Immersion positively affects learning in virtual reality games compared to equally interactive 2d games

Citation: Thompson, Meredith, Uz-Bilgin, Cigdem, Tutwiler, M Shane, Anteneh, Melat, Meija, Josephine Camille et al. 2021. "Immersion positively affects learning in virtual reality games compared to equally interactive 2d games." Information and Learning Science, 122 (7/8).

As Published: 10.1108/ILS-12-2020-0252

Publisher: Emerald

Persistent URL: https://hdl.handle.net/1721.1/144420

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Immersion Positively Affects Learning in Virtual Reality Games Compared to Equally Interactive 2D Games

Meredith Thompson, Cigdem Uz-Bilgin, M. Shane Tutwiler, Melat Anteneh, Josephine Camille Meija, Annie Wang,
Philip Tan, Richard Eberhart, Dan Roy, Judy Perry, Eric Klopfer

Final version appears in https://www.emerald.com/insight/publication/issn/2398-5348

Purpose: *This study isolates the effect of immersion on players' learning in a virtual reality-based game about cellular biology by comparing two versions of the game with the same level of interactivity. We identify immersion and additional interactivity as two key affordances of VR as a learning tool. A number of research studies compare VR with two-dimensional or minimally interactive media; this study focuses on the effect of immersion as a result of the head mounted display (HMD).*

Study design/ methodology/ approach: *In the game, players diagnose a cell by exploring a virtual cell and search for clues that indicate one of five possible types of cystic fibrosis. Fifty-one adults completed all aspects of the study. Players took pre and post assessments and drew pictures of cells and translation before and after the game. Players were randomly assigned to play the game with the head mounted display (stereoscopic view) or without the headset (non-stereoscopic view). Players were interviewed about their drawings and experiences at the end of the session.*

Findings: *Players in both groups improved in their knowledge of the cell environment and the process of translation. Players who experienced the immersive stereoscopic view had a more positive learning effect in the content assessment, and stronger improvement in their mental models of the process of translation between pre- and post-drawings compared to players who played the two-dimensional game.*

Originality/ Value: *This study suggests that immersion alone has a positive effect on conceptual understanding, especially in helping learners understand spatial environments and processes.*

Keywords: Virtual Reality, Stereoscopic, Immersion, Interactivity, Cellular Biology, Game Based Learning

1. Introduction

The increasing accessibility and decreasing cost of Virtual Reality (VR) make VR more feasible for educational settings (Castaneda, Cachony & Swanson, 2017). In this study, we use Onyesolu & Ese (2011)'s definition of VR, "the simulation of a real or imagined environment that can be experienced visually in the three dimensions of width, height, and depth and that may additionally provide an interactive experience visually in full real-time motion with sound and possibly with tactile and other forms of feedback" (p. 53). Immersive VR can take the form of a head mounted display (HMD) or a full room CAVE set up; in contrast, non-immersive VR and virtual environments (VE) are viewed on a two-dimensional flatscreen such as a computer monitor (Jensen & Konradsen, 2018). Immersive VR holds great potential for education both for engaging learners and the ability to expand the context of learning to include places too dangerous, inconvenient, or expensive to experience without VR (Bailenson, 2018). Research to date suggests that the media-rich environment of VR can pique students' interest (Makranksy, Andreasen, Baceviciute & Mayer, 2020); however, some attribute this interest to the novelty of VR (Merchant et al., 2014). Moving beyond the novelty effect will require strong evidence that the attributes of VR can continue to motivate learners and effectively facilitate learning. Skeptics assert that VR offers little more than videos or computer experiences, however, mounting research indicates that VR learning experiences create a lasting impact on learners (author, 2019). Furthermore, VR can be integrated as part of a series of approaches, including Augmented Reality (Liou, Yang, Chen & Tarng, 2017). VR has a number of different attributes, we focus on two important attributes of VR: the immersion created by the head mounted display (HMD), and the interactivity enabled through the hand controllers (Johnson-Glenberg, 2018). Understanding how VR influences learners requires a systematic investigation of both attributes of the technology. While current research on learning through VR shows mixed results (Makransky, Terkildsend & Mayer, 2019; Parong & Mayer, 2017), many current studies of VR conflate the contributions of VR by varying both the users' perspective (flatscreen or viewed through a VR headset and the level of interactivity of the experience (such as comparing a virtual slide presentation to interactive experience, e.g., Parong & Mayer, 2017). In order to isolate the effect of the VR based display, we compare two similarly interactive stimuli (Gauthier & Jenkinson, 2017), varying only the type of display. By having players play identical versions of the game with and without the head mounted display (HMD), this study seeks to isolate the effect of immersion resulting from the stereoscopic lens set up on the HMD. We have the following two research questions.

(RQ1) How does a stereoscopic view of a virtual cell help individuals learn about the cellular environment compared to a non-stereoscopic view when interactivity is the same?
(RQ2) How does a stereoscopic view impact players' understanding of the cellular process of translation of RNA to amino acid chains compared to a non-stereoscopic view?

Understanding how to apply VR in education requires understanding how the affordances of VR impact learning. This study focuses on two affordances of VR: the immersion of the learner in a virtual world through the head mounted display (HMD) and the interaction learners have through the head tracking and hand controllers. This study contributes to the existing research on VR and learning by isolating the effect of immersion from interactivity on individuals who play a VR game with the HMD (Immersive) and without the HMD (non-immersive), and in considering both factual knowledge about cells and knowledge of the process of translation.

2. Literature review

We will first provide an overview of the current understanding of VR and learning based on findings from recent literature reviews of VR and learning. Then we will explore existing research on stereoscopic views and learning. Finally, we will review existing games within the context of biology and cell biology.

2.1 Reviews of immersive virtual reality in education

Hamilton, McKechnie & Edgerton (2020) reviewed experimental studies that compared immersive VR to non-immersive digital environments between 2013 and 2019. They identified 29 experimental studies that included immersive VR, were above high school level, and included measures of learning outcomes. Most studies included STEM topics, used multiple choice questions and a non-stereoscopic/posttest design, and focused on short term learning outcomes. The studies included in this review indicated that VR was well suited to topics that were highly visual and included experiential learning.

Jensen & Konradsen (2018) investigated the use of immersive VR in education and training. They searched literature between 2013 and 2017 and focused on 21 studies. They found that certain attributes were important for effective VR learning experiences. For example, low interaction VR experiences were less distraction, and may be better for building cognitive skills such as recall of information. While fidelity is important in VR simulations, skills learned in simulations help with future simulation performance and don't necessarily transfer to the real world. When considering cognitive skills, they sound that HMDs were better for developing spatial awareness, and the 2 dimensional alternatives are more effective in fact-based information.

Pellas, Dengel, and Christopolous (2020) investigated VR looking at the impact of VR on learning outcomes from an educational technology standpoint. They looked at 31 studies between 2009-2019, examining features of the studies, the impact of VR on learning outcomes, and the equipment used in the VR design. Most studies used hands on and active learning experiences, and used discovery as the mechanism for learning more often than presentation and collaboration. Science was the most frequent topic between science, technology, engineering, and math. They found contradictory results when comparing VR and 2D options, but found that CAVE (room scale) approaches tended to have positive results when compared to 2D. VR supported learning by doing, active learning, and self-regulated learning. However, the current set of studies focused on user testing, so additional research with improved measures (surveys) are needed to confirm learning gains.

Overall, these reviews of the current research suggest: 1. STEM, especially science, is a frequent topic in VR 2. VR experiences appear to be motivating to learners 3. Current research supports learning in skills and in spatial understanding 4. Results are mixed as to whether cognitive skills improve through VR 5. Additional research is needed to understand the impacts of VR on learning. This study addresses the gaps in understanding the types of cognitive skills that VR can address, and the impact of VR on learning factual and procedural knowledge.

2.2. Stereoscopic view and learning

Stereoscopic technology produces a perception of three dimensions by providing two slightly different perspectives on the same objects (Afifi, 2016). While the widespread use of VR headsets in classrooms is fairly recent, stereoscopic views can be produced without a VR head mounted display. Stereoscopic 3D movie viewers can produce a three-dimensional effect from a two-dimensional image, and stereoscopic viewers have been used to investigate maps, including searching for rocket construction sites in World War II. Within VR, the illusion of three dimensions is produced through a HMD for personal VR use, or through specialized glasses in CAVE setups. We review some of the current research on the impact of stereoscopic views on learning.

Parmar et al. (2016) developed an interactive circuitry simulation (IBAS) designed to teach college students about various electrical measuring equipment in order to compare stereoscopic and non-stereoscopic virtual realities. Participants using the stereoscopic view adjusted to the task faster than subjects in the non-stereoscopic condition and performed better on knowledge, application, analysis, and evaluation measures (Parmar et al., 2016). A study comparing the effectiveness of non-stereoscopic and stereoscopic displays in teaching geography found that students in the stereoscopic condition performed better on an academic achievement test and reported having had a more positive experience than students in the non-stereoscopic condition (Taştı & Avci, 2020). Undergraduate students that watched a stereoscopic origami demonstration performed better on a paper folding test than those who learned by watching a non-stereoscopic instructional video (Dan & Reiner, 2018)

Stereoscopic learning is not ideal for teaching all subjects and can detract from learning if not implemented correctly (Ferdig, Blank, Kratcoski, & Clements, 2015). Roettl & Terlutter's (2018) study found that participants using stereoscopic technology performed worse on recall and recognition memory tasks than those using either 2D or 3D displays. Makransky, Terkildsen, & Mayer's (2019) study comparing stereoscopic and non-stereoscopic displays for learning about mammalian transient protein expression found that participants who used the stereoscopic display had lower scores on a biology knowledge test than those who used the non-stereoscopic display. Students in the stereoscopic condition also had more trouble applying the information they had learned to new scenarios as measured by a transfer test (Makransky, Terkildsen, & Mayer, 2019). Participants in the stereoscopic condition of Parmar et al.'s (2016) study took more time to complete the exercises than those in non-stereoscopic condition.

Stereoscopic-based instruction can provide students with more than just content knowledge. Loup-Escande, Jamet, Ragot, Erhel, and Michinov (2017) compared learning outcomes and experiences of an interactive mechanics lesson between students using monoscopic and stereoscopic displays. Although there was no significant difference in memorization scores, students who used the stereoscopic display had a higher successful task completion rate (100% task completion for stereoscopic versus 90% task completion for non-stereoscopic) and reported feeling a greater sense of immersion, flow, and affective quality than students who used a non-stereoscopic display (Loup-Escande, Jamet, Ragot, Erhel, & Michinov, 2017). In a study comparing a stereoscopic and two non-stereoscopic (3D and 2D) versions of the same game, the stereoscopic display produced a greater sense of spatial presence than either non-stereoscopic display (Roettl & Terlutter, 2018). In Makransky, Terkildsen, & Mayer's (2019) study using non-stereoscopic and stereoscopic displays with and without text narration for learning about mammalian transient protein expression, students in both stereoscopic conditions had higher presence scores than students in either of the non-stereoscopic condition.

VR can provide students with a more enjoyable, motivating learning experience. In their study comparing stereoscopic and non-stereoscopic displays for learning biology, Ferdig, Blank, Kratcoski, & Clements (2015) found that participants preferred 3D learning over 2D learning for visualizing and understanding DNA. When learning about the structure and function of ATP-Synthase, participants in the stereoscopic condition responded more positively on an experience questionnaire than those in the non-stereoscopic condition (Huk, Steinke, & Floto, 2010). In Taştı & Avcı's (2020) study, students in the stereoscopic group reported having a generally more positive experience than students in the non-stereoscopic group as measured by an animation opinion scale.

Proper implementation and curriculum design is important for increasing the effectiveness of stereoscopic learning and reducing learners' extraneous cognitive load. Lesson modifications that can improve learning in VR include the addition of pre-training (Meyer, Omdahl, & Makransky, 2019), offloading cognition by providing students with visual cues (Huk, Steinke, & Floto, 2010; Weisberg & Newcombe, 2017), and the careful selection of supplementary 3D images and accompanying explanations (Ferdig, Blank, Kratcoski, & Clements, 2015). Learning with stereoscopic displays can reduce cognitive load when there are skills training, procedural, or spatial elements to the subject being taught, such as box origami (Concannon, Esmail & Roduta Roberte, 2019; Dan & Reiner, 2018). However, attempting to learn memorization-heavy topics in VR can result in a higher cognitive load (Makransky, Terkildsen, & Mayer 2019; Remmele, Weiers, & Martens, 2015; Roettl & Terlutter, 2018). Stereoscopic displays can cause dizziness, nausea, and general discomfort, which impedes learning, especially in users who are unfamiliar with stereoscopic technology (Kaplan-Rakowski 2019; Parmar et al., 2016; Roettl & Terlutter, 2018).

2.2. Related Work: VR in Biology simulations, games, and research

Applications of VR have expanded beyond gaming and flight and surgical training to applications in learning and empirical research (Slater & Sanchez-Vives, 2016). Science, technology, engineering, and mathematics (STEM) is a frequent domain for VR learning games (Pellas, Denge & Christopoulos, 2020). A number of VR experiences have focused on cellular biology. Virtual tours of cells are simulated in *Journey to the Center of a Cell* (Journey, 2016) and *Cellscape: VR Biology* and (DeWitt, 2016). Players catch a glimpse of how researchers are using CRISPR to fight sickle cell anemia through Immersive Genome Institute's *CrisprVR: Exploring Sickle Cell and Genome Editing.* Educational technology companies are also developing biology themed VR experiences. Labster offers 30 different virtual labs related to biology (Labster). *VR Sharecare* includes simulations, videos, VR and AR experiences about cells and therapies (https://www.sharecare.com/pages/vr). zSpace has a series of AR experiences that include topics such as Cells and Viruses (https://zspace.com/edu/content/subjects/life-science/cells-and-viruses/5636).

Biology-focused games introduce additional interactions with the abstract and complex environment of cells. In *A Journey Inside a Cell,* players learn about cells by defending the cell against an invading virus (Luden.io, n.d.). Museum visitors at the Pacific Science Center played *miTOSSis* by picking up virtual cells at different stages of mitosis and matching them to a diagram on a virtual wall by tossing them against the wall (Allen Institute, 2017). In the game *Mission Biotech*, players become a researcher in a virtual laboratory and learn about biology through conducting virtual experiments (https://www.virtualheroes.com/portfolio/education/mission-biotech/).

There is a growing body of research around the efficacy of VR games and simulations, although the results of the studies are mixed. VR animations were more effective than traditional lecture in helping students understand how particles can be internalized in cells (Johnson et al., 2017). However, students who learned about cells using a slide deck fared better on the assessment than students who used the VR simulation *In Cell: Journey Inside A Cell* (Parong & Mayer, 2018). Parong and Mayer found that text prompts improved recall across both VR and non-VR groups. Giving the students an overview of the cell before the VR experiences, or "pretraining", improved learners' recall of a virtual cell experience (Mayer, Omdahl & Makransky (2019). In studying VR Bio3D, a game designed to allow engineering students to learn about protein synthesis (Ibarra-Herrera et al, 2019). Students found the game-based version of VR Bio3D more motivating for learning than a narrative version.

2.3 Multimedia Theory of Learning

The cognitive theory of multimedia learning (CTML) states that the brain processes text and images first separately (Dual Coding theory), that each channel has limited capacity (similar to cognitive load theory), and that learning is the process of bringing the information together and making sense of it (Mayer, 2014). According to this theory, the goal of multimedia design is to limit extraneous processing that is not related to the learning objective, to manage the essential processing required to achieve the learning objective, and to foster generative processing (sensemaking) beyond the learning objective.

According to the theory, five cognitive processes take place in multimedia learning: selecting relevant words from speech or text, selecting relevant images, organizing the words into an understandable representation, organizing the pictures into an understandable representation, and combining both of these representations with prior knowledge (Mayer, 2005). The CTML works in tandem with Sweller's cognitive load theory (Sweller, Ayres, & Kalyuga, 2011), which also acknowledges "separate channels for dealing with auditory and visual material" (Sweller, 1999) and acknowledges working memory as one of the essential elements in learning. Multimedia learning makes use of the three different memory stores: sensory memory, working memory, and long-term memory. In the CTML, the working memory acts as the hub for all multimedia learning. Inside of the working memory, the brain receives raw material and must use the five cognitive processes to make sense of the information given, selecting images and words to organize and integrate them.

However, the working memory can only store so much information. When the learner's cognitive processing requires more than the available capacity, the learner undergoes cognitive overload (Moreno & Mayer, 2003). Moreno and Mayer identify three different kinds of cognitive processes that work together to help a person learn: essential processing, incidental processing, and representational holding. Essential processing makes sense of the material presented through the five cognitive processes of the CTML. Incidental processing is any processing not part of the learning material but is still present, such as background music during a presentation. Representational holding is the process of storing information in the working memory over a period of time. Cognitive overload takes place when the power required to process the presented material exceeds the abilities of the learner. Moreno and Mayer (2003) identify five different overload scenarios: one channel essential processing overload (in either visual or verbal), essential processing overload in both channels, essential processing and incidental processing overload by extraneous information, essential processing and incidental processing overload by confusing presentation, and essential processing and representational holding overload.

In the case of virtual reality (VR), learners typically experience significantly higher cognitive load and as a result may learn less factual knowledge in comparison to those who do not use VR (Makransky, Mayer, & Terkildsen, 2019). Due to an increased amount of sensory information, learners experience essential processing overload in VR compared to a less immersive counterpart (Richards & Taylor, 2015). In order to combat cognitive load, Mayer and Moreno propose a variety of solutions depending on the kind of load. In previous studies, VR has been an overload on either or both of the visual and verbal channels. The suggested load-reducing methods are off-loading, segmenting, and pretraining (Moreno & Mayer, 2003). Off-loading consists of redistributing processing from visual to auditory or auditory to visual (Moreno & Mayer, 2003). Segmenting better apportions time between information (Moreno & Mayer, 2003). Pretraining provides an overview of important names and characteristics before the main learning takes place (Moreno & Mayer, 2003). Other technologies such as augmented reality (AR) also provide affordances that can help learners understand challenging concepts in science (Yoon, Anderson, Lin, Elinich, 2017). This study explores what facets of the technology help learners, and the types of topics that are well suited to VR.

3. Method

3.1. Biology Game (name removed)

This study investigates an educational game designed to help students learn cellular biology called Biology Game *Cellverse*. First, we will provide some scientific background about cystic fibrosis that is helpful in understanding the game. Then, we will describe the game goal, the virtual game space, the game mechanics, the tools available to players, the guidance players receive, and how new parts of the game are revealed during game play.

Here, we will provide some additional background that the students do not see in the game but is helpful to understand to clarify game structure. Cystic fibrosis (CF) is a genetic disease affecting the lungs resulting from malformation of a protein called the cystic fibrosis transmembrane conductance regulator (CFTR). Improperly formed CFTR proteins disrupt the flow of ions across the cell membrane, halting the sweeping motion of the cilia that would normally prevent mucus buildup in the lungs. Different CF classes are triggered by different DNA mutations, and affect the CFTR protein at different points in the protein creation and distribution process: (1) protein production (class I), (2) protein processing (class II), (3) protein channel malformation (class III), (4) conduction mutations (class IV), and (5) insufficient protein at the cell membrane (class IV) (Cystic Fibrosis Foundation, n.d.). Each mutated sequence produces a detectable clue in the cell, such as shortened mRNA strands (class I), or blocked protein channels at the membrane (class III). Finding the clues and linking them to different classes of CF engages students in interrogating the process of translation of RNA to amino acids and introduces them to the complex, dynamic environment of the virtual cell. Recently, scientists found that a type of cell called an ionocyte (Montoro et al., 2018) plays a large role in CF. Like many cells, ionocytes are not round; they are oblong and have "projections" emanating from the center. The cell in *Cellverse* is modeled after an ionocyte.

The goal of the game is to explore, diagnose, and select a therapy for a virtual cell. Players begin by reviewing a comic strip "narrative", where they are introduced to the game: they are working with a scientist named Dr. Blob to help determine what type of cystic fibrosis a young patient has, so time is of the essence. The entire game includes a single player segment and a collaborative segment. In this study, we focus on the single player segment on translation.

When the player first enters the virtual cell, they are in a projection of the cell, which is a horn like feature that is a useful entry point as it is a simpler environment with fewer organelles. Players are introduced to FR3ND, a non-player character (NPC) who accompanies the player during the game. FR3ND shows players features of the hand controllers, including how to move forward, how to select organelles and how to call up a virtual clipboard to learn about and collect samples of the organelles, as shown in Figure 1, and how to bring up a dashboard guide, as shown in Figure 2. Once the player completes the tutorial, they start the inquiry, and are given a checklist at the bottom of the dashboard. The first step is to "look for the organelle with translating, bound ribosomes". The second step is to "Check for signs that translation stops early", with two options, either "sampling the CFTR amino acid chain to determine its length" or "sample a codon on the mRNA chain to check for early stop codons". The final step in the checklist is to "verify your evidence". The player then navigates around the cell environment trying to find the organelle that fits this description. If the player takes too long in looking for the organelle, FR3ND offers hints to guide the player to the Rough Endoplasmic Reticulum.

FIGURE 1

Figure 1. FR3ND surrounded by proteins in the microscale view

Once at the RER, FR3ND suggests moving closer to the organelle and selecting the "B" button on the hand controller to move into a "nano view". In the nano view, the player can see the RER with more detail, and at a scale where the tRNA (dark blue) are buzzing around, while long green chains of mRNA wind through the ribosomes (red). As in the micro level view, players can move around the environment and select organelles and identify them using the clipboard, and sample them by pressing the button at the bottom of the clipboard. Sampling them adds information to the "evidence" column (see upper right hand side of the dashboard in Figure 2). FR3ND suggests to the player that they are small enough to go through the translocation channels on the RER. If they do, they will see that the other side is not as busy, and only contains amino acid chains assembled by translating mRNA chains into amino acids at the ribosomes using tRNA.

FIGURE 2

Figure 2. Dashboard showing checklist of tasks completed at bottom and evidence collected at right in the nanoscale view

A few studies have documented the evolution of the game over time. In a study of content knowledge, researchers found that students gain more accurate overall mental models of the cell, but did not improve in their understanding of size and scale or in their factual knowledge of cell organelles (Thompson et al, 2020). Players with the VR HMD in the collaborative version of the game used mostly ego-centered references while describing the environment to their collaborators outside of VR. In their conversations, their descriptions of the VR environment suggested that they experienced the virtual cell with a high degree of presence, which is the feeling that they are in the environment. Players also gained an awareness of the cell as a three-dimensional entity (Uz-Bilgin, Anteneh & Thompson, 2020). This inquiry will focus on optimizing the complex visualization of the virtual cell by exploring the types of information where learning is enhanced through the stereoscopic viewpoint. This study will also apply aspects of the multimedia theory of learning to determine how best to scaffold those learning experiences.

In this study, we used two Oculus Rift S headsets. The Oculus Rift S has a resolution of 2560x1400 (1280x1440 per eye) and a refresh rate of 80 Hz. The input includes 6 degrees of freedom using inside-out tracking through 5 built-in cameras. Players used 2nd generation Oculus Touch motion tracked controllers. The participant in the stereoscopic version of the game wore the headset on their head and held the game controllers in both hands and played the game while sitting on a chair. The participants in the non- stereoscopic version of the game sat in the same chair and held the hand controllers, but had the HMD placed on the desk in front of them.

3.2. Participants

Players were recruited through emails and posts on listservs and social media sites at a university; the sample included adults (18+) who were students, staff, and friends of students at the university.

Participants were 51 adults with different nationalities including American, Chinese, Estonian, French, Indian, Israelis, Italian, Spanish, Taiwanese, Ukrainians, and Zambian. Participants were of different ethnicities and mostly identified as Asian and White. Their age ranged from 18 to 65 years old, with a median age of 26. Participants' other demographic information including gender, VR experience, game play experience, and ethnicity are shown below (see Table 1). 71% reported that they had one VR experience, and 37% identified as game players.

TABLE 1

Table 1: Demographic Information of Participants

Fourteen participants (28%) had taken a Biology class this year while 15 participants (30%) reported had not taken a Biology course in at least 8 years. As shown in Table 2, 27 participants (44%) felt either somewhat (14/ 28%) or extremely (13/26%) not confident in their biology knowledge. Fourteen participants (28%) reported feeling somewhat confident in their biology knowledge, while 1 individual reported being extremely confident. Nine participants (18%) reported a "neutral" stance about their biology confidence.

TABLE 2

Table 2: Biology Experience of Participants

All procedures in the study were classified as exempt in human subjects review (COUHES 17090995534).

3.3. Research Design

The study is a mixed method concurrent design (Creswell & Clark, 2017). The quantitative segment is a randomized comparison study focused on the media effect of a stereoscopic viewpoint on learning where two groups each experienced a different version of the game. The qualitative segment includes analysis of participants' drawings, observation notes of participants during game play, and review of interview responses throughout the session. Drawings are used as a way to understand participants' mental models of cells. Mental models are defined as "ideas that represent the construction of imaginative understanding and visualization in the minds of students used to describe phenomena" (Majid & Suyono, 2018, p. 244). Mental models can be externalized as conceptual models through drawings of a system (Dauer et al., 2019). The focus of the analysis for this paper is on the pre and post assessments and pre and post drawings of the process of transcription.

3.4 Instruments

Two data collection instruments were used in the study: an online questionnaire, and open-ended pages for participants' drawings of cell processes.

The online questionnaire gathered information about biology knowledge (pre and post gameplay), background questions about game experience, VR experience, biology experience, self-described interest in science (pre-gameplay only), and participant demographics. The biology knowledge assessment consisted of ten questions that were developed based on existing standardized measures (e.g., National Assessment of Educational Progress), about cell organelle function, location, and cell processes. Preliminary inspection of the item correlation matrix indicated that more than one latent trait may be driving patterns of response. The reliability of the raw scores was initially low (Cronbach's alpha = 0.40) but was acceptable after correcting for negative correlations (i.e., allowing for a second latent factor), with a final Cronbach's alpha of 0.70 for the post assessment and 0.60 for the pre assessment.

Participants' drawings were gathered on two standard sized sheets of paper. The pre-game play cell drawing page stated "Please take three minutes to draw a cell. Draw what you remember, label anything you can, and we will talk through it at the end." at the top of the paper, and "Sketch anything you remember about the process of translation. How are ribosomes, ER, mRNA, and amino acids involved?". The post-game play drawing page stated "Please take five minutes to sketch a cell *based on what you saw in the game*. Draw what you remember, label anything you can, and we will talk through it at the end" at the top half of the paper and "Now sketch your ideas about the process of translation *based on what you saw in the game*. How are ribosomes, ER, mRNA, and amino acids involved?" at the bottom half of the paper.

3.4. Data collection/ Procedures

Data collection for the study occurred between February and March of 2020. We gathered data from 60 participants and received completed data from 51 participants. Participants were randomly assigned to one of two versions of the 3D game: a stereoscopic version where the player wore an HMD and a non-stereoscopic version where the player did not wear an HMD. Each session took 60 minutes, and began with a consent form and study description (3 minutes), drawing a cell and drawing the process of translation (5 minutes), completing a pre assessment (5 minutes), playing the game (30 minutes), completing performance tests (5 minutes), completing the post assessment (5 minutes), then drawing another depiction of a cell and of translation, and answering questions in a post interview (7 minutes). Participants completed a delayed post assessment one week after their game session. Participants who completed all parts of the study were emailed a $20.00 egift card to Amazon.com.

3.5 Measures

3.5.1 Outcomes

Outcomes measure change in understanding of cell biology and processes covered by the game. Players responses to the cell biology assessment questions were used to surface student understanding as a latent variable, which we label "theta", using the ltm package in the R statistical program (Rizourplous, 2006). Descriptive analyses of the item correlations and reliability indicated that two latent traits were present in the data, thus we used a multidimensional item response theory (MIRT) model to explain the patterns of response, as shown in table 3 (Immekus, Snyder, & Ralston, 2019). The two outcome variables (PRIMARY and SECONDARY) measure the primary and secondary ability levels indexed by the latent traits (thetas) in our MIRT model. Across the sample, the primary ability score had a mean of -.07 (sd=.70), while the secondary ability score had a mean of .10 (sd=.50).

3.5.2 Predictor variables

Our main predictor variable, TREAT, was a dichotomous variable coded as 1 if the student was in the stereoscopic (VR) condition or 0 if the student was in the non-stereoscopic screen-based condition. This variable was entered as a main effect in our model.

3.5.3 Covariates

Three exogenous variables were used to adjust for past student experience with video games (1-4) and VR technology (0-2), and the region of the world in which they were primarily educated (North America, Europe, South Asian, Middle East, Africa, South East Asia), which may result in variability in prior exposure to these concepts. These measures were treated as random variables (intercepts) in the final fitted model.

TABLE 3

Table 3. Model comparison between one (fit) and two (fit2) dimensional latent trait models.

We found no substantive difference in pre assessment scores between the stereoscopic and non-stereoscopic subgroups, indicating that our randomization was successful. We note in Table 3 that students in the stereoscopic group had higher primary and secondary ability scores, compared to their peers in the non-stereoscopic condition. In addition, students in the VR group had slightly higher experience playing video games, on average, compared to their peers, though they had equivalent experience using virtual reality.

TABLE 4

Table 4. Descriptive statistics of the pre and post assessment questions and participant demographics.

3.6 Data Analytic Plan

Based on our sample size and subsequent desire to reduce over-fitting our statistical models to the data, we opted to fit a Bayesian multilevel model (Gelman et al., 2020). Bayesian modeling and inference generally proceeds along 4 steps: 1) specify a joint distribution of the outcomes (likelihood & marginal priors), 2) draw from a posterior distribution via

Markov Chain Monte Carlo sampling, 3) evaluate model fit, and 4) analyze manipulations of predictors and visualize the results (Muth, Oravecz, & Gabry, 2018). We executed these steps primarily via the rstanarm package in the R statistical programing language (Goodrich et al., 2020). Our hypothesized data generating model was:

EQUATION 1

In this model, the student's ability scores (primary or secondary) are normally distributed, with a given mean and standard deviation. The mean is a linear function of a fixed intercept, random intercepts for past game experience (j), virtual reality experience (k), and the geographic region from which the student hails (l), and a slope coefficient for the treatment condition into which the student was randomly assigned. The prior distribution on the slope coefficient is weakly regularizing (Gelman et al., 2020). We answer our research question by investigating the magnitude, direction, and precision of the posterior prediction of the estimated slope coefficient in models of the primary and secondary measures of knowledge. For example, if the mean of the posterior for the estimated slope coefficient was positive and larger than the standard deviation of the posterior, we would state that students who used VR demonstrated higher knowledge, adjusting for prior game and VR experience, as well as region of origin.

4. Results

4.1 RQ1 How does a stereoscopic view of a virtual cell help individuals learn about the cellular environment compared to a non-stereoscopic view?

Our measures of the cellular environment included the results from the biology assessment questions. For the biology assessment questions, we standardized our estimates of the primary and secondary ability scores (mean=0, sd=1) to help aid in the interpretation of our findings. Our models converged successfully with no divergent transitions post warm-up and strong evidence of adequate mixing via visual examination of trace-plots. All effective sample sizes were greater than 10% of the total sample size, Monte Carlo standard errors were all less than 10% of the posterior standard deviations, and all R-hat statistics were less than 1.1. A visual inspection of our predicted probability plot showed that the model-estimated trends were adequately similar to the observed trend from the data. Our final findings were robust when fitted with both narrow and wide priors. Our statistical code and tables of the main findings can be found in the online supplemental material.

FIGURE 3

Figure 3. Posterior estimates of VR (TREAT) and Screen (Intercept) usage on the primary (Model 1) and secondary (Model 2) measures of understanding, adjusting for game playing experience, prior VR usage, and region. Dark lines are 50% credible intervals, thin lines are 89% credible intervals, open points are posterior medians.

We note in Figure 3 that the posterior median estimate of the primary ability level of the non-stereoscopic group (Intercept, model 1) was approximately -0.3 (sd=0.50), while the stereoscopic group was estimated to be approximately 0.3 units higher (sd=0.40). Similarly, the posterior median estimate of the secondary ability level of the screen-based group (Intercept, model 2) was approximately 0 (sd=0.5) whereas the VR group was estimated to be approximately 0.3 units higher (sd=0.5). Overall, given the data, the stereoscopic group consistently demonstrated higher ability scores as compared to the non-stereoscopic group, though the estimates are statistically noisy.

We highlight this trend between VR usage and ability level in Figure 4. Adjusting for past game experience, VR experience, and region of primary education, we note that students in the stereoscopic (with HMD) condition have primary ability scores that are about 0.3 sd units higher than their peers in the non-stereoscopic (no HMD) version (left), and secondary ability scores that were approximately 0.3 sd units higher than their screen-based peers (right). These measures of difference are somewhat uncertain, due in part to the small sample size of this pilot study. Overall, these results suggest that a stereoscopic view was associated with higher learning gains on the cell biology content knowledge assessment than a non-stereoscopic view.

FIGURE 4

Figure 4. Model estimated differences between Screen and VR groups on the primary (left) and secondary (right) measures of performance, adjusting for game playing experience, prior VR usage, and region. Dark lines are the model average effect, light lines are estimates of uncertainty based on draws form the posterior distribution.

4.2 RQ2: How does a stereoscopic view impact players' understanding of the cellular processes compared to a non-stereoscopic view?

Participants' drawings of cells and translation are our evidence of players' understanding of the complex cellular environment. As mentioned previously, participants were asked to draw translation and to include mRNA, tRNA, ribosomes, and amino acids in their drawing. The presence of any of those items in the drawing received a score of 1, and scores of all items were combined for an overall drawing score. Two examples of pre and post translation drawings are included as Figure 5 and Figure 6.

FIGURE 5

Figure 5. Example of two stereoscopic participants' drawings of the process of translation pre-game play (top) and post-game play (bottom)

FIGURE 6

Figure 6. Example of two stereoscopic participants' drawings of the process of translation pre-game play (top) and post-game play (bottom)

We examined participants' drawings of the translation process that they learned in the game, including their representation of the roles specific organelles in the translation process. As shown in Figure 7, participants' translation scores increased after the game play (Pre-test *MD*:3, *CI*:2.04~3.39, Post-test *MD*:5, *CI*:3.98~5.13). When stereoscopy and non- stereoscopy conditions are compared, participants in the stereoscopy condition have a tendency to have higher translation scores (Yes *MD*: 4, *CI*: 3.37~4.59 and No *MD*:3, *CI*: 2.60~4.03), as shown in Figure 8.

FIGURE 7

Figure 7. Pre-test and post test scores of translation drawings

FIGURE 8

Figure 8. Scores of translation drawings disaggregated by condition (stereoscopic versus non-stereoscopic)

In general, both stereoscopic and non-stereoscopic drawings improved between pre and post, showing more specific and accurate details from the game, such as the translocation channels, or the ribosome in two segments. When disaggregated by treatment type, players who experienced the stereoscopic view of the game had higher gains in scores of their diagrams of the translation process than players who had the non-stereoscopic view.

In addition to completing their post assessments, players also provided feedback at the end of the game. Players in both groups found the game engaging, but densely packed with information. Many players noted that the game included a lot of reading during the game and suggested providing some of the information beforehand as a way to frame the game. Players that had not taken biology in a few years appreciated having the clipboard as a prompt for prior knowledge. Players with a stronger baseline knowledge of biology were better able to connect their prior knowledge to their game experience, which became apparent in their drawings and their interviews. Players who did not have a strong background knowledge understood that the cell was more complex than they remembered but did not show any gains in their understanding of cells or of the process of translation. When asked to describe the process of translation, one participant wrote "no clue" on the pre assessment and "still no clue, but there was a lot going on" on the post assessment.

5. Discussion

Results of this study provide evidence of the types of learning goals that are well suited to the stereoscopic viewpoint of a VR game. *Cellverse)* converts a normally two-dimensional exercise of understanding cell structure and organelle function into a three-dimensional task of finding clues in the cell to diagnose the cell. Task-based learning is supported by stereoscopic viewpoints in understanding physical tasks such as assembling circuits (Parmar et al. 2016), understanding geographical landscapes (Taştı & Avci, 2020), and origami folding (Dan & Reiner, 2018), a finding also echoed in reviews of VR studies (Jensen & Konradsen, 2020). The task of exploring and diagnosing the cell may have been facilitated by the stereoscopic than the non-stereoscopic format of the game because players were able to develop a mental model of the cell environment by moving through it. This feeling of "being there", also known as presence, was helpful in learning the process of mammalian protein expression (Makransky, Terkildsen & Mayer, 2019), and also appears to have a positive impact on players' understanding of the process of translation in the *Cellverse* game. Stereoscopic viewpoints may support learning associated with tasks by reducing cognitive load of individuals learning those tasks (Mayer, 2014).

All players made gains in their understanding of the cellular environment as evidenced through their drawings, yet the players in the stereoscopic view made greater gains in their representing of the process of translation. The first person viewpoint helped players see translation as a spatial process, which is concurrent with research on VR (Jensen & Konradsen, 2020; Uz-Bilgin, Anteneh & Thompson, 2020). All players witnessed the process of translation and were able to interact with the environment. However, in the stereoscopic view, players did not just witness the process, but were actively engaged in moving around the environment, watching the ribosomes come together and move apart, and even moving through the translocation channel. As with other VR experiences, being situated in the cell was very engaging to learners (Makranksy, Andreasen, Baceviciute & Mayer, 2020). Many of the participants had limited experience with VR, so this interest may have been influenced by the novelty of the medium (Merchant et al., 2014). However, the biology assessment questions showed higher gains when controlling for past VR experience. Together the drawings and the assessment suggest that learning due to the VR setup went beyond novelty.

Both groups showed gains from pre to post, however players in the stereoscopic condition had stronger gains than students in the non-stereoscopic view after game experience, VR experience, and region of primary education were taken into account. The low level of reliability in the results could be a result of the small sample size. The results also may suggest that neither treatment was particularly helpful in having players learn facts about the cell. CTML suggests that individuals have a limited capacity to process and store information (Moreno & Mayer, 2003). While learners are able to gain a holistic understanding of the cell environment as a result of the environment, they are not able to learn specific factual details. Other studies have also concluded that topics that require memorization of facts are not well suited to stereoscopic environments (Makransky, Terkildsend & Mayer, 2019; Remmele, Weires & Martens, 2015; Roettl & Terlutter, 2018, Jensen & Konradsen, 2020. This finding is useful in determining the learning goals for a VR experience, and for understanding how to support players' learning during a VR experience. The technique of providing "pretraining" may help players learn and retain information about the cell (Meyer, Omdahl & Makransky, 2019; Moreno & Mayer, 2009), and a few of our participants noted in the post interviews that having a reference sheet would have helped remind them of current knowledge and build upon existing knowledge during the game. Offloading cognition by having the clipboard as a reference during the game was regarded as helpful, however, all participants noted that there was a lot of text, especially in the dashboard view (see figures 2 and 3). This amount of information was challenging for players to digest while experiencing the dynamic view of the virtual cell. Future versions of the game may also move the information into a pretraining document, or segment the information by introducing different information at different points in the game, as Parong & Mayer (2017) did in their study of a VR learning game. Another way of segmenting information is to distribute the information between players in a non-VR setting, a strategy we use in the collaborative version of the game (Thompson et al., 2018).

6. Conclusion

Immersive visualizations afforded by VR have the potential to help learners develop better understanding of complex topics. While VR hardware has decreased in cost, production costs for VR remain high, thus it is useful to know what types of learning and topics are well suited to VR. The results of this study indicate that the stereoscopic view afforded by the head-mounted display is associated with a positive effect on learning in VR.

This study also highlights the challenge of measuring knowledge gains in virtual reality. Traditionally, knowledge gains are measured in pre- and post-assessments and concept inventories. However, the types of learning best achieved in VR are less specific in factual knowledge and more holistic. While measures of factual knowledge showed lower reliability, evidence from drawings suggested that players with the stereoscopic view were associated with higher gains in their understanding of the process of translation between pre and post drawings. As designers and educators tap into the affordances of VR for learning, new ways of assessing learning gains will also have to be developed and validated.

While this study focuses on the role of stereoscopic visualization in biology, learning how to leverage technology in support of learning has the potential to inform future curriculum other domains. We believe that VR is one of many possible technological tools, and different tools provide different affordances for learning (Liou et al., 2017). As VR becomes more feasible as a learning tool, additional research will be needed to understand whether VR is effective and if so, how it is best integrated into educational experiences.

Acknowledgement

We thank all of the study participants, the student programmers and designers. We acknowledge funding from the Oculus Foundation and MIT Integrated Learning Initiative (MITili).

References

3D Visualisation Aesthetics Lab UNSW, 2017. Journey to the Centre of the Cell - 3DVAL.
Discovery Institute, 2009. Journey Inside The Cell.

Afifi, M. 2017, May 23. Basics of stereoscopic imaging in virtual and augmented reality systems. Retrieved November 15, 2020, from https://medium.com/@mahmoudnafifi/basics-of-stereoscopic-imaging-6f69a7916cfd

Allen Institute 2019. “New VR Game Lets You Learn about Cells by Tossing Them.” 19 Feb. 2019, alleninstitute.org/what-we-do/cell-science/news-press/articles/new-vr-game-lets-you-learn-about-cells-tossing-them.

Bailenson, J., 2018. *Experience on demand: What virtual reality is, how it works, and what it can do*. WW Norton & Company.

Castaneda, L., Cechony, A., & Swanson, T. 2017. *Implications of Virtual Reality in Applied Educational Settings*. https://www.foundry10.org/research/the-implications-of-vr-in-applied-educational-settings

Concannon, B. J., Esmail, S., & Roduta Roberts, M. 2019. Head-Mounted Display Virtual Reality in Post-Secondary Education and Skill Training: A Systematic Review. In Frontiers in Education (Vol. 4, p. 80). Frontiers.

Cystic Fibrosis Foundation. n.d.. Types of CFTR Mutations. Retrieved July 15, 2017, from https://www.cff.org/What-is-CF/Genetics/Types-of-CFTR-Mutations/

Dan, A., & Reiner, M. (2018). Reduced mental load in learning a motor visual task with virtual 3D method. *Journal of Computer Assisted Learning*, *34*(1), 84-93.

Dauer, J. T., Bergan-Roller, H. E., King, G. P., Kjose, M. K., Galt, N. J., & Helikar, T. (2019). Changes in students’ mental models from computational modeling of gene regulatory networks. *International Journal of STEM Education*. https://doi.org/10.1186/s40594-019-0193-0

Ferdig, R., Blank, J., Kratcoski, A., & Clements, R. (2015). Using stereoscopy to teach complex biological concepts. *Advances in physiology education*, *39*(3), 205-208.

Gauthier, A., & Jenkinson, J. (2017). Serious game leverages productive negativity to facilitate conceptual change in undergraduate molecular biology: A mixed-methods randomized controlled trial. *International Journal of Game-Based Learning (IJGBL)*, *7*(2), 20-34.

Gelman, A., Hill, J., & Vehtari, A. (2020). Regression and other stories. Cambridge University Press.

Goodrich, B., Gabry, J., Ali, I. and Brilleman, S., 2018. rstanarm: Bayesian applied regression modeling via Stan. R package version 2.17. 4. *Online< http://mc-stan. org.*.

Huk, T., Steinke, M., & Floto, C. (2010). The educational value of visual cues and 3D-representational format in a computer animation under restricted and realistic conditions. *Instructional Science*, *38*(5), 455-469.

Ibarra-Herrera, C.C., Carrizosa, A., Yunes-Rojas, J.A. *et al*. Design of an app based on gamification and storytelling as a tool for biology courses. *Int J Interact Des Manuf* **13,** 1271–1282 (2019). https://doi-org.libproxy.mit.edu/10.1007/s12008-019-00600-8

Immekus, J. C., K. Snyder, and P. A. Ralston. (2019). "Multidimensional Item Response Theory for Factor Structure Assessment in Educational Psychology Research." Frontiers in Education 4: 45. doi:10.3389/feduc.2019.00045.

Innovative Genomics Institute, n.d. CRISPR-VR: Explore Sickle Cell and Genome Editing. Innovative Genomics Institute (IGI). URL https://innovativegenomics.org/vr/ (accessed 5.27.21).

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

Johnson-Glenberg, M. C. (2018). Immersive VR and education: Embodied design principles that include gesture and hand controls. *Frontiers in Robotics and AI*, *5*, 81.

Kaplan-Rakowski, R. (2019). The Effect of Stereoscopic Three-Dimensional Images on Vocabulary Learning. *Contemporary Educational Technology*, 2019, 10(4), 324-337 DOI: https://doi.org/10.30935/cet.634172

Labster, n.d. Labster | 200+ virtual labs for universities and high schools [WWW Document]. URL https://www.labster.com/ (accessed 5.27.21).

Liou, H. H., Yang, S. J., Chen, S. Y., & Tarng, W. (2017). The influences of the 2D image-based augmented reality and virtual reality on student learning. *Journal of Educational Technology & Society*, *20*(3), 110-121.

Loup-Escande, E., Jamet, E., Ragot, M., Erhel, S., & Michinov, N. (2017). Effects of stereoscopic display on learning and user experience in an educational virtual environment. *International Journal of Human–Computer Interaction*, *33*(2), 115-122.

Majid, A., & Suyono, S. (2018, September). Misconception Analysis Based On Students Mental Model In Atom Structure Materials. In *Seminar Nasional Kimia-National Seminar on Chemistry (SNK 2018)* (pp. 244-247). Atlantis Press. https://doi.org/10.2991/snk-18.2018.53

Makransky, G., Andreasen, N. K, Baceviciute S., & Mayer R. M. 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*. DOI: 10.1037/edu0000473

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

Mayer, R. E. (2014). Incorporating motivation into multimedia learning. *Learning and Instruction*, *29*, 171-173

Mayer, R., and Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52.

Merchant, Z., Goetz, E.T., Cifuentes, L., Keeney-Kennicutt, W. and Davis, T.J., 2014. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, *70*, pp.29-40.

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*, 103603. Article link

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. https://doi.org/10.1037//0022-0663.94.3.598.

Montoro, D.T., Haber, A.L., Biton, M. et al. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* 560, 319–324 (2018).

Muth, C., Oravecz, Z., & Gabry, J. (2018). User-friendly Bayesian regression modeling: A tutorial with rstanarm and shinystan. Quantitative Methods for Psychology, 14(2), 99-119.

Onyesolu, M.O. and Eze, F.U., 2011. Understanding virtual reality technology: advances and applications. *Adv. Comput. Sci. Eng*, pp.53-70.

Parmar, D., Bertrand, J., Babu, S. V., Madathil, K., Zelaya, M., Wang, T., ... & Frady, K. (2016). A comparative evaluation of viewing metaphors on psychophysical skills education in an interactive virtual environment. *Virtual Reality*, *20*(3), 141-157.

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

Pellas, N., Dengel, A., & Christopoulos, A. (2020). A scoping review of Immersive Virtual Reality in STEM Education. *IEEE Transactions on Learning Technologies*. *13*(4), 748-761.

Remmele, M., Weiers, K., & Martens, A. (2015). Stereoscopic 3D's impact on constructing spatial hands-on representations. *Computers & Education*, *85*, 74-83. Article link

Richards, D. and 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*, pp.157-171.

Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, 17(5), 1-25.

Roettl, J., & Terlutter, R. (2018). The same video game in 2D, 3D or virtual reality–How does technology impact game evaluation and brand placements? *PloS one*, *13*(7), e0200724. Article link

ShareCare Reality Lab, n.d. VR [WWW Document]. Sharecare. URL https://www.sharecare.com/pages/vr (accessed 5.27.21).

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, *3*, 74.

Steve Bambury VR, 2017. Cellscape VR Biology Guided Tour.
Sweller, J. (1999). Instructional design. In *Australian educational review*.

Sweller, J., Ayres, P. and Kalyuga, S., 2011. Measuring cognitive load. In *Cognitive load theory* (pp. 71-85). Springer, New York, NY.

Taştı, M. B., & Avcı, Ü. (2020). Examination of using monoscopic three-dimensional (M3D) and stereoscopic three-dimensional (S3D) animation on students. *Education and Information Technologies*, 1-26.

Thompson, M., Olivas-Holguin, H., Wang, A., Fan, J., Pan, K., Vargas, D., Gerr, J. (2018). Rules, roles, and resources: Strategies to promote collaboration in virtual reality contexts. *CHI 2018 Novel Interaction Techniques for Collaboration in VR Workshop*.

Thompson, M., Uz Bilgin, C., Wang, A., Anteneh, M., Roy, D., Klopfer, E. (2020). Influence of virtual reality on high school students' conceptions of cells. *Journal of Universal Computer Science, 26*(8).

Uz-Bilgin, C., Thompson, M. & Anteneh, M. Exploring How Role and Background Influence Gameplay through Analysis of Spatial Dialogue in Collaborative Problem-Solving Games. *Journal of Science Education and Technology 29*, 813–826 (2020). https://doi.org/10.1007/s10956-020-09861-5

Virtual Heroes, 2020. Mission Biotech | Virtual Heroes | ARA | Serious Games | Game-based Training | VR Learning – ARA - Serious Games | Game-based Training | VR Learning | Modeling and Simulation | Augmented Reality | Educational Simulations. URL https://www.virtualheroes.com/portfolio/education/mission-biotech/ (accessed 5.27.21).

Yoon, S., Anderson, E., Lin, J., & Elinich, K. (2017). How augmented reality enables conceptual understanding of challenging science content. *Journal of Educational Technology & Society*, *20*(1), 156-168.

zSpace, n.d. Cells and Mitochondria [WWW Document]. URL https://zspace.com/edu/subjects/life-science/cells-and-viruses/7353 (accessed 5.27.21a).

zSpace, n.d. Collection: Cells [WWW Document]. URL https://zspace.com/edu/subjects/life-science/cells-and-viruses/7686 (accessed 5.27.21b).

zSpace, n.d. Cystic Fibrosis [WWW Document]. URL https://zspace.com/edu/subjects/life-science/anatomy/5475 (accessed 5.27.21c).