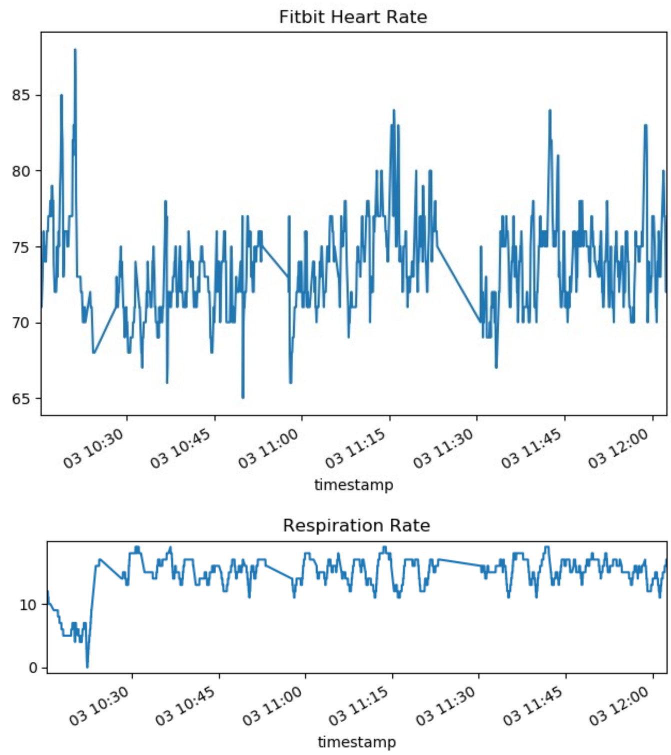
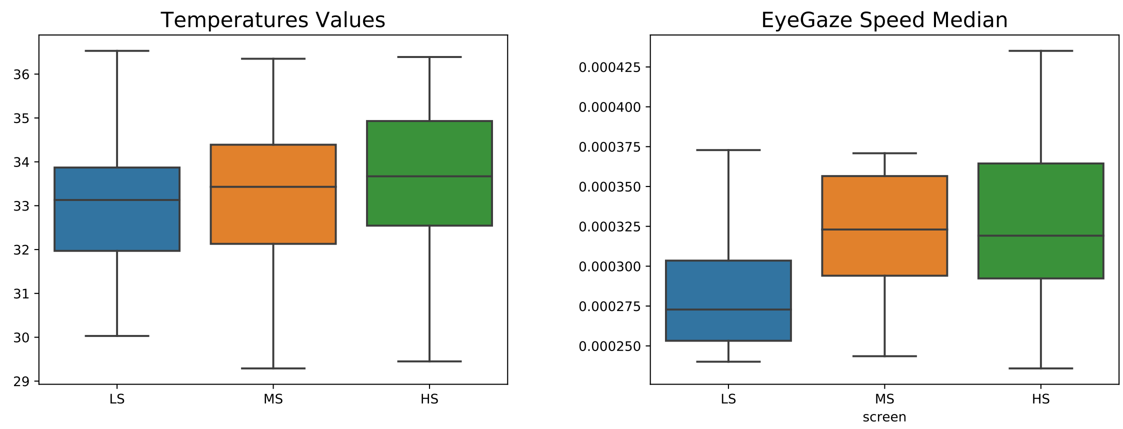


Figure 5: Heart rate and respiration rate obtained for a complete set of 6 blocks by a random user.

The rest of the physiological information was obtained in the same way. Then, an analysis is performed to analyze the changes of the physiological variables with respect to the speed of the primary task. The variables that show a higher change when the speed is increased are the body temperature and the eye gaze speed. Figure 6 (a) shows how the body temperature



(a) (b)

Figure 6: Change of physiological variables when the primary task speed is the lowest (LS), intermediate (MS) and the highest speed (HS): (a) Temperature, (b) Eye gaze

increases while the primary task becomes more complex ; additionally, in Figure 6 (b) the median eye gaze speed is shown which value increases specially between the low speed and the intermediate speed of the primary task.

4. Discussion

The primary goal of this study was to develop a multitasking game that allows to measure the impact of concurrent multitasking in CL. This goal has been achieved, being able to perform the game and to obtain both, subjective and objective information, from all the participants involved in the pilot study. The analysis performed focused on the impact on all the variables when the primary task complexity is changed increasing the speed. With respect to the subjective information, is has been found a clear evolution when the complexity of the primary task increased. This means that the participants felt the complexity when the speed of the primary task increased. Additionally, the NASA-TLX and the Paas questionnaire provided the same information with a high correlation value. This confirms that the Paas questionnaire, a uni-dimensional scale provides similar information to the NASA-TLX questionnaire. From all the physiological data obtained, it is perceived that there is an increase in all the vital signs when the speed of the primary game complexity is increased. The eye gaze movements gathered from the Hololens have been used to calculate the speed of the movements of each of the participants. According to Figure 6 (b), it is perceived that there is an important difference between the median speed of the eye gaze when the speed of the primary game increases. Even though, there is no statistical significance, the p-value obtained is 0.058. Nevertheless, the eye movement increases as the CL is increasing including a more demanding primary task. With respect to the highest speed of the primary game, it is noted that the median value of the eye gaze does not improve but the dispersion of the values increases, showing that some participants have increased their eye gaze speed compared to a less demanding primary task. With respect to the temperature, this increase is also perceived in Figure 6 (b) even though, no statistical difference is obtained. The heart rate, the respiratory rate and the galvanic skin response are also affected but, whereas a change is clearly noticed between the lowest and the intermediate speeds; it is not always perceived for the highest speed.

5. Conclusions

This work has designed and implemented a simulation game to measure CL when there is a concurrent multitasking. Additionally, real-time physiological variables have been obtained during the experimental tests. These tests allowed to confirm that all the structure built to gather, send and process vital signs of the participants worked correctly while performing the multitasking game. Moreover, the subjective questionnaires were requested finding that, from the two tools used, both of them seem appropriate to measure CL. With respect to future work, the number of participants must increase in order to have a higher amount of data that will allow to further analyze vital signs impact and correlation.

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 Politécnica de Madrid under the call "First Projects".

References

3D, U., 2023. https://unity.com/es.

Anderson, J. R., 2007. How can the human mind occur? How Can the Human Mind Occur in the Physical Universe? 1, 237–248.

Anmarkrud, O., Andresen, A., Braten, I., 2019. Cognitive load and working memory in multimedia learning: Conceptual and measurement issues. Educational Psychologist 54, 61–83.
DOI: 10.1080/00461520.2018.1554484

Ayres, P., 2006. Using subjective measures to detect variations of intrinsic cognitive load within problems. Learning and Instruction 16, 389–400.
DOI: 10.1016/j.learninstruc.2006.09.001

Ayres, P., 2020. Something old, something new from cognitive load theory. Computers in Human Behavior 113, 106503.
DOI: 10.1016/j.chb.2020.106503

Biondi, F. N., Saberi, B., Graf, F., Cort, J., Pillai, P., Balasingam, B., 2023. Distracted worker: Using pupil size and blink rate to detect cognitive load during manufacturing tasks. Applied Ergonomics 106, 103867.
DOI: 10.1016/j.apergo.2022.103867

Camina, E., Güell, F., 2017. The neuroanatomical, neurophysiological and psychological basis of memory: Current models and their origins. Frontiers in Pharmacology 8, 1–16.

Chen, S., Epps, J., 2013. Automatic classification of eye activity for cognitive load measurement with emotion interference. Computer Methods and Programs in Biomedicine 110, 111–124.

Cierniak, G., Scheiter, K., Gerjets, P., 2009. 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, 315–324.
DOI: 10.1016/j.chb.2008.12.020

Cohen, J., LaRue, C., Cohen, H. H., 2017. Attention interrupted: Cognitive distraction & workplace safety. Professional Safety 62, 28–34.
DOI: https://www.jstor.org/stable/10.2307/48687645

Colom, R., Martínez-Molina, A., Shih, P., Santacreu, J., 2010. Intelligence. Computers in Human Behavior 38, 543–551.

Cooper, G., 1990. Cognitive load theory as an aid for instructional design. Australasian Journal of Educational Technology 6, 1–10.

Empatica, 2023. https://www.empatica.com/research/e4/.

Fitbit, 2023. https://www.fitbit.com/global/es/products/smartwatches/versa.

Gimino, A., 2002. Students' investment of mental effort. Proceedings on the annual meeting of the American Educational Research Association 1, 1–10.

Gore, N., 2023. https://humansystems.arc.nasa.gov/groups/TLX/.

Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., Brou, R., 2009. Intelligence. Applied Cognitive Psychology 24, 1149–1167.

Hart, S., 2006. Nasa-task load index (nasa-tlx); 20 years later. PsycEXTRA Dataset 1, 1–6.
DOI: 10.1037/e577632012-009

Hart, S., Staveland, L. E., 1988. Development of nasa-tlx (task load index): Results of empirical and theoretical research. Advances in Psychology 1, 139–183.
DOI: 10.1016/s0166-4115(08)62386-9

Hidalgo-Muñoz, A., Béquet, A., Astier-Juvenon, M., Pépin, G., Fort, A., Jallais, C., Tattegrain, H., Gabaude, C., 2019. Respiration and heart rate modulation due to competing cognitive tasks while driving. Frontiers in Human Neuroscience 12, 1–8.

Hudlicka, E., 2003. To feel or not to feel: The role of affect in human–computer interaction. International Journal of Human-Computer Studies 59, 1–32.

Kimura, T., Takemura, N., Nakashima, Y., Kobori, H., Nagahara, H., Numao, M., Shinohara, K., 2020. Warmer environments increase implicit mental workload even if learning efficiency is enhanced. Frontiers in Psychology 11, 1–14.

Leppink, J., Paas, F., der Vleuten, C. V., Gog, T. V., Merriënboer, J. J. V., 2013. Development of an instrument for measuring different types of cognitive load. Behavior Research Methods 45, 1058–1072.
DOI: 10.3758/s13428-013-0334-1

Makransky, G., Terkildsen, T. S., Mayer, R. E., 2019. Role of subjective and objective measures of cognitive processing during learning in explaining the spatial contiguity effect. Learning and Instruction 61, 23–34.
DOI: https://doi.org/10.1016/j.learninstruc.2018.12.001

Medtronic, 2023. https://www.zephyranywhere.com/system/components.

Microsoft, 2023. https://www.microsoft.com/en-us/hololens/.

Minkley, N., Xu, K. M., Krell, M., 2021. Analyzing relationships between causal and assessment factors of cognitive load: Associations between objective and subjective measures of cognitive load, stress, interest, and self-concept. Frontiers in Education 6, 1–15.
DOI: 10.3389/feduc.2021.632907

Mutlu-Bayraktar, D., Cosgun, V., Altan, T., 2019. Cognitive load in multimedia learning environments: A systematic review. Computers & Education 141, 103618.
DOI: 10.1016/j.compedu.2019.103618

Örun, O., Akbulut, Y., 2019. Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention. Computers in Human Behavior 92, 216–229.

Paas, F., 1992. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. Journal of Educational Psychology 84, 429–434.
DOI: 10.1037/0022-0663.84.4.429

Paas, F., Sweller, J., 2014. Implications of cognitive load theory for multimedia learning. The Cambridge Handbook of Multimedia Learning 1, 27–42.
DOI: 10.1017/cbo9781139547369.004

Solhjoo, S., Haigney, M., McBee, E., van Merrienboer, J., Schuwirth, L., Artino, A., Battista, A., Ratcliffe, T., Lee, H., Durning, S., 2019. Heart rate and heart rate variability correlate with clinical reasoning performance and self-reported measures of cognitive load. Scientific Reports 9, 1–9.

Swani, L., Tyagi, R., 2017. Dockerization (replacement of vms=. International Research Journal of Engineering and Technology 4, 1–6.

Sweller, J., van Merriënboer, J. J. G., Paas, F., 2019. Cognitive architecture and instructional design: 20 years later. Educational Psychology Review 31, 261–292.
DOI: 10.1007/s10648-019-09465-5

van Merriënboer, J., Kirschner, P., 2018. Ten steps to complex learning: a systematic approach to four-component instructional design. Routledge, New York.