Heliyon 10 (2024) e31717

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Comparing desktop 3D virtual reality with web 2.0 interfaces: Identifying key factors behind enhanced user capabilities

Anna Sudár [a,b,c,*], Ádám B. Csapó [a,b,c]

[a] *Corvinus Institute for Advanced Studies, Corvinus University of Budapest, Fővám tér 8, Budapest, 1093, Hungary*
[b] *Institute of Data Analytics and Information Systems, Corvinus University of Budapest, Fővám tér 8, Budapest, 1093, Hungary*
[c] *Hungarian Research Network, Piarista u. 4, Budapest, 1052, Hungary*

ARTICLE INFO

Keywords:
3D environments
Desktop VR
Human cognitive capabilities
Spatial workflow management
Digital workflow effectiveness

ABSTRACT

The aim of this paper is to investigate how commonly used 2D digital layouts can be transformed into 3-dimensional dashboards with the effect of reducing cognitive load. To this end, we compared user performance metrics, pupil dilation data as well as subject-reported qualitative measures in a Web 2.0-based 2D scenario and two different versions of a desktop 3D virtual reality scenario. All three scenarios focused on a use case involving the most prevalent 2D digital formats and designs encountered in digital education, making use of e.g. textual information (PDF files, PPT files), images and videos. Based on the assumption that cognitive load differences can be validated based on pupillometry measurements, we showed that it is possible to develop 3D virtual reality scenarios where users experience less cognitive load while achieving the same performance metrics as in commonly used 2D environments. At the same time, our experiment also showed that such improvements do not come automatically; instead, 3D workflows that require less locomotion – even at the expense of increased camera rotations – seem to result in more effective cognitive load reduction.

1. Introduction

A large body of scientific research, along with the rapid evolution of practical approaches has shown that virtual reality (VR) holds a unique transformative potential in the way people learn, whether as children and young adults in school, or as adults in a professional context [73,13,97,36,53,16,57]. First and foremost, the ability to visualize spatial structures in three dimensions, regardless of scale, offers a key improvement over traditional text-based descriptions or 2-dimensional drawings. Second, the fact that VR allows such structures to be dynamically interacted with helps support constructive learning approaches [35,29], increasing users' motivation to learn in general [107,55]. A third, and perhaps less often cited advantage of VR is that it provides users with the ability to lay out 2D documents with more flexibility, allowing for more information to be viewed at the same time in a spatial arrangement that strongly reflects associative and semantic relationships [90,19,54].

At the same time, despite considerable research attention having been dedicated to VR ergonomics in specific use-case scenarios (see e.g. [21,28]), few general principles have been formulated with respect to the effects of VR design on user capabilities in practical digital scenarios. This observation is related both to objective capabilities (e.g., ability to learn and carry out knowledge-based digital workflows) as well as to more subjective, latent factors (e.g., willingness to make the effort required for learning). In this paper, we

---

[*] Corresponding author at: Corvinus Institute for Advanced Studies, Corvinus University of Budapest, Fővám tér 8, Budapest, 1093, Hungary.
E-mail addresses: anna.sudar@uni-corvinus.hu (A. Sudár), adambalazs.csapo@uni-corvinus.hu (Á.B. Csapó).

https://doi.org/10.1016/j.heliyon.2024.e31717
Received 19 March 2024; Received in revised form 1 May 2024; Accepted 21 May 2024

focus specifically on such questions, by assessing users' effectiveness as well as cognitive load experienced within an ecologically valid learning scenario in 2 versions of a desktop VR environment, as well as in a Web-based 2D environment. We demonstrate through these experiments that although the objective performance of the users is similar in the 3 cases, in one of the 3D cases, users are able to achieve this performance at a significantly lower cognitive load compared to the other two cases. The differences between the 3D environments, then, allow us to formulate specific design recommendations for the development of educational and work-related 3D virtual spaces.

The novelty of this paper is two-fold. First, although several studies have demonstrated the capability of VR to provide more information to users at lower cognitive load than 2D environments e.g. [98,42], few studies have tested this in the context of a practical use-case scenario with multi-modal information being presented to users [66,56,45]. Second, even fewer studies have tested the same use-case in different VR environments to uncover the decisive factors that can lead to the success or failure of a VR solution [24]. In this paper, we show that the cognitive load experienced by users is significantly lower in one of the two VR scenarios, which allows us to draw novel conclusions as to the factors which influence the effectiveness of VR applications in knowledge-based digital workflows. We also note that a large majority of existing studies focus on immersive VR applications e.g. [52,47], whereas our study centers around desktop VR solutions [1,46]. While the former can be appealing in simulating environments where physical interactions are paramount, the latter offer a more affordable and practical alternative for digital workflows, and can therefore potentially be applied in a somewhat broader scope of applications.

The paper is structured as follows. In Section 2, we provide an overview of some of the key motivating factors behind this work, including the benefits of VR in general and in education, and those parts of cognitive load theory which are relevant to this work. In Section 3, we present the experimental framework which we have developed and utilized to obtain the results reported in this paper. Finally, the experiments themselves are presented in Sections 4 and 5. Note that the first of the two experiments (comparing one of the two 3D environments with a 2D scenario, described in Section 4) has already been reported in detail in [89]; however, in this work, we have complemented that experiment with a follow-up experiment involving a different 3D space and a greater number of test subjects. Together, the two experiments enable us to perform a principled comparison between users' performance and cognitive load experienced in a Web-based 2D environment and two different desktop VR scenarios. Results of this comparison, and a discussion of the implications are provided in Sections 6 and 7.

2. Basic concepts

In this section, we provide an overview of the background behind this research – both in terms of VR technology in general and in an educational context, as well as in terms of the relevant aspects of cognitive load theory.

2.1. *Benefits of desktop 3D virtual reality over classical 2D solutions*

Virtual reality (VR) is a technological framework that aims to present users with simulated 3D environments. It typically involves the use of specialized equipment, such as headsets or goggles (head-mounted displays, HMDs), that present a three-dimensional virtual world in front of the user's eyes. Such immersive displays are designed to create a sense of presence by tracking the movements and head orientation of the user, as a counterpart to desktop 3D applications that generally employ more abstract forms of locomotion and offer a lower degree of (physical) immersion [85].

Fully immersive VR has a lot of advantages and is also preferable to desktop 3D environments in certain fields, such as healthcare and the life sciences, where high visual and environmental fidelity, and a convincing first-person experience are essential. In other fields, such as education, there can be less of a need for full immersion; further, given the potentially inhibitive cost of immersive equipment, combined with the possible discomfort caused by motion sickness on lower-end devices, desktop-based 3D visualizations may offer a more practical alternative [40]. As a result, desktop VR solutions are gaining increasing traction, and several studies have confirmed that even in the case of desktop VR, a sense of presence can still be experienced, even if not to the same extent as in fully immersive VRs (see e.g. [7,69]).

Numerous scientific studies have explored the benefits and comparative effectiveness of desktop virtual realities over classical computing interfaces in terms of reduction of mental strain [34,33], enhanced memory retrieval [6,32,48,23,18,74,86,46], improvement of digital guidance in workflows [88,87] and improved possibilities for content management [79,81,80]. However, few if any previous works have proposed a general theory as to how such benefits can be achieved in a systematic and dependable way.

The focus of this research is on the connection of virtual reality to cognitive load, and on the question of how specific design choices can lead to reliable differences in cognitive load experienced by users. Although our primary focus is on desktop VR spaces, conclusions drawn from our experiments may carry over to more immersive variants of VR as well.

2.2. *Virtual reality for educational purposes*

Virtual environments have emerged as a dynamic tool in modern education, providing a multifaceted approach aimed at enriching learning experiences.

In the past decade, many researchers have addressed the question of how VR can be used to good effect in education [76,41,25,29,108]; thus, solutions have been developed for different levels of education, including primary education [83,91], secondary education [65,8,14], higher education [73,63,58], and private education centers [51].

It has been argued that VR in education is beneficial because such systems enable learners to engage with content in real time, and encourage them to actively apply their knowledge in simulated scenarios, which results in a deeper understanding as well as better retention. Additionally, VR can transcend geographical barriers and promote collaborative learning by allowing learners to communicate in real time and share their experiences irrespective of geographical distance. By exchanging ideas, problem-solving together, and collectively constructing knowledge, learners enhance the learning process within virtual environments [77].

Skill acquisition is a central focus within virtual environments, which can provide better support for both cognitive and practical competencies [62,22]. Whether learners are engaged in procedural or declarative learning contexts, virtual environments stimulate higher mental effort, thus accelerating skill mastery through immersive engagement [22]. To mitigate cognitive overload, textual annotations can play a crucial role within virtual environments [50,1]. By highlighting relevant information amidst the visual stimuli, annotations can help guide learners' attention and enhance comprehension, ultimately optimizing learning outcomes [1]. In addition, the integration of game elements further enhances the immersive learning experience within virtual environments. By incorporating performance metrics, personalized features, and social dynamics, game elements foster motivation and engagement within these environments [31].

2.3. *Cognitive load theory*

A highly relevant theory in educational psychology, Cognitive Load Theory (CLT), seeks to describe and demonstrate how the human cognitive system processes information and what this means for instructional design. The goal of cognitive load theory in this context is to clarify how increased information processing demands during learning tasks can impact students' capacity for taking in and storing new information in long-term memory [95,96,93,72].

To be able to learn something, humans need to process new information and this requires the use of working memory, which is well-known for being limited in its capacity and for its general ability to process no more than a few information elements at a given time [17,99,100]. In cases where learners exceed these limitations, cognitive overload occurs [98]. Generally, this happens as a result of inadequate teaching strategies, or arises due to the presence of additional external distractions [95]. In order to be able to design and create appropriate educational environments and learning frameworks, it is therefore necessary to understand the types and effects of cognitive load.

2.3.1. *Categories and effects of cognitive load*

Sweller and his colleagues identified 10 cognitive load-related effects, which arise to a different extent within different fields such as digital workspace design, education, user interfaces and communication [93]. Three of these effects are particularly relevant to this paper, given that the reduction of cognitive load is especially pertinent to the design of digital content and environments:

- The *worked example (problem completion) effect* is a phenomenon whereby students who are given working examples to study perform better on future tests than those who are required to solve the same problems independently.
- The *split-attention effect* occurs when the learners have to split their attention (spatially or temporally) between at least two related information sources which cannot individually be understood without mental integration. Generally speaking, more effective learning can be observed in the case of instructional strategies where the materials are in an integrated format as opposed to a split-attention format.
- The *modality effect* is closely related to the split-attention effect and relies on the structure of working memory, which processes visual and auditory information on different channels. Based on this structure, it could be advantageous and can reduce the extraneous cognitive load if, for example, the textual information that accompanies images and animations is presented in an auditory form. In this case, the information that is presented separately can also be interpreted separately, without integration.

We note that although some of the findings of this paper can be explained in terms of these effects, the experimental results themselves do not rely on cognitive load theory itself, other than being derived based on pupil dilation measurements (for more details, see section 2.3.2).

2.3.2. *Measuring cognitive load*

Historically, cognitive load was conceptualized rather than directly measured. Breakthroughs during the last four decades, however, have led to both indirect and direct methods for assessing cognitive load and its impacts.

Indirect measures include computational models developed by Sweller and his colleagues [94]. These models quantify distinctions between candidate proxies for cognitive load, such as analyzing knowledge acquisition or learning performance. While useful, these measures lack continuous monitoring capabilities throughout the learning process [9,72,11,67].

Direct measures can be further categorized into subjective and objective approaches. Subjective measures rely on self-reported task difficulty and cognitive load through specialized questionnaires, such as Paas's subjective rating scale [68] and the NASA TLX [30]. Despite offering a viable approach in diverse situations, such questionnaires may nevertheless suffer from subjective bias and low data resolution [4,9,102,15].

In contrast, objective measures, often seen as the gold standard, provide continuous and high-resolution data. These measures encompass physiological data like eye-tracking and pupillometry [37,109,27,106,82,71], electrodermal activity [12,49] and EEG [3,26], heart-rate measurement [84,103] as well as performance metrics like secondary task performance [10,20,70]. Such objective measures offer a more reliable alternative to subjective self-reports, enhancing our understanding of cognitive load in various contexts.

2.3.3. *Cognitive pupillometry*

In this study, we focus on pupil dilation as an objective measure, which has been recognized for decades as a reliable indicator of cognitive effort. Originally observed by Kahneman and Beatty over 60 years ago [39], pupil dilation has been linked to cognitive load particularly in tasks requiring memory. This association remains relevant today, with numerous studies confirming the relationship between pupil size and cognitive demands [38,101,104,43].

Pupil dilation reflects three types of stimuli: it constricts in response to light and near objects, and dilates with increased cognitive activity, such as arousal or mental effort. Controlled by the sympathetic nervous system, the iris dilator muscle governs pupil size, explaining its enlargement during heightened arousal [59].

Advancements in technology, particularly affordable eye trackers, have facilitated precise measurement of pupil dilation, offering enhanced temporal resolution and accuracy in detecting even subtle changes in pupil diameter [105,78]. The size of the human pupil typically ranges from approximately 2 to 8 mm in diameter [61,105]. In recent times, measuring pupil dilation has become comparatively straightforward. The availability of affordable eye trackers has made it feasible to obtain sufficient temporal resolution and accuracy for detecting even minor alterations in pupil diameter.

2.3.4. *Cognitive load in virtual spaces*

As discussed earlier, there is a limit to the amount of information the brain can process at a given time. Cognitive load theory suggests that if this limit is exceeded, performance will be reduced and mental strain will increase [95,93,96,72].

Aside from the many benefits of virtual reality, its use oftentimes requires active engagement and presents users with an active and rich immersive experience, therefore potentially leading to increased demands on cognitive processing. However, it can be assumed that with careful design, such factors can be controlled even in virtual reality spaces.

In recent years, many researchers have made efforts to measure the cognitive load associated with learning, task-solving, and problem-solving in virtual and augmented reality. Several measurements have shown that it is possible to reduce cognitive load by using such 3D technologies [98,42,24,44], while other works have measured an opposite effect [2]. Still other researchers have found no significant differences in cognitive load between VR and non-VR scenarios (e.g., [5]). Based on this, the case could be made that much depends on the particular use-case and the particular technological solution that is being used. At the same time, arriving at a deeper understanding of the key factors that could be used to predict cognitive load would nevertheless be important, especially in cases where the results of a virtual reality experiment are relied upon to determine whether and how a real-world environment should be built [5,24]. Such measurements are also highly relevant to the world of education, within which online learning and remote training have gained increasing traction in the past years.

Prior to this work, we expected that the impact of content arrangement and distracting visual elements would have a significant effect on our results. Thus, we hypothesized that when it came to designing virtual spaces with integrated 2D content, the modality effect, and also a popular effect connected to multimedia learning called the seductive details effect – according to which interesting yet irrelevant details, known as seductive details, are unnecessary for achieving the instructional objective [60,75,92] – would have a key role to play in explaining the cognitive load experienced by users [64]. For this reason, we aimed to manipulate the number of visual distractions and the geometries of content arrangements in our experiments.

3. Experimental design for this study

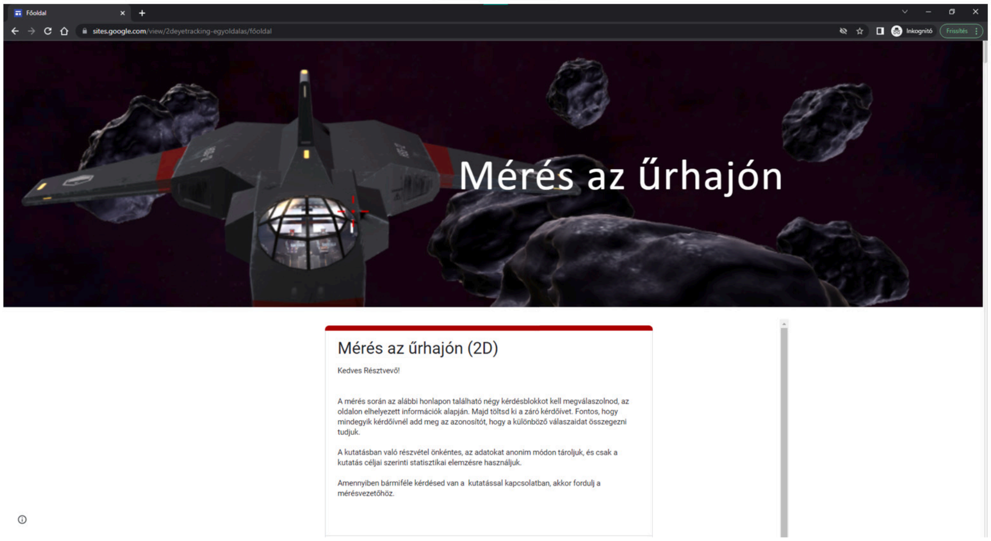
In order to assess differences in the effectiveness of user workflows, as well as the cognitive load they entail in 2D digital interfaces and different kinds of 3D environments, we chose to develop a specific workflow and a set of experimental guidelines that could be replicated in both 2D and 3D environments. The experiments revolved around subjects having to read through / look through various kinds of digital content relevant to a topic, and to then answer a set of questionnaires related to the materials they studied.

3.1. *Experimental setup*

In the 2D case, we used a classical Google Sites page (https://sites.google.com/view/2deyetracking-egyoldalas/f%C5%91oldal, accessed on 10 March, 2024), loaded into a Chrome Browser, which included all of the learning materials serially embedded into it, interleaved with the questionnaires (also in the form of embedded Google Forms within the Google Sites page). Here, the documents pertaining to each of four different topics and the associated questionnaire were embedded strictly in order; hence, users could study 3 documents—PDF files, images or YouTube videos—and then fill out the associated questionnaire for each of the topics, before proceeding to the next topic (Fig. 1).

In the 3D case, the same documents and questionnaires were laid out in two different 3D spaces within the MaxWhere platform (https://maxwhere.com) – a desktop 3D platform that allows for 3D spaces to be created with freely arranged display panels (so-called “smartboards” in MaxWhere jargon). Importantly, such smartboards can contain any kind of document that a desktop browser could normally display (e.g., webpages, PDF files, images, audio-video files).

Whereas in the 2D case, the learning materials were serially embedded into a Google form and users had to scroll up and down to view different documents and to access the questionnaires, in 3D, free navigation was allowed and expected. Two examples of how documents were laid out in 3D can be seen on Figs. 2 and 3.



Mérés az űrhajón (2D)

Kedves Résztvevő!

A mérés során az alábbi honlapon található négy kérdésblokkot kell megválaszolnod, az oldalon elhelyezett információk alapján. Majd töltsd ki a záró kérdőívet. Fontos, hogy mindegyik kérdőívnél add meg az azonosítót, hogy a különböző válaszaidat összegezni tudjuk.

A kutatásban való részvétel önkéntes, az adatokat anonim módon tároljuk, és csak a kutatás céljai szerinti statisztikai elemzésre használjuk.

Amennyiben bármiféle kérdésed van a kutatással kapcsolatban, akkor fordulj a mérésvezetőhöz.

Fig. 1. Layout and design of the 2D webpage using which the 2D scenario experiments were conducted.

It is important to note that the exact same documents and questionnaires which appeared embedded into the Google Sites page in the 2D case were presented to users in the 3D case, with no content added or removed in either case. The only difference was that the content appeared serially, from top to bottom in the 2D case, whereas it was presented in a spatial layout in the 3D case. The relationship between the 2D and 3D cases is shown in Fig. 4.

3.2. Tasks and scenarios

In order to model participants' abilities in independent topic processing and comprehension across different environments, participants were given question answering tasks across four distinct thematic groups within the domain of astronomy. Astronomy was chosen as the topic of focus based on the expectation that the participants would have little, if any familiarity with it, ensuring equitable task engagement irrespective of prior knowledge. The study design incorporated common task and content types encountered in both educational practice and digital contexts, ensuring ecological validity across educational levels. Detailed task instructions were provided to participants at the outset to mitigate experimenter influence, followed by non-intervention during task completion. All materials were presented in Hungarian to align with participants' language proficiency. Materials included PDF documents containing comprehensible professional content, videos sourced from the Hungarian-language National Geographic channel on YouTube, and custom-created images elucidating various astronomical phenomena.

The workflow to be carried out by test subjects consisted of reading and / or viewing learning materials pertaining to 4 different subtopics within the field of astronomy ("*Universe*", "*Planets*", "*Satellites*" and "*Space Research*"), and answering a questionnaire with respect to each of the subtopics. Each of the subtopics were introduced to the subjects through 3 different documents – including PDF files, images, or YouTube videos.

The questionnaires corresponding to each of the subtopics consisted of true-or-false questions, multiple choice questions and questions requiring short answers of one or two words. Three examples of typical questions are:

- True or false?—Black holes can be observed based on the gravitational effects they have on surrounding gases, dust and stars;
- What are the rings of Saturn made of (select all that apply)?—ice, rocks, space debris, gases, asteroids, and/or moons.
- Why were Hubble's mirrors polished at night?

3.3. Data collection and measurements

The results of the questionnaires were compared in terms of the percentage of questions answered correctly and the time taken to answer all of the questionnaires. However, in the 3D case, time spent with navigation between the thematic groups was discounted from the latter metric, so that the actual time spent on the perusal of the content and the filling out of the questionnaires could be compared directly. Hence, completion time data was logged based on an analysis of the screen recordings, with the timer being started when the participant stopped in front of a content group and interacted with a display, and stopped when they clicked on the submit button on the questionnaire associated with the content group.

In the meantime, eye gaze and pupil dilation data were also recorded to assess test subjects' focus of attention and cognitive load experienced. The latter measurements were obtained using the EyeTribe eye tracker (https://theeyetribe.com/dev.theeyetribe.com/dev.theeyetribe.com/general/index.html, accessed on 10 March, 2024) and its accompanying OGAMA (Open Gaze and Mouse Analyzer) Version 5.1 software.

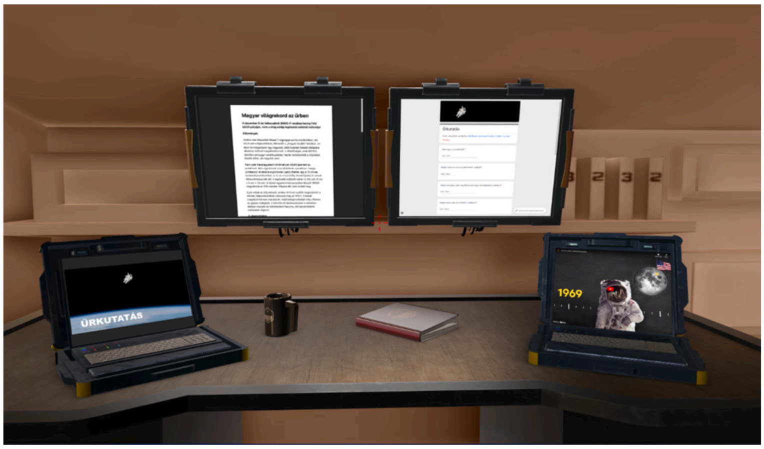


Fig. 2. Spatial arrangement of 2D content in a 3D virtual space.

Fig. 3. Spatial arrangement of 2D content in a second 3D virtual space.

In addition, the frequency of certain interaction patterns was recorded, including when:

- Users viewed content on a group of smartboards simultaneously from some distance ("*holistic overview mode*")
- Users alternated focus between different smartboards while remaining in a stationary position but frequently changing their camera orientation ("*alternating mode*");

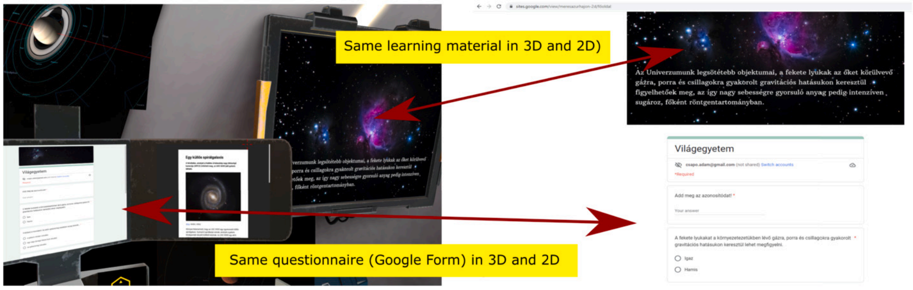
Finally, in a post-experiment questionnaire, subjects were asked for demographic data, as well as presented with questions about their digital leisure habits, and about their subjective assessments of immersion in the virtual space (in the 3D case). Subjects were also asked to rate the difficulty of the 4 different subtopics and the most difficult questionnaire from all four. For those subjects who completed the tests on the two-dimensional interface, the final questionnaire did not include questions on the 3D space.

3.4. Procedure

Prior to the experiments, written consent was obtained from all participants to use the collected data as follows. All data collected was anonymized and used solely for the purpose of the statistical analyses reported in this paper. All of the experiments were conducted in accordance with the ethical principles laid out in the Declaration of Helsinki.

At the start of the session, test subjects carrying out the required task in 3D indicated whether they were familiar with the MaxWhere software and, if so, approximately how much time they had spent using the software. Participants who were not familiar with the software spent approximately 30 minutes learning about it and acquiring basic user skills. Basic knowledge includes confident VR navigation using the software, activating and deactivating the smartboards (display panels) and interacting with the content that is displayed on them. Mastery of confident use of the software was assessed by the test administrators.

Following this, the participants were seated in a quiet room in front of a laptop computer. The room was dimly lit without any direct light source so as not to introduce unwanted artifacts into the eye tracking data. For each test subject, the eye-tracker was calibrated prior to the experiment.



Az Univerzumunk legsötétebb objektumai, a fekete lyukak az őket körülvevő gázra, porra és csillagokra gyakorolt gravitációs hatásukon keresztül figyelhetőek meg, az így nagy sebességre gyorsuló anyag pedig intenzíven sugároz, főként röntgentartományban.

Világegyetem

csapo.adam@gmail.com (not shared) Switch accounts

*Required

Add meg az azonosítódat! *

Your answer

A fekete lyukakat a környezetezetükben lévő gázra, porra és csillagokra gyakorolt gravitációs hatásukon keresztül lehet megfigyelni. *

○ Igaz
○ Hamis

Fig. 4. This figure shows by example that the same documents and questionnaires were used in both the 3D case (left-hand side) and in the 2D case (right-hand side). No materials/questionnaires were added to or removed from the experiment in either case. Here, we can see that the PDF document on the topic of the "Universe" appears on the tilted panel at the back in the 3D case as well as on the upper half of this specific view of the 2D case. Whereas in the 2D case, the questionnaire appears directly below the learning material, in the 3D space, it can be found on the left-hand side of the screenshot.

Fig. 5. Layout and design of the 3D space in which the first experiment was conducted, as reported in [89].

Following the calibration, the test administrator explained the task to be carried out. The order in which the questionnaires were filled out, with the selection of the first and the last questionnaire, was left up to the participants (in the 2D case, this was less of an issue, although no strict order was enforced in the 2D case either, regardless of the fact that the documents on the subtopics and the corresponding questionnaires followed each other in a serial order).

4. Comparison of 2D environment with the first desktop VR space

In a first experiment, we compared measurement data obtained from the 2D scenario and one particular 3D space, based on the experimental framework described in Section 3, and as reported in [89]. The space in which the experiment was carried out was a spaceship-themed environment, as shown on Fig. 5.

4.1. Participants

In this experiment, a total of 14 test subjects participated carried out the task in 3D. However, video data were corrupted for one person. The results from a further four subjects had to be discarded due to there being breaks in pupil dilation measurements midway during the experiment. The remaining nine subjects (three women, six men) were aged between 17–55 years, with a mean age of 32.5 years (SD: 14.15).

A total of seven test subjects (four women, three men) participated in the 2D measurement, and were aged between 25–33 years, with a mean age of 27.83 (SD: 2.93).

The mean age for all participants was 30.84 (11.39) years. All participants were neurotypical Hungarian native speakers who participated in the experiment on a voluntary basis. Informed consent was obtained from all participants prior to the experiment, which was carried out based on and in accordance with the institutional endorsement of the authors' affiliation. All of the data collected during the experiment was anonymized and used exclusively as input to the statistical analyses detailed in later parts of this paper.

4.2. Results

When comparing results from the 2D scenario and this 3D scenario, the data showed that the correctness of answers provided by subjects in the 2D and 3D case was very similar, with no statistically significant differences. At the same time, within the topic of Satellites, subjects completed the questionnaire significantly faster in 3D than in 2D (while there was no significant difference in completion times within the other topics). These results failed to confirm, in a general sense, the hypothesis that subjects would perform better in 3D than in 2D (save for the completion time in the case of the Satellites topic)—although they certainly did not perform worse.

The fact that the 3D environment included a high volume of visual clutter led us to the conclusion that reducing this clutter — perhaps by using a more minimalistic 3D space with less need for navigation, other than rotation of view — could lead to higher yields in performance in the 3D case.

Regarding subjective evaluations of difficulty and pupil diameter measurements, it was observed that there was no correlation between the two. This was a surprising result. Leaving aside the possibility that the tasks were actually more difficult when the subjects perceived them to be easier, this counter-intuitive finding may have also been due to the amount of visual clutter, and thus, a degree of excitement experienced by the test subjects.

5. Follow-up experiment in a second VR environment

Our second experiment was carried out using the same 2D scenario (however, with more test subjects) and in a second 3D VR space which contained a circular arrangement of content clusters (Fig. 6). In the 3D virtual space, the contents were placed on a total of 21 display panels. The questionnaires were placed in the middle of each cluster, above them was the title slide, and to the right, left and bottom were the documents that contained the answers to the questions contained in the questionnaires.

5.1. Motivations behind follow-up study

Based on our tentative conclusions from the first experiment (Section 4), and our prior experience in designing 3D spaces with custom layouts (see also [90]), we performed the same experiment using a different VR space and a larger number of test subjects. Our goal was to use a second virtual space in the follow-up study that would more adequately address the following principles based on the idea that in the case of virtual spaces containing many visual details, the pupil measurements may have been particularly wide as a result of heightened arousal:

- The need for navigation and the amount of visual clutter should be limited in order to lower cognitive load
- The size of the documents should indicate their importance
- The arrangement of documents into clusters should allow the user to conclude that the content placed there form a content unit
- By creating the same layout for every subtopic, such that the content types are the same in corresponding locations in each case (with only the particular topic differing), users should be able to understand semantic relationships more quickly and at a lower cognitive load
- The digital content inside the virtual space should be displayed vertically or on panels that are only minimally tilted, as users have been shown to prefer vertical surfaces for documents [90].

Thus, for our follow-up study reported in later parts of this paper, we opted for a virtual space in which there were no unnecessary disruptors, which was much smaller, in which less navigation was needed, and in which the display panels (“smartboards”) holding the content were also more purposefully organized.

5.2. Participants

A total of 40 test subjects participated in this experiment, with an average age of 26.57 (SD: 6.96), and an equal distribution of 20 women and 20 men.

All of the participants were Hungarian native speakers and individuals with neurotypical development, who participated in the experiment on a voluntary basis. Informed consent was obtained from all participants prior to the experiment, which was carried out based on and in accordance with the institutional endorsement of the authors’ affiliation. All of the data collected during the experiment was anonymized and used exclusively as input to the statistical analyses detailed in later parts of this paper.

Participants were divided into two groups. 20 subjects (14 women, 6 men), average age of 27.35 (SD: 6.72) participated in the measurement in the 3D virtual space. 5 out of the 20 subjects wore glasses during the experiment. 20 subjects (14 men, 6 women) participated in the 2D measurement, average age of 25.8 (7.29), and 2 of them wore glasses during the measurement.

5.3. Data collection and measurements

The methods employed in the second experiment were the same as in the first experiment, as described earlier in Section 3 on the framework used for our experiments.

Fig. 6. Exterior of the 3D space in which the 3D case of the second experiment was conducted.

In particular, the subjects participating in the 2D and 3D also measurement had to provide feedback on which subtopic they found to be the most difficult in retrospect, as well as some demographic data (gender, age, highest education) – as described in Section 3.3. In addition to the data listed above, the participants in the 3D measurement also had to fill out a series of questions related to their navigation experiences, as well as an IPQ questionnaire consisting of 14 questions, in which they had to evaluate their General Presence, Spatial Presence, Involvement, and Experienced Realism experiences on a 7-point Likert scale. However, this data is not the subject of this research.

Similarly to the earlier experiment, descriptive statistics were calculated to summarize and describe the main features of the data. Statistics that were collected included the completion time, pupil dilation, and difficulty of the questionnaires. Completion times were measured based on screen recordings of the experiment (here, again, discounting the locomotion between content groups), and pupil dilation was measured using the EyeTribe eye tracker and its accompanying OGAMA Version 5.1 software.

On the basis of these statistics, group differences were assessed using an Independent Samples T-Test to determine the differences in completion time and in pupil size in 2D and the second 3D scenario. The Mann-Whitney U test was used to determine the differences between the final scores of the groups and the pupil dilation differences between the two 3D scenarios. Spearman correlation analysis was utilized to examine relationships between the difficulty of the questionnaires and pupil dilation as well as completion time. All statistical analyses were conducted using SPSS and JASP, with a significance level set at 0.05.

5.4. Key hypotheses

Prior to conducting our experiment, we formulated the following key hypotheses:

1. Subjective assessments of lower cognitive load, along with lower pupil dilation would characterize the 3D case as opposed to the 2D case;
2. Questionnaires would generally be filled out faster and with more correct answers in 3D compared to 2D.
3. Compared to the previous measurement, the current 3D group would perform the task with a lower pupil dilatation than the previous 3D group;
4. Subjective assessments, by the test subjects, of the difficulty of questionnaires would correlate with pupil dilation and correctness of answers;
5. Lower cognitive load would correlate with a faster completion time.

Table 1
Descriptives of the pupil dilation data of the 2D and 3D groups.

	Group	N	Mean	SD	SE	Coefficient of variation
Pupil Dilatation	3D	20	19.464	2.902	0.649	0.149
	2D	20	22.204	3.971	0.888	0.179

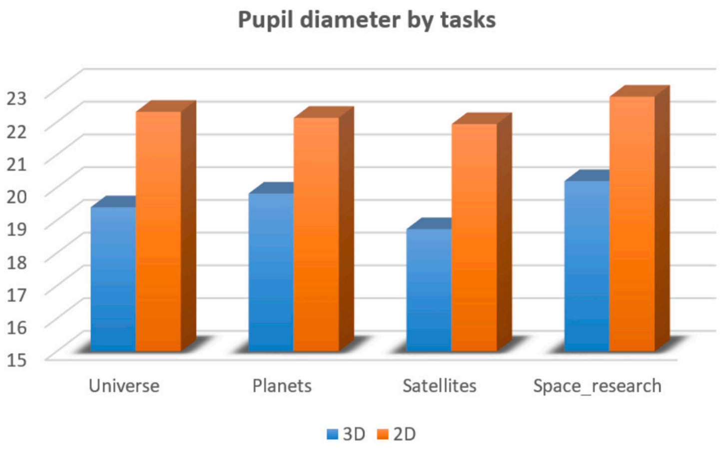


Fig. 7. Pupil diameter sizes grouped by tasks.

Table 2
Descriptives of the pupil dilation data of the previous 3D and current 3D (referred to as "3D sublimus" group in the table) groups.

	Group	N	Mean	SD	SE	Coefficient of variation
PupilTasks	3D sublimus	20	19.464	2.902	0.649	0.149
	3D spaceship	9	26.203	4.371	1.457	0.167

6. Results

We used the Mann-Whitney U test to determine whether there were differences in the final scores of the four tasks. The test revealed no statistical difference in the final scores between the 3D (M = 22.387, SD = 2.167) and 2D (M = 22.125, SD = 1.879) groups.

We used the Independent Samples T-Test to determine whether there were differences in completion times in the context of the four tasks. The test revealed no statistical differences in the completion times of the four tasks between the 3D (M = 36.7, SD = 6.959) and 2D (M = 37.8, SD = 9.059) groups.

Regarding pupil dilation, the group descriptives and differences between the two groups are displayed in Table 1 and in Fig. 7. An Independent Samples T-Test revealed that the group which performed the measurement in 2D had significantly larger pupil dilation (M = 22.204, SD = 3.971) than the 3D group (M = 19.464, SD = 2.902), (t(34.792) = -2.492, $p < 0.018$), (Cohen's d = 0.788).

The results showed that the group which filled out the questionnaire in the 3D space (N = 20) had smaller pupil dilation with the mean rank of this group being 11.30, while the group which used the Spaceship 3D space for the same task (N = 9) had a mean rank of 22.23. A Mann-Whitney U test revealed that this difference was statistically significant, U = 16, $p < 0.001$, r = 0.65 (Table 2 and Fig. 8).

The Spearman coefficient obtained did not show a correlation between the perceived difficulty of the questionnaires and pupil dilation; nor was any correlation detected between pupil dilation and completion time.

7. Discussion

In the main experiment presented in this paper, 40 subjects were examined, divided into 2 groups (3D experimental, 2D control). One group of 20 subjects had to fill out four questionnaires on a classical 2D dimensional Web-based interface, such that the answers to the questions were found in related documents integrated into the same webpage. The other group of 20 subjects had to perform the same tasks in a 3D virtual space, with the difference that the related documents did not follow one another linearly, as in the case of the 2D webpage, but were instead spatially spread out in a 3D space surrounding the questionnaires.

Among the cognitive load effects mentioned in the introduction, the split-attention and the modality effect typically play an important role in the case of online teaching materials, as teachers often share information in the form of documents that are related

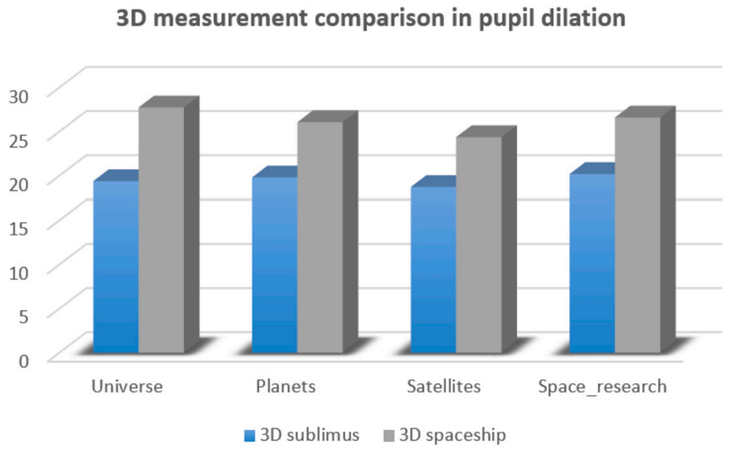


Fig. 8. Comparison of the pupil diameter size between the previous 3D research results and the current ones (referred to as the "3D sublimus" case).

in content but are separate from each other. The aim of our research, then, was to create a virtual environment that can reduce the cognitive load for students, as well as soften the impact of these two effects.

The results of our pupillometry measurements showed a significant difference between the two groups, as expected. Therefore, we are able to conclude that based on the results of prior research, it was possible to create an educational virtual space that enabled subjects to study the same learning materials at a reduced cognitive load.

In the case of the space we developed, users were able to interpret the learning materials as belonging to well-defined thematic units or blocks, despite the information coming from different documents and modalities. Based on the data we collected, we observed that the subjects often positioned themselves in such a way that they were able to view multiple documents simultaneously from a suitable distance and orientation. We refer to this as an *overview type mode of interaction* that is not possible to achieve with the same flexibility on classical 2D user interfaces, but which nevertheless mirrors the way people often interact with documents in real life.

In addition, we observed a mixing of modalities such that, for example, subjects played the videos in the virtual space while navigating to the questionnaires to fill out the relevant questions as they were listening to the answers to them. While this could also be done in 2D, it is less natural than in 3D, where users simply have to "turn their heads" to temporarily focus on a different display panel.

Since the virtual space lacked any specific decorative elements, it was assumed that users could focus on the content and on solving the task at hand. Even so, there were no significant differences in completion time and overall score between the two groups.

Comparing these results to an earlier experiment reported in [89] and summarized in Section 4, in which the 2D scenario was the same and the VR scenario involved a more spacious environment with more visual disruptors, we were able to conclude that indeed the factors of less visual clutter and a preference towards camera rotations as opposed to spatial locomotion were instrumental in reducing cognitive load. Particularly in the prior experiment, increased visual clutter may have accentuated the split-attention effect.

When it comes to assessing cognitive load within VR, past literature has predominantly focused on immersive VR solutions [52,47] and in some cases, researchers have compared these measurements to 2D interfaces as well [1,46]. The majority of these studies have reported results that point in the same direction as the results presented in this paper. However, the results we have presented in this paper are unique in that our comparison focused on a 2D environment and two desktop VR environments (thereby leading to a comparison of the two 3D environments as well). The results presented here may also allow us to draw more general conclusions, given that they were derived based on everyday digital workflows (i.e., studying and answering questions based on multimedia content) rather than based on any specific laboratory environment, which is typical in this field [24,42,98,14]. To the best of our knowledge, there have been no previous studies measuring cognitive load in the context of such a commonly encountered digital task and in a comparable desktop 3D environment. Further, our results suggest that with the application of certain design principles, 3D spaces can be developed that have a higher chance of reducing cognitive load.

Limitations of the study include the relatively fewer number of test subjects in our preliminary experiment, based on which we implicitly compared the first VR scenario to our second, enhanced VR scenario. Nevertheless, our comparison between the 2D environment and the second VR scenario, as well as between the 2D environment and the first VR scenario still holds as valid.

Further research would also be necessary to determine whether the negative influence of disruptors carries over to scenarios where long-term memorization is required and e.g. the principles of the memory palace technique could come into play. In order to investigate this, we plan to use not only eye-tracking but also to carry out EEG measurements in the future.

The results reported in this paper allow us to conclude that desktop VR environments can serve as a good alternative to 2D interfaces in distance learning, and it is worth considering their introduction into the broader public education system.

8. Conclusions

In this paper, we focused on comparing the effectiveness and cognitive load associated with different desktop 3D virtual reality learning scenarios and a 2D Web-based e-learning scenario. Based on the observation that researchers have previously come to partially conflicting conclusions about the relative effectiveness of virtual reality compared to classical 2D interfaces, our goal was to experimentally investigate the effects of different spatial parameters such as the amount of visual clutter and the types of navigation methods required. During the process, we used the metric of pupil dilation as an objective measure, and we also linked our results to some of the core effects within cognitive load theory.

Results obtained confirmed our hypothesis that fewer decorative elements in a VR space, and a relative reliance on camera rotations as opposed to camera movement when navigating between content can lead to the reduction of cognitive load, both compared to other VR scenarios and to classical Web 2.0 based 2D scenarios. Further, the 3D space allowed users to exhibit interative behaviors that are more akin to physical interactions with printed documents, e.g. by positioning themselves relative to multiple documents at the same time.

Our findings suggest that well-designed desktop VR environments can offer a promising alternative for learning as well as digital work, with the potential for enhanced cognitive outcomes. Through active learning, collaborative engagement, skill acquisition, and the integration of game-like elements, virtual environments can lead to transformative educational practices that can significantly shape the future of learning.

Funding

This research received no external funding.

Ethics approval

This study was conducted in accordance with the Declaration of Helsinki, and was approved by the Doctoral School of Multidisciplinary Engineering Sciences (MMTDI) at the Széchenyi István University on October 10, 2022. Electronic informed consent was obtained from all the participants. The questionnaires were anonymized, and participants were free to opt out of participation in the study whenever they were uncomfortable.

Consent to participate

Each participant read and accepted the declaration of consent before the measurements.

Consent for publication

Each participant read and accepted the declaration of consent before the measurements and accepted that the data provided can be used and published in scientific form.

Code availability

Not applicable.

CRediT authorship contribution statement

Anna Sudár: Writing – review & editing, Writing – original draft, Validation, Methodology, Conceptualization. **Ádám B. Csapó:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and materials

Data will be made available on request.

Acknowledgements

The research presented in this paper was supported by the Hungarian Research Network (HUN-REN), and was carried out within the HUN-REN Cognitive Mapping of Decision Support Systems research group.

References

[1] P. Albus, A. Vogt, T. Seufert, Signaling in virtual reality influences learning outcome and cognitive load, Comput. Educ. 166 (2021) 104154.

[2] S.A.W. Andersen, P.T. Mikkelsen, L. Konge, P. Cayé-Thomasen, M.S. Sørensen, The effect of implementing cognitive load theory-based design principles in virtual reality simulation training of surgical skills: a randomized controlled trial, Adv. Simul. 1 (2016) 1–8, https://doi.org/10.1186/s41077-016-0022-1.

[3] E.W. Anderson, K.C. Potter, L.E. Matzen, J.F. Shepherd, G.A. Preston, C.T. Silva, A user study of visualization effectiveness using eeg and cognitive load, in: Computer Graphics Forum, Wiley Online Library, 2011, pp. 791–800.

[4] Ø. Anmarkrud, A. Andresen, I. Bråten, Cognitive load and working memory in multimedia learning: conceptual and measurement issues, Educ. Psychol. 54 (2019) 61–83, https://doi.org/10.1080/00461520.2018.15544848.

[5] A. Armougum, E. Orriols, A. Gaston-Bellegarde, C. Joie-La Marle, P. Piolino, Virtual reality: a new method to investigate cognitive load during navigation, J. Environ. Psychol. 65 (2019) 101338, https://doi.org/10.1016/j.jenvp.2019.101338.

[6] B. Berki, 2d advertising in 3d virtual spaces, Acta Polytech. Hung. 15 (2018) 175–190, https://doi.org/10.12700/APH.15.3.2018.3.10.

[7] B. Berki, Experiencing the sense of presence within an educational desktop virtual reality, Acta Polytech. Hung. 17 (2020) 255–265, https://doi.org/10.12700/APH.17.2.2020.2.14.

[8] D. Bogusevschi, C. Muntean, G.M. Muntean, Teaching and learning physics using 3d virtual learning environment: a case study of combined virtual reality and virtual laboratory in secondary school, J. Comput. Math. Sci. Teach. 39 (2020) 5–18.

[9] R. Brunken, J.L. Plass, D. Leutner, Direct measurement of cognitive load in multimedia learning, Educ. Psychol. 38 (2003) 53–61, https://doi.org/10.1207/S15326985EP3801_7.

[10] R. Brünken, J.L. Plass, D. Leutner, Assessment of cognitive load in multimedia learning with dual-task methodology: auditory load and modality effects, Instr. Sci. (2004) 115–132, https://doi.org/10.1023/B:TRUC.0000021812.96911.c5.

[11] R.E. Brünken, J.L. Plass, R.E. Moreno, Current issues and open questions in cognitive load research, https://doi.org/10.1017/CBO9780511844744.014, 2010.

[12] M. Buchwald, S. Kupiński, A. Bykowski, J. Marcinkowska, D. Ratajczyk, M. Jukiewicz, Electrodermal activity as a measure of cognitive load: a methodological approach, in: 2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), IEEE, 2019, pp. 175–179.

[13] D. Checa, A. Bustillo, A review of immersive virtual reality serious games to enhance learning and training, Multimed. Tools Appl. 79 (2020) 5501–5527, https://doi.org/10.1007/s11042-019-08348-9.

[14] A. Christopoulos, S. Mystakidis, E. Cachafeiro, M.J. Laakso, Escaping the cell: virtual reality escape rooms in biology education, Behav. Inf. Technol. (2022) 1–18, https://doi.org/10.1080/0144929X.2022.2079560.

[15] G. Cierniak, K. Scheiter, P. Gerjets, Explaining the split-attention effect: is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load?, Comput. Hum. Behav. 25 (2009) 315–324, https://doi.org/10.1016/j.chb.2008.12.020.

[16] M. Coban, Y.I. Bolat, I. Goksu, The potential of immersive virtual reality to enhance learning: a meta-analysis, Educ. Res. Rev. 36 (2022) 100452, https://doi.org/10.1016/j.edurev.2022.100452.

[17] N. Cowan, Working Memory Capacity, Psychology Press, 2012, https://doi.org/10.4324/9780203342398.

[18] Á.B. Csapó, I. Horvath, P. Galambos, P. Baranyi, Vr as a medium of communication: from memory palaces to comprehensive memory management, in: 2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), IEEE, 2018, pp. 000389–000394.

[19] P. Daeijavad, F. Maurer, Layouts of 3d data visualizations small multiples around users in immersive environments, in: 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), IEEE, 2022, pp. 258–261.

[20] K.E. DeLeeuw, R.E. Mayer, A comparison of three measures of cognitive load: evidence for separable measures of intrinsic, extraneous, and germane load, J. Educ. Psychol. 100 (2008) 223, https://doi.org/10.1037/0022-0663.100.1.223.

[21] N. Dozio, F. Marcolin, G.W. Scurati, L. Ulrich, F. Nonis, E. Vezzetti, G. Marsocci, A. La Rosa, F. Ferrise, A design methodology for affective virtual reality, Int. J. Hum.-Comput. Stud. 162 (2022) 102791, https://doi.org/10.1016/j.ijhcs.2022.102791.

[22] I. Dubovi, Cognitive and emotional engagement while learning with vr: the perspective of multimodal methodology, Comput. Educ. 183 (2022) 104495.

[23] E. Fassbender, W. Heiden, The virtual memory palace, J. Comput. Inf. Syst. 2 (2006) 457–464.

[24] J.G. Frederiksen, S.M.D. Sørensen, L. Konge, M.B.S. Svendsen, M. Nobel-Jørgensen, F. Bjerrum, S.A.W. Andersen, Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: a randomized trial, Surg. Endosc. 34 (2020) 1244–1252, https://doi.org/10.1007/s00464-019-06887-8.

[25] L. Freina, M. Ott, A literature review on immersive virtual reality in education: state of the art and perspectives, in: The International Scientific Conference Elearning and Software for Education, 2015, pp. 10–1007.

[26] N. Friedman, T. Fekete, K. Gal, O. Shriki, Eeg-based prediction of cognitive load in intelligence tests, Front. Human Neurosci. 13 (2019) 191, https://doi.org/10.3389/fnhum.2019.00191.

[27] R. Gavas, D. Chatterjee, A. Sinha, Estimation of cognitive load based on the pupil size dilation, in: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2017, pp. 1499–1504.

[28] N. Glaser, M. Schmidt, Systematic literature review of virtual reality intervention design patterns for individuals with autism spectrum disorders, Int. J. Hum.-Comput. Interact. 38 (2022) 753–788, https://doi.org/10.1080/10447318.2021.1970433.

[29] D. Hamilton, J. McKechnie, E. Edgerton, C. Wilson, Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design, J. Comput. Educ. 8 (2021) 1–32, https://doi.org/10.1007/s40692-020-00169-2.

[30] S.G. Hart, L.E. Staveland, Development of nasa-tlx (task load index): results of empirical and theoretical research, in: P.A. Hancock, N. Meshkati (Eds.), Advances in Psychology, in: Human Mental Workload. North-Holland., vol. 52, 1988, pp. 139–183.

[31] Y. Hong, N. Saab, W. Admiraal, Approaches and game elements used to tailor digital gamification for learning: a systematic literature review, Comput. Educ. 105000 (2024).

[32] I. Horváth, Behaviors and capabilities of generation ce students in 3d vr, in: 2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), IEEE, 2019, pp. 491–494.

[33] I. Horváth, Maxwhere 3d capabilities contributing to the enhanced efficiency of the trello 2d management software, Acta Polytech. Hung. 16 (2019) 55–71, https://doi.org/10.12700/APH.16.6.2019.6.5.

[34] I. Horváth, A. Sudár, Factors contributing to the enhanced performance of the maxwhere 3d vr platform in the distribution of digital information, Acta Polytech. Hung. 15 (2018) 149–173, https://doi.org/10.12700/APH.15.3.2018.3.9.

[35] C.L. Huang, Y.F. Luo, S.C. Yang, C.M. Lu, A.S. Chen, Influence of students' learning style, sense of presence, and cognitive load on learning outcomes in an immersive virtual reality learning environment, J. Educ. Comput. Res. 58 (2020) 596–615, https://doi.org/10.1177/07356331198674.

[36] M. Javaid, A. Haleem, Virtual reality applications toward medical field, Clin. Epidemiol. Glob. Health 8 (2020) 600–605, https://doi.org/10.1016/j.cegh.2019.12.010.

[37] A.W. Joseph, R. Murugesh, Potential eye tracking metrics and indicators to measure cognitive load in human-computer interaction research, J. Sci. Res. 64 (2020) 168–175, https://doi.org/10.37398/JSR.2020.640137.

[38] D. Kahneman, Attention and Effort, Prentice-Hall, 1973.

[39] D. Kahneman, J. Beatty, Pupil diameter and load on memory, Science 154 (1966) 1583–1585, https://doi.org/10.1126/science.154.3756.1583.

[40] D. Kamińska, T. Sapiński, S. Wiak, T. Tikk, R.E. Haamer, E. Avots, A. Helmi, C. Ozcinar, G. Anbarjafari, Virtual reality and its applications in education, Surv. Inf. 10 (2019) 318.

[41] S. Kavanagh, A. Luxton-Reilly, B. Wuensche, B. Plimmer, A systematic review of virtual reality in education, Themes Sci. Technol. Educ. 10 (2017) 85–119.
[42] S. Küçük, S. Kapakin, Y. Göktaş, Learning anatomy via mobile augmented reality: effects on achievement and cognitive load, Anat. Sci. Educ. 9 (2016) 411–421, https://doi.org/10.1002/ase.1603.
[43] B. Laeng, S. Sirois, G. Gredebäck, Pupillometry: a window to the preconscious?, Perspect. Psychol. Sci. 7 (2012) 18–27, https://doi.org/10.1177/1745691611427305.
[44] A.F. Lai, C.H. Chen, G.Y. Lee, An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory, Br. J. Educ. Technol. 50 (2019) 232–247, https://doi.org/10.1111/bjet.12716.
[45] B. Lampert, A. Pongrácz, J. Sipos, A. Vehrer, I. Horvath, Maxwhere vr-learning improves effectiveness over classical tools of e-learning, Acta Polytech. Hung. 15 (2018) 125–147.
[46] E.A.L. Lee, K.W. Wong, Learning with desktop virtual reality: low spatial ability learners are more positively affected, Comput. Educ. 79 (2014) 49–58, https://doi.org/10.1016/j.compedu.2014.07.010.
[47] J.Y. Lee, N. de Jong, J. Donkers, H. Jarodzka, J.J. van Merriënboer, Measuring cognitive load in virtual reality training via pupillometry, IEEE Trans. Learn. Technol. (2023).
[48] E.L. Legge, C.R. Madan, E.T. Ng, J.B. Caplan, Building a memory palace in minutes: equivalent memory performance using virtual versus conventional environments with the method of loci, Acta Psychol. 141 (2012) 380–390, https://doi.org/10.1016/j.actpsy.2012.09.002.
[49] P. Li, Y. Li, Y. Yao, C. Wu, B. Nie, S.E. Li, Sensitivity of electrodermal activity features for driver arousal measurement in cognitive load: the application in automated driving systems, IEEE Trans. Intell. Transp. Syst. 23 (2021) 14954–14967, https://doi.org/10.1109/TITS.2021.3135266.
[50] W. Li, Q. Feng, X. Zhu, Q. Yu, Q. Wang, Effect of summarizing scaffolding and textual cues on learning performance, mental model, and cognitive load in a virtual reality environment: an experimental study, Comput. Educ. 200 (2023) 104793.
[51] T.J. Lin, Y.J. Lan, Language learning in virtual reality environments: past, present, and future, J. Educ. Technol. Soc. 18 (2015) 486–497.
[52] R. Liu, L. Wang, T.A. Koszalka, K. Wan, Effects of immersive virtual reality classrooms on students' academic achievement, motivation and cognitive load in science lessons, J. Comput. Assist. Learn. 38 (2022) 1422–1433.
[53] S.M.C. Loureiro, J. Guerreiro, S. Eloy, D. Langaro, P. Panchapakesan, Understanding the use of virtual reality in marketing: a text mining-based review, J. Bus. Res. 100 (2019) 514–530, https://doi.org/10.1016/j.jbusres.2018.10.055.
[54] W. Luo, A. Lehmann, H. Widengren, R. Dachselt, Where should we put it? Layout and placement strategies of documents in augmented reality for collaborative sensemaking, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 2022, pp. 1–16, https://doi.org/10.1145/3491102.3501946.
[55] G. Makransky, S. Borre-Gude, R.E. Mayer, Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments, J. Comput. Assist. Learn. 35 (2019) 691–707, https://doi.org/10.1111/jcal.12375.
[56] G. Makransky, R.E. Mayer, Benefits of taking a virtual field trip in immersive virtual reality: evidence for the immersion principle in multimedia learning, Educ. Psychol. Rev. 34 (2022) 1771–1798.
[57] G. Makransky, G.B. Petersen, The cognitive affective model of immersive learning (camil): a theoretical research-based model of learning in immersive virtual reality, Educ. Psychol. Rev. (2021) 1–22, https://doi.org/10.1007/s10648-020-09586-2.
[58] B. Marks, J. Thomas, Adoption of virtual reality technology in higher education: an evaluation of five teaching semesters in a purpose-designed laboratory, Educ. Inf. Technol. 27 (2022) 1287–1305, https://doi.org/10.1007/s10639-021-10653-6.
[59] S. Mathôt, Pupillometry: psychology, physiology, and function, J. Cogn. 1 (2018), https://doi.org/10.5334/joc.18.
[60] R.E. Mayer, The Cambridge Handbook of Multimedia Learning, Cambridge University Press, 2005.
[61] D. McDougal, P. Gamlin, 1.26 - pupillary control pathways, in: R.H. Masland, T.D. Albright, T.D. Albright, R.H. Masland, P. Dallos, D. Oertel, S. Firestein, G.K. Beauchamp, M. Catherine Bushnell, A.I. Basbaum, J.H. Kaas, E.P. Gardner (Eds.), The Senses: A Comprehensive Reference, Academic Press, New York, 2008, pp. 521–536.
[62] E. McGovern, G. Moreira, C. Luna-Nevarez, An application of virtual reality in education: can this technology enhance the quality of students' learning experience?, J. Educ. Bus. 95 (2020) 490–496.
[63] Z. Merchant, E.T. Goetz, L. Cifuentes, W. Keeney-Kennicutt, T.J. Davis, Effectiveness of virtual reality-based instruction on students' learning outcomes in k-12 and higher education: a meta-analysis, Comput. Educ. 70 (2014) 29–40, https://doi.org/10.1016/j.compedu.2013.07.033.
[64] D. Mutlu-Bayraktar, V. Cosgun, T. Altan, Cognitive load in multimedia learning environments: a systematic review, Comput. Educ. 141 (2019) 103618.
[65] E. Nersesian, A. Spryszynski, M.J. Lee, Integration of virtual reality in secondary stem education, in: 2019 IEEE Integrated STEM Education Conference (ISEC), IEEE, 2019, pp. 83–90.
[66] A.V. Oje, N.J. Hunsu, D. May, Virtual reality assisted engineering education: a multimedia learning perspective, Comput. Educ.: X Real. 3 (2023) 100033.
[67] F. Paas, A. Renkl, J. Sweller, Cognitive load theory and instructional design: recent developments, Educ. Psychol. 38 (2003) 1–4, https://doi.org/10.1207/S15326985EP3801_1.
[68] F.G. Paas, Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach, J. Educ. Psychol. 84 (1992) 429, https://doi.org/10.1037/0022-0663.84.4.429.
[69] F. Pallavicini, A. Pepe, M.E. Minissi, Gaming in virtual reality: what changes in terms of usability, emotional response and sense of presence compared to non-immersive video games?, Simul. Gaming 50 (2019) 136–159, https://doi.org/10.1177/1046878119831420.
[70] B. Park, L. Knörzer, J.L. Plass, R. Brünken, Emotional design and positive emotions in multimedia learning: an eyetracking study on the use of anthropomorphisms, Comput. Educ. 86 (2015) 30–42, https://doi.org/10.1016/j.compedu.2015.02.016.
[71] L. Perkhofer, O. Lehner, Using gaze behavior to measure cognitive load, in: Information Systems and Neuroscience: NeuroIS Retreat 2018, Springer, 2019, pp. 73–83.
[72] J.L. Plass, R. Moreno, R. Brünken, Cognitive load theory, https://doi.org/10.1017/CBO9780511844744, 2010.
[73] J. Radianti, T.A. Majchrzak, J. Fromm, I. Wohlgenannt, A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda, Comput. Educ. 147 (2020) 103778, https://doi.org/10.1016/j.compedu.2019.103778.
[74] E.D. Ragan, A. Sowndararajan, R. Kopper, D.A. Bowman, The effects of higher levels of immersion on procedure memorization performance and implications for educational virtual environments, Presence Teleoperators Virtual Environ. 19 (2010) 527–543, https://doi.org/10.1162/pres_a_00016.
[75] G.D. Rey, A review of research and a meta-analysis of the seductive detail effect, Educ. Res. Rev. 7 (2012) 216–237.
[76] M.A. Rojas-Sánchez, P.R. Palos-Sánchez, J.A. Folgado-Fernández, Systematic literature review and bibliometric analysis on virtual reality and education, Educ. Inf. Technol. 28 (2023) 155–192, https://doi.org/10.1007/s10639-022-11167-5.
[77] P. Rospigliosi, Metaverse or simulacra? Roblox, minecraft, meta and the turn to virtual reality for education, socialisation and work, 2022.
[78] E.R. Samuels, E. Szabadi, Functional neuroanatomy of the noradrenergic locus coeruleus: its roles in the regulation of arousal and autonomic function part I: principles of functional organisation, Curr. Neuropharmacol. 6 (2008) 235–253.
[79] T. Setti, A.B. Csapo, A canonical set of operations for editing dashboard layouts in virtual reality, Front. Comput. Sci. 3 (2021) 659600, https://doi.org/10.3389/fcomp.2021.659600.
[80] T. Setti, Á.B. Csapó, Outlines of a graph-tensor based adaptive associative search model for Internet of digital reality applications, in: 2022 IEEE 1st International Conference on Internet of Digital Reality (IoD), IEEE, 2022, pp. 000049–000054.
[81] T. Setti, Á.B. Csapó, Quantifying the effectiveness of project-based editing operations in virtual reality, in: 2022 1st IEEE International Conference on Cognitive Aspects of Virtual Reality (CVR), IEEE, 2022, pp. 000049–000054.

[82] C. Sibley, J. Coyne, C. Baldwin, Pupil dilation as an index of learning, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications Sage CA, Los Angeles, CA, 2011, pp. 237–241.

[83] J. Sipilä, Benefits and Challenges of Adopting Virtual Reality in Primary Education, Oulu University Library, 2020, p. 21.

[84] S. Solhjoo, M.C. Haigney, E. McBee, J.J. van Merrienboer, L. Schuwirth, A.R. Artino, A. Battista, T.A. Ratcliffe, H.D. Lee, S.J. Durning, Heart rate and heart rate variability correlate with clinical reasoning performance and self-reported measures of cognitive load, Sci. Rep. 9 (2019) 1–9, https://doi.org/10.1038/s41598-019-50280-3.

[85] B. Sousa Santos, P. Dias, A. Pimentel, J.W. Baggerman, C. Ferreira, S. Silva, J. Madeira, Head-mounted display versus desktop for 3d navigation in virtual reality: a user study, Multimed. Tools Appl. 41 (2009) 161–181.

[86] P. Srivastava, A. Rimzhim, P. Vijay, S. Singh, S. Chandra, Desktop vr is better than non-ambulatory hmd vr for spatial learning, Front. Robot. AI 6 (2019) 50, https://doi.org/10.3389/frobt.2019.00050.

[87] A. Sudár, A. Csapó, An mcmc-based method for clustering display panels with the goal of generating navigation paths in 3d, in: Proceedings of the 2021 12th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), 2021, pp. 23–25, Online.

[88] A. Sudár, Á. Csapó, Interaction patterns of spatial navigation and smartboard use in vr workspaces, in: Accentuated Innovations in Cognitive Info-Communication, Springer, 2022, pp. 149–166.

[89] A. Sudár, Á.B. Csapó, Descriptive markers for the cognitive profiling of desktop 3d spaces, Electronics 12 (2023) 448, https://doi.org/10.3390/electronics12020448.

[90] A. Sudár, Á.B. Csapó, Elicitation of content layout preferences in virtual 3d spaces based on a free layout creation task, Electronics 12 (2023) 2078, https://doi.org/10.3390/electronics12092078.

[91] D. Sulisworo, V.Y. Erviana, B. Robiin, Y. Sepriansyah, A. Soleh, et al., The feasibility of enhancing environmental awareness using virtual reality 3d in the primary education, Educ. Res. Int. 2022 (2022), https://doi.org/10.1155/2022/4811544.

[92] N. Sundararajan, O. Adesope, Keep it coherent: a meta-analysis of the seductive details effect, Educ. Psychol. Rev. 32 (2020) 707–734.

[93] J. Sweller, Chapter Two - Cognitive Load Theory, Psychology of Learning and Motivation, vol. 55, Academic Press, 2011, pp. 37–76.

[94] J. Sweller, Measuring cognitive load, Perspect. Med. Educ. 7 (2018) 1–2, https://doi.org/10.1007/s40037-017-0395-4.

[95] J. Sweller, J.J. van Merriënboer, F. Paas, Cognitive architecture and instructional design: 20 years later, Educ. Psychol. Rev. 31 (2019) 261–292, https://doi.org/10.1007/s10648-019-09465-5.

[96] J. Sweller, J.J. Van Merrienboer, F.G. Paas, Cognitive architecture and instructional design, Educ. Psychol. Rev. (1998) 251–296, https://doi.org/10.1023/A:1022193728205.

[97] J. Tham, A.H. Duin, L. Gee, N. Ernst, B. Abdelqader, M. McGrath, Understanding virtual reality: presence, embodiment, and professional practice, IEEE Trans. Prof. Commun. 61 (2018) 178–195, https://doi.org/10.1109/TPC.2018.2804238.

[98] M. Thees, S. Kapp, M.P. Strzys, F. Beil, P. Lukowicz, J. Kuhn, Effects of augmented reality on learning and cognitive load in university physics laboratory courses, Comput. Hum. Behav. 108 (2020) 106316, https://doi.org/10.1016/j.chb.2020.106316.

[99] M.L. Turner, R.W. Engle, Is working memory capacity task dependent?, J. Mem. Lang. 28 (1989) 127–154, https://doi.org/10.1016/0749-596X(89)90040-5.

[100] N. Unsworth, R.W. Engle, The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory, Psychol. Rev. 114 (2007) 104–132, https://doi.org/10.1037/0033-295X.114.1.104.

[101] P.W. Van Gerven, F. Paas, J.J. Van Merriënboer, H.G. Schmidt, Memory load and the cognitive pupillary response in aging, Psychophysiology 41 (2004) 167–174, https://doi.org/10.1111/j.1469-8986.2003.00148.x.

[102] T. Van Gog, F. Kirschner, L. Kester, F. Paas, Timing and frequency of mental effort measurement: evidence in favour of repeated measures, Appl. Cogn. Psychol. 26 (2012) 833–839, https://doi.org/10.1002/acp.2883.

[103] P. Vanneste, A. Raes, J. Morton, K. Bombeke, B.B. Van Acker, C. Larmuseau, F. Depaepe, W. Van den Noortgate, Towards measuring cognitive load through multimodal physiological data, Cogn. Technol. Work 23 (2021) 567–585, https://doi.org/10.1007/s10111-020-00641-0.

[104] M.L.H. Võ, A.M. Jacobs, L. Kuchinke, M. Hofmann, M. Conrad, A. Schacht, F. Hutzler, The coupling of emotion and cognition in the eye: introducing the pupil old/new effect, Psychophysiology 45 (2008) 130–140, https://doi.org/10.1111/j.1469-8986.2007.00606.x.

[105] C.A. Wang, D.P. Munoz, A circuit for pupil orienting responses: implications for cognitive modulation of pupil size, Curr. Opin. Neurobiol. 33 (2015) 134–140, https://doi.org/10.1016/j.conb.2015.03.018.

[106] P. van der Wel, H. Van Steenbergen, Pupil dilation as an index of effort in cognitive control tasks: a review, Psychon. Bull. Rev. 25 (2018) 2005–2015, https://doi.org/10.3758/s13423-018-1432-y.

[107] N. Wenk, J. Penalver-Andres, K. Buetler, T. Nef, R.M. Müri, L. Marchal-Crespo, Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment, Virtual Real. 27 (2023) 307–331, https://doi.org/10.1007/s10055-021-00565-8.

[108] B. Wu, X. Yu, X. Gu, Effectiveness of immersive virtual reality using head-mounted displays on learning performance: a meta-analysis, Br. J. Educ. Technol. 51 (2020) 1991–2005, https://doi.org/10.1111/bjet.13023.

[109] J. Zagermann, U. Pfeil, H. Reiterer, Measuring cognitive load using eye tracking technology in visual computing, in: Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization, Association for Computing Machinery, New York, NY, USA, 2016, pp. 78–85.