Investigating the redundancy principle in immersive virtual reality environments: An eye-tracking and EEG study

Article in Journal of Computer Assisted Learning · February 2022

DOI: 10.1111/jcal.12595

4 authors:

Sarune Baceviciute
University of Copenhagen

Gordon Lucas
University of Copenhagen

Thomas Terkildsen
Aalborg University

Guido Makransky
University of Copenhagen

Received: 7 April 2021 | Revised: 18 June 2021 | Accepted: 11 July 2021
DOI: 10.1111/jcal.12595

ARTICLE

Journal of Computer Assisted Learning WILEY

Investigating the redundancy principle in immersive virtual reality environments: An eye-tracking and EEG study

Sarune Baceviciute | Gordon Lucas | Thomas Terkildsen | Guido Makransky

Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Correspondence
Sarune Baceviciute, University of Copenhagen, Øster Farimagsgade 2A, 1353 København K, Denmark.
Email: sb@psy.ku.dk

Funding information
Innovationsfonden

Abstract

Background: The increased availability of immersive virtual reality (IVR) has led to a surge of immersive technology applications in education. Nevertheless, very little is known about how to effectively design instruction for this new media, so that it would benefit learning and associated cognitive processing.

Objectives: This experiment explores if and how traditional instructional design principles from 2D media translate to IVR. Specifically, it focuses on studying the underlying mechanisms of the redundancy-principle, which states that presenting the same information concurrently in two different sensory channels can cause cognitive overload and might impede learning.

Methods: A total of 73 participants learned through a specifically-designed educational IVR application in three versions: (1) auditory representation format, (2) written representation format, and (3) a redundancy format (i.e. both written and auditory formats). The study utilized advanced psychophysiological methods of Electroencephalography (EEG) and eye-tracking (ET), learning measures and self-report scales.

Results and Conclusions: Results show that participants in the redundancy condition performed equally well on retention and transfer post-tests. Similarly, results from the subjective measures, EEG and ET suggest that redundant content was not found to be more cognitively demanding than written content alone.

Implications: Findings suggest that the redundancy effect might not generalize to VR as originally anticipated in 2D media research, providing direct implications to the design of IVR tools for education.

KEYWORDS
EEG, eye-tracking, immersive virtual reality, learning, redundancy principle

1 | INTRODUCTION

Educators and instructional designers around the globe are in search of new and alternative ways to engage and educate the new generation of students. Considering the recent popularity of immersive virtual reality (IVR) and acknowledging its captivating nature, it is not surprising that this technology is becoming more frequently used in various educational contexts (Raditanti et al., 2020). IVR tools have, for instance, already been incorporated in the teaching of curricula at high school and university levels (Makransky et al., 2021; Jones, 2018). IVR is also emerging in the training of professionals in organizational settings (Butussi & Chittaro, 2018; Chittaro & Buttussi, 2015; Muller Queiroz et al., 2018).

Incipient research investigating digital learning suggests that IVR can function as a powerful motivational aid (Makransky & Lilleholt, 2018, Makransky & Petersen, 2019; Chittaro & Buttussi, 2015; Huang et al., 2020). A recent meta-analysis by Wu et al. (2020) also found an advantage of IVR lessons compared to less-immersive learning approaches on learning outcomes. Cummings and

Bailenson (2016) define immersion as an objective measure of the vividness offered by a system, and the extent to which the system is capable of shutting out the outside world. Therefore, IVR lessons accessed through head mounted displays (HMDs) are often referred to as immersive lessons, and lessons accessed through traditional 2D monitors are often referred to as less immersive media or non-immersive media (Wu et al., 2020). The immersion principle in multimedia learning (Makransky, 2021) and the cognitive affective theory of immersive learning (CAMIL; Makransky & Petersen, 2021) describe how the fundamental driver of increased learning outcomes in immersive media is the use of instructional design principles that are effective in immersive lessons. Latest research has also shown that how well IVR promotes learning is greatly dependent on how IVR-specific content has been designed (Meyer et al., 2019; Baceviciute et al., 2020; Makransky, 2021; Luo et al., 2021; Parong & Mayer, 2018). In this direction, recent reviews have highlighted several gaps in IVR based educational research and propose that future research should: (1) Use learning theories to guide IVR based application development and research (Raditanti et al., 2020); (2) Shift attention from VR technology to VR-based instructional design with a redefined focus on the effective integration of technology and theory (Luo et al., 2021); and (3) Use more diversified research designs and methods to improve the rigour and relevance (Luo et al., 2021).

The CAMIL provides a theoretical framework for understanding and investigating learning in immersive environments such as IVR. The CAMIL identifies presence and agency as the two main affordances of learning in immersive environments builds on existing learning and motivational theories to describe how presence and agency influence learning through several affective and cognitive factors such as interest, motivation, self-efficacy, embodiment, cognitive load, and self-regulation (Makransky & Petersen, 2021). The model describes that it is not the medium of IVR that causes specific learning outcomes, but rather the instructional methods used in IVR that will constitute its effectiveness. The CAMIL builds on empirical evidence that media interacts with method, meaning that learning methods affect learning, but certain methods are more or less relevant in IVR. For instance, research has identified instructional methods such as the pre-training principle (Meyer et al., 2019, Petersen et al., 2020), and generative learning strategies such as enactment (Makransky et al., 2021), and summarization (Klingenberg et al., 2020) to be more effective in more immersive compared to less immersive learning environments. Such findings therefore suggest that it is important to conduct research that specifically investigates how instructional design principles developed in 2D media generalize to immersive learning environments, rather than conducting media comparison studies that confound instructional design factors (Makransky et al., 2019b, Baceviciute et al., 2020). This knowledge is necessary so that instructional designers can develop effective learning material for IVR and related learning technologies.

The current experiment investigates issues related to written and auditory informational representations in educational IVR environments (IVREs), as these continue to be used as the primary vehicles for representing learning content not only in non-immersive, but also in immersive media (Baceviciute et al., 2021). Specifically, we focus on the redundancy principle from the cognitive theory of multimedia learning (CTML), which states that presenting the same information concurrently in two different sensory channels (i.e., auditory and visual) can cause cognitive overload and might impede learning (Mayer, 2014, 2020). Understanding the impact and underlying mechanisms of visual and auditory redundancy is important because instructional designers are typically faced with instructional design decisions related to effective learning information representations in immersive educational applications. Although there is evidence for the redundancy principle in 2D media (Adesope & Nesbit, 2012), the articles that have investigated the redundancy principle in IVR (Makransky et al., 2019b; Moreno & Mayer, 2002) have not found evidence for the principle. Existing results suggest that redundant information in immersive lessons could potentially have beneficial as well as detrimental consequences to learning. The redundancy principle was thus selected to be investigated in this study because there is inconsistency between the evidence for the principle when comparing 2D and immersive environments. Furthermore, providing redundant information may be specifically relevant in IVR settings where learners can view and interact with many elements in an immersive 360-degree environment. This is fundamentally different from learning with a 2D monitor where learners have a visual overview of an entire environment. In the current study, we use advanced psychophysiological methods, including electroencephalography (EEG) and eye-tracking (ET), learning measures, and self-reported scales to gain a better understanding of the underlying mechanisms of the redundancy principle in immersive learning.

2 | BACKGROUND

2.1 | IVR for learning and education

IVR can be conceptualized in various ways. In this article we refer to it as a complex media system that on the one hand consists of a unique technological setup, which encompasses sensory immersion made available through a head mounted display (HMD; Howard, 2019), and on the other – of immersive content that capitalizes on technological immersion to represent pedagogy (Mikropoulos & Natsis, 2011). While IVR is still not an integrated learning tool, the last decade has seen the technology become widely explored in various educational contexts spurred in part by its captivating nature and ability to separate the learner from external distractions (Raditanti et al., 2020). IVR has, for example, been used to supplement teaching at school (Petersen et al., 2020, Makransky et al., 2021); while others have also used it for informal learning (Christensen & Knezek, 2016). IVR has also been applied in various educational levels: from K-12 instruction to higher education (Makransky et al., 2019a, Makransky et al., 2021; Jones, 2018; Luo et al., 2021) to professional training in industrial contexts (Butussi & Chittaro, 2018; Chittaro & Buttussi, 2015; Muller Queiroz et al., 2018; Tang et al., 2020). Applications of IVR also span across different fields; however due to the unique ability of the technology to facilitate the visualization of complex phenomena that is hard to access or to explain without technological support and very specialized tools (Jensen & Konradsen, 2018; Johnson-Glenberg, 2019), IVR has become especially popular in STEM education (e.g., biology, physics and math; see Raditanti et al., 2020).

Following this emergence of IVR in education, educational psychology and instructional design researchers have begun to examine whether immersive technology can in fact benefit learning. Evidence supports its motivational benefits, suggesting that students enjoy learning digitally more than traditional methods (Makransky & Lilleholt, 2018; Makransky & Petersen, 2019; Makransky & Petersen, 2019; Makransky et al., 2020; Bogusevschi et al., 2019), and that educational content is perceived as more engaging when presented in an immersive format (Makransky et al., 2019b; Parong & Mayer, 2018). Furthermore a meta-analysis by Wu et al. (2020) provides evidence that immersive technologies have a small positive effect on knowledge acquisition as well as skill development compared to more traditional media. This is supported by the meta-analysis by Luo et al., 2021 who also found a medium effect for HMD-based lessons. There is however, variance regarding the effectiveness of IVR for learning, and several studies have identified negative implications of using IVR in education. Some, for example have discussed the isolating nature of IVRs (Mütterlein & Hess, 2017), while other studies have found it to lead to extraneous cognitive load (CL; Makransky et al., 2019b; Richards & Taylor, 2015).

One challenge is that many studies take a purely techno-centric approach to IVR based learning, which does not consider that IVR also incorporates educational content that needs to be strategically designed and evaluated to promote pedagogy (Baceviciute et al., 2020; Fowler, 2015; Jensen & Konradsen, 2018; Mikropoulos & Natsis, 2011). Recent research in this direction has started to produce empirical evidence for the importance of instructional design in IVR. One study, for example, exported a non-immersive VR simulation to an immersive format without optimization, and showed that direct translation of content from 2D media to 3D can lead to lower learning and a heightened CL to the learner (Makransky et al., 2019b). In a follow-up study, no diminishing effects were found on learning when translating learning content from 2D to 3D with respect to unique affordances of VR (Baceviciute et al., 2021). The authors concluded that for IVR to be successful in education, instruction and learning content needs to be specifically designed to fit the affordances of immersive technology. Similarly, prior research found that auditory informational representations were not as effective as written representations when comparing learning outcomes of retention, self-efficacy, intrinsic CL and extraneous attention (Baceviciute et al., 2020). EEG frequency comparisons performed in the study suggested that auditory informational representations were also not as cognitively stimulating (Baceviciute et al., 2020). Other studies that have investigated the importance of instructional design in IVR have found differences in learning effectiveness when using different pedagogical agents in IVR (Makransky et al., 2019c). Studies have also identified the importance of using scaffolding strategies such as pre-training (Meyer et al., 2019; Petersen et al., 2020), as well as generative strategies of summarizing (Parong & Mayer, 2018) and enacting after an IVR lesson (Makransky et al., 2021). These results not only suggest that the design of learning content is imperative for learning efficacy of IVR, but also show that traditional instructional design principles from non-immersive media might not always directly translate to IVR applications, necessitating further and more in-depth investigations into instructional learning content design in this medium.

2.2 | The redundancy principle in multimedia learning

Contrary to the intuitive belief that presenting the same information in various formats enhances learning, the redundancy principle states that redundant information inhibits learning (Mayer, 2014, 2020). This finding has been observed in numerous studies (Craig et al., 2002; Gerjets et al., 2009; Kalyuga et al., 2004; Mayer et al., 2001) and is based on the Cognitive Load Theory (CLT; Sweller, 2011) and CTML. These theories explain that the redundancy effect occurs due to an increase in extraneous CL that arises due to concurrent processing of redundant information. The need to process redundant information sources generates strong demands on the learners' working memory (WM), and thus cognitive resources are not spent on learning. Processing novel information is heavily constrained by working memory capacity and duration, and without rehearsal can only be stored in short term memory for a brief period of time. As such, according to CTML, instructional design should aim to minimize any unnecessary WM load in the presentation of novel information. Based on this, the redundancy principle formulated in the CTML (Mayer, 2014; Mayer & Johnson, 2008) states that redundant information should generally be avoided during learning, since '*people learn better when the same information is not presented in more than one format*' (Mayer, 2014, pp. 19–20).

What information is redundant, however, might depend on the learning context, as well as the learners' expertise (Mayer, 2014). As an example, in complex learning scenarios novice learners might use concurrent information representations as supporting explanatory material. However, as their levels of expertise increase and the need for additional explanation decreases, this information will eventually become redundant. A meta-analysis carried out by Adesope and Nesbit (2012) summarized the data of 57 studies to estimate effect sizes comparing combined auditory and written redundancy conditions to either written-only or auditory-only representations. Their analysis shows that across all studies redundancy slightly improves learning outcomes (Hedges $g = 0.15$). For example, redundancy conditions had no advantage compared to written-only conditions ($g = -0.04$). On the other hand, redundancy enhanced learning when contrasted with auditory-only representation ($g = 0.29$). This advantage stems mostly from studies where correspondence between the auditory and written text was low ($g = 0.99$), rather than high, ($g = 0.21$). The prevalence of the redundancy effect was further moderated by factors such as learners' prior knowledge, their freedom in pacing the learning content, or the simultaneous presentation of other visual information, such as animations and diagrams (Adesope & Nesbit, 2012). While this meta-analysis did not specifically investigate the redundancy principle in IVR, its findings suggest that a general applicability of the principle cannot be supported across different media and educational contexts.

Few research studies have examined the redundancy principle in IVR. Moreno and Mayer (2002) investigated the redundancy effect in a VR simulation across two different media conditions (i.e., IVR, and desktop VR) and three different method conditions (i.e., auditory text, written text, and redundancy). There was no difference between the redundancy and auditory conditions on the outcomes of retention and transfer, but both conditions significantly outperformed the text-only condition on these outcomes. The authors concluded that the findings are inconsistent with prior studies on redundancy (Moreno & Mayer, 2002). Their interpretation is that it is possible that students in the redundancy condition may have focused on the auditory narration alone. The authors reasoned that this might be a consequence of the experiential nature of the IVRE, making learners less likely to read a text box if they can obtain the same information by listening to a narration. However, as Moreno and Mayer (2002) did not have access to gaze data, their interpretation could not be explored and corroborated. In a recent media and methods experiment (Makransky et al., 2019b) also investigated the redundancy effect across desktop and immersive versions of VR simulations. In accordance with the previous study, the authors failed to find evidence for the redundancy principle across both media conditions. These initial findings suggest that the redundancy principle, originally conceived in 2D media, might not be extendable to IVR, but the mechanisms underlying these findings are not clear.

No studies have investigated whether learners primarily read or listen to text when learning in redundancy conditions in IVR. Therefore, in the current study we want to examine whether learners in the redundancy condition attend more to the auditory or written information using ET. This would provide valuable information about the underlying processes that take place when attending to and learning from different information representation methods in IVR. Addressing gaps in existing literature, another aim of this study is to gain greater insight into the cognitive demands imposed on the learner when learning with redundant information. In CLT (Sweller, 2011; Sweller et al., 2011), three dimensions of CL have been proposed: Intrinsic CL (i.e., intrinsic difficulty of the topic/learning material), extraneous CL (i.e., CL imposed by factors external to the learning material, e.g., instructions, explanations), and germane CL (i.e., effort that is required for learning). Traditionally, CL has been assessed using singular self-report items (Ayres, 2006; Cierniak et al., 2009; Paas, 1992; Salomon, 1984). To combat the lack of a uniformly used scale, Leppink et al. (2013) developed and validated a CL scale which measures CL demands more reliably. The self-reported items, however, have limitations (such as self-report bias) which do not provide the full insight of cognitive processing during learning (Makransky et al., 2019b). To supplement the self-report items, this study also attempts to measure CL with EEG and ET.

2.3 | Using EEG to measure cognitive load during learning

Several studies have explored the use of EEG as an effective online measure of cognition during learning across media, including IVR (Antonenko et al., 2010; Makransky et al., 2019b, Baceviciute et al., 2020; Baceviciute et al., 2021; Örün & Akbulut, 2019). In particular, frequency-based analyses of EEG data have recently seen traction as an unobtrusive measure that can be used during learning (Antonenko & Keil, 2017; Baceviciute et al., 2020; Baceviciute et al., 2021; Scharinger, 2018). Previous experimental and theoretical work has focused on oscillations in the Theta and Alpha frequency bands. These have been consistently demonstrated to be sensitive to the changes in cognitive processes, such as attention and WM load, which are relevant for novel information acquisition (Antonenko & Keil, 2017; Brouwer et al., 2012). Generally, increases in Theta activation (4–8 Hz) have been previously linked to increased mental effort (Klimesch, 1999). More specifically, Theta frequency activity in frontal areas, has been linked to working memory capacity across several studies (Puma et al., 2018). In these studies, increasing levels of spectral power in the Theta band is proposed to reflect increasing WM load (Mühl et al., 2015). Parietal Theta, on the other hand, has been linked to effective long-term memory encoding, suggesting that increases in parietal Theta could be later linked to successful memory retrieval, which is vital for learning (Osipova et al., 2006). Given that redundancy of learning information is theoretically believed to be more difficult as it elicits higher levels of extraneous WM load, such literature suggests that the redundancy format would have higher levels of frontal Theta in comparison with the other conditions.

Oscillatory activation in the Alpha frequency band (8–12 Hz) has been previously linked to changes in attentional processes (Frey et al., 2014). Generally, Alpha frequency activation is known to decrease with attentional engagement (i.e., in wake states), and increase in states of low cortical arousal (i.e., during sleep) (Antonenko & Keil, 2017; Klimesch, 1999). Lower levels of Alpha power could therefore be expected in redundancy conditions, given that redundant information requires more CL since the inputs from both sensory modalities would require more attentional resources, and thereby increase CL.

2.4 | Eye tracking during learning

van Gog and Scheiter (2010) discussed the use of ET as an additional tool to study the learning process, particularly for research with multimedia learning. ET allows researchers to look beyond performance measures to study what media or representations are visually attended to by learners, thus giving insight into the origin of well-known effects such as the redundancy effect or the modality effect. Note, however, that ET offers no explanation of why participants are attending to stimuli in a certain order or duration (van Gog & Scheiter, 2010). One example of how ET was used in the framework of CTML is the study by Schmidt-Weigand et al. (2010), which investigated the modality effect with animations wherein explanatory text was either written or auditory. They found evidence for the split-attention effect in the written condition. Crucial insight was gained by viewing tie measure determined by ET (i.e., extracted from fixation and saccade durations), which revealed how participants in the writing conditions would begin reading but then are forced to divide their attention between the text and the animation. While retention, transfer and visual memory task scores did not differ between the two groups, ET showed how participants in the written text condition spent most of their time on task fixating on the written text rather than the animation (Schmidt-Weigand et al., 2010). In their study of the redundancy effect in multimedia web pages, Liu et al. (2011) also observed this preference for the written text over the explanatory image material. The authors found significantly more and longer fixations in the written text condition than in the auditory condition. However, the redundancy condition group spent less time fixating on the text than the written text only group. In a similar methodology, De Koning et al. (2010) employed ET to measure visual attention allocation via relative fixation times on relevant areas of interest (AOIs). Total fixation times on AOIs were theorized to be an indication of greater cognitive processing, and as such longer time spent viewing was thus generally predicted to cause greater learning (De Koning et al., 2010). A review of the use of ET in research on learning has since reinforced this notion (Lai et al., 2013).

Even though gaze measures (i.e., fixation length and duration) are predominant in ET, other ET measurements have also been investigated in WM load and reading studies. For example, blinking has been proposed to be indicative of mental load (Holland & Tarlow, 1972), and researchers observed that blinking decreases during cognitive processing and memory workload (Holland & Tarlow, 1975). This was explained by the connection of the visual mental operations and the visual perceptual system. As such, blinking might be suppressed to enhance visual processing. Stern and Skelly (1984) tested experimentally whether blinking rate and duration vary depending on task demand and task modality. In two experiments it was shown that blink rate is significantly affected by task demand, with higher task demand causing a lower blinking rate. Furthermore, performing a visual task led to a lower blinking rate than performing an auditory task. In the context of textual-auditory redundancy, the expectancy therefore would be for visually richer representations (i.e., those involving written text) to produce lower blink rates than auditory representations. More recently, a systematic review showed the usefulness of blinking as a measurement of mental load and mental fatigue (Martins & Carvalho, 2015). Specifically, Martins and Carvalho (2015) found an inverse relationship of task difficulty and blinking, that is, higher difficulty results in less blinking. Since redundancy of information is thought to be more cognitively loading that non-redundant information, we could therefore assume lower blink rates with concurrent information representations rather than when attending to non-redundant learning content.

Although less investigated, saccadic eye movements (i.e., the voluntary movement of an eye between two fixation pints) have also previously been reported as another ET measure to successfully capture differences in WM load and cognitive processing. Prior studies have, for instance, already related increases in velocity and length of saccadic eye movement to higher task difficulty and conversely that decreases in saccade velocity might indicate tiredness and lower task performance (Zagermann et al., 2016). Assuming that redundancy of information increases CL, we would expect higher saccadic movement when learning with redundant content. In reading research, saccadic eye movements have for the most part been investigated over meaningless word strings, providing little support for learning-relevant investigations (Boland, 2004). No comparative studies have been produced in listening research.

2.5 | Research Questions

Building on prior research from instructional design, IVR and learning, as well as novel psychophysiological measurement techniques, we aim to investigate the following four research questions in this study:

- **RQ 1**: How does redundant information influence the learning outcomes of retention and transfer in IVR?
- **RQ 2**: Are redundant information representations perceived to be more or less cognitively demanding than non-redundant information representations when assessed with self-reported CL measures in IVR?
- **RQ 3**: How do cognitive processing demands differ when learning with redundant and non-redundant information representations in IVR? How these differences are reflected in EEG Theta and Alpha frequency band activations?
- **RQ 4**: Are there any differences in visual attention, as observed by ET, when learning with redundant and non-redundant information formats in IVR? Do participants pay more attention to learning irrelevant stimuli in redundant or in non-redundant information?

3 | METHODS

3.1 | Participants

In total, 73 fluent English-speaking and normal-sighted participants (44 female) without prior knowledge of the presented topic and not diagnosed with any neurological illness or a learning disorder partook in the experiment. Participants were 19–41 years old ($M = 23.97$, $SD = 3.78$) and were recruited via university mailing lists and social media channels. Partaking in the study was voluntary. Participants signed an informed consent form prior to the experiment. Permission for conducting the study was obtained from the institutional board. Due to errors during ET and EEG data collection (e.g., incomplete data sets, faulty calibration procedures), data of several participants was excluded from certain analyses in this study. The final sample size included in the ET data analysis is 68 participants, and in the EEG data analysis is 63 participants.

FIGURE 1 Different learning content representations used for the written condition (left), auditory condition (middle) and redundancy condition (right) [Colour figure can be viewed at wileyonlinelibrary.com]

3.2 | Experimental design

Research questions (Section 2.5) were investigated using a between-subjects design with three experimental conditions wherein learning material presented was identical, but its representation varied (see Figure 1). In the first condition ($N = 25$, 15 female) information was represented as read-out-load text (auditory condition); in the second condition ($N = 24$, 14 female) the same material was displayed as written text on an overlay reading interface (written condition). Participants in the last condition (redundancy condition) received both written and auditory learning content representations from the first two conditions ($N = 24$, 15 female). Group assignment was randomized prior to arrival of participants through the use of unique participant IDs. Demographics, prior knowledge, and reading habits were assessed via a pre-test survey. Learning outcome variables and CL measures were collected immediately after the IVR learning experience by subjecting participants to a post-test. Psychophysiological cognitive learning measures (ET and EEG) were recorded during the entire learning experience, not including the pre- or post-test.

3.3 | Experimental procedures

Each participant was tested individually in an experimental psychology lab. The experimental procedure was as follows (~90 min): (1) participant briefing, (2) signing an informed consent form, (3) mounting of the EEG headset, (4) EEG signal quality and impedance test, (5) pre-test survey, (6) introduction to the VR controls, (7) VR HMD mounting, (8) EEG signal quality and impedance test (9) VR learning experience (~15 min), (10) dismounting of the EEG and the VR HMD, (11) post-test survey, (12) participant debriefing. During the VR learning experience, the participants were seated and asked to avoid unnecessary movements to maximize the quality of the psychophysiological recordings. Participants were rewarded for their participation with a gift card valued at approximately 15 Euros. The procedure was semi-automated with the help of the iMotions experiment facilitation software.

3.4 | Materials

Experimental materials consisted of an IVR simulation, a pre-test and a post-test survey, and psychophysiological measurements (i.e., ET and EEG).

3.4.1 | IVR simulation

The Unity3D game development engine was employed to develop the IVR simulation used in this study. The simulation was run on the HTC Vive VR system. To represent current virtual learning content remediation trends (see Baceviciute et al., 2021) and to control for information delivery format, the IVR simulation was designed to consist of two main components: explicit learning content represented in three different formats (see Figure 1), and an IVE in which those formats were embedded in.

The IVE in the simulation was developed to represent a virtual hospital room in order to establish semantic relations with the learning content used in the study. To simulate a hospital room scenario, the IVE was equipped with several, archetypal props (i.e., hospital cabinets, a painting, a TV screen, etc.), and a soundscape matching the environmental setting. Two virtual characters, a doctor and a patient, also populated the simulation. Although explicit learning content was contained to three explicit learning content representations, the IVE helped to contextualize learning content (see Baceviciute et al., 2021). The participant's character was not embodied by a virtual avatar. In the simulation the participant was seated on a virtual chair. The simulation started with the doctor avatar entering the room. Prior to the display of the learning content, three information snippets were provided to introduce the participants to the controls of the simulation and to explain the experimental task.

Explicit learning content used in the simulation was an expository science text on the topic of Sarcoma cancer. All of the learning content was developed based on an information pamphlet provided by a national cancer society, designed to inform the general public and thus assumed no prior knowledge of the topic. At the start of the simulation, participants were tasked to gather information on Sarcoma cancer, as if they were to retell the information to a friend after the experience. Adapted learning content was split into 24 snippets of text with the length of 300–400 characters, each of which delivered a unique piece of information. Following experimental study design, three different representation means were developed for representing content snippets. For the written condition a static overlay interface showing the text was superimposed on the scene (Figure 1). In the auditory condition, identical learning content was played back as a non-diegetic voice over. The voice over was produced by recording a voice actor reading out written snippets of text. Audio was delivered to the participants via built-in HTC Vive headphones. In the redundancy condition both representations were present, therefore the audio was played back at the same time as the text was presented to be read on the overlay interface. Throughout all experimental conditions the order and semantic representation of the snippets was kept identical. In the two conditions that included written text representations, visual features (i.e., font type, line spacing, etc.) and formatting (i.e., paragraph structure, indentation, etc.) of the text were also kept consistent. After each snippet, the participants signalled that they finished processing the information by pressing a button on the HTC VR controller. The appearance of the subsequent snippet of text was triggered by a second button press. Although the participants were able to control the pace of appearance of the snippets, learning content presentations was for the most part sequential, that is, participants could not stop, rewind, or replay a given snippet. Triggers recorded by button presses later served the secondary purpose of epoching EEG and ET signals.

3.4.2 | Pre-test survey

The purpose of the pre-test was to capture demographic information, current reading habits, and the level of prior knowledge about Sarcoma cancer. The prior knowledge test contained seven questions on the topic of Sarcoma cancer: one 5-point Likert scale question assessing the overall familiarity with Sarcoma cancer (i.e., '*Please indicate how familiar would you consider yourself to be with the topic of Sarcoma cancer*'), and six yes/no questions regarding the specific concepts related to the learning material (e.g., '*I know what the two most common types of sarcoma cancer are*'). A total prior knowledge score was calculated by adding all prior knowledge items together. Additional survey questions asked participants to report their current mental state and any use of psychoactive drugs (i.e., caffeine, nicotine and alcohol) on the day of the experiment.

3.4.3 | Learning assessment instruments

To answer RQ 1 (Section 2.5) two tests were customarily designed to quantify participants' learning outcomes: a knowledge retention test consisting of 24 multiple-choice questions (one for each snippet from the simulation), and a knowledge transfer test consisting of three open-ended questions. The tests were based on methods previously used in similar studies (e.g., Makransky et al., 2019a, 2019b; Baceviciute et al., 2020). The goal of the retention test was to measure how well the participants retained the information conveyed in the snippets (e.g., *Snippet text: Bone sarcoma occurs in the body's bone tissue, especially around the shoulder, knee or hip joints. Question: Which bones are most commonly affected by bone sarcoma? Multiple choice: (A) Bone sarcomas often occur around the shoulder, knee or hip joints [correct answer], (B) Bone sarcomas often occur in the bones around the feet or hands, (C) Bone sarcomas often occur in or around the elbows or wrists, (D) Bone sarcomas often occur around the chest and/or the back bones*). The transfer test, on the other hand, required that the participants used the knowledge from the overall learning experience and applied it to a novel context, measuring comprehension of the learnt material (e.g., *Imagine the scenario – you are an oncologist and your patient, who is diagnosed with Sarcoma cancer, is not responding to your treatment plan, what would your next steps be and why?*). Learners were given 3 min to respond to each question. The knowledge transfer test was administered first, followed by the knowledge retention test. The knowledge transfer test was coded by two independent evaluators. These graders anonymously scored each item by summing up all correctly stated components (1–4 points per answer). Afterwards, both evaluators were invited to an open discussion panel, where they settled any discrepancies in their scores. A participant's final transfer score was then calculated by summing the scores of the three questions (maximum of 12 points). An individual's score on the knowledge retention test was determined by simply adding the correctly answered multiple-choice items together (maximum of 24 points).

3.4.4 | Self-reported cognitive load scales

Two measures were used to assess participants' self-reported CL experienced during the immersive VR learning simulation (RQ 2). The first measure was composed of four widely used individual items in CL research: an item from Paas (1992) focusing on overall mental effort invested during learning, an item from Ayres (2006), probing perceived difficulty of the learning content, an item from Cierniak et al. (2009) measuring the perceived difficulty of the provided textual format, and an item from Salomon (1984) where participants reported how well they concentrated during the learning experience. All items were scored on a 9-point Likert scale. Secondly, we employed a 10-item validated CL instrument developed by Leppink et al. (2013). This instrument was comprised of three items for measuring intrinsic CL, three items measuring extraneous CL, and four items measuring germane CL (Leppink et al., 2013). Participants reported their answers on 5-point Likert scales.

3.4.5 | EEG measurement

To further gain insight into cognitive processing during learning (RQ3, Section 2.5), participants' EEG data was collected using the Advanced Brain Monitoring (ABM) X-10, wireless 9-channel EEG. This device samples brain data at a rate of 256 hz. The Ag/AgCl electrodes were placed at Fz, F3, F4, Cz, C3, C4, POz, P3, P4 and referenced to two connected mastoids, with impedance levels maintained below 10 kΩ. EEG data was synchronized with the presentation of the learning material using the ABM external Sync Unit (ESU) and Cedrus Stim Tracker. Data collection and storage was handled via iMotions biometric data acquisition software.

EEG data pre-processing was conducted using Matlab's EEGlab toolkit. First, the raw EEG data was filtered with a high-pass filter (0.5 Hz) and a low-pass filter (100 Hz). The automatic channel rejection tool from EEGlab was used to reject channels with improbable signal distributions (probability z-scores above 5). All electrodes were re-referenced to average references and line noise was removed at

50 and 100 Hz using a CleanLine filter. Subsequently, manual visual inspection was performed wherein all irregular noise activity, such as short bursts stemming from muscle activity, was removed. Independent component analysis (ICA) was further used to remove artefacts stemming from eye-movements and blinks. Artefact removal procedures were semi-automated by combining thorough visual EEG data analysis and the MARA algorithm (Multiple Artifact Rejection Algorithm). Lastly, to isolate the sections when the participants were engaging with the learning material, the continuous stream of EEG data was epoched using triggers generated by the button presses produced by the participants.

EEG Power Spectral Density (PSD) estimates were calculated using the discrete Fourier transform (DFT) with a Hanning window of 1 s width and 50% overlap, enabled by the NeuroSpec toolbox for MATLAB (Halliday et al., 1995). The resulting data was normalized and log-transformed in order to minimize skewness in the dataset and to standardize unit variance. Following prior work (e.g., Baceviciute et al., 2021; Baceviciute et al., 2020; Klimesch, 1999), for each frequency band a mean peak frequency estimate was calculated in SPSS. The following limits were applied: 4–7 Hz for Theta and 8–13 Hz for Alpha.

3.4.6 | Eye tracking (ET) data collection and analysis materials

In order to investigate RQ4, we employed a HTC Vive with Tobii Pro eye tracking retrofit hardware, which was digitized at 80 Hz. Before starting the learning experience, each participant performed a five-point gaze calibration task designed by Tobii, specifically for use in VR (Tobii, 2020). This task would be re-run until the calibration outcome provided by the Tobii SDK showed that a good or excellent calibration had been achieved. A good calibration required a mean distance of measured gaze from the target calibration point to be less than 40 pixels, whereas the mean difference threshold for achieving an excellent calibration was less than 20 pixels. All participants managed to calibrate within these thresholds.

In this study we particularly focused on collecting real-time gaze data (i.e., fixation and saccades) and on determining participant's blink-rate during the learning experience. These measures were collected for the overall learning experience, as well as for three dynamic AOIs specified for this study (see Figure 2). The first AOI covered the doctor character, enabling tracking of how much participants focused on the virtual agent during the learning experience. The second AOI contained the overlay reading interface and was thus only present in the interface and redundancy conditions. This AOI was used to measure how much time participants spent reading, as well as to estimate the cognitive effort put into reading. The last AOI was placed over the environmental props collectively and was used to measure observation of the environment and extraneous attention paid to task-irrelevant objects.

The ET data was processed using an I-VT filter for gaze analysis and the gaze-data was mapped to the three pre-defined AOIs. As a means of investigating where participants directed their gaze and attention during the simulation, we investigated the time spent looking at the AOIs. Further, we separated the raw data of the eye-tracker into blinks, fixations and saccades. Counts were normalized to an average per minute to account for the variable time in the simulation. We compared the overall blinking rate and the blinking rate while looking at the interface AOI. To further compare reading styles between the written and redundancy conditions, we looked at various metrics regarding their eye-movements. The four measures were saccades per minute, average saccade amplitude, average saccade distance, and average saccade duration. These were calculated for the entire simulation and the interface AOI. Furthermore, we investigated data regarding fixations for the entire simulation and for each respective AOI. Two metrics were derived: average fixation count per minute and average fixation duration.

3.4.7 | Extraneous attention measure

To further understand visual attention demands when learning with different information representation displays in IVR (RQ 4), an extraneous visual attention measure was employed. Six open-ended questions were asked to probe the participants' attention to task irrelevant stimuli (i.e., painting, clock, TV screen, and patient number). The questions were focused on assessing if the participants could remember specific details about these peripheral objects in the environment (e.g., *Question: 'There was a painting hanging across from you in the hospital room - which object was drawn on the painting*?' *Answer*: '*Flower/Leaf/Plant*'). The number of correct answers was totalled to a final 'extraneous attention measure' (maximum score of 13).

4 | RESULTS

A comparison of the three groups on the retention and transfer scores, extraneous attention measure, CL items and scales, EEG frequency band averages, and ET measures were calculated using one-way analyses of variance (ANOVAs) in IBM SPSS 2019. In case of significant differences, a Tukey's post-hoc t-test was performed. Effect sizes were estimated by calculating Cohen's Delta. Significance level was set to 0.05 for all analyses.

4.1 | Did the groups differ on basic characteristics?

Before investigating the four research questions, we determined whether the three experimental groups differed on basic characteristics. Analyses revealed no significant differences between the groups in prior knowledge, $F_{(2,70)} = 0.502$, $p = 0.608$, reading habits, $F_{(2,70)} = 0.352$, $p = 0.705$, or in familiarity with VR, $F_{(2,70)} = 0.635$, $p = 0.533$. Further, a Chi-square test was used to investigate differences in the proportion of men and women between the groups. No significant differences were found in gender distribution, $X^2_{(2, N = 73)} = 0.088$, $p = 0.957$. As such, the results indicate that there were no significant differences between the learners in the three

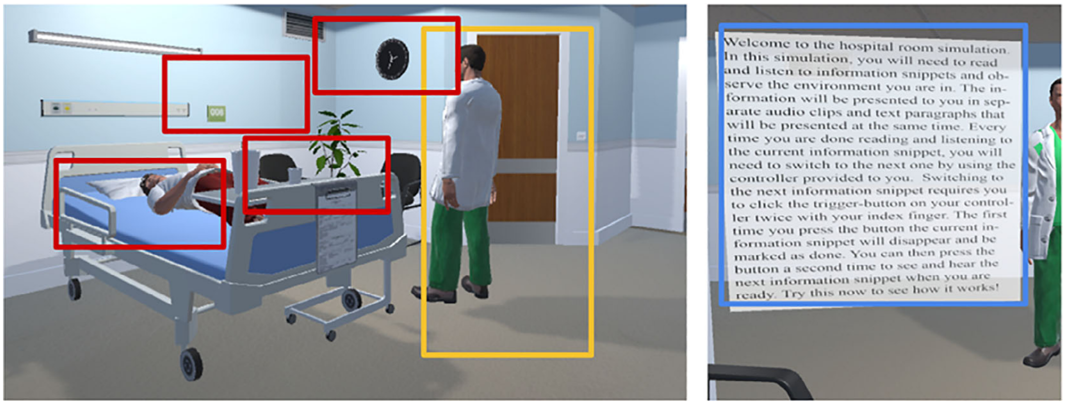

FIGURE 2 AOIs used for ET in this study. Yellow area defines the doctor character AOI, red areas – extraneous attention props AOI, and blue area - the overlay reading interface AOI [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 ANOVA results of post-test survey measures comparing auditory, written and redundancy conditions

	Auditory		Written		Redundancy		ANOVA		
	M	*SD*	*M*	*SD*	*M*	*SD*	*F*	*df*	*p*
Retention	15.48	*3.75*	18.67	*2.30*	18.88	*2.68*	10.011	*72*	**0.000****
Transfer	5.48	*2.18*	6.29	*1.78*	5.96	*1.52*	1.191	*72*	**0.310**
Mental effort	6.12	*1.17*	6.29	*1.46*	5.92	*1.10*	0.541	*72*	**0.585**
Content diff.	5.64	*1.63*	4.96	*1.12*	4.33	*1.61*	4.819	*72*	**0.011***
Form diff.	4.68	*1.68*	5.25	*1.73*	4.13	*1.54*	2.790	*72*	**0.068**
Concentration	6.24	*1.23*	6.79	*1.35*	6.42	*1.44*	1.071	*72*	**0.348**
Intrinsic CL	3.40	*0.71*	3.19	*0.79*	3.24	*0.96*	0.431	*72*	**0.651**
Extraneous CL	2.71	*0.94*	3.14	*0.99*	2.36	*0.80*	4.330	*72*	**0.017***
Germane CL	3.36	*0.86*	3.52	*1.08*	3.73	*0.55*	1.147	*72*	**0.323**
Ex. attention	5.04	*2.05*	3.83	*1.66*	2.88	*2.15*	7.46	*72*	**0.001***

**p* < 0.05, ***p* < 0.001.

groups on prior knowledge, basic characteristics and gender composition prior to the experiment.

4.2 | RQ 1: Did redundancy influence learning outcomes of retention and transfer?

The first objective (RQ1) of this study was to investigate whether different representations of text in an IVR learning environment affect participants' learning outcomes, as reflected by a knowledge retention test and a knowledge transfer test. As can be seen in Table 1, we found a significant difference between the groups in knowledge retention, ($F_{(2,70)} = 10.011$, $p < 0.001$). Post-hoc analysis revealed that the auditory ($M = 15.48$, $SD = 3.75$) condition scored significantly lower than written ($M = 18.67$, $SD = 2.30$, $p = 0.001$, $d = 1.0$) and redundancy ($M = 18.88$, $SD = 2.68$, $p < 0.001$, $d = 1.0$) groups. There was no significant difference between the written and redundancy groups ($p = 0.968$). We therefore conclude that participants in the auditory condition remembered less information than those in the written or redundancy conditions.

A further ANOVA analyses revealed no significant differences in transfer test scores between the experimental groups ($F_{(2,70)} = 1.191$, $p = 0.310$). That is, participants in the auditory ($M = 5.48$, $SD = 2.18$), written ($M = 6.29$, $SD = 1.78$), and redundancy ($M = 5.96$, $SD = 1.52$) conditions did not differ significantly on their ability to apply the knowledge to a new context as assessed in the transfer test. In conclusion, the redundancy group performed equally well as the written group on both learning outcomes; and performed better than the auditory group on the outcome of retention. This is a major empirical finding of this paper.

4.3 | RQ 2: Did redundancy impact self-reported cognitive load?

The second goal of the present study was to determine how auditory, written, or redundant text representation influences the CL of learners in VR. ANOVA results for all CL items and scales included in this study are summarized in Table 1. No significant differences were found on the items measuring mental effort, $F_{(2,70)} = 0.541$, $p = 0.585$, form difficulty, $F_{(2,70)} = 2.790$, $p = 0.068$, or concentration, $F_{(2,70)} = 1.934$, $p = 0.348$. A significant difference was found for content difficulty, $F_{(2,70)} = 4.819$, $p = 0.011$, where post-hoc analysis revealed that participants in the auditory condition ($M = 5.64$, $SD = 1.63$) rated the content difficulty significantly higher than in the redundancy group ($M = 4.33$, $SD = 1.61$, $p = 0.008$, $d = 0.80$). No significant differences were found between the written condition and the other two conditions.

In addition to these individual items, we measured CL with the scale from Leppink et al. (2013). We found no significant differences in self-reported Intrinsic CL, $F_{(2,70)} = 0.431$, $p = 0.651$, or Germane CL, $F_{(2,70)} = 1.147$, $p = 0.323$. However, there was a significant difference in Extraneous CL, $F_{(2,70)} = 4.330$, $p = 0.017$. Post-hoc analysis showed significantly lower scores in the redundancy ($M = 2.36$, $SD = 0.80$) condition compared to the written condition ($M = 3.14$, $SD = 0.99$, $p = 0.012$, $d = 0.86$). No significant differences were observed between the auditory group and the other experimental groups. We thus conclude that self-reported extraneous CL was lower in the redundancy group compared with the written group, and that content was perceived to be more difficult in the auditory condition than in the redundancy condition.

4.4 | RQ 3: Did cognitive demands differ between the groups, as observed by EEG measures?

Another aim of this study was to understand if cognitive processing demands differ when learning with redundant and non-redundant information representations in IVR (RQ 3). To this end we investigated between-group differences in mean EEG power. For each of the frequency bands (i.e., Theta, Alpha), a one-way ANOVAs compared three experimental groups on mean peak frequencies for each electrode (Table 2, Figure 3). For mean Theta frequencies a significant difference between the groups was observed on every single electrode (F3, F4, C3, C4, P3, P4, Fz, Cz, POz), $p = [1^{-10}; 0.042]$. The significant differences remained for six of the electrodes (F3, P3, P4, Fz, Cz, POz), after accounting for multiple comparisons using a Bonferroni correction ($0.05/9 = 0.0056$). Post-hoc comparisons indicated that significant differences are found between the auditory and redundancy, and auditory and written conditions, suggesting lowest cognitive demands in the auditory condition. The written condition showed no significant differences when compared to the redundancy in Theta, suggesting no significant difference in cognitive demands when comparing these conditions. No significant differences between the groups in mean Alpha band activity were detected.

4.5 | RQ 4: Are there any differences in visual attention between conditions?

To understand visual attention allocation (RQ 4), this study investigated between-group differences in several ET measurements: blinks, fixations and saccades. Group means and ANOVA statistics of all ET variables are summarized in Table 3. Notably, all comparisons regarding the overlay AOI only concern two groups (i.e., written and redundancy).

TABLE 2 ANOVA results of EEG Theta and Alpha measures comparing auditory, written and redundancy conditions

	Auditory		Written		Redundancy		ANOVA		
	M	*SD*	*M*	*SD*	*M*	*SD*	*F*	*df*	*p*
Theta F3	−0.35	*0.23*	−0.15	*0.14*	−0.16	*0.18*	7.769	*62*	**0.001***
Theta Fz	−0.23	*0.14*	−0.04	*0.11*	−0.06	*0.14*	13.586	*62*	**0.000****
Theta F4	−0.35	*0.21*	−0.19	*0.17*	−0.17	*0.20*	5.153	*62*	**0.009***
Theta C3	−0.31	*0.14*	−0.22	*0.18*	−0.18	*0.18*	3.353	*62*	**0.042***
Theta Cz	−0.15	*0.15*	0.01	*0.09*	0.01	*0.11*	12.602	*62*	**0.000****
Theta C4	−0.33	*0.13*	−0.22	*0.14*	−0.21	*0.20*	4.006	*62*	**0.023***
Theta P3	−0.26	*0.12*	−0.11	*0.12*	−0.09	*0.13*	12.930	*62*	**0.000****
Theta POz	−0.22	*0.15*	0.01	*0.10*	0.04	*0.10*	31.247	*62*	**0.000****
Theta P4	−0.25	*0.12*	−0.07	*0.09*	−0.07	*0.14*	15.922	*62*	**0.000****
Alpha F3	−0.53	*0.27*	−0.39	*0.19*	−0.45	*0.27*	1.797	*62*	**0.175**
Alpha Fz	−0.47	*0.20*	−0.35	*0.15*	−0.43	*0.15*	2.736	*62*	**0.073**
Alpha F4	−0.53	*0.27*	−0.43	*0.20*	−0.44	*0.26*	1.048	*62*	**0.357**
Alpha C3	−0.38	*0.23*	−0.43	*0.23*	−0.42	*0.22*	0.419	*62*	**0.659**
Alpha Cz	−0.39	*0.25*	−0.32	*0.14*	−0.37	*0.10*	0.774	*62*	**0.466**
Alpha C4	−0.38	*0.22*	−0.45	*0.18*	−0.44	*0.21*	0.733	*62*	**0.485**
Alpha P3	−0.32	*0.22*	−0.31	*0.17*	−0.35	*0.17*	0.298	*62*	**0.743**
Alpha POz	−0.33	*0.23*	−0.26	*0.11*	−0.32	*0.13*	0.953	*62*	**0.391**
Alpha P4	−0.29	*0.20*	−0.27	*0.17*	−0.33	*0.14*	0.578	*62*	**0.564**

*p < 0.05, ** p < 0.001.

To gain an insight into which parts of the simulation the participants attended to, the percentage of time spent looking at the three AOIs were compared. These percentages were derived by summarizing participants' fixations and their durations: comparing the total viewing duration with the duration for each AOI specifically. Significant differences in viewing durations were found for all three AOIs. Firstly, for the doctor AOI ($F_{(2,65)} = 635.766$, $p < 0.001$), a post-hoc test showed a significant difference between auditory ($M = 78.59$, $SD = 14.10$) as compared to the written ($M = 0.36$, $SD = 0.27$, $p < 0.001$, $d = 7.84$) and redundancy ($M = 1.73$, $SD = 2.10$, $p < 0.001$, $d = 7.62$) conditions. Yet, no significant difference was observed between the written and redundancy ($p = 0.861$) groups. This shows that participants in the auditory condition spent most of their time observing the doctor character, while in the learners in the written and redundancy conditions did not attend to the doctor character as much. Secondly, the redundancy ($M = 95.78$, $SD = 4.36$) group spent significantly less time than the written ($M = 98.12$, $SD = 1.28$) group viewing the overlay AOI ($F_{(1,41)} = 5.829$, $p = 0.020$, $d = 0.73$). Nevertheless, these results illustrate that in both conditions participants spent an average of over 95% of the time on viewing the text,

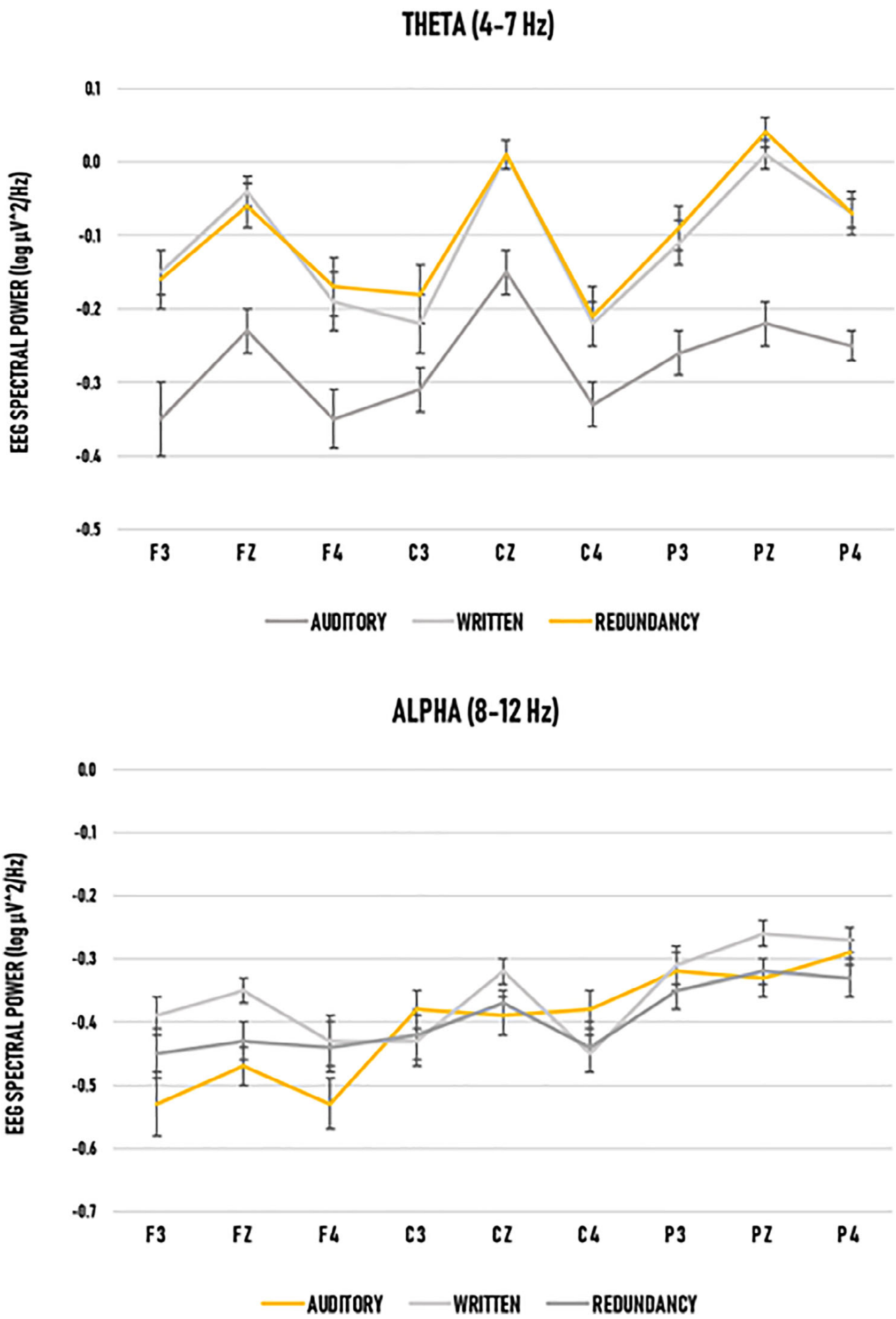

FIGURE 3 EEG power comparisons between conditions for all electrode positions for all participants in Theta (top) and Alpha (bottom) frequency bands [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 3 ANOVA results of ET measures comparing auditory, written and redundancy conditions

	Auditory		Written		Redundancy		ANOVA		
	M	*SD*	*M*	*SD*	*M*	*SD*	*F*	*df*	*p*
% of time spent in an AOI									
% Doctor	78.59	*14.10*	0.36	*0.27*	1.73	*2.10*	635.766	*67*	**0.000****
% Interface			98.12	*1.28*	95.78	*4.36*	5.829	*42*	**0.020***
% Extr. attention	9.69	*7.46*	0.68	*0.61*	0.47	*0.53*	31.642	*67*	**0.000****
Overall fixation counts/min									
All fixations/min	56.01	*22.53*	203.95	*19.65*	194.26	*14.25*	434.053	*67*	**0.000****
All saccades/min	71.16	*28.13*	218.85	*24.23*	242.80	*67.75*	107.63	*67*	**0.000****
All blinks/min	14.18	*12.15*	3.93	*4.17*	7.58	*6.05*	8.933	*67*	**0.000****
Interface AOI measures									
Int. fixations/min	–	–	202.34	*19.24*	191.78	*14.52*	4.096	*42*	**0.050**
Int. saccades/min	–	–	215.68	*23.50*	237.89	*65.29*	2.243	*42*	**0.142**
Int. blinks/min	–	–	3.98	*4.26*	7.75	*6.31*	5.302	*42*	**0.026***

*$p < 0.05$, **$p < 0.001$.

suggesting that students in the redundancy condition spent most of their time reading. Lastly, a significant difference for the extraneous task-irrelevant objects AOI ($F_{(2,65)} = 31.642$, $p < 0.001$) was observed with the auditory ($M = 9.69$, $SD = 7.46$) group spending significantly more time gazing at the task-irrelevant stimuli than the written ($M = 0.68$, $SD = 0.61$, $p < 0.001$, $d = 1.70$) or redundancy ($M = 0.47$, $SD = 0.53$, $p < 0.001$, $d = 1.74$) groups. The difference between the written and redundancy groups was not significant ($p = 0.987$). To summarize, the learners in the written and redundancy conditions spent most of the time reading the text, whereas the learners in the auditory condition spent time attending to the doctor character as well as the task-irrelevant stimuli. Data from blinks and saccades further illustrate whether participants in the written and redundancy conditions spent their time reading.

The group comparisons for fixations and saccades were conducted between all three groups and between the written and redundancy groups for the overlay AOI specifically. Notably, over the course of the simulation there were significant differences in fixations per minute ($F_{(2,65)} = 434.053$, $p < 0.001$). These differences occurred because the auditory ($M = 56.01$, $SD = 22.52$) group had significantly fewer, but longer fixations than either written ($M = 203.95$, $SD = 19.65$, $p < 0.001$, $d = 7.00$) or redundancy ($M = 194.26$, $SD = 14.25$, $p < 0.001$, $d = 7.34$) groups. The difference in fixations on the overlay interface was marginally not significant ($F_{(1,41)} = 4.096$, $p = 0.050$). Additionally, we observed significant differences in overall saccade count ($F_{(2,65)} = 107.63$, $p < 0.001$). The post-hoc comparison revealed that participants in the auditory ($M = 71.16$, $SD = 28.13$) condition moved their eyes significantly less than those in the written ($M = 218.85$, $SD = 24.23$, $p < 0.001$, $d = 5.60$) or redundancy group ($M = 242.80$, $SD = 67.75$, $p < 0.001$, $d = 3.31$), with no significant difference between written and redundancy conditions ($p = 0.176$). Saccades inside the overlay AOI showed no significant difference between written or redundancy either ($F_{(1,41)} = 2.243$, $p = 0.142$). These findings illustrate further that participants in the auditory condition were focused on the doctor and listened, whereas the learners in the remaining two conditions read the text on the interface. The gaze patterns for the Overlay AOI were not significantly different between written or redundancy representations, which suggests they were reading in a similar manner.

Finally, we observed a significant difference for average blinks per minute ($F_{(2,65)} = 8.933$, $p < 0.001$), where a further post-hoc investigation revealed a significant difference between the auditory ($M = 14.18$, $SD = 12.15$) and both the written ($M = 3.93$, $SD = 4.17$, $p < 0.001$, $d = 1.13$) and redundancy ($M = 7.58$, $SD = 6.05$, $p = 0.028$, $d = 0.69$) groups, and a non-significant difference between written and redundancy ($p = 0.340$). However, the difference in average blinks per minute for the interface AOI between the written ($M = 3.98$, $SD = 4.26$) and redundancy ($M = 7.75$, $SD = 6.31$) was significant, $F_{(1,41)} = 5.302$, $p = 0.026$, $d = 0.69$. This means that participants in the written condition blinked on average less while gazing at the overlay interface than participants in the redundancy condition. Since eye blinks typically decrease when reading, this indicates that participants in the redundancy condition read less than in the written condition; however, they still spent significantly more time reading than participants in the auditory condition.

Exploring RQ 4 further, ANOVA results comparing the extraneous attention measure scores between groups is shown in Table 1. This data reveals a significant difference between the three conditions ($F_{(2,65)} = 7.459$, $p < 0.001$). Post-hoc analysis showed that significant differences occurred between the auditory ($M = 5.04$, $SD = 2.05$) and redundancy ($M = 2.88$, $SD = 2.15$) conditions ($p = 0.001$, $d = 1.03$). This provides evidence that participants in the auditory condition retained more task-irrelevant information that was present in the environment than those in the redundancy group. No significant differences were found between written ($M = 3.83$, $SD = 1.66$) and auditory ($p = 0.88$), nor between written and redundancy conditions ($p = 0.217$).

5 | DISCUSSION

5.1 | Empirical contributions

The first major finding in this study relates to RQ 1, which investigated the effects of redundancy on learning outcome measures of knowledge retention and transfer. Contrary to traditional assumptions summarized in CTML about the redundancy principle in non-immersive 2D media, our results showed no decrease in learning outcomes when learning information was presented in a redundant format in IVR. These results indicate that learners remembered facts and were able to utilize knowledge learned with the same efficiency in redundant information representations as in non-redundant information representations. Our findings highlighting the advantage of redundancy representations over auditory representations go hand-in-hand with the conclusions summarized in a meta-analysis by Adesope and Nesbit (2012). Even though their findings were in the realm of 2D media, given that similar results for redundancy were found in low prior knowledge learners, in system-paced learning materials, and picture free-materials, it could be argued that all of these situations represent more complex learning environments, drawing parallels to IVR. This might imply that redundancy of learning content in more complex learning environments (e.g., IVR) could in fact be beneficial for learning, as opposed to redundancy in customary and less complex media systems (e.g., power point presentations, book illustrations).

Furthermore, highlighting differences in learning outcomes between auditory and written information representations, our study replicates results obtained of prior research (Baceviciute et al., 2020), wherein auditory information was likewise found to be inferior to written information in terms of knowledge retention, but not knowledge transfer. Referencing Mayer (2014, 2020) and Baceviciute et al., (2020), attribute this finding to the transient nature of auditory information. According to the authors, when learning with auditory content, participants might not have been able to engage in WM processes as successfully as in conditions involving textual representations, where the participants were able to more easily repeat and integrate information. They argue that in complex environments, such as IVR, there might be a greater need to anchor learning than in simpler 2D learning scenarios (Baceviciute et al., 2020).

In regards to self-reported CL outcomes addressed in RQ2, results show that redundant information representations were not perceived to be more cognitively demanding than non-redundant information representations, as observed with both single-item CL items and with the validated Leppink et al. (2013) instrument. In fact, with the latter measure, redundant content was found to be least extraneously loading (significantly when compared to written representations). Since no differences between written and redundant information representations were observed in learning outcomes, this shows that in this study, the participants might have used corresponding information representations more as an aid, rather than perceiving them as an additive strain to their learning. In addition to that, supporting findings reported by Baceviciute et al., (2020), our results show that learning content presented in an auditory representation format was perceived to be the most difficult from which to learn as compared to other formats. This once again can be attributed to the transient nature of auditory content, which might influence learner's perceptions of that content despite the fact that no content manipulations were actually introduced in the experiment.

Another major finding of this study comes from the obtained EEG estimates for the Theta frequency band. Specifically, we observed significant differences between the redundancy information representation format and the auditory representation format, and between the written representation format and the auditory representation format in the Theta band. Since overall higher Theta activation is normally associated with increased cognitive load, our results hint that redundancy and written conditions required more mental effort from the participants when learning in those formats. Previous work in 2D media has hypothesized that the need to combine redundant information sources generates strong demands on the learner's WM capacity, and therefore it is more difficult for students to remember the information acquired (Mayer, 2014, 2020). From our EEG results we see that as compared to auditory information processing, the participants did invest more cognitive capacity in redundant information processing. However, since there was no difference in the EEG Theta band activity between redundant format and written-only format, we can assume that the difference in cognitive processing observed when compared to the auditory condition was not attributed to information redundancy per se, but is rather a difference that can be ascribed to the high cognitive demands imposed by written information. In this direction, Baceviciute et al., (2020) have also found that reading (as compared to listening) yields overall higher levels of mental workload, suggesting that reading might simply be a more cognitively demanding process than listening. Interestingly, in this study we did not find any significant differences between conditions in the alpha frequency band, although it has typically also been described as a reliable measure of cognitive demands (Klimesch, 1999). Prior literature reports that changes in theta but not alpha can be associated with impairments in WM (e.g., Goodman et al., 2019). In the current study, this could suggest that written content is not necessarily more cognitively loading, when compared to auditory content, but that it does impose additional demands on the learner's WM load during learning.

Another major contribution of this study stems from the viewing duration results for the interface AOI obtained by the ET measures (RQ 4). Results indicate that participants in both redundancy and written conditions spent more than 95% of their time fixated on the interface AOI – a virtual element that was used to display text in the IVR environment. Contrary to what was previously assumed by Moreno and Mayer's (2002) study which hypothesized that learners listen and do not read under redundancy conditions, this shows that participants still spent most of their time reading content, when both information representation formats were available. This fact is also supported by our results obtained from fixation and saccade measures, which both showed significant differences between the auditory format and both written representation formats, but not between the redundancy and written conditions. Similarly, prior literature has also reported lower blink rates during visual information processing (Stern & Skelly, 1984) which was also observed in our study, once again suggesting that in redundancy participants continue to engage in the process of reading. These results support previous findings produced by Schmidt-Weigand et al. (2010) and Liu et al. (2011), who suggest that when text is placed in front of learners it encapsulates the majority of their attentional resources, not leaving much attentional capacity to engage in other activity (e.g., engage in animations or images). This is supported by the results from the extraneous attention measure, as well as time spent on extraneous task-irrelevant objects AOI. These results showed that the learners in the auditory condition engage in environmental observations significantly more than learners in the conditions involving text, which supports the cognitively demanding nature of a reading task. Significantly higher saccadic eye movement for both reading conditions found in this study also speaks to this claim.

Even though most of the participant's time and attentive resources were spent on reading written content, we did observe significant effects in viewing times between redundancy and written conditions, differentiating written only and written-auditory information representations. Firstly, results show that participants in the redundancy condition spent significantly less time fixating on the interface AOI than in the written condition. In addition to that, higher blink rates were found in the redundancy condition. Both of these findings hint that participants did read less in the redundancy condition. This, together with the lack of difference observed between the two textual conditions in the learning results, as well as in the EEG results, suggests some of their cognitive resources from the visual modality were most likely successfully offloaded to the auditory modality.

Lastly, another finding in this study comes from the viewing duration results for the doctor character AOI, which showed significantly longer viewing duration times for this AOI in the auditory condition as compared with two other conditions. Interestingly, even though the audio recorded in this simulation was not tied to the doctor character, this implies that participants in the auditory condition were using this character as an anchor point for grounding their attention while listening to the auditory information. This confirms the assumptions made by Baceviciute et al. (2020), which suggested that in complex learning environments there might be a psychological need to ground transient auditory information.

5.2 | Limitations and future work

In this study our focus was set on investigating written-auditory redundancy. However, future studies should investigate different forms of information redundancy, as it is not clear if findings obtained in this study would generalize to more diverse contexts. In this study we purposefully did not embed any learning information in the surrounding IVRE. However, considering that presence in a simulated world is perhaps among the most powerful affordances offered by IVREs (Makransky et al., 2021), future studies should consider how learning information could visually be embedded in an IVRE. This would allow researchers to investigate different forms of information redundancy and explore how picture/text redundancy, traditionally described in 2D media, can be generalized in IVREs. In general, CAMIL describes how presence and agency are the main affordances of learning in IVR. By investigating the redundancy principle in this study, we focus on the role of cognitive load and information processing when learning in IVR. CAMIL also describes how presence and agency can lead to more learning through high levels of embodiment. The level of interaction in the IVR used in this study was quite limited, therefore, it did not fully take advantage of the affordance of agency, or embodiment which is possible in IVR. Therefore, future research should consider how instructional design features (such as redundant information) generalize to more interactive learning environments that make better use high levels of presence and agency which are the main affordances of learning in IVR.

Furthermore, since we observed that some of the information was successfully offloaded to the auditory channel, it could be useful to investigate different written-auditory information couplings, with varying degrees of auditory-written text correspondence (for instance, if only some information was presented in a written format, or in an auditory format). These would lean more towards the signalling effect described by CTML, wherein information presented in two different modalities is not fully redundant, but instead is used for emphasizing and cuing information processing in the other modality. Investigations with varying degrees of text-audio correspondence have already been proposed by the review study carried out by Adesope and Nesbit (2012) for redundancy in 2D media. Less textually-dense redundancy conditions could especially be relevant for IVR, where textual representations are typically deemed to be impractical, and not fully encompassing true power of the immersive media.

In this study, relatively short paragraphs of text were used for the investigation. Some studies have argued that text length might influence the redundancy effect (Mayer, 2014), suggesting that future studies should include investigations on how text length might influence redundant information processing. In a similar vein, some studies have suggested that prior knowledge of the learner might influence the redundancy principle. Specifically, redundancy effects are said to be heightened in novice learners, as they need to utilize more cognitive processing capacity due to the novelty of learned information (Mayer, 2014). In our study we controlled for prior knowledge, as all participants were novice learners. Nevertheless, information that we used was relatively simple, targeted towards a general-population of learners. It could therefore, be interesting and pertinent to investigate if and how redundancy effects generalize to IVR, with increasing information complexity, and when comparing novice and advanced learners.

Considering our ET results, which emphasized the cognitively loading nature of textual information; as well as our EEG findings, which indicated higher WM demands in both written conditions, we can make a general assumption that the moment that there is written information placed in front of learners, they will spend time on it and read it. This might not be the case in non-learning scenarios, or in scenarios where text is not the essence of the IVR situation. Future studies should therefore investigate whether these findings translate to situations wherein written text plays a supporting role, rather than being at the core of learning. Similarly, this study was solely focused on healthy learner population and did not consider learners of different learning backgrounds and styles. As such, future research should also investigate how underprivileged learner populations (e.g., learner's with special needs, and learning disorders, such as ADHD, ADD, Dyslexia, etc.), as well as learners with different learning backgrounds and styles process written and auditory information in IVR.

Nevertheless, considering the unique affordances of IVR, such as presence and agency, (Makransky et al., 2019b, Makransky & Petersen, 2021; Jensen & Konradsen, 2018; Mikropoulos & Natsis, 2011), and practical complexities surrounding the development of this technology, we invite future researchers and instructional designers to extend their investigations beyond traditional textual and auditory information representations, and focus more on studying the efficacy of visual, embodied and dynamic representation forms, that might be more suited for this complex new learning medium.

6 | CONCLUSION

This article summarized a between-subjects experiment, investigating the redundancy principle in an IVR environment for learning. Results for learning outcomes and various self-reported and psychophysiological measures of CL indicate that the redundancy principle might not generalize to immersive technology as originally anticipated in non-immersive media research. Instead, findings show that when attending to redundant learning content in immersive environments, learners use less cognitive processing capacity without compromising learning efficacy. The results therefore imply that redundancy of learning content in more complex learning environments such as IVR could in fact be beneficial for learning. This finding also suggests that instructional design principles, originally discovered in traditional 2D media, might not directly translate to IVR, calling for further research in the field of instructional design for immersive media systems.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jcal.12595.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Sarune Baceviciute https://orcid.org/0000-0002-9995-8045
Gordon Lucas https://orcid.org/0000-0002-5626-6890
Guido Makransky https://orcid.org/0000-0003-1862-7824

REFERENCES

Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: A meta-analysis. *Journal of Educational Psychology*, *104*(1), 250.

Antonenko, P., Paas, F., Grabner, R., & Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*(4), 425–438.

Antonenko, P. D., & Keil, A. (2017). Assessing working memory dynamics with electroencephalography: Implications for research on cognitive load. In I. R. Z. Zheng (Ed.), *Cognitive load measurement and application* (pp. 93–111). Routledge.

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction*, *16*(5), 389–400.

Baceviciute, S., Mottelson, A., Terkildsen, T., & Makransky, G. (2020, April). Investigating representation of text and audio in educational VR using learning outcomes and EEG. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–13).

Baceviciute, S., Terkildsen, T., & Makransky, G. (2021). Remediating learning from non-immersive to immersive media: Using EEG to investigate the effects of environmental embeddedness on reading in Virtual Reality. *Computers & Education*, *164*, 104122.

Bogusevschi, D., Muntean, C. H., & Muntean, G. M. (2019). Earth course: Knowledge acquisition in technology enhanced learning STEM education in primary school. In *EdMedia+ Innovate learning* (pp. 1243–1252). Association for the Advancement of Computing in Education (AACE).

Boland, J. E. (2004). Linking eye movements to sentence comprehension in reading and listening. In I. C. Manuel & J. C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond* (pp. 51–76). Pychology Press.

Brouwer, A. M., Hogervorst, M. A., Van Erp, J. B., Heffelaar, T., Zimmerman, P. H., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural Engineering*, *9*(4), 045008.

Butussi, F., & Chittaro, L. (2018). Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Transactions on Visualization and Computer Graphics*, *24*(2), 1063–1076.

Chittaro, L., & Buttussi, F. (2015). Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. *IEEE Transactions on Visualization and Computer Graphics*, *21*(4), 529–538.

Christensen, R., & Knezek, G. (2016). Blending formal and informal learning through an online immersive game environment: Contrasts in interactions by middle school boys and girls. In *Society for Information Technology & Teacher Education International Conference* (pp. 535–540). Association for the Advancement of Computing in Education (AACE).

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, *25*(2), 315–324.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology*, *94*(2), 428.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272–309.

De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, *20*(2), 111–122.

Fowler, C. (2015). Virtual reality and learning: Where is the pedagogy? *British Journal of Educational Technology*, *46*(2), 412–422.

Frey, J., Mühl, C., Lotte, F., & Hachet, M. (2014). *Review of the use of electroencephalography as an evaluation method for human-computer* *interaction*. In PhyCS 2014 - International Conference on Physiological Computing Systems (2014). Lisbon, Portugal: SCITEPRESS.

Gerjets, P., Scheiter, K., Opfermann, M., Hesse, F. W., & Eysink, T. H. (2009). Learning with hypermedia: The influence of representational formats and different levels of learner control on performance and learning behavior. *Computers in Human Behavior*, *25*(2), 360–370.

Goodman, M., Zomorrodi, R., Kumar, S., Barr, M., Daskalakis, Z., Blumberger, D., Fischer, C., Flint, A., Mah, L., Herrmann, N., & Pollock, B. (2019). Changes in theta but not alpha modulation are associated with working memory impairments in Alzheimer's dementia and mild cognitive impairment. *Biological Psychiatry*, *85*(10), 213–214.

Halliday, D. M., Rosenberg, J. R., Amjad, A. M., Breeze, P., Conway, B. A., & Farmer, S. F. (1995). A framework for the analysis of mixed time series/point process data-theory and application to the study of physiological tremor, single motor unit discharges and electromyograms. *Progress in Biophysics and Molecular Biology*, *64*(2), 237–278.

Holland, M. K., & Tarlow, G. (1972). Blinking and mental load. *Psychological Reports*, *31*(1), 119–127.

Holland, M. K., & Tarlow, G. (1975). Blinking and thinking. *Perceptual and Motor Skills*, *41*(2), 403–406.

Howard, M. (2019). Virtual reality interventions for personal development: A meta-analysis of hardware and software. *Human–Computer Interaction*, *34*(3), 205–239.

Huang, W., Roscoe, R. D., Johnson-Glenberg, M. C., & Craig, S. D. (2020). Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion. *Journal of Computer Assisted Learning*, *37*(3), 745–758.

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515–1529.

Johnson-Glenberg, M. C. (2019). The necessary nine: Design principles for embodied VR and active STEM education. In *Learning in a digital world* (pp. 83–112). Springer.

Jones, N. (2018). *Simulated labs are booming* (pp. S5–S7). Nature Outlook: Science and Technology Education.

Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors*, *46*(3), 567–581.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2-3), 169–195.

Klingenberg, S., Jørgensen, M. L., Dandanell, G., Skriver, K., Mottelson, A., & Makransky, G. (2020). Investigating the effect of teaching as a generative learning strategy when learning through desktop and immersive VR: A media and methods experiment. *British Journal of Educational Technology*, *51*(6), 2115–2138.

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W., … Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90–115.

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., & Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072.

Liu, H. C., Lai, M. L., & Chuang, H. H. (2011). Using eye-tracking technology to investigate the redundant effect of multimedia web pages on viewers' cognitive processes. *Computers in Human Behavior*, *27*(6), 2410–2417.

Luo, H., Li, G., Feng, Q., Yang, Y., & Zuo, M. (2021). Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019. *Journal of Computer Assisted Learning*, *37*(3), 887–901.

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, *66*(5), 1141–1164.

Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019a). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning*, *35*(6), 691–707.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019b). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225–236.

Makranksy, G., Wismer, P., & Mayer, R. (2019c). A Gender Matching Effect in Learning with Pedagogical Agents in an Immersive Virtual Reality Science Simulation. *Journal of Computer Assisted Learning*, *35*(3), 349–358.

Makransky, G., & Petersen, G. (2019). Investigating the process of learning with desktop virtual reality: a structural equation modeling approach. *Computers & Education*, *134*, 15–30.

Makransky, G., Petersen, G. B., & Klingenberg, S. (2020). Can an immersive virtual reality simulation increase students' interest and career aspirations in science? *British Journal of Educational Technology*, *51*(6), 2079–2097.

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 1–22.

Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology*, *113*(4), 719–735.

Makransky, G. (2021). The immersion principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed.). Cambdridge University Press.

Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, *140*, 103603.

Martins, R., & Carvalho, J. M. (2015). *Eye blinking as an indicator of fatigue and mental load - A systememature review*. In Occupational Safety and Hygiene III, p. 10.

Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning*. Cambridge University Press.

Mayer, R. E. (2020). *Multimedia learning*. Cambridge University Press.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*(1), 187.

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, *100*(2), 380.

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education*, *56*(3), 769–780.

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology*, *94*(1), 156–163.

Mühl, C., Heylen, D., & Nijholt, A. (2015). Affective brain-computer interfaces: Neuroscientific approaches to affect detection. In *Oxford handbook of affective computing* (pp. 217–232). Oxford University Press.

Muller Queiroz, A. C., Moreira Nascimento, A., Tori, R., Brashear Alejandro, T., Veloso de Melo, V., de Souza Meirelles, F., & da Silva Leme, M. I. (2018). Immersive virtual environments in corporate education and training. In *Twenty-fourth Americas Conference on Information Systems* (pp. 1–10). Association for Information Systems.

Mütterlein, J., & Hess, T. (2017). *Immersion, presence, interactivity: Towards a joint understanding of factors influencing virtual reality acceptance and use*. In Twenty-third Americas Conference on Information Systems. Association for Information Systems, Boston, MA.

Örün, Ö., & Akbulut, Y. (2019). Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention. *Computers in Human Behavior*, *92*, 216–229.

Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *Journal of Neuroscience*, *26*(28), 7523–7531.

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*, 429–434.

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785.

Petersen, G. B., Klingenberg, S., Mayer, R. E., & Makransky, G. (2020). The virtual field trip: Investigating how to optimize immersive virtual learning in climate change education. *British Journal of Educational Technology*, *51*(6), 2099–2115.

Puma, S., Matton, N., Paubel, P. V., Raufaste, É., & El-Yagoubi, R. (2018). Using theta and alpha band power to assess cognitive workload in multitasking environments. *International Journal of Psychophysiology*, *123*, 111–120.

Raditanti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778.

Richards, D., & Taylor, M. (2015). A Comparison of learning gains when using a 2D simulation tool versus a 3D virtual world: An experiment to find the right representation involving the Marginal Value Theorem. *Computers & Education*, *86*, 157–171.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributes. *Journal of Educational Psychology*, *78*, 647–658.

Scharinger, C. (2018). Fixation-related EEG frequency band power analysis: A promising methodology for studying instructional design effects of multimedia learning material. *Frontline Learning Research*, *6*(3), 56–71.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system-and self-paced instruction in multimedia learning. *Learning and Instruction*, *20*(2), 100–110.

Stern, J. A., & Skelly, J. J. (1984). The eye blink and workload considerations. In *Proceedings of the Human Factors Society Annual Meeting* (Vol. 28, pp. 942–944). SAGE Publications.

Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37–76). Academic Press.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Measuring cognitive load. In *Cognitive load theory* (pp. 71–85). Springer.

Tang, Y. M., Ng, G. W. Y., Chia, N. H., So, E. H. K., Wu, C. H., & Ip, W. H. (2020). Application of virtual reality (VR) technology for medical practitioners in type and screen (T&S) training. *Journal of Computer Assisted Learning*, *7*(2), 359–369.

Tobii. (2020). *Tobii PRO SDK calibration*. http://developer.tobiipro.com/commonconcepts/calibration.html

van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *2*(20), 95–99.

Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology*, *51*(6), 1991–2005.

Zagermann, J., Pfeil, U., & Reiterer, H. (2016). 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* (pp. 78–85). ACM.

How to cite this article: Baceviciute, S., Lucas, G., Terkildsen, T., & Makransky, G. (2021). Investigating the redundancy principle in immersive virtual reality environments: An eye-tracking and EEG study. *Journal of Computer Assisted Learning*, 1–17. https://doi.org/10.1111/jcal.12595