Investigating the Effect of Teaching as a Generative Learning Strategy when Learning through Desktop and Immersive VR: A Media and Methods Experiment

Accepted in the British Journal of Educational Technology in the special issue Immersive Virtual Reality in Education on August 12, 2020

Sara Klingenberg[1], Maria L. M. Jørgensen[2], Gert Dandanell[2], Karen Skriver[2], Aske Mottelson[1], Guido Makransky[1]

1. Department of Psychology, University of Copenhagen, Copenhagen, Denmark.

2. Department of Biology, University of Copenhagen, Copenhagen, Denmark

Corresponding Author: Guido Makransky, University of Copenhagen

Address: Øster Farimagsgade 2A, 1353 København K, gm@psy.ku.dk

Sara Klingenberg (sara.klingenberg@sund.ku.dk) is a Research Assistant at the Virtual Learning Lab at the Department of Psychology at the University of Copenhagen.

Maria L. M. Jørgensen (mjoergensen@bio.ku.dk) is a Postdoc at the Department of Biology at the University of Copenhagen.

Gert Dandanell (dandanell@bio.ku.dk) is an Associate Professor at the Department of Biology at the University of Copenhagen.

Karen Skriver (kskriver@bio.ku.dk) is a Professor at the Department of Biology at the University of Copenhagen and leader of a research group focusing on gene regulatory networks.

Aske Mottelson (amot@psy.ku.dk) is a Postdoc at the Virtual Learning Lab at the Department of Psychology at the University of Copenhagen.

Guido Makransky (gm@psy.ku.dk) is an Associate Professor at the Department of Psychology at the University of Copenhagen and leader of the Virtual Learning Lab.

This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. The final article will be available, upon publication, via its DOI: 10.1111/bjet.13029

Abstract

Immersive virtual reality (IVR) simulations for education have been found to increase affective outcomes compared to traditional media, but the effects on learning are mixed. As reflection has previously shown to enhance learning in traditional media, we investigated the efficacy of appropriate reflection exercises for IVR. In a 2x2 mixed-methods experiment 89 (61 female) undergraduate biochemistry students learned about the electron transport chain through desktop virtual reality (DVR) and IVR (media conditions). Approximately half of each group engaged in a subsequent generative learning strategy (GLS) of teaching in pairs (method conditions). A significant interaction between media and methods illustrated that the GLS of teaching significantly improved transfer (d=1.26), retention (d=0.60), and self-efficacy (d=0.82) when learning through IVR, but not DVR. In the second part of the study, students switched media conditions, and the experiment was repeated. This time, significant main effects favoring the IVR group on the outcomes of intrinsic motivation (d=0.16), perceived enjoyment (d=0.94), and presence (d=1.29) were observed, indicating that students preferred IVR after having experienced both media conditions. The results support the view that methods enable media that affect learning, and that the GLS of teaching is specifically relevant for IVR.

Keywords: immersive virtual reality, media versus methods, generative learning strategies, biochemistry education, head-mounted displays, learning

Practitioner notes

What is already known about this topic

- Previous research has found a media effect with Immersive Virtual Reality (IVR) in education leading to better motivational outcomes compared to less immersive media, but effects on learning outcomes are mixed.
- There is evidence that Generative Learning Strategies (GLSs) such as summarizing and enacting can increase learning in IVR.
- There is also evidence that some instructional methods, such as pre-training, may be beneficial for learning in IVR.

What this paper adds

- Evidence that the GLS of teaching improves self-efficacy, retention, and transfer in educational IVR.
- An interaction effect between media (DVR/IVR) and method (GLS/no-GLS) on self-efficacy, retention, and transfer supporting the theoretical view that method enables media.
- No difference in perceived enjoyment, motivation, and presence for students who were new to learning through these media (DVR/IVR), but differences became significant when students learned through the other media first with students preferring IVR.

Implications for practice and/or policy

- Since IVR learning experiences can be highly engaging but also cognitively demanding, it is beneficial to introduce reflection exercises after an IVR learning experience to ensure that students reflect over the material and integrate it with their long-term memory.

- One effective solution is to engage students in the GLS of teaching after an IVR simulation, thereby prompting them to select relevant information, organize it into a coherent structure, and elaborate on it by incorporating it with their existing knowledge.

Introduction

A contemporary challenge in biochemistry education is to present the abstract nature of chemical and molecular processes in a way that students can relate to, comprehend, and understand (McClean et al., 2005; Schönborn & Anderson, 2006). This challenge is especially pronounced when teaching students about complex intracellular processes that take place on a submicroscopic level, which limits interaction and visualization in the real world (Schönborn & Anderson, 2006). As a result, students might conduct traditional laboratory experiments, yet find the intracellular processes intangible or difficult to relate to (Tsivitanidou et al., 2018). Correspondingly, teaching materials, such as blackboard drawings, textbooks, and illustrations limit possibilities for interaction. They often fail to capture temporal and spatial relationships of molecular processes, and students therefore frequently struggle to interpret the depicted models (McClean et al., 2005). Therefore, a challenge remains to create a link between students' theoretical knowledge about biochemical processes and laboratory experiments (Tsivitanidou et al., 2018).

A way of addressing this challenge is through the use of Virtual Reality (VR). VR is "a computer-mediated simulation that is three-dimensional, multisensory, and interactive so that the user's experience is "as if" inhabiting and acting within an external environment" (Burbules, 2006, p. 37). VR provides unique learning opportunities, as simulations allow students to act as if they were in the real world while interacting with otherwise intangible or inaccessible objects (Bower, 2017; Mikropoulos & Natsis, 2011). VR makes it possible for students to experience a different world or reality that might otherwise be too dangerous, expensive, or impossible in the real world (Dalgarno & Lee, 2010; Freina & Ott, 2015). For example, students can experience the intracellular processes that take place in biochemical experiments in real-time, and they can perform laboratory experiments under varying conditions, which would be impossible due to time and costs of physical experiments in most university settings (De Jong et al., 2013; Jones, 2018). Also, in VR, students can make mistakes in a controlled environment without aversive costs or safety effects (Jensen & Konradsen, 2018; Merchant et al., 2014). Additionally, students can progress through exercises at their own pace, allowing them to spend more time on particularly relevant tasks or questions (Mayer & Chandler, 2001). Last, students can receive immediate individual feedback from a virtual agent during the experiment, which is difficult in most classroom settings where one teacher is responsible for assisting several students (Makransky et al., 2019b). These unique characteristics have been linked to learning affordances including enhanced spatial knowledge representation, increased motivation and engagement, improved contextualization of learning as well as opportunities for experiential learning (Bower, 2017; Dalgarno & Lee, 2010). Therefore, VR is particularly relevant for learning experiences that cannot easily be studied in a traditional classroom setting (Bailenson, 2018), such as a biochemistry lesson on the electron transport chain (ETC) used in the current study. Through high-quality three-dimensional representations, VR allows students to explore the cascade of complex interactions between molecules along the ETC in a way that is not possible in a laboratory experiment in the real world or using a textbook. In this way, VR can enhance students' ability to connect experimental procedures with the intracellular processes (Fitzgerald & Riva, 2001), thereby addressing current challenges in biochemistry education.

Objective

We will briefly introduce some of the most prominent perspectives in the field of research on existing media, which serve as the foundation for our main objectives. One perspective is that "the medium is the message", which refers to the general belief that the medium itself, rather than the content, is essential for successfully delivering an instructional message (Lee, 1999; McLuhan, 1964). The first objective of this study is therefore to investigate the presence of a media effect in conducting a university-level biochemistry lesson. We examine the outcomes of delivering a lesson on the ETC through immersive virtual reality (IVR) using a head-mounted display (HMD) compared to presenting the lesson through desktop virtual reality (DVR) using a traditional two-dimensional monitor. In the following, we will therefore differentiate between DVR and IVR, and use VR as a term referring to both.

An alternative perspective upheld by learning theorists such as Clark (1994) is that there is no media effect on learning; that the chosen instructional method rather than the medium, is important for learning. The second objective of this study is therefore to investigate the presence of a method effect in conducting a biochemistry lesson. We investigate the consequences of using the generative learning strategy (GLS; Fiorella & Mayer, 2016) of teaching in pairs following the virtual lesson compared to a condition where no GLS was used.

Yet another viewpoint, supported by recent results from media and methods experiments in the field of IVR, suggests an interaction between media and methods in settings where the affordances of the medium are specifically positively or negatively influenced by the instructional method (e.g., Makransky et al., 2020a; Meyer et al., 2019). The third objective of this study is therefore to investigate if there is an interaction effect between media and methods when conducting a biochemistry lesson in IVR or DVR with or without the GLS of teaching.

The rationale behind using the GLS of teaching is twofold. According to the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2014), multimedia learning is affected by the cognitive load characteristics of the instructional content. This is particularly relevant for high immersive media, such as IVR, since students might be overwhelmed by the many stimuli presented by the media. Several studies have suggested that acquiring new information in IVR can lead to cognitive overload (e.g., Makransky et al., 2017b; Meyer et al., 2019; Moreno & Mayer, 2002; Parong & Mayer, 2018). According to CTML, cognitive overload happens as a result of an increase in intrinsic cognitive load due to the difficulty of the learning material combined with an increase in extraneous cognitive load due to the many stimuli afforded by the medium (Mayer, 2014; Makransky et al., 2017b; 2020a). From this perspective, introducing a GLS can reduce students' cognitive load on working memory and improve learning outcomes by encouraging them to reflect on the learning content. This is supported by research suggesting that the GLS of teaching can significantly improve learning outcomes among students (e.g., Fiorella & Mayer, 2013; Kobayashi, 2019a). Furthermore, recent systematic reviews of IVR applications in higher education emphasize the importance of grounding IVR applications in existing learning theories that offer guidelines on motivations, learning process, and learning outcomes (Hamilton et al., 2020; Radianti et al., 2020). We therefore used the GLS of teaching to provide a strong theoretical learning framework for the instructional method of this study.

In short, the current study addresses three questions in relation to education: how IVR differs from DVR, how to successfully implement IVR/DVR, and how media and methods interact. Our assessment is based on measuring six constructs that have been identified as important for the process of learning in VR; including presence, perceived enjoyment, intrinsic motivation, self-efficacy, retention, and transfer (Dalgarno & Lee, 2010; Lee et al., 2010;

Makransky & Lilleholt, 2018; Makransky & Petersen, 2019; Weiss et al., 2006). We measure presence, the psychological experience of "being there", as it is a fundamental aspect of understanding VR experience, and because it has previously been linked to learning (Cummings & Bailenson, 2016; Jelfs & Whitelock, 2000; Makransky & Lilleholt, 2018; Shin, 2017; Slater & Wilbur, 1997). Additionally, we measure self-efficacy, that is students' beliefs about their abilities to perform certain actions, which can affect motivational and learning outcomes (Bandura, 1997). Studies suggest that this is particularly relevant when learning in VR, as interacting and receiving immediate feedback in a high-fidelity environment can provide students with mastery experiences that enhance their self-efficacy (Gegenfurtner et al., 2014; Makransky et al., 2020b). This assists learning through inquiry, which has been suggested to enhance a deeper understanding of the learning material (Concannon et al., 2019; Pedaste et al., 2015). Based on Mayer's (2008) principles for design and assessment of multimedia instruction, we measure learning outcomes through a retention test that focuses on remembering, and a transfer test that focuses on understanding. According to Ainley and Armatas (2006, p. 376), a relatively consistent picture has emerged, showing that specifically, these two types of learning are differentially affected by participants' experience with specific kinds of instructional components. One proposed reason for this difference is that transfer learning requires deeper processing, understood as a deeper engagement with the content in order to make sense of it, and apply it to new situations (Ainley & Armatas, 2006). This is related to motivational processes, such as the involvement and engagement of the participants (Ainley & Armatas, 2006), which support the measurement of intrinsic motivation and perceived enjoyment.

We investigated these outcomes by employing the instructional method of teaching in pairs as a GLS for approximately half of the participants in a 2x2 between-subjects design through IVR and DVR. Ultimately, results from this study can provide evidence for how the GLS of teaching in pairs can be used in conjunction with VR technology to enhance learning. Furthermore, a media by methods experiment can shed light on how media and methods affect learning with VR technology. To our knowledge, no study has previously combined the GLS of teaching with IVR to investigate the effects of methods and media on learning outcomes.

Literature review

In the following, we will refer to DVR and IVR. DVR is a computer-generated three-dimensional virtual space experienced through standard audio-visual equipment (i.e., a PC with a two-dimensional monitor; Kaplan-Rakowski & Gruber, 2019). In this study, IVR refers to an interactive 360° three-dimensional simulation accessed through an HMD that provides head and position tracking. An HMD can render a different image for each eye, creating visual cues for depth perception, and it increases the size of the visual field compared to a monitor. IVR compared to DVR has higher technical fidelity, which affords enhanced interaction, display/immersion, and scenario/realism of the simulation (Buttussi & Chittaro, 2018). Here, the term *immersion* relates to the ability of the medium to present a vivid virtual environment while shutting out physical reality (Cummings & Bailenson, 2016). Thus, IVR affords the creation of three-dimensional spatial representations, realistic multisensory channels for interaction, immersion of the user in the virtual environment, and direct manipulation with objects in the virtual laboratory (Mikropoulos & Natsis, 2011; Sherman & Craig, 2003). Consequently, students can achieve highly realistic first-person experiences of the virtual learning environment, which has been associated with affordances such as enhanced presence (e.g., Makransky et al., 2017b; Moreno & Mayer, 2002), self-efficacy (e.g., Meyer et al., 2019), motivation (e.g., Chittaro & Buttussi, 2015; Parong & Mayer, 2018), and enjoyment (e.g., Bogusevschi et al., 2019; Meyer et al., 2019).

The Generative Learning Strategy of Teaching

Generative learning emphasizes the active participation of the student in the learning process (Wittrock, 1974). Fiorella and Mayer (2016) describe this as a way of actively engaging in and processing new information to enhance a deeper understanding of the learning content (i.e., comprehension) and to be able to apply this knowledge to other contexts (i.e., transfer). This involves *selecting* the most relevant information, *organizing* it into a coherent mental representation, and *integrating* it within pre-existing knowledge. Therefore, the GLS of teaching is consistent with the active processing assumption of the CTML which proposes that students actively engage in cognitive processing to select relevant information, organize it into a coherent structure, and integrate it with prior knowledge (Mayer, 2014).

Fiorella and Mayer (2016) differentiate between eight GLSs: summarizing, mapping, drawing, imagining, self-testing, self-explaining, enacting, and teaching. This study uses learning by teaching, which is defined as the act of explaining newly acquired information to help others learn (Fiorella & Mayer, 2016). It is suggested that learning outcomes are dependent on the accuracy of the information, quality of explanations, and reflections generated during teaching (Roscoe & Chi, 2008). Work by Fiorella and Mayer (2013) found improved learning outcomes following teaching compared to no teaching, even when students only prepared to teach without actually performing the activity. However, students who taught outperformed both groups in a follow-up test one week after the intervention. Results from meta-analyses by Kobayashi (2019a, 2019b) highlight the importance of interaction in the GLS of teaching as the student engages in deeper reflection when answering meaningful questions asked by the audience. This is particularly relevant when teaching a fellow student, as both students know the material and can provide relevant feedback (Duran, 2017; Kobayashi, 2019a). Others have proposed that interaction could be a motivating factor which enhances students' self-efficacy and self-esteem, although research findings supporting this argument have been mixed (Kobayashi, 2019b; Rienovita et al., 2018). Kobayashi (2019b) also found that informing students of the subsequent task of teaching enhanced learning outcomes; that is, students who knew what to expect outperformed those who did not (Coleman et al., 1997). Overall, these findings demonstrate that the GLS of teaching can enhance learning outcomes, particularly if students know what to expect and can interact during teaching.

The Role of Media on Learning in Virtual Reality

From a media perspective, we would expect to find a main effect of media on the outcomes in this study (i.e., a difference in learning outcomes between IVR and DVR conditions). Several studies have investigated the effects of IVR compared to less immersive media (Dalgarno & Lee, 2010; Freina & Ott, 2015). Overall, results indicate that lessons in IVR lead to better affective and motivational outcomes compared to similar lessons in less immersive media such as DVR (Makransky et al., 2019a). However, understanding the cognitive factors that influence learning in VR is complex, and comparisons of learning in IVR and less immersive media are inconsistent. Some studies have found IVR to enhance learning (Alhalabi, 2016; Chittaro & Buttussi, 2015; Lamb et al., 2018; Makransky et al., 2019a; Webster, 2016), while others suggest that IVR may have mixed effects (Jensen & Konradsen, 2018), no effects (Leder et al., 2019; Moreno & Mayer, 2002), or even negative effects on knowledge and transfer (Makransky et al., 2017b; Parong & Mayer, 2018; Richards & Taylor, 2015) compared to less immersive media. These mixed results suggest that even though a medium with high fidelity such as IVR in some cases improves learning outcomes, these effects might vary with the nature of the task (Buttussi & Chittaro, 2018; Han, 2019; Jensen & Konradsen, 2018). Despite the mixed results, the media perspective would predict a significant effect of media with IVR leading to different learning outcomes compared to DVR in both method conditions.

The Role of Method on Learning in Virtual Reality

Another perspective is that the instructional method, rather than the medium, modulates learning (Clark, 1994). Following this perspective, we would expect a main effect on learning outcomes with students engaging in the GLS of teaching performing better in both media conditions. Results from a number of studies support this view by showing that instructional methods that have been identified in less immersive media generalize to IVR. These include the segmentation principle (Parong & Mayer, 2018), the modality principle (Moreno & Mayer, 2002), the personalization principle (Makransky et al., 2019b; Moreno & Mayer, 2004), and the pre-training principle (Petersen et al., 2020). The methods perspective on learning in VR would predict a main effect of methods where the GLS of teaching increases learning outcomes across both media conditions.

Method Enables Media when Learning in Virtual Reality

A third perspective posits that different instructional methods enable learning in IVR. From this viewpoint, we would predict an interaction between media and methods, indicating that the GLS of teaching is more effective in IVR than DVR. A number of recent studies have demonstrated the benefits of combining an IVR lesson with different forms of scaffolding. For example, a media (IVR or video) and methods (with or without pre-training) experiment by Meyer et al. (2019) found an interaction effect with pre-training increasing knowledge, transfer, and self-efficacy in the IVR condition, but not in the video condition. The authors conclude that the pre-training sufficiently limited cognitive load in the IVR lesson, thereby allowing the affective affordances of learning in IVR to increase generative learning resulting in better learning outcomes. Furthermore, they reason that pre-training did not result in an improvement in the video condition because the initial cognitive load was not as high as in the IVR condition.

Similarly, Makransky et al. (2020a) found that a science lesson in IVR resulted in significantly higher presence and enjoyment compared to experiencing a similar video lesson, but there were no differences between the media on the outcomes of declarative knowledge, procedural knowledge, or transfer. In a follow-up experiment, they added the GLS of enactment, which consists of carrying out the learned procedures with concrete manipulatives, following the lesson (Fiorella & Mayer, 2016). They found an interaction effect with the GLS of enactment leading to significantly better procedural knowledge and transfer outcomes in the IVR condition compared to the video condition. The authors reason that the GLS of enactment was particularly helpful when learning in IVR because these highly engaging learning experiences may result in students not spending enough resources on reflecting over the learned content when there is no integrated or follow-up GLS activity. Based on this perspective, we would predict an interaction between media (IVR/DVR) and method (GLS/no-GLS), with the GLS of teaching specifically increasing learning outcomes in IVR.

Summary of Theory and Research Questions

Based on the theoretical and empirical research outlined above, we have developed a number of research questions that investigate the three perspectives of how media and methods impact learning in IVR and DVR. The questions are based on evaluating six constructs: *transfer*, defined as students' ability to use the learning material in a new context (Mayer, 2008); *retention*, defined as students ability to recall the learning material much the same way as it was presented during instruction (Mayer, 2008); *self-efficacy*, defined as students' perceived capabilities for learning or performing actions (Bandura, 1997); *motivation*, defined intrinsically as students' behaving out of their own interests, often accompanied by feelings of enjoyment (Ryan & Deci, 2017); *enjoyment*, defined as students perceiving the activity as enjoyable in its own right (Davis et al., 1992; Tokel & İsler, 2015) and *presence*, defined as students' experience of being in a place even when they are physically situated in another (Makransky et al., 2017b).

Research Question 1: How do media and methods influence the outcomes of transfer, retention, and self-efficacy in an immediate post-test?

Research Question 2: How do media and methods influence the outcomes of intrinsic motivation, perceived enjoyment, and presence in an immediate post-test?

Research Question 3: What is the effect of media and methods when students are asked to re-use the simulation with the alternative media on the outcomes of transfer, retention, self-efficacy, presence, perceived enjoyment, and intrinsic motivation?

Methods

Participants

The sample consisted of 89 participants (61 females, 28 males, 0 non-binary) from a large European university. The learning intervention was part of a mandatory class required for all first-year undergraduate biochemistry students. Students ranged in age from 19 to 36 (M = 21.34, SD = 2.18).

Research design

We employed a 2x2 mixed-methods design in a natural classroom setting. In the first part of the study, students were randomly assigned to one of two method conditions (GLS/no-GLS) and one of two initial media conditions (IVR/DVR). In the GLS conditions, students would, in pairs, conduct teaching with one student initially occupying the role as the teacher for five minutes in the audience of the other student, whose role was to pose relevant comments and questions. Then, students switched roles for another five minutes. Students in the no-GLS condition bypassed this step and continued directly from the media intervention to the post-test with no intermediate activity. The IVR and DVR conditions refer to experiencing the virtual lesson on either an HMD or a desktop PC.

Due to teachers wanting to provide all students with the opportunity to use both media conditions, a follow-up intervention was conducted on the same day where students switched media conditions and the experiment was repeated. Thus, we investigated two independent variables (method and media), where media was administered within subjects and method was administered between subjects. Consequently, students were divided into four experimental conditions: *(a)* IVR with the GLS of teaching followed by DVR with the GLS of teaching (N = 23), *(b)* DVR with the GLS of teaching followed by IVR with the GLS of teaching (N = 20), *(c)* only IVR followed by only DVR (N = 24), and *(d)* only DVR followed by only IVR (N = 22). Students engaging in the GLS were informed that they had to teach core concepts from the simulation to a classmate after completing the simulation.

Procedure

Initially, all students were assembled in a classroom where they received an introduction to the experiment. Then, they received randomized ID numbers, which assigned them to one of the four experimental conditions *(a-d)* (see Figure I).

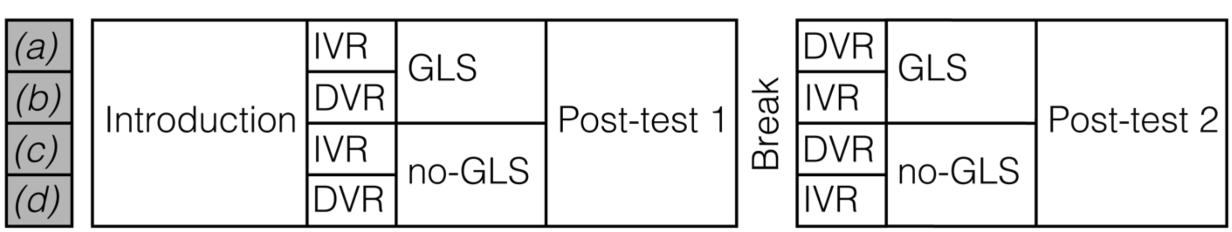

Figure I: Overview of the experimental procedure and the four experimental conditions: *(a), (b), (c)*, and *(d)*.

To avoid distracting students in the alternative media condition, students in the initial DVR conditions *(b)* and *(d)* were led into a different classroom where they engaged in the simulation on their own laptops. Meanwhile, students in the initial IVR conditions *(a)* and *(c)* received a five-minute oral instruction on how to set up and use the HMDs before entering the simulation. There were enough HMDs that every student in the IVR condition could use one simultaneously. Two to three lab instructors were available in each classroom to supervise the sessions, equivalent to ten students sharing two instructors. Immediately after the simulation, students in conditions *(c)* and *(d)* were allocated to separate classrooms to complete post-test 1. Meanwhile, students in conditions *(a)* and *(b)* stayed in their respective classrooms and engaged in the GLS of in teaching in pairs for 10 minutes. Then they completed the same post-test as students in conditions *(c)* and *(d)*.

Figure II: Students in the IVR-condition. Photo: Leif Bolding.

After a break of one hour, all students returned to complete the second part of the experiment. Here, they engaged in the media condition different from the one they completed previously; students in conditions *(b)* and *(d)* used IVR, while students in condition *(a)* and *(c)* used DVR. The following procedure was similar to the one described in the first part of the study. Ultimately, all students finished post-test 2. The total duration of the experiment was about 4 hours, and experiments took place in the course of two days.

Materials

The VR simulation

The learning intervention consisted of the “Electron Transport Chain Virtual Simulation” developed by the EdTech company Labster (see Figure III for screenshots of the simulation or Labster, 2020, for an introduction to the simulation). Research on educational technology has previously used similar VR simulations (e.g., Coleman & Smith, 2019; de Vries & May, 2019; Makransky et al., 2019a). The simulation used in this study encourages inquiry-based learning where students learn from their own mistakes (Pedaste et al., 2015) and based on Mayer's (2008) knowledge taxonomy, it promotes four types of knowledge, including: facts (e.g., the definition of the ETC), concepts (e.g., understanding the importance and uses of photosynthesis), procedures (e.g., the step by step process of conducting the experiment), and beliefs (e.g., building self-efficacy by providing positive feedback after successfully completing a task). Furthermore, it is designed in a way that allows students to experience the same simulation in IVR and DVR.

The IVR simulation was administered on a Lenovo Mirage Solo headset, which has a built-in screen, world tracking, and includes a hand-held controller and headphones. Interactivity in the simulation occurred through movements of the head and use of the controller, allowing students to explore a 360° virtual environment at their own pace. The DVR simulation was accessed through students’ own laptops, mouse/touchpads, keyboards, and their own or provided headphones.

The simulation took students on average 36 minutes (SD=8.77) to complete. The main learning objectives were to teach students about photosynthesis, the properties of light and colorful pigments, and the ETC (see Appendix B for details). In the simulation, students could explore the intracellular processes and perform complex laboratory experiments in the fully equipped virtual laboratory. Throughout the simulation, students interacted with a virtual agent and a lab pad where they received instructions and multiple-choice questions with explanatory feedback to prime their metacognition. Students had to answer the questions in order to progress through the experiment.

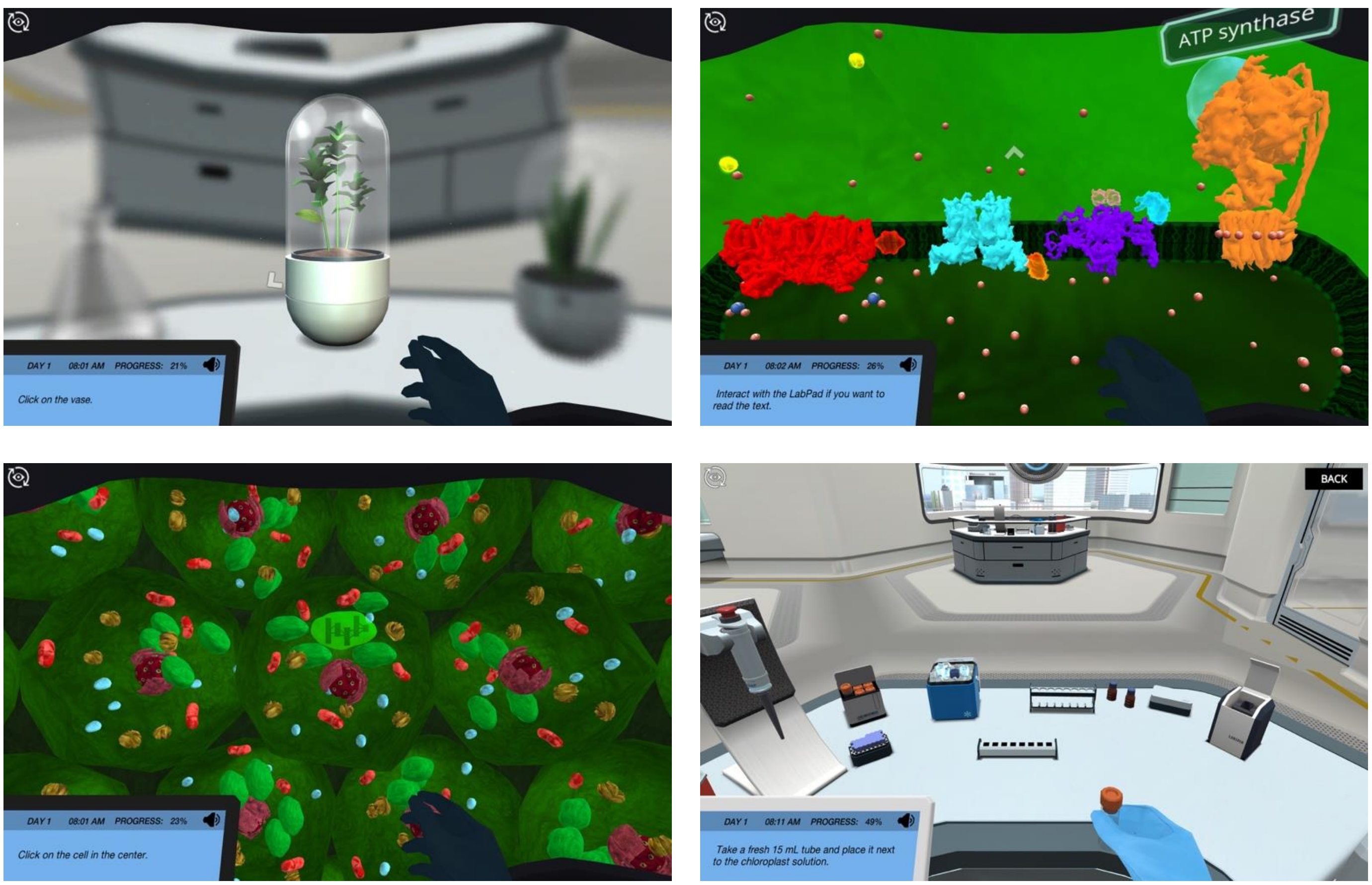

Figure III: Screenshots of The Electron Transport Chain Virtual Simulation taken from Labster (2020).

Post-test measures

Post-test 1 was identical to post-test 2, consisting of demographic items concerning gender and age and scales to measure retention, transfer, self-efficacy, motivation, perceived enjoyment, and presence (see List of Items in Appendix A).

Results

Initially, an analysis was conducted to investigate if the students in the four conditions differed on basic characteristics of age and gender. A one-way ANOVA indicated that the groups did not differ significantly on age, $F_{(3,85)} = 1.387$, $p = .252$. Furthermore, a chi-square test indicated that the groups did not differ significantly in the proportion of males and females, $\chi 2$ (df = 3, $N = 89$) $= 4.892$, $p = .180$.

Results for RQ 1: Influence of Media and Methods on the Outcomes of Transfer, Retention, and Self-efficacy

We investigated RQ 1 through two factorial ANOVAs with media (IVR/DVR) and methods (GLS/no-GLS) as independent variables, with transfer, retention, and self-efficacy as dependent variables. For transfer, the first row of Table 1 shows that the main effect of media $F_{(1,85)} = 1.417$, $p = .237$ was not significant, indicating that there was no significant difference between the knowledge transfer when learning through IVR and DVR. However, there was a significant main effect for methods $F_{(1,85)} = 6.100$, $p = .016$ indicating that the GLS of teaching increased transfer, and a significant interaction $F_{(1,85)} = 6.771$, $p = .011$. Post-hoc analyses using independent samples t-tests separately for each media condition showed a significant effect of enactment in IVR $t_{(45)} = 4.339$, $p < .0001$, d = 1.26, but not in DVR $t_{(40)} = -.080$, $p = .937$, d = .02.

For retention, the main effect for media $F_{(1,85)} = .063$, $p = .802$, and main effect for methods $F_{(1,85)} = .783$, $p = .379$ were not significant. However, the interaction between media and methods was significant $F_{(1,85)} = 4.127$, $p = .045$. Post-hoc analyses using independent samples t-tests independently for each media condition showed a marginally significant effect of enactment in IVR $t_{(45)} = 2.015$, $p = .050$, d = .60, but not in DVR $t_{(40)} = -.843$, $p = .404$, d = 0.26.

Similarly, for self-efficacy there was no main effect for media $F_{(1,85)} = .564$, $p = .455$, or method $F_{(1,85)} = 3.367$, $p = .070$. However, the interaction between media and methods was significant $F_{(1,85)} = 4.213$, $p = .043$. Post-hoc analyses using independent samples t-tests again showed a significant effect of enactment in IVR $t_{(45)} = 2.781$, $p = .008$, d = .82, but not in DVR $t_{(40)} = -.153$, $p = .879$, d = .05. Taken together these results show consistent interaction effects between media and methods indicating that the GLS of teaching was specifically effective in the IVR learning condition for the outcomes of transfer, retention, and self-efficacy, but not in the DVR condition.

Table 1: Means and standard deviations for the dependent variables measured in post-test 1. Significance values for the media by method ANOVAS are presented in the final three columns.

	Post-test 1 Results						
	IVR		**DVR**		**Media**	**Method**	**Interaction**
GLS of Teaching	**GLS**	**no-GLS**	**GLS**	**no-GLS**	**p-value**	**p-value**	**p-value**
Transfer	3.48 (2.17)	1.04 (1.65)	2.80 (2.73)	2.86 (2.46)	.237	**.016**	**.011**
Retention	12.52 (2.35)	10.83 (3.29)	11.20 (2.67)	11.86 (2.44)	.802	.379	**.045**
Self-efficacy	3.56 (.54)	3.15 (.68)	3.29 (.58)	3.32 (.61)	.455	.070	**.043**
Motivation	3.76 (.46)	3.48 (.65)	3.51 (.47)	3.4 (.68)	.194	.114	.486
Enjoyment	3.97 (.67)	3.59 (.97)	3.95 (.46)	3.88 (.81)	.425	.175	.354
Presence	3.42 (.61)	3.34 (.66)	3.24 (.67)	3.01 (.62)	.064	.263	.570

Results for RQ 2: Influence of Media and Methods on motivation, enjoyment, and presence

RQ 2 was investigated with two factorial ANOVAs with media (IVR/DVR) and method (GLS/no-GLS) as independent variables, and intrinsic motivation, perceived enjoyment, and presence as dependent variables. For motivation, row four in Table 1 shows that there was not a significant main effect for media $F_{(1,85)} = 1.717$, $p = .194$, or methods $F_{(1,85)} = 2.546$, $p = .114$, or a significant interaction between media and methods $F_{(1,85)} = .489$, $p = .486$. Similarly, for perceived enjoyment, row five in Table 1 illustrates that there was not a significant main effect for media $F_{(1,85)} = .643$, $p = .425$, or methods $F_{(1,85)} = 1.875$, $p = .175$, or a significant interaction between media and methods $F_{(1,85)} = .867$, $p = .354$. Finally, for presence the final row of Table 1 demonstrates that there was not a significant main effect for media $F_{(1,85)} = 3.523$, $p = .064$, or methods $F_{(1,85)} = 1.270$, $p = .263$, or a significant interaction between media and methods $F_{(1,85)} = .325$, $p = .570$. Taken together these results illustrate that there were surprisingly no differences between the media conditions on the outcomes of intrinsic motivation, perceived enjoyment, and presence, and that there were further no differences between the method conditions or interactions on these variables.

Results for RQ 3: Influence of Media and Methods when Students Switched Media Conditions

RQ3 was investigated with two factorial ANCOVAs. These included media (IVR/DVR) and methods (GLS/no-GLS) as independent variables, the score from post-test 2 on each outcome (transfer, retention, self-efficacy, motivation, enjoyment, and presence) as the dependent variables, and the score for the respective variable in post-test 1 as the covariate. That is, we investigated the effect of media and methods on respondents' post-test 2 score while accounting for the score they obtained in post-test 1 as a covariate for each of the six outcomes in the study.

The same pattern of results which are presented in the top portion of Table 2 was observed for the outcomes of transfer, retention, and self-efficacy. For transfer, there was not a significant main effect for media $F_{(1,83)} = 1.308$, $p = .256$, or methods $F_{(1,83)} = 2.94$, $p = .090$, or a significant interaction $F_{(1,83)} = 1.394$, $p = .241$. For retention there was not a significant main effect for media $F_{(1,72)}$ = 1.071, p = .304, or methods $F_{(1,72)}$ = .051, p = .822, or a significant interaction $F_{(1,72)}$ = 1.671, p = .200. For self-efficacy there was also not a significant main effect for media $F_{(1,73)}$ = .011, p = .918, or methods $F_{(1,73)}$ = .080, p = .779, or a significant interaction $F_{(1,73)}$ = 2.816, p = .144.

We observed a different pattern for the outcomes of intrinsic motivation, perceived enjoyment, and presence, where there was a significant main effect for media across the three outcomes. For intrinsic motivation, a significant main effect for media was observed, indicating that students were more intrinsically motivated when using IVR in the second intervention compared to DVR $F_{(1,73)}$ = 5.766, p = .019, d = 0.16, after accounting for their post-test 1 motivation score. However, the main effect for methods $F_{(1,73)}$ = 1.044, p = .310 and the interaction $F_{(1,73)}$ = .971, p = .328 were not significant. For perceived enjoyment, a significant main effect for media also indicated that students enjoyed using the IVR simulation in post-test 2 compared to the DVR version $F_{(1,73)}$ = 16.748, p < .001, d = 0.94, after accounting for their post-test 1 enjoyment score. However, the main effect for methods $F_{(1,73)}$ = .600, p = .441, and the interaction $F_{(1,73)}$ = 1.173, p = .282 were not significant. For presence, a significant main effect for media also indicated that students reported being significantly more present when using the IVR simulation in post-test 2 compared to the DVR simulation $F_{(1,73)}$ = 43.326, p < .001, d = 1.29, after accounting for their post-test 1 presence score. However, the main effect for method $F_{(1,73)}$ = 1.055, p = .308 and interaction $F_{(1,73)}$ = .552, p = .460 were not significant. Overall, these results indicate that students had higher intrinsic motivation, perceived enjoyment, and presence when using the IVR compared to the DVR version of the simulation, but that this only appeared in post-test 2 once students had used the alternative version of the simulation and a frame of reference.

Table 2: Means and standard deviations for the dependent variables measured in post-test 2. Significance values for the media by method ANCOVAS where the post-test 1 result for each variable is included as the covariate are presented in the final columns.

Post-test 2 Test Results							
	IVR		DVR		Media	Method	Interaction
GLS of Teaching	GLS	no-GLS	GLS	no-GLS	p-value	p-value	p-value
Transfer	2.75 (2.49)	2.52 (2.98)	4.26 (2.89)	1.17 (1.95)	.256	.090	.241
Retention	12.06 (2.24)	13.18 (2.21)	13.23 (1.57)	11.85 (2.96)	.304	.822	.200
Self-efficacy	3.43 (.55)	3.64 (.56)	3.81 (.57)	3.38 (.68)	.918	.779	.144
Motivation	3.60 (.45)	3.42 (.74)	3.61 (.59)	3.41 (.659	**.019**	.310	.328
Enjoyment	4.19 (.45)	4.20 (.62)	3.71 (.87)	3.32 (.88)	**<.001**	.441	.282
Presence	3.72 (.58)	3.59 (.81)	2.89 (.62)	2.75 (.71)	**<.001**	.308	.460

Discussion and conclusions

Empirical Contributions

The main empirical findings in this study are consistent interaction effects between media and methods, indicating that the GLS of teaching was specifically effective in IVR on the outcomes of transfer, retention, and self-efficacy, but not in DVR.

No differences were observed between the media conditions on the outcomes of intrinsic motivation, perceived enjoyment, and presence in post-test 1. Furthermore, we found no differences between the method conditions or interactions on these variables. This finding is inconsistent with most previous research that has identified media differences in favor of IVR compared to less immersive media on outcomes of intrinsic motivation (e.g., Makransky et al., 2019a), perceived enjoyment (e.g., Makransky et al., 2019a), and presence (e.g., Makransky et al., 2017b). However, students' average score measuring outcomes were relatively high both in the IVR and the DVR group (i.e., in intrinsic motivation between 3.4 and 3.76 on a scale from 1-

5 and in perceived enjoyment between 3.59 and 3.97 on a scale from 1-5) indicating that they generally enjoyed both media conditions.

Lastly, post-test 2 indicated that students had higher intrinsic motivation, perceived enjoyment, and presence when using IVR compared to DVR, once they had used the alternative version of the simulation. This suggests that students preferred IVR when they had a frame of reference after trying both media conditions. Finally, no differences were found on outcomes of transfer, retention, and self-efficacy in post-test 2.

Theoretical Contributions

The major theoretical implication from this study supports the perspective that method can enable media affordances. We found an interaction between media and method on the outcomes of transfer, retention, and self-efficacy. This suggests that while using IVR compared to DVR may not improve learning among students, combining an IVR lesson with the GLS of teaching can enhance learning significantly.

This finding is consistent with principles from CTML (Mayer, 2014), suggesting that acquiring new information in IVR can increase students' cognitive load to an extent that exceeds working-memory capacity, as experiences in IVR are highly engaging without necessarily facilitating appropriate metacognitive processing and reflection. Applying the GLS of teaching to an IVR lesson encourages students to reflect over the learned material and facilitates the active process of selecting, organizing, and integrating information (Fiorella & Mayer, 2016). The GLS of teaching may in this way work as supportive scaffolding, which enables the affordances and potential for using IVR in educational settings. Hence, in this study, the medium of IVR could make it easier for students to relate to the ETC because the simulation provided them an opportunity to explore an otherwise 'invisible' phenomena. However, the GLS of teaching was needed in order for students to optimally integrate the new information within their long-term memory.

Practical Considerations

The empirical findings from this study highlight the importance of the instructional method when implementing IVR simulations in educational settings. These results are, therefore, relevant to practitioners who wish to use IVR in an educational context. Specifically, we demonstrate that IVR can enhance learning outcomes as well as self-efficacy when implemented with the GLS of teaching in pairs; a method which is applicable to most learning scenarios. It should be noted that when the GLS of teaching was not used, the DVR group performed equally or better than the IVR group on post-test outcomes of transfer, retention, and self-efficacy. This emphasizes the importance of combining an IVR lesson with appropriate instructional methods. It also suggests that an IVR lesson should not be considered a replacement for all learning activities but rather that given the right method, it can enhance learning in particular educational settings.

Limitations and Future Research

The GLS used in this study is complicated to investigate in an experimental setting, as many factors are involved in the act of teaching. Students may have learned from the act itself, from listening to a fellow student, or from the social interaction that occurred during this exchange. Future research should, therefore, attempt to disentangle these factors to investigate the underlying processes that resulted in the learning gains. Furthermore, studies could investigate if grouping strategy affects learning outcomes when using the GLS of teaching in VR.

This study is conducted on the basis of a single IVR lesson that took place over the course of a full day. This was in order to match the time spent on a similar real-life laboratory experiment of the same content. Most research investigating the value of IVR in education uses short interventions below 20 minutes that are not integrated into a course (e.g., Meyer et al., 2019; Parong & Mayer, 2018). The IVR experience used in this study took students in average 36 minutes to complete and was an integrated part of a higher education course. Thus, the current study builds on one of the more integrated and longer experimental designs compared to most of the current literature in the field. However, virtual experiments give students the opportunity to obtain results immediately rather than having to wait for physical laboratory results, which are typically not available straightaway. This might explain why the duration might be somewhat shorter than a similar physical lab experiment.

The intervention was not repeated, as it was part of a first-year university biochemistry course in which students learn about basic intracellular functions, such as the ETC, only once. Future research could focus on longitudinal studies to investigate the long-term effects of a single IVR lesson or the effects of multiple IVR lessons as part of an extended education program.

There is abundant literature about cybersickness when using IVR (e.g., LaViola, 2000; Rebenitsch & Owen, 2016). Cybersickness can occur due to discrepancy between the vestibular and visual senses, for example when someone in a stationary position experiences locomotion in a virtual environment (LaViola, 2000). Initiatives to reduce cybersickness in the current study include high-quality equipment to prevent display issues, providing user control, and avoiding locomotion why students could remain seated throughout the entire simulation. Furthermore, students were told that they could stop the learning session if they experienced any type of discomfort. No students stopped the lesson, which indicates that this was not an issue. However, we did not measure cybersickness. Therefore, future studies could investigate the effects of cybersickness, as this issue is important to be aware of when using IVR lessons in classrooms.

Some students experienced technical problems in the IVR condition, which may have impaired the outcomes of intrinsic motivation, perceived enjoyment, and presence. However, in the IVR condition, every ten students were assisted by two or three lab instructors to limit the effects of technical difficulties. Nevertheless, other studies could also investigate if these outcomes are dependent on students being new to IVR or accustomed to the medium. Future studies could also implement IVR with the GLS of teaching in subjects other than biochemistry to investigate the generalizability of the findings.

Finally, based on the finding that media interacts with methods, more research should focus on investigating what instructional design principles generalize to more immersive media such as IVR. Assuming that applying a GLS to an IVR lesson improves learning outcomes due to students reflecting over the learning content, a starting point would be to examine if other GLSs outlined by Fiorella and Mayer (2016) can improve learning outcomes in similar ways.

Acknowledgments

We would like to thank all of the students who participated in the study at the Department of Biology at University of Copenhagen.

Statements on Ethics

This project was approved by the University Ethics Committee.

Open Data & Conflicts

No conflicts of interest to declare. Data available upon request.

References

Ainley, M., & Armatas, C. (2006). Motivational Perspectives on Students' Responses to Learning in Virtual Learning Environments. In J. Weiss, J. Nolan, J. Hunsinger, & P. Trifonas (Eds.), *International Handbook of Virtual Learning Environments* (pp. 365–394). Springer Netherlands.

Alhalabi, W. S. (2016). Virtual reality systems enhance students' achievements in engineering education. *Behaviour & Information Technology, 35*(11), 919–925. https://doi.org/10.1080/0144929X.2016.1212931

Bailenson, J. N. (2018). *Experience on Demand: What Virtual Reality Is, how it Works, and what it Can Do.* Norton & Company.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W H Freeman/Times Books/ Henry Holt & Co.

Bogusevschi, D., Muntean, C., & Muntean, G.-M. (2019). Teaching and Learning Physics using 3D Virtual Learning Environment: A Case Study of Combined Virtual Reality and Virtual Laboratory in Secondary School. In K. Graziano (Ed.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2019* (pp. 721–728). Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/p/207721

Bower, M. (2017). *Design of technology-enhanced learning: Integrating research and practice.* http://ebookcentral.proquest.com

Burbules, N. C. (2006). Rethinking the Virtual. In Joel Weiss, J. Nolan, J. Hunsinder, & P. Trifonas (Eds.), *Foundations of Virtual Learning Environments* (pp. 37–58). Springer. https://doi.org/https://doi.org/10.1007/978-1-4020-3803-7

Buttussi, Fabio, & 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. https://doi.org/10.1109/TVCG.2017.2653117

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. https://doi.org/http://dx.doi.org/10.1109/TVCG.2015.2391853

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, *42*(2), 21–29. https://doi.org/10.1007/BF02299088

Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The Effect of Instructional Explanations on Learning from Scientific Texts. *Journal of the Learning Sciences*, *6*(4), 347–365. https://doi.org/10.1207/s15327809jls0604_1

Coleman, S. K., & Smith, C. L. (2019). Evaluating the benefits of virtual training for bioscience students. *Higher Education Pedagogies*, *4*(1), 287–299. https://doi.org/10.1080/23752696.2019.1599689

Concannon, B. J., Esmail, S., & Roduta Roberts, M. (2019). Head-Mounted Display Virtual Reality in Post-secondary Education and Skill Training. *Frontiers in Education*, *4*. https://doi.org/10.3389/feduc.2019.00080

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. https://doi.org/10.1080/15213269.2015.1015740

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10–32.

https://doi.org/10.1111/j.1467-8535.2009.01038.x

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Social Psychology*, *22*(14), 1111–1132. https://doi.org/10.1111/j.1559-1816.1992.tb00945.x

De Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science*, *340*(6130), 305–308. https://doi.org/10.1126/science.1230579

de Vries, L. E., & May, M. (2019). Virtual laboratory simulation in the education of laboratory technicians–motivation and study intensity. *Biochemistry and Molecular Biology Education*, *47*(3), 257–262. https://doi.org/10.1002/bmb.21221

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating Internalization: The Self-Determination Theory Perspective. *Journal of Personality*, *62*(1), 119–142. https://doi.org/10.1111/j.1467-6494.1994.tb00797.x

Duran, D. (2017). Learning-by-teaching. Evidence and implications as a pedagogical mechanism. *Innovations in Education and Teaching International*, *54*(5), 476–484. https://doi.org/10.1080/14703297.2016.1156011

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, *38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Fiorella, L., & Mayer, R. E. (2016). Eight Ways to Promote Generative Learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

Fitzgerald, M., & Riva, G. (2001). Virtual reality. In L. Beolchi (Ed.), *Telemedicine glossary* (pp. 327–329). European Commission-DG INFSO.

Freina, L., & Ott, M. (2015). A literature review on immersive virtual reality in education. *The International Scientific Conference ELearning and Software for Education*, 133–141. https://doi.org/10.12753/2066-026X-15-020

Gegenfurtner, A., Quesada-Pallarès, C., & Knogler, M. (2014). Digital simulation-based training: A meta-analysis. *British Journal of Educational Technology*, *45*(6), 1097–1114. https://doi.org/10.1111/bjet.12188

Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2020). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*. https://doi.org/10.1007/s40692-020-00169-2

Han, I. (2019). Immersive virtual field trips in education: A mixed-methods study on elementary students' presence and perceived learning. *British Journal of Educational Technology*, *n/a*(n/a). https://doi.org/10.1111/bjet.12842

Jelfs, A., & Whitelock, D. (2000). The notion of presence in virtual learning environments: what makes the environment "real." *British Journal of Educational Technology*, *31*(2), 145–152. https://doi.org/10.1111/1467-8535.00145

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. https://doi.org/10.1007/s10639-017-9676-0

Jones, N. (2018). The Virtual Lab: Can a simulated laboratory experience provide the same benefits for students as access to a real-world lab? *Nature*, *562*, S5–S7.

Kaplan-Rakowski, R., & Gruber, A. (2019). Low-immersion versus high-immersion virtual reality: Definitions, classification, and examples with a foreign language focus. *Proceedings* *of the Innovation in Language Learning International Conference 2019*.

Kobayashi, K. (2019a). Interactivity: A Potential Determinant of Learning by Preparing to Teach and Teaching. *Frontiers in Psychology*, *9*, 2755. https://doi.org/10.3389/fpsyg.2018.02755

Kobayashi, K. (2019b). Learning by Preparing-to-Teach and Teaching: A Meta-Analysis. *Japanese Psychological Research*, *61*(3), 192–203. https://doi.org/10.1111/jpr.12221

Labster. (2020). *Electron Transport Chain Virtual Lab Simulation*. https://www.labster.com/simulations/electron-transport-chain/

Lamb, R., Antonenko, P., Etopio, E., & Seccia, A. (2018). Comparison of virtual reality and hands on activities in science education via functional near infrared spectroscopy. *Computers & Education*, *124*, 14–26. https://doi.org/https://doi.org/10.1016/j.compedu.2018.05.014

LaViola, J. J. (2000). A Discussion of Cybersickness in Virtual Environments. *SIGCHI Bull.*, *32*(1), 47–56. https://doi.org/10.1145/333329.333344

Leder, J., Horlitz, T., Puschmann, P., Wittstock, V., & Schütz, A. (2019). Comparing immersive virtual reality and powerpoint as methods for delivering safety training: Impacts on risk perception, learning, and decision making. *Safety Science*, *111*, 271–286. https://doi.org/https://doi.org/10.1016/j.ssci.2018.07.021

Lee, E. A.-L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, *55*(4), 1424–1442. https://doi.org/10.1016/j.compedu.2010.06.006

Lee, J. (1999). Effectiveness of Computer-Based Instructional Simulation: A Meta Analysis. *International Journal of Instructional Media*, *26*(1), 71–85. https://www.learntechlib.org/p/85364

Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2020). 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*. https://doi.org/10.1037/edu0000473

Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning*, *35*(6), 691–707. https://doi.org/10.1111/jcal.12375

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), 1–24. https://doi.org/10.1007/s11423-018-9581-2

Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior*, *72*, 276–285. https://doi.org/10.1016/j.chb.2017.02.066

Makransky, G., & Petersen, G. B. (2019). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. *Computers & Education*, *134*(July 2018), 15–30. https://doi.org/10.1016/j.compedu.2019.02.002

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*, *n/a*(n/a). https://doi.org/10.1111/bjet.12954

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2017). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *May*, 0–1. https://doi.org/10.1016/j.learninstruc.2017.12.007

Makransky, G., Wismer, P., & Mayer, R. E. (2019). 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. https://doi.org/10.1111/jcal.12335

Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Pearson Merrill Prentice Hall.

Mayer, R. E. (2014). *Cognitive theory of multimedia learning. The Cambridge handbook of multimedia learning* (2nd.). New York: Cambridge University Press.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, *93*(2), 390–397. https://doi.org/10.1037/0022-0663.93.2.390

McClean, P., Johnson, C., Rogers, R., Daniels, L., Reber, J., Slator, B. M., Terpstra, J., & White, A. (2005). Molecular and cellular biology animations: development and impact on student learning. *Cell Biology Education*, *4*(2), 169–179. https://doi.org/10.1187/cbe.04-07-0047

McLuhan, M. (1964). *The medium is the message*. MIT Press.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers and Education*, *70*, 29–40. https://doi.org/10.1016/j.compedu.2013.07.033

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 and Education*, *140*(June), 103603. https://doi.org/10.1016/j.compedu.2019.103603

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

https://doi.org/10.1016/j.compedu.2010.10.020

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, *94*(3), 598–610. https://doi.org/10.1037/0022-0663.94.3.598

Parong, J., & Mayer, R. E. (2018). Learning Science in Immersive Virtual Reality. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000241

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003

Petersen, G. B., Klingenberg, S., Mayer, R. E., & Makransky, G. (n.d.). The Virtual Field Trip: Investigating How to Optimize Immersive Virtual Learning in Climate Change Education. *British Journal of Educational Technology*. https://doi.org/https://doi.org/10.1111/bjet.12991

Pintrich, P. R., Smith, D. A. F., Garcia, T., & Mckeachie, W. J. (1993). Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (Mslq). *Educational and Psychological Measurement*, *53*(3), 801–813. https://doi.org/10.1177/0013164493053003024

Radianti, 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. https://doi.org/10.1016/j.compedu.2019.103778

Rebenitsch, L., & Owen, C. (2016). Review on cybersickness in applications and visual displays.

Virtual Reality, *20*(2), 101–125. https://doi.org/10.1007/s10055-016-0285-9

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 and Education*, *86*, 157–171. https://doi.org/10.1016/j.compedu.2015.03.009

Rienovita, E., Taniguchi, M., Kawahara, M., Hayashi, Y., & Takeuchi, Y. (2018). Implementation of interactive peer learning environment enhances learners´ self-steem and self-efficacy. *International Journal of Learning Technologies and Learning Environments*, *1*(1), 1–24. http://www.iaiai.org/journals/index.php/IJLTLE/article/view/227

Roscoe, R. D., & Chi, M. T. H. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science*, *36*(4), 321–350. https://doi.org/10.1007/s11251-007-9034-5

Ryan, R. M., & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. Guilford Publications. https://books.google.dk/books?id=Bc_DDAAAQBAJ

Schönborn, K. J., & Anderson, T. R. (2006). The importance of visual literacy in the education of biochemists*. *Biochemistry and Molecular Biology Education*, *34*(2), 94–102. https://doi.org/10.1002/bmb.2006.49403402094

Sherman, W. R., & Craig, A. B. (2003). *Understanding Virtual Reality: Interface, Application, and Design*. Morgan Kaufmann Publishers.

Shin, D.-H. (2017). The role of affordance in the experience of virtual reality learning: Technological and affective affordances in virtual reality. *Telematics and Informatics*, *34*(8), 1826–1836. https://doi.org/https://doi.org/10.1016/j.tele.2017.05.013

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. In *Presence: Teleoperators and Virtual Environments* (pp. 603–616). MIT Press.

Tokel, S. T., & İsler, V. (2015). Acceptance of virtual worlds as learning space. *Innovations in Education and Teaching International*, *52*(3), 254–264. https://doi.org/10.1080/14703297.2013.820139

Tsivitanidou, O. E., Gray, P., Rybska, E., Louca, L., & Constantinou, C. P. (2018). *Professional Development for Inquiry-Based Science Teaching and Learning*. Springer International Publishing. https://books.google.dk/books?id=Gf5sDwAAQBAJ

Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, *24*(6), 1319–1333. https://doi.org/10.1080/10494820.2014.994533

Weiss, Joel, Nolan, J., Hunsinger, J., & Trifonas, P. (Eds.). (2006). *The International Handbook of Virtual Learning Environments*. Springer Netherlands. https://doi.org/10.1007/978-1-4020-3803-7

Wittrock, M. C. (1974). A Generative Model of Mathematics Learning. *Journal for Research in Mathematics Education*, *5*(4), 181–196. https://doi.org/10.2307/748845

Appendix A

List of Items

LABEL	ITEM	Source and Cronbach's Alpha Reliability
Motivation		Adapted from Deci, Eghrari, Patrick, & Leone, 1994.
Mot_1	I enjoy working with Photosynthesis	
Mot_2	Electron transport chain activities are fun to perform	
Mot_3	Photosynthesis is boring	
Mot_4	Electron Transport Chain does not hold my attention at all	Cronbach's alpha = 0.81 (post-test 1) and 0.81 (post-test 2)
Mot_5	I would describe photolysis of water and electron transport as very interesting	
Self-efficacy		Adapted from Pintrich, Smith, Garcia, & Mckeachie, 1993.
SE_1	I am confident and can understand the basic concepts of Electron transport chain	
SE_2	I am confident that I understand the most complex concepts related to Photosynthesis	Cronbach's alpha = 0.81 (post-test 1) and 0.83 (post-test 2)
SE_3	I am confident that I can do an excellent job on the assignments and tests in this course	
SE_4	I expect to do well in this course	
SE_5	I am certain that I can master the skills being taught in this course	
Perceived Enjoyment		Adapted from Tokel & İsler, 2015.
Enj_1	I find using virtual reality/computer simulations enjoyable	Cronbach's alpha = 0.89 (post-test 1) and 0.87 (post-test 2)
Enj_2	Using virtual reality/computer simulations is pleasant	
Enj_3	I have fun using virtual reality/computer simulations	
Presence		Adapted from Makransky, Lilleholt, & Aaby, 2017.
Pres_1	The virtual environment seemed real to me.	
Pres_2	I had a sense of acting in the virtual environment, rather than operating something from outside.	
Pres_3	My experience in the virtual environment seemed consistent with my experiences in the real world	Cronbach's alpha = 0.74 (post-test 1) and 0.85 (post-test 2)
Pres_4	While I was in the virtual environment, I had a sense of "being there".	
Pres_5	I was completely captivated by the virtual world.	
Retention		Cronbach's alpha = 0.71 (post-test 1) and 0.67 (post-test 2)
Q1	Light is composed of particles called:	
Q2	Which photosystem does the light-dependent reaction begin with?	
Q3	Where are the molecules of the electron transport chain found in plant cells?	

Q4	A biological redox reaction always involves:	
Q5	The electrons for the light-dependent reactions comes from	
Q6	What molecule reduces NADP+ to NADPH in photosystem I?	
Q7	Chlorophyll gives leaves a green color because it	
Q8	What is the primary function of the light-dependent reactions of photosynthesis?	
Q9	What are the products of the light reaction used for in the dark reactions?	
Q10	What is the final electron acceptor in the light reaction?	
Q11	What molecule absorbs sunlight for photosynthesis?	
Q12	Which of the following statements are true for plastoquinone	
Q13	How many membranes surround the chloroplast?	
Q14	When oxygen is released as a result of photosynthesis, it is a by-product of which of the following?	
Q15	A plant has a unique photosynthetic pigment. The leaves of this plant appear to be reddish yellow. What wavelengths of visible light are being absorbed by this pigment?	
Q16	The figure below shows the absorption spectrum for chlorophyll a and the spectrum for the activity of photosynthesis. Why are they different?	
Transfer		Each question is scored on a scale from 0 to 4 based on a rubric.
Transfer1	Explain in details photosystem I	
Transfer2	Explain in details photosystem II	

Appendix B

The Electron Transport Chain Virtual Simulation

The main objective was to teach students about photosynthesis, the properties of light and colorful pigments, and the ETC. In the virtual environment, students had to help a group of engineers to investigate if dark algae could work as a sustainable source of energy. Thus, students' task was to explore whether green light could be utilized to perform photosynthesis in the dark algae. They had to test their hypothesis by performing experiments in the virtual lab with the help of a virtual pedagogical agent. Initially, students used the Hill reaction and spectrophotometry to measure the energy created by photosynthesis inside the algae. To fully understand this process, students explored the inside of a chloroplast cell and observed the steps in the ETC by interacting with different molecules. Then, students returned to the virtual laboratory where they used their newly acquired knowledge to complete the experiment. During the simulation, students interacted with the virtual agent and a lab pad, where they received instructions and multiple-choice questions with explanatory feedback to prime their meta cognition (see Figure II for screenshots of the simulation).