The Cognitive Effect of Spatial Contiguity in Procedural Training Using Mixed Reality

Sara Gutiérrez, Industrial University of Santander, Colombia
Luis Bautista, Industrial University of Santander, Colombia
Fernanda Maradei, Industrial University of Santander, Colombia
Julián Núñez, Industrial University of Santander, Colombia

The European Conference on Arts, Design & Education 2022
Official Conference Proceedings

Abstract
New technologies such as Mixed Reality (MR) are often used in procedural training to provide information to the trainee. When a trainee is faced with a visual scene with learning material, MR shows the information as a set of visual and/or auditory resources. Thus, the apprentice must perceive which resources are directly associated with the task to be executed. For this, the visual attention guide components are important, which have an impact on the improvement of information processing and on the optimization of cognitive resources. The objective of this study is to determine the cognitive effect of one of the visual attention techniques called Spatial Contiguity on students engaged in a procedural training scenario using Mixed Reality. Through a systematic literature review (SLR), related to the use of graphical user interfaces visual elements or techniques to guide visual attention in training, it was found that there are certain techniques of visual elements that guide the user's attention. Spatial contiguity, color codification, movement, and blinking are the most used. In organizing observed information, cognitive processes beyond attention influence the direction and duration of eye movements; because of this to analyze visual attention processes in the study Eye-tracking is implemented, which is a tool used to record the eye movements of subjects while they perform tasks, this allowed to measure cognitive processing of stimuli from learning materials presented. Statistically significant differences are expected to be found in measures of cognitive processing, based on different forms of spatial contiguity presentations.

Keywords: Spatial Contiguity, Procedural Training, Mixed Reality, Cognitive Load, Eye Tracking

1. Introduction

Learning involves a cognitive process in which an individual obtains and puts into practice knowledge [1] and procedural training refers to the training of how to perform a specific skill or task, and is considered knowledge related to methods, procedures, or operation of equipment [2], in relation to this, cognitive load measures have been shown to be an important factor in establishing a relationship between the learning performance of students and cognitive load as an indicator of their mental efforts. Eye-tracking data provides information about eye movements, the areas where people focus their attention, the information they ignore, and the objects that bother them [3]. The purpose of the developed experiment is to determine the cognitive effect of spatial contiguity in procedural training using mixed reality (MR) based on cognitive processing measures of the student during a learning activity through the implementation of the principle of multimedia learning design: Spatial Contiguity [4], showing the presentation of learning material through graphics and text which are separated from each other within the visual field of 27°, 45°, 60° or 75°; this by using Eye-Tracking.

The experiment is based on comparing four groups of four participants each who learned with knee anatomy material for taking direct measures of cognitive processing, such as the duration of fixation, the number of fixations and saccade velocity. For this study, we start from the assumption analyzed by Johnson and Mayer [5], where it is indicated that the principle of spatial contiguity establishes that people learn better when the corresponding words and images are presented close to each other instead of far from each other in the visual scene observed by the student. Therefore, it is inferred that the cognitive processing measures will show a statistically significant difference for the material shown through different separation between the graphic and the text in the procedural training, in turn, the objective measures of cognitive processing will be lower for the material presented in a contiguous configuration. The results showed, first, for the objective measure of the number of fixations lower values for the learning material presented with discontiguous content and for the objective measure of total time of the task development, higher values for the learning material presented with discontiguous content, and second, for the measurements of the duration of fixations and saccade velocity, lower values for learning material presented with contiguous content.

2. Theoretical background

2.1 Cognitive load

It refers to the resources used by working memory and that affect the student while performing a cognitive task [6][7]. According to cognitive load theory (CLT), there are three types of cognitive load that interfere with learning: (a) intrinsic load caused by the inherent complexity of instructional information, (b) germane load directly related to schema construction and automation, and (c) extraneous load caused by instructional elements that are unnecessary for learning [8]. The intrinsic and extraneous load are additive, together they determine the total cognitive load imposed by the learning material, which determines the working memory resources necessary to process the information. Currently, researchers and professionals of visual computing want to reduce the extraneous cognitive load so that most of the working memory resources can be dedicated to learning through the correct design of instructional material [5].

2.2 Mixed Reality in procedural training

In the article “What Is Mixed Reality, Anyway? Considering the Boundaries of Mixed Reality in the Context of Robots”, the authors refer to MR as an application of human-computer interaction to combine virtual and real-world elements [9]. Likewise, Melanie J. Maas and Janette M. Hughes in their bibliographic review "Virtual, augmented and mixed reality in K–12 education: a review of the literature" comment that MRI allows to offer a means to incorporate the complete body with real elements and virtual through the continuum of reality [10], additionally, Enrico Costanza, Andreas Kuns and Morten Fjeld mention that MR systems are designed in such a way that they give users the illusion that digital elements they are in the same space as physicists [11].

Due to the technological development that has been presented today, MR applications are increasingly implemented in human-machine interfaces, education and training. In the document “Mixed Reality in Learning Factories” [12], the authors comment that the use of RM can improve the delivery of knowledge and skills, since it helps to understand processes, data, methods and systems. The implementation of this technology favors production processes due to increased productivity, reduced downtime and improved employee safety [13]. Additionally, another field in which valuable advances have been made is in medicine, especially in surgical training, as for example in an implementation carried out in the Department of Orthopedics at Tongji University, Sahngai, China, whose findings can be seen in the document "Mixed Reality-Based Preoperative Planning for Training of Percutaneous Transforaminal Endoscopic Discectomy: A Feasibility Study" [14], where the objective was to explore the effect of preoperative planning using MR in training of percutaneous transforaminal endoscopic discectomy, the researchers concluded that an effective and repeatable training method was needed to help inexperienced surgeons, so they relied on MR technology.

2.3 Evaluation of cognitive load using Eye-Tracking

Objective cognitive load measurement methods are of great importance for research on learning, since they measure cognitive load while it occurs [15], that is, while the participant is observing the stimulus, they do not present interruption in the learning processes for the assessment of cognitive activity and load. One of the objective methods for evaluating cognitive load is Eye-Tracking, which focuses on capturing eye movements [16]. Eye-Tracking allows for very detailed analysis, as it provides a deep insight into the processing of human information regarding the allocation of visual attention and cognitive activity in the process and integration of learning information presented in text or image [15].

2.4 Measurement of cognitive load using Eye-Tracking

Eye-Tracking allows the identification of fixations, saccades, pupil dilation and blinking. These movements provide evidence of voluntary and open visual attention, because the goal of eye movement measurement and analysis is to obtain information about the attentive behavior of the viewer [17]. However, this study will focus on the analysis of fixations and saccades, as these are the basic unit of data for most Eye-Tracking analyzes [18]. The increase in fixation duration, the number of fixations and saccade velocity show a higher level of cognitive load, which indicates a greater effort in processing the learning material and greater attention resources [19].

2.4.1 Fixations

They are voluntary movements and correspond to a focused state in which the eye remains immobile for a period of time, lasting from 200-300 milliseconds to several seconds. The number of fixations indicates the number of times a user looked at a given area of the stimulus. Duration of fixation has been related to the level of cognitive processing with a high duration of fixation indicating greater tension on working memory [20].

2.5 Spatial Contiguity Principle

Spatial Contiguity Principle indicates that students learn better when the corresponding words and pictures are presented close to each other than when they are presented far from each other, since students do not have to use cognitive resources to do a large visual search in the scene and are more likely to be retain the information presented in the working memory at the moment [4]. The learning material can be presented in different ways, one of them is a version with integrated content or contiguous, in which the words and images are presented as close as possible or guiding the student from the text to the graphic or vice versa, so that encourages students to build mental connections with each other. Students do not have to search to find a graphic that corresponds to displayed text; therefore, they can dedicate their cognitive resources to active learning processes, including building connections between words and pictures. In a presentation with discontiguous or non-integrated content, words and images are far from each other in the scene, as when text appears in a section of the visual scene and graphic in a separate section. Therefore, it is inferred that when estimating cognitive processing measures using Eye-Tracking for material presented with contiguous content, these will be lower than for material with discontiguous content [5].

In relation to the additional configurations, a study carried out by Ramona E. Su y Brian P. Bailey [21] will be taken into account, who studied different screen configurations and how the physical separation between screens affected performance, subjective workload and satisfaction, in the study carried out, they reached the conclusion that the screens should be placed at an angle of 45 ° or less to each other, as this allows the user to see the screens more easily, so for the proposed study, it will be taken into account for the contiguous presentation configuration at most a width of 45° of separation between the graphic information and the text, and as discontiguous a maximum width of 75°.

3. Method

3.1 Ethical Implications

The study requires participants to use a head-mounted Eye-Tracking device and a Mixed Reality Glasses (Hololens 2), however it is considered non-invasive and the risk to participants is classified as minimal. Each participant received and signed an informed consent document. To preserve the confidentiality of the information, they were previously identified with coded serial numbers.

3.2 Experiment

An experimental design was developed, and the elements can be observed in Table 1. The presentation of the learning material was done through the Hololens 2 and the capture of cognitive load data using Eye-Tracking was proposed in a controlled environment, where the participants were randomly assigned a learning material with a type of content, in order to measure the load cognitive and make a comparison between the four ways of presenting the learning material.

Participants	16 participants (9 female and 7 male)
Taks	Procedural training activity about the anatomy of the knee
Stimuli	Random assignment of stimuli with different separation: 27°, 45°, 60 and 75°
Independent Variables	- Procedural learning activity with 27° separation between graphics and text. - Procedural learning activity with 45° separation between graphics and text. - Procedural learning activity with 60° separation between graphics and text. - Procedural learning activity with 75° separation between graphics and text.
Dependent Variables	Eye-Tracking metrics: fixation duration.

Table 1: Experiment Design

The proposed study was established as a quasi-experimental research project, it is a prospective, cross-sectional study, and finally, it is proposed with a descriptive scope.

3.3 Apparatus

A setup was determined on a MR glasses (Hololens 2) shown in the Figure 1 to display the stimulus as the Eye-Tracking data was captured. A head-mounted ocular tracker SMI shown in the Figure 2 was used, implemented to allow free mobility of the participants. SMI software was used to set up the experiment and compile the results: Experiment Center and BeGazeTM 3.7.

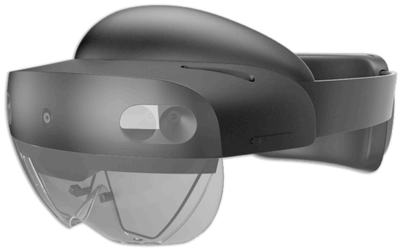

Figure 1: Microsoft Hololens 2

Figure 2: SMI Eye-tracking Glasses

3.4 Participants

This study was developed with volunteer participants who had no experience or knowledge related to the capture of measures for the evaluation of cognitive load through Eye-Tracking. In total, 16 participants, 9 female and 7 male subjects between the ages of 18 and 30 participated, who participated under the same environmental conditions.

3.5 Stimuli

In this study, four stimuli were used which present the learning material in four different forms, where information is provided about the knee and how the perforations of the portals are made to introduce the tools during an arthroscopic intervention of knee. In the first stimulus, the graphic information is separated from the text by 27° of amplitude in the visual field, in the second the separation is 45°, in the third it is 60° and finally in the fourth it is 75°. The learning material shown can be observed below in the Figure 3.

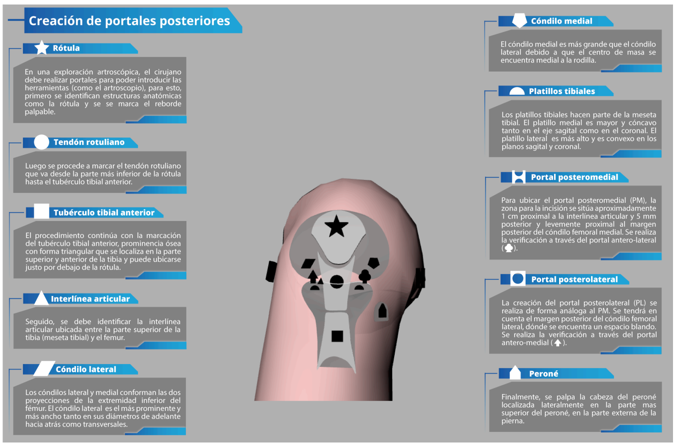

Figure 3: Procedural training activity about the anatomy of the knee

3.6 Procedure

Before the development of the test, each participant is informed about the risks and conditions of the test, and later they were asked to complete a survey about previous knowledge on the learning topic, through a Likert scale, where the score of "1" referred to not having knowledge at all with the subject, and "5" having a lot of knowledge about the subject. For this pretest, all of the participants affirmed not having knowledge related to the subject of the learning material. During the development of the test, the participant puts on the MR and SMI glasses and the Eye-Tracking system is calibrated; later, the participant is located in front of real knee model where the 3D information is superimposed, where the stimulus assigned in a random way is shown and the Eye-Tracking data is captured while the participant performs the recognition of the stimulus. The participants did not receive any type of financial reward and took part in the study freely and voluntarily. Below in Figure 4, a participant is shown taking the test.

Figure 4: Participant taking the test

3.7 Hyphotesis

For the hypothesis below, the variables mentioned in Table 1 were taken into account.

•H1: There is a significant difference in the fixation duration and saccade velocity for the learning material presented with different forms of spatial contiguity presentations.

3.8 Definition of variables

3.8.1 Independent Variables

As mentioned above, four presentations are used with learning material, in the first presentation, the graphic information is separated from the text by 27° of amplitude in the visual field, in the second the separation is 45°, in the third it is 60° and finally in the fourth it is 75°.

3.8.2 Dependent Variables

•Duration of fixations: Time in which the participant focuses the eye and it remains immobile.

4. Data analysis and results

The experimental design is considered as unifactorial, since it aims to study the influence of the independent variable on the response variable. The equation Yi presents the model.

$$Yi=\mu+\tau\_i+\varepsilon\_i$$

4.1 Descriptive analysis

By general inspection of the data, the results shown in Table 2 are observed.

Stimuli	Average of Duration of fixations [s]
27	0,270
45	0,277
60	0,280
75	0,277

Table 2: Experiment data

An analysis of descriptive statistics related to the duration of the fixations was developed, where it can be observed that the mean measure is greater for the learning material displayed in 60° with a value of 0.28[s] compared to the value obtained for the material displayed in 27° of 0.27 [s], which indicates that the longer the fixations last, there will be a greater effort on the working memory for the 60° configuration. According to the data obtained, it can be inferred that the visualization of discontiguous content could generate greater effort on working memory, as well as greater difficulty in the task.

4.2 Statistical analysis

4.2.1 Durations of Fixations

A normality test of the data of Duration of Fixations was developed, for this a Shapiro- Wilk normality test was executed, for which the following statistical hypotheses are established:

• H0: The distribution of the data is normal value significance p-value> 0.05
• H1: The distribution of the data is different from the normal p-value <0.05

For the analysis of the Durations of Fixations data, normality test yielded a p-value = 0.0, so H0 is rejected and H1 is accepted, which indicates that the distribution of the data it is different from normal.

Therefore, the data will be handled as non-parametric, using the Kruskal-Wallis test and establishing the following statistical hypotheses:

•H0: There are no statistically significant differences in the data p-value> 0.05
•H1: There are statistically significant differences in the data p-value <0.05

This result yields a p-value = 0.000, therefore, it is determined that there are statistically significant differences between the comparisons of pairs of stimuli. An Error Diagram of Duration of Fixations was made through visual analysis to determine which are the possible significant differences, as can be seen in Figure 5. A logarithmic curve behavior was expected but a sine curve behavior is obtained as a result, which does not show a constant behavior of error growth or decrease.

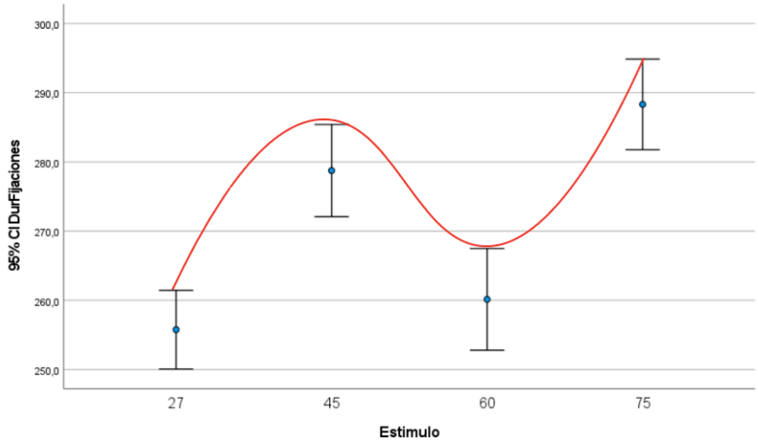

Figure 5: Error Diagram of Duration of Fixations

To determine in which of the pairwise comparisons there is a statistically significant difference, a Post-Hoc Tukey analysis was carried out. The p values obtained for each group of couples can be seen in Table 3. In the analyzed data it can be seen that the p-value <0.005 for all groups, so we can say that there are statistically significant differences between all groups.

Compared groups	p-value
60-27	0.006
60-45	0.000
60-75	0.000
27-45	<0.001
27-75	0.000
45-75	0.002

Table 3: P-value for Post-Hoc analysis of Duration of Fixations

5. Discussion

During the development of the investigation, a study developed by M. Andrzejewska and A. Skawińska [22] was found, in which two program codes were analyzed, significant differences were found for the subjective measures of mean duration of fixation, in two experimental conditions analyzed. The results obtained showed that these values for the objective cognitive load processing measures were associated with the development time of the task, obtaining measures that signified high cognitive load in the experimental conditions that required more time for the development of the activity. Other results of a study developed by T. Zu, J. Hutson, L. C. Loschky, and N. Sanjay Rebello[23] analyzed measures based on Eye-Tracking, since it was intended to make a comparison between two forms of presentation of learning content, according to the results obtained, the percentage of time spent was the objective measure which was mainly associated with extraneous cognitive load, it was found that the percentage of time spent observing animation by students in a non-redundant condition (low cognitive load) was significantly higher than that of those in a redundant condition (high cognitive load). Which, according to the authors, means that students invest more attentional resources to observe the animation and establish connections between definitions with a non-redundant condition, instead of using a redundant condition, since in this there is little time of permanence in the animations which present the lesson demonstration.

Taking into account the approaches made, the results of the present study differ with the theories and experiments developed by Mayer, Johnson, M. Andrzejewska and A. Skawińska and T. Zu, J. Hutson, L. C. Loschky, and N. Sanjay Rebello [5][23] since the results show that the presentation that has the highest levels in the duration of the fixations was the 60° configuration, and the one that showed the lowest levels was the 27° configuration, which does not show a homogeneous behavior in the variable. In contrast, the studies cited showed a homogeneous behavior of the data is presented regarding the spatial contiguity principle, which presumably ensures that the application to multimedia material leads to cognitive processing measures that indicate a lower cognitive load, than those obtained when the spatial contiguity principle is not applied. This could be due to two aspects, firstly, the size of the sample, since this was study with 16 participants and the studies analyzed for the development of this experiment had a minimum sample of 24 participants, secondly, the nature of the activity, due to the complexity of the task, since, in the studies developed on the subject, tasks with low complexity content are presented, or related to the area of expertise of the students.

6. Conclusion

In the study developed, there was no accurate results of the cognitive effect of spatial contiguity or evidence that the principle affected the objective measure of Duration of Fixations, that is, the presentations of the learning material, did not show exact results that indicated a homogeneous behavior of the dependent variable, since a logarithmic curve behavior was expected as a result, but a sinusoidal curve behavior is obtained as a result, which does not show a constant behavior of growth or decrease of the error. Additionally, there was no similar behavior of the variable for the treatments with a condition of closeness between graphics and text or between the pair of treatments with a condition of distance between graphics and text, which also indicates that the variable has a different behavior than previously analyzed in other studies.

7. Limitations and future work

The experiment developed was limited by the number of participants, which would have been beneficial to have a larger datasheet. One possible alternative interpretation of the results of this study is that the heterogeneous behavior of the fixation duration variable could attributed to the demands of the task, because it was a topic with high complexity of information, that is, a learning material could have been used of which the students were not aware, but which presented simpler information. Another limitation of the study is that only one set of learning material was used (information about the anatomy of the knee), so future research is needed to determine if the findings of the cognitive processing measures continue the behavior established in the cognitive theory of multimedia learning and additionally, if the studies are generalized to other sets of learning materials.

References

[1] U. Patayon, J. M. Gallegos, P. Mack, R. Bacabis, and C. Vicente, “Signaling and Pacing: A Comparative Study on Evidence Based Stimuli using an Eye Tracking Device,” *Procedia Comput. Sci.*, vol. 179, no. 2019, pp. 313–320, 2021, doi: 10.1016/j.procs.2021.01.011

[2] A. Sowndararajan, R. Wang, and D. A. Bowman, “Quantifying the benefits of immersion for procedural training,” *IPT/EDT 2008 - Immersive Proj. Technol. Disp. Technol. Work.*, 2008, doi: 10.1145/1394669.1394672

[3] D. Mutlu-Bayraktar, P. Ozel, F. Altindis, and B. Yilmaz, “Relationship between objective and subjective cognitive load measurements in multimedia learning,” *Interact. Learn. Environ.*, vol. 0, no. 0, pp. 1–13, 2020, doi: 10.1080/10494820.2020.1833042

[4] R. E. Mayer, “Spatial Contiguity Principle,” *Multimed. Learn.*, pp. 81–95, 2012, doi: 10.1017/cbo9781139164603.006

[5] C. I. Johnson and R. E. Mayer, “An eye movement analysis of the spatial contiguity effect in multimedia learning,” *J. Exp. Psychol. Appl.*, vol. 18, no. 2, pp. 178–191, 2012, doi: 10.1037/a0026923

[6] F. G. Paas, J. J. Van Merriënboer, and J. J. Adam, “Measurement of cognitive load in instructional research.,” *Percept. Mot. Skills*, vol. 79, no. 1 Pt 2, pp. 419–430, 1994, doi: 10.2466/pms.1994.79.1.419

[7] L. Perkhofer and O. Lehner, *Using gaze behavior to measure cognitive load*, vol. 29. Springer International Publishing, 2019.

[8] J. Sweller, J. J. G. van Merriënboer, and F. Paas, “Cognitive Architecture and Instructional Design: 20 Years Later,” *Educ. Psychol. Rev.*, vol. 31, no. 2, pp. 261–292, 2019, doi: 10.1007/s10648-019-09465-5

[9] M. M. Reverte-Vinaixa *et al.*, “Revision anterior cruciate ligament reconstruction using tibial or hamstring tendon allografts.,” *J. Orthop. Surg. (Hong Kong)*, vol. 22, no. 1, pp. 60–64, 2014, [Online]. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24781616

[10] M. J. Maas and J. M. Hughes, “Virtual, augmented and mixed reality in K–12 education: a review of the literature,” *Technol. Pedagog. Educ.*, vol. 29, no. 2, pp. 231–249, 2020, doi: 10.1080/1475939X.2020.1737210

[11] E. Costanza, A. Kunz, and M. Fjeld, “Mixed reality: A survey,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5440 LNCS, pp. 47–68, 2009, doi: 10.1007/978-3-642-00437-7_3

[12] M. Juraschek, L. Büth, G. Posselt, and C. Herrmann, “Mixed Reality in Learning Factories,” *Procedia Manuf.*, vol. 23, no. 2017, pp. 153–158, 2018, doi: 10.1016/j.promfg.2018.04.009

[13] E. Bottani *et al.*, “Wearable and interactive mixed reality solutions for fault diagnosis and assistance in manufacturing systems: Implementation and testing in an aseptic bottling line,” *Comput. Ind.*, vol. 128, p. 103429, 2021, doi: 10.1016/j.compind.2021.103429.

[14] H. Yu, Z. Zhou, X. Lei, H. Liu, G. Fan, and S. He, “Mixed Reality−Based Preoperative Planning for Training of Percutaneous Transforaminal Endoscopic Discectomy: A Feasibility Study,” *World Neurosurg.*, vol. 129, pp. e767–e775, 2019, doi: 10.1016/j.wneu.2019.06.020

[15] A. Korbach, R. Brünken, and B. Park, “Measurement of cognitive load in multimedia learning: a comparison of different objective measures,” *Instr. Sci.*, vol. 45, no. 4, pp. 515–536, 2017, doi: 10.1007/s11251-017-9413-5

[16] K. Rayner, *Eye movements and attention in reading, scene perception, and visual search*, vol. 62, no. 8. 2009

[17] A. T. Duchowski, *Eye Tracking Methodology*. 2017.

[18] B. T. Carter and S. G. Luke, “Best practices in eye tracking research,” *Int. J. Psychophysiol.*, vol. 155, no. October 2019, pp. 49–62, 2020, doi: 10.1016/j.ijpsycho.2020.05.010

[19] M. Keskin, K. Ooms, A. O. Dogru, and P. De Maeyer, “Exploring the Cognitive Load of Expert and Novice Map Users Using EEG and Eye Tracking,” *ISPRS Int. J. Geo-Information*, vol. 9, no. 7, 2020, doi: 10.3390/ijgi9070429

[20] J. Zagermann, U. Pfeil, and H. Reiterer, “Measuring cognitive load using eye tracking technology in visual computing,” *ACM Int. Conf. Proceeding Ser.*, vol. 24-October, pp. 78–85, 2016, doi: 10.1145/2993901.2993908

[21] R. E. Su and B. P. Bailey, “Put them where? Towards guidelines for positioning large displays in interactive workspaces,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 3585 LNCS, pp. 337–349, 2005, doi: 10.1007/11555261_29

[22] M. Andrzejewska and A. Skawińska, “Examining Students’ Intrinsic Cognitive Load During Program Comprehension – An Eye Tracking Approach,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12164 LNAI, no. Icl, pp. 25–30, 2020, doi: 10.1007/978-3-030-52240-7_5

[23] T. Zu, J. Hutson, L. C. Loschky, and N. Sanjay Rebello, “Use of eye-tracking technology to investigate cognitive load theory,” *arXiv*, pp. 472–475, 2018, doi: 10.1119/perc.2017.pr.113

Contact email: sara2208474@correo.uis.edu.co
lueduba@uis.educ.co
mafermar@uis.edu.co
julian2182367@correo.uis.edu.co