Authenticity, Interactivity, and Collaboration in Virtual Reality Games: Best Practices and Lessons Learned

Annie Wang[1], Meredith Thompson[1*], Cigdem Uz-Bilgin[1], Eric Klopfer[1]

[1]The Education Arcade, Comparative Media Studies & Writing, MIT, Cambridge, MA, USA

*** Correspondence:**
Corresponding Author
meredith@mit.edu

Keywords: virtual reality1, immersive virtual environments2, learning games3, biology4, collaborations5, STEM education6

Abstract

Virtual reality has become an increasingly important topic in the field of education research, going from a tool of interest to a tool of practice. In this paper, we document and summarize the studies associated with our four-year design project, Collaborative Learning Environments in Virtual Reality (CLEVR). Our goal is to share the lessons we gleaned from the design and development of the game so that others may learn from our experiences as they are designing, developing, and testing VR for learning. We translate "lessons learned" from our user studies into "best practices" when developing authentic, interactive, and collaborative experiences in VR. We learned that authentic representations can enhance learning in virtual environments but come at a cost of increased time and resources in development. Interactive experiences can motivate learning and enable users to understand spatial relationships in ways that two dimensional representations cannot. Collaboration in VR can be used to alleviate some of the cognitive load inherent in VR environments, and VR can serve as a context for collaborative problem solving with the appropriate distribution of roles and resources. The paper concludes with a summation of best practices intended to inform future VR designers and researchers.

1 Introduction

Virtual reality can bring new perspectives to classroom learning. In the last 20 years, immersive VR has become an increasingly common topic in the field of education research (Hew and Cheung, 2010; Merchant et al., 2012; Ahn et al., 2017) as the technology becomes more viable for classroom use (Castaneda, Cechony, & Swanson, 2020), prompting educators to explore how to leverage VR for educational purposes. The accessibility of VR has increased as the overall cost of VR has decreased in recent years (Korbey, 2017). However, more research is necessary to move beyond the "novelty" of VR (Merchant et al, 2014) and understand its full potential in K-12 learning. Increasing access has supported a growth in the number of studies of VR and learning; however, additional research is needed on longer term learning outcomes (Jensen & Konradsen, 2018; Pellas, Dengel & Christoupoulos, 2020), especially with projects that extend beyond one-time implementations of VR experiences (Merchant et al, 2014; Hamilton, McKechnie & Edgerton, 2020). This paper addresses the need for extended studies of VR projects by documenting a set of studies on one multi-year design project, Collaborative Learning Environments in Virtual Reality, or CLEVR. The CLEVR team designed, developed, and deployed *Cellverse,* a game designed to help introductory high school students learn cellular biology. In this article, we discuss lessons learned in our design, development, testing, and analysis so designers and educators can learn from them.

At the beginning of the CLEVR development process, we described our intentions for the game in an article titled "Authenticity, Interactivity, and Collaboration in VR Learning Games" (Thompson et al, 2018). This article outlined our theoretical frameworks for the game, initial design, and planned trajectory. Since the start of the project (2017), *Cellverse* has developed significantly in breadth, depth, and focus. Moreover, we approach game development through a framework of Design-based Research, or DBR (Ameel & Reeves, 2008; Sandoval & Bell, 2004). Ongoing user testing, studies with various types of users, and reviews by subject matter experts have enabled us to collect valuable qualitative and quantitative data that have enhanced our understanding of how to incorporate authenticity, interactivity, and collaboration in VR learning games (Thompson et al, 2018; Wang et al, 2019; Uz-Bilgin & Thompson, 2021 ; Uz-Bilgin, Anteneh, & Thompson, 2020; Uz-Bilgin, Anteneh, & Thompson, 2021; Thompson et al, 2020; Wang, 2020).

We begin this manuscript by defining authenticity, interactivity, and collaboration in context to VR, followed by several theories and frameworks essential to understanding learning in VR, including the Cognitive Theory of Multimedia Learning and the Cognitive Affective Model of Immersive Learning. (VR, in this article, should always be assumed to refer only to *immersive* VR). We then introduce the CLEVR project and *Cellverse*, the game that was ultimately produced from CLEVR research. This is followed with a critical analysis of our *Cellverse* studies between Summer 2017 - Spring 2020, describing both lessons learned and best practices of VR in learning. The recommendations made for best practices arise from both our research results and from our practical experiences creating, testing, facilitating, and studying the game. The manuscript concludes with a discussion of the advantages and challenges for authenticity, interactivity, and collaboration in educational VR.

2 Authenticity, Interactivity, and Collaboration

2.1 Authenticity

Our goal for an authentic game had three levels of authenticity: authenticity of narrative, authenticity of environment, and authenticity of action. Authenticity of narrative is critical for promoting interest and motivation, but is not inherently tied to VR; authenticity of action and environment, conversely are closely tied to VR's affordances. We will briefly discuss all three types of authenticity in this section. Authenticity refers to the ability for VR to produce and render scenarios, experiences, and processes that closely resemble real life (Thompson et al., 2018). Such an affordance is unique to the technology due to its multisensory qualities; VR stimulates the user's sense of sight, sound, and can even include smell and touch. This sensory engagement allows the user to virtually experience environments that may be too distant, expensive, or dangerous to approach otherwise (Bailenson, 2018).

One dimension of authenticity is the *authenticity of environment*. Biology in particular makes for particularly fertile ground for depicting the authenticity of environment through VR. As mentioned above, individuals do not have accurate views of cells, in part because of how cells are depicted in biology education. This need for authentic virtual environments (VEs) may tap into a critical need in K-12 education, particularly within the sciences. Teaching realistic systems requires authenticity in order to ensure that students are able to gain accurate mental models of critical topics (Jacobsen, 2017). The *authenticity of the environment* also helps establish a sense of presence in the virtual environment. Slater and Sanchez-Vives (2017) suggest that presence is related to the user's "place illusion" of the VE and the perception of "plausibility" of interactions. Makransky & Peterson (2020) discuss representational fidelity in realism, smoothness of interaction, and consistency felt by the user in the interactions with the VE, three variables originally proposed by Dalgarno and Lee (2010). In the case of an educational game, *authenticity of environment* was one of our learning goals. The virtual environment of *Cellverse* attempts to represent cells as they exist in nature: three-dimensional, active, and densely packed. Using feedback from subject matter experts, we iteratively redesigned the cellular environment to reflect cutting-edge research on cell internal structure. We also incorporated ongoing research using databases and resources designed for scientists -- for example, the biological quantitative database B1onumb3rs (Wang et al., 2019) -- in order to provide an accurate presentation of the relative size of organelles and the density of organelles and proteins within the cell.

Authenticity of action means that, as much as possible, the actions available to the player within the game reflect actual techniques available to scientists the player takes within the game resemble actions that individuals can do in real life. This closely ties to the definition of situated learning, as the player is able to join a community of practice by “doing what the experts do”. While this is not a requirement for VEs, we prioritized authenticity of action to ensure that our game introduced the capabilities of biologists and did not create or amplify misconceptions about biology.

Finally, *authenticity of narrative* contributes to the users degree of buy-in for the virtual environment. Johnson-Glenberg & Megowan-Romanowicz (2017) found that narrative increased users’ interest in the experience. Users can be primed for social interactions in the VE by watching an engaging conversation between two agents in the virtual world (Daher et al., 2017). Authenticity of narrative ties together the action and environment to create a more powerful learning experience in XR.

The three levels of authenticity complement each other in building a sense of presence and agency in educational games, as depicted in Figure 1.

INSERT FIGURE 1

Figure 1: The Theory of Authenticity in XR (TAX)

2.2 Interactivity

In the context of VR, interactivity is closely tied to *immersion*. Immersion is a function of VR hardware, creating the illusion of physical presence in a non-physical world (Slater & Sanchez-Vives, 2017). Dede et al. (2017) argue that immersion is essential to motivation and learning in VR. Whereas *presence* reflects psychological feeling, immersion is the technology or practical application that creates presence. Slater & Sanchez-Vives (2016) suggest that the more seamless the underlying technology, the more potential for immersion that exists – this is encapsulated in the various levels of technological sophistication, including but not limited to haptic feedback or in the degrees of freedom available to the user. The high level of interactivity possible with virtual reality is recognized as a key affordance that sets VR apart from other technologies, such as film and video (Lindgren & Johnson-Glenberg, 2013; Makransky & Petersen, 2020).

On the other hand, since a user both provides and is provided with information, interactivity requires a combination of hardware and careful design to be successfully implemented. Interactivity occurs when a user affects virtual objects or avatars, prompting changes in the VE. Interactivity embedded within a VR experience enables the user to communicate with the VE, by using buttons, manipulations, gestures, or other modalities to produce feedback from their virtual surroundings. Embodiment, such as gesture and movement, has been linked to positive learning outcomes in improved learning in physics (Johnson-Glenberg & Megowan-Romanowicz, 2017) developing a better understanding of electricity (Johnson-Glenberg, 2017), helping students learn laboratory skills (Lindgren et al, 2016); helping medical students learn anatomy (Jang et al, 2017), as well as helping scientists prepare samples for microscopy (Leinen et al, 2015) and testing compounds for new pharmaceutical drugs (Yuan, Chan & Hu, 2017).

Interaction may have multiple definitions depending on context, and is defined in this manuscript as the level of responsiveness the VE provides to a user. Johnson-Glenburg (2017) outlines three constructs that contribute to the degree of embodiment: sensori-motor engagement, gestural congruence, and sense of immersion. Sensorimotor engagement can offload cognition to enable the user to learn more complex topics (Weisberg & Newcombe, 2017). Gestures that match the learning objectives can reinforce learning and facilitate the initial uptake of ideas (Pouw et al, 2014). Immersion also supports embodiment; more sophisticated virtual environments can give the user more options for agency within the environment, which makes the user an active part of the virtual environment. From macroscopic interactions such as moving across the virtual space and controlling what is in their field of view to microscopic interactions such as waving your hand or looking in a mirror, the user's actions prompt a virtual response. This action and response cycle draws the user into the virtual experience (Wang, 2020). As the type of manipulation and the type of response can vary, designers must consider how interaction can enhance learning goals (Bailenson, 2018; Johnson-Glenburg, 2017).

2.3 Collaboration

Virtual environments provide new venues for collaboration between individuals. Collaborative problem-solving is considered essential for the future of work and is deemed a vital "21st-century learning" skill (Fiore, 2017). Collaborative, goal-oriented activities create what Johnson and Johnson (1989) call "positive interdependence" among team members, wherein individuals in a group rely on each other's strengths to achieve their goal (Laal, 2013). Previous research has identified principles of collaborative learning that may be integrated into VR experiences including interdependence, thoughtful group formation, individual accountability, and attention to social skill development (Cuseo, 1997; Lee, 2009).

Establishing rules and developing distinct roles for users are both useful ways of encouraging collaboration within VR environments (Uz-Bilgin et al., 2020). Earlier studies have described the benefits of establishing collaborative roles in VR. Jensen and Konradsen (2018) used games to create rules for social interaction and roles for individuals in virtual problem-based activities. Defined roles also helped visitors engage with a VR museum exhibit experience on an aircraft carrier (Zhou et al., 2016). Finally, middle school students in the EvoRoom VR environment benefited from clear roles in gathering and sharing information with their peers (Lui and Slotta, 2014). VR may also encourage individuals uncomfortable with leadership to be proactive and assume roles with more responsibility. Slater et al. (2000) found that users participating in a VR activity using a head-mounted display were more likely to willingly take on a leadership role than when they were involved in with the same activity within an in-person group.

3 Learning Theories and Theoretical Framework

Here, we will define the learning theories and frameworks we draw upon in this project and in this article. The learning theories that are particularly salient to virtual reality are situated learning, spatial learning, embodied learning, and embedded learning.

3.1 Learning Theories: Situated, Spatial, Embodied, Embedded

Situated learning theory suggests that optimal learning occurs when the learner is able to experience the activities and environment in as authentic of a context as possible. Understanding the context and activities of the area being studied allows students to experience "legitimate peripheral participation" and inducts the learner into a community of practice (Wenger, 1998). Part of this induction is in the expert making their thinking visible to the learner. For example, in an introductory research course for freshmen, students become cognitive apprentices when the professor makes implicit ideas about research explicit (Thompson, Pastorino, Lee & Lipton, 2015). The ability to transfer skills learned in a VE to real-life experiences is a result of the similarities between the learning environment and the actual environment (Dede et al., 2017), or when students are able to "do what the experts do" by emulating real-life scientific techniques. VR experiences involving situated learning are popular within the sciences, particularly virtual laboratories that enable users to iteratively practice essential skills in "lab-like" VEs without requiring real-world resources (Chiu et al., 2015; Lindgren et al., 2016).

Spatial learning refers to both learning to navigate a real or artificial (VR-rendered) space. Spatial learning is helpful for individuals in navigating their everyday lives, but has also been identified as an important skill in learning STEM. For example, size and scale are important concepts to understand within STEM learning domains, but can be challenging to conceptualize for learners (Jones et al., 2003). Because VR allows the opportunity for users to directly manipulate virtual objects, it may be used to enhance learners' perception of relative size and scale.

The theory of *embodied learning* states that connecting physical action to learning objectives creates deeper learning (Kiefer and Trumpp, 2012). Through embodied learning, knowledge is cemented as memory through the body's repeated interactions with the physical environment (Lindgren & Johnson-Glenberg, 2013). Previous research suggests that the multimodal nature of VR may make it optimal for facilitating information retrieval in 3D spaces, thus strengthening users' mental models (Dede et. al, 2017; Johnson-Glenberg, 2018). Providing a 3D virtual environment for users to experience abstract concepts may produce more effective learning than in 2D models, in biology (Tan & Waugh, 2014), physics (Johnson-Glenberg, 2018) and chemistry (Lindgren et al, 2016; Chiu, DeJaehjer & Caho, 2015).

While embodied learning supports user cognition through physical movement, *embedded learning* supports user cognition through features that are part of the virtual environment. A signpost on a highway is a real-world example of embedding learning into an environment; instead of being forced to memorize highway numbers while navigating a road, a driver can simply recall their location by glancing at the words on a passing sign. Embedding cognitive activity within the environment frees up mental capacity by storing extraneous (non-essential) information into accessible actions or tools (Weisberg & Newcombe, 2017), instead of overloading limited mental resources (Pouw et al, 2014).

3.2 Cognitive Theory of Multimedia Learning

Introduced by Mayer (1998), the Cognitive Theory of Multimedia Learning (CTML) describes two main sensory channels for memory processing: visual (sight) and auditory (sound). These two input methods are processed separately by the mind and do not overlap with each other. According to CTML, the brain processes information through a series of steps: filtering information, organizing it, integrating it into previous knowledge or schema, and finally processing it into long-term memory storage. According to CTML, when cognitive processing exceeds a user's mental capacity, "essential overload is experienced, inhibiting learning" (Meyer et. al 2019).

VR creates higher cognitive load among users compared to other forms of media, which may impede memory recall and memorization (Parmar et al. 2016; Makransky et al. 2019; Roettl & Terlutter 2018). Critics argue that VR produces comparatively poorer learning outcomes because the medium is overwhelming to users (Moreno & Mayer, 2002). These critics point to evidence that VR is a poor medium for imparting declarative or factual, static knowledge (Mayer, 2019). Findings from ongoing cognitive science research in VR learning seem congruent with CTML; when users doing VR-based tasks were compared with users working on the same task on a non-immersive platform, the VR users reported higher enjoyment but revealed lower levels of gained declarative knowledge (Parong &

Mayer, 2018; Makransky, Terkildsen, & Mayer, 2019). Mayer et al. (2019) suggest that the heavier cognitive load inherent to VR prevents users from processing incoming facts into long-term memory, thus preventing effective learning.

The rich sensory experience afforded by VR comes at a cost. Designers should be aware that cognitive load informs all design principles of VR, so learning designers must temper their VR experience to avoid overwhelming users with excessive cognitive load. That being said, this manuscript argues against the suggestion that VR consistently makes for poor learning experiences simply because it produces high cognitive load. As Mayer, Omdahl, & Makransky (2019) have argued, VR may not be a good medium for transferring declarative (fact-based) knowledge -- however, the numerous examples listed in this document suggest that it is useful for other forms of learning. Declarative knowledge, while central to the American education system, is not the only type of knowledge essential for 21st-century learners. We must also note that not all cognitive load is "bad," as cognitive load is inherent to any learning material (Paas et al, 2003). Cognitive load can be *essential* to the learning process, *generative* based on what the learner is learning, or *extraneous* and thus hinder learning. In designing with the information-rich and sensory-stimulating technology of VR, designers need to be purposeful in maximizing essential and generative cognitive load and minimizing extraneous cognitive load (Mayer, 2020).

3.3 The Cognitive Affective Model of Immersive Learning (CAMIL)

More recently, Makransky & Petersen (2021) proposed a model of learning designed for immersive learning: the Cognitive Affective Model of Immersive Learning (CAMIL). Much of the early research on learning in VR focused on whether or not VR helped learning compared to other media. A central premise in CAMIL is that media and method interact; maximizing the effectiveness of an immersive learning experience requires an understanding of the affordances of that medium and how to tap into (manipulate) those affordances. Presence and agency are the two main psychological affordances of the medium of VR (Johnson-Glenberg, 2018; Makransky & Petersen, 2021) and so "instructional methods that enrich learning through higher presence or agency will specifically increase learning through immersive technology" (Makransky & Petersen, 2021, p. 6). Presence and agency are linked to the level of immersion, the degrees of interactivity (control factors), and the degree of representational fidelity of the experience. CAMIL states that presence and agency impact six factors that influence learning: interest, intrinsic motivation, self-efficacy, embodiment, cognitive load, and self-regulation. Calibrating presence and agency in VR environments impacts each of those learning outcomes.

In order to better understand how VR can best be harnessed for learning, we must understand how researchers can participate in the VR design process and how scholarly research on VR-based learning can inform ongoing development of educational games and simulations. Our suggestions are discussed through the lens of the CLEVR Project and its resulting game *Cellverse*.

4 Collaborative Learning in Virtual Reality (CLEVR) and Cellverse

CLEVR is a research collaboration between the MIT Education Arcade and the MIT Game Lab. It is funded by Oculus Education and has been developed by an interdisciplinary team of researchers, game designers, programmers, and artists.

Cellverse, the game produced through the CLEVR Project, has been developed as both a single and two-player game that explores concepts of cell biology, particularly cell organelles and cell processes. Our team used the Next Generation Science Standards (NGSS, 2013) as a baseline for *Cellverse's* educational content to orient learning goals for high school-age student users. The software was built using Unity 3D and is supported by the Oculus Rift system. We used a design-based research methodology (Collins et al, 2004), where we conducted tests and interviews with users and experts throughout the design process.

We sought a narrative that would focus on the DNA-to-RNA-to-protein process. We met with biologists to explore different diseases that could support the game narrative. We chose cystic fibrosis (CF) because it was the first genetic disease that can be treated through FDA-approved gene therapy. This helped support our goal of *authenticity of action*, as one of our initial goals for the game was to end by creating a specific gene sequence to fix the faulty sequence causing CF in the patient. CF is caused by disruptions at one of a few points in the process of protein synthesis; each of these disruptions is caused by different genetic sequences and is best addressed with a targeted treatment. In the single-player narrative of *Cellverse*, the player is a student intern using a remote-controlled microbot to navigate through a human lung cell. The cell, like its human host, has CF; the players must find clues in the cell structure, organelles, and processes to diagnose and recommend treatment that is suited to the class of CF that matches the clues. The player's goal is to explore the cell's internal structure and observe the cellular process of translation to figure out which form of CF is affecting the cell in order to provide the unnamed patient with the most effective medical treatment. In the game, players view the cell using a machine of microscopic size, a "microbot", and an even smaller "nanobot", rather than shrinking down to the cellular scale. Here we maintain *authenticity of action* as microscopic and nanoscopic robots are already being developed, and using those as a probe of the living cell is more realistic than making a person smaller (e.g., Venugopian, et al., 2020).

VR remains a novel experience for many people - to reduce the risk of extraneous cognitive load overwhelming users right away, players begin with a tutorial that places their microbot into a remote, sparsely populated area of the lung cell. This was an intentional decision on the part of the designers, as we wanted the players to focus on game mechanics during the initial part of the game rather than become distracted or overwhelmed by their surroundings. The structure of the cell that causes CF, an ionocyte, has projections that contain fewer organelles. The feature of this type of cell lends itself well to the goal of segmenting the introduction to the game while maintaining *authenticity of environment*. Players are immediately greeted by a non-player character (NPC) named FR3ND, a robot who teaches the player the basics of head movement (e.g. that they have a 360 degree view of the cell), selection (e.g. of organelles), and navigation (e.g. point to your destination and press "A"). The tutorial gradually guides the player from their starting location into the densely populated main "body" of the cell, where the tutorial ends and the game begins.

To accomplish the task of identifying forms of CF, the player is equipped with a number of tools and informational tips that allow them to shift between different levels of scale (microscopic and nanoscopic scale), read descriptions of selected organelles, collect virtual samples, review different classes of CF, and determine whether the clues meet the description of the class of CF. The "clipboard," for example, is a tool that is attached to the player's virtual left hand. Players can select organelles around them with their right hand, and a description of the organelle's functions will instantly appear on the clipboard. They may also use the clipboard to sample organelles in order to collect clues using a "Sample" button at the bottom of the clipboard, as shown in Figure 2. Finally, players are capable of shifting their viewing robot between microscale (the microbot's original scale) and nanoscale (a smaller nanobot) in order to view particles of different sizes. By approaching the rough endoplasmic reticulum (ER), they can activate a nanobot that enables them to "shrink" to nanoscale and observe macromolecules (e.g. RNA and amino acids) that would not be visible at the microscopic (or micro) scale.

INSERT FIGURE 2 HERE

Fig. 2: Screenshot of clipboard tool showing a user sampling glutamine, an amino acid.

The virtual environment of the multiplayer narrative of CLEVR *Cellverse* is identical to the above description, but otherwise varies drastically. The multiplayer game is a cross-platform experience involving two players and offers greater challenge and complexity than the single-player version. One player, the "Explorer," wears the HMD and is tasked with navigating through the cell's virtual environment and viewing cell functions up close. Unlike the single-player experience, the Explorer is not provided with as much textual information on organelles or cellular disorders - access to this information is granted to the second player, or the "Navigator." The Navigator is equipped with a touchscreen tablet interface that provides a limited "bird's-eye" view of the same cellular environment. The Explorer and Navigator have to combine their complementary roles and resources in order to accomplish their task, creating positive interdependence between the users (Thompson et al, 2019).

Cellverse has been in development since Summer of 2017 and has undergone numerous iterations, which we have discussed in other publications (Wang et al, 2019; Uz Bilgin & Thompson, 2020; Uz Bilgin, Anteneh, & Thompson, 2020; Thompson et al, 2020; Wang, 2020). In this paper, we look across all of the studies and papers to synthesize our experiences as lessons learned and best practices," in designing learning games that include authenticity, interactivity, and collaboration.

5 Studies and Methods

Our *Cellverse* user study encompasses 3+ years and many user tests, each with their own goals, target user groups, collaborators, and data collection methods. One aspect of each study design was to include a diverse body of participants, with a wide range of ages, backgrounds, and previous access to VR. Each of these articles we have published about *Cellverse* draws from four main studies we conducted during the project, which we describe below. We have organized the research questions and findings for each article into a table format and have indicated the data source for each article to one of the four examples below.

1. User testing (2017-2018) – Using a design-based research framework, we ran ongoing user testing with subject matter experts (11), adult volunteers (35), and teachers (8), 54 people in total, between 2017 and 2019. These user tests occurred once every 8 weeks and included individuals that were invited to test different games and simulations being developed for educators. During the user test, individuals answered pre and post surveys, created cell drawings before and after using *Cellverse*, and were interviewed at the end. Data were also gathered from observation notes gathered while the users used *Cellverse*.
2. Qualitative Studies (2018) – In the summer and fall of 2018, we conducted two qualitative studies of the collaborative version of the game. Participants completed pre and post surveys, created cell drawings before and after they played the game, and were interviewed at the end of the game. All participants were videotaped. Video recordings were transcribed and analyzed using qualitative coding and epistemic network analysis. These studies included a study of 8 pairs of STEM teachers, and a study of 4 pairs of K-12 students (2 from middle school, 2 from high school) and 4 pairs of high school graduates in a biotechnology workforce development program.
3. Quantitative study (2019) – In the fall of 2019, we conducted a quantitative study at two urban high schools near the Boston area. One hundred and fifty-three students participated in the study. All students completed a pre and post survey about their knowledge of cellular biology and CF. The post survey questions also included scales about presence, mental workload, and spatial skills. All students drew pictures of a cell before and after they played the game. They were given 25 minutes in the VE, where they were told to figure out what was wrong with the cell. Data were analyzed using descriptive statistics and inferential statistics.

4. Quantitative study (2020) – In the spring of 2020, we conducted a quantitative study of adults. Sixty-one people participated in the study. Participants were randomly assigned to one of two interventions: playing *Cellverse* in the head-mounted display (HMD) with hand controllers or playing the game viewing the game on a flatscreen with hand controllers. All participants completed a pre and post survey about their knowledge of cellular biology. Post survey questions also included scales about presence, mental workload, and spatial skills. All participants drew pictures of a cell and of the process of translation before and after they experienced the game. After participants were set up and given 5 minutes to explore, they were asked to find three organelles in the cell, and the researcher timed how long it took them to find those organelles. They were given 25 minutes in the virtual environment, where they were told to figure out what was wrong with the cell. After they were finished, they were asked to find the same organelles, and the time it took to find them was recorded. Participants engaged in a short interview at the end of the session where they described their drawings of a cell and of the process of translation and also provided feedback about the game.

These four data collection activities are the foundation for the research studies and experiences described in this paper. We link the data collection activities, research questions, and findings in Table 1.

INSERT TABLE 1 HERE

6 Lessons Learned and Best Practices

In designing and studying the game, we gained additional insight into the opportunities and challenges of creating an authentic, interactive, and collaborative game. Below we describe how we incorporated those three features into the design, what we learned, and the resulting best practices for design.

6.1 Authenticity

While creating *Cellverse*, we aimed for authenticity in all aspects of the game. The forms of authenticity that emerged through *Cellverse* can be explained by the theory of authenticity in XR (TAX): authenticity of narrative; authenticity of in-universe actions; and authenticity of environment. We will briefly discuss how these three types of authenticity are reflected in the game.

Cellverse is a game designed for teaching cellular biology, so we began design by prioritizing the creation of an authentic environment by creating an accurate representation of the cell. In many biology textbooks and learning materials, cells are portrayed in a flat, schematic-type format: static, generic, round, one-dimensional, and mostly empty (Thompson et al., 2020). Furthermore, relative density and positions of organelles are generally not illustrated, resulting in representations that have only one or two (as opposed to a more realistic number of) mitochondria or ribosomes. When designing the game, we prioritized *authenticity of environment* by doing extensive research on the environment inside cells. We consulted professors, scientists, and doctors who were subject matter experts (SMEs) SMEs for advice on where to find this type of information. They pointed the team to resources such as B1ONUMB3RS and the Cystic Fibrosis Foundation's website. The commitment to authenticity came at a cost; midway through the project a study linked a brand new cell to CF (Montoro et al., 2018). These ionocytes had some major differences between regular cells, and we dedicated extra time to recreate the cell environment in response to these new findings.

Early in the design process, we made the decision to adhere to *authenticity of action*, aligning the actions in the game and the virtual world with existing capabilities in science. As an educational game, we aimed to introduce students to the types of manipulations scientists could actually use on cellular environments. We chose CF because it was the first disease with an FDA approved genetic therapy (Office of the Commissioner, 2019). This means that players could do what scientists do – maintaining *authenticity of action*. Players could identify the class of CF, the associated genetic sequence, and customize a therapy for the patient. Organelles are identified using a microscopic technique called SLAMMing, where wavelengths of light interacting with organelles of different density appear in different colors (You, Tu & Cheney, 2018). We were able to maintain *authenticity of action* with ongoing connections SMEs including professors, scientists, and researchers who regularly work with human cells. These SMEs were influential in the game design and informed player actions within the narrative, including but not limited to traversing the environment, viewing important cellular processes, and collecting important context clues in order to develop a plan of action for treating the cell with real-life medical techniques. SME feedback also helped us shape player experience with dynamic cellular processes, as CF is a genetic disorder that is intrinsically tied to errors in the protein synthesis process. The malformation of CFTR, the protein responsible for CF, and the resulting 5 classes of CF demonstrate breakdowns at different parts of protein synthesis. Using a real-life disorder to demonstrate such a microscopic function maintains *authenticity of narrative*, and provides an authentic example of the importance of protein synthesis, as well as an authentic answer to students' perennial question of "why do we need to know this?".

After playing *Cellverse*, a majority of user participants remarked that the cellular environment was more complex, dynamic, and densely packed than they expected (Thompson et al, 2020). Viewing and exploring the cellular environment improved players' conceptions of cells; participants' drawings after they completed the game were more complex and included organelles that were not in their initial drawings (Thompson et al, 2020; Uz-Bilgin & Thompson et al, 2021). The appearance of new organelles suggests that playing the game triggered players' memories about organelles they had learned about in the past. Furthermore, players experienced a change in the way they conceived of a cellular environment. Players remarked that the game changed cells from a topic they read about and passively observed to something that they engaged with as an active learning experience. Players made stronger connections between the organelles and their functions in the cell in the process of translation, which was a focus of the game. Players drawings of the process of translation improved in their representations of ribosomes, their documentation of the process of RNA to amino acid chains, and their representation of the endoplasmic reticulum (Thompson et al, 2020).

6.1.1 Authenticity: Lessons Learned

Prioritizing high authenticity can result in learning gains, but those gains are contingent on the learning goal and on the attributes of the learners. The type of learning that VR lends itself best to is not always the easiest to measure, which prompted us to further evaluate what it means to "improve" in knowledge of cellular biology. Traditional measures of improvement are simpler to collect and analyze and often focus on factual knowledge. While learning gains in factual knowledge were small, players did gain a holistic understanding of cells. Players' drawings after *Cellverse* indicated that playing the VR based game was associated with more authentic mental models of cells among novice, intermediate, and experts in biology (Thompson et al, 2020; Uz Bilgin et al, 2020). Numerous mentions of the experience being "hands-on" and "interactive" imply that playing *Cellverse* produced embodied learning for some players, connecting users' physical actions to learning objectives. Statistically significant improvements in recall of organelles and processes in drawings and interview responses suggest that organelle labels and information that were integrated into the environment (e.g., the clipboard tool) facilitated the recall process, which could be evidence of embedded cognition in practice (Pouw et al, 2014). However, our evidence suggests that background knowledge appears to be critical to improving learning outcomes. More background knowledge of biology and more experience with VR were associated with increased improvement in cell and translation drawings from pre- to post-game (Thompson et al, 2020).

Authenticity of environment impacts the degree of cognitive load experienced by the user in the information-rich VR environment. Extraneous cognitive load can impede learning (Mayer, 2019), but not all cognitive load is necessarily negative; essential cognitive load and generative cognitive load is created naturally by learning material and can be conducive to learning. One strategy in alleviating cognitive load in virtual environments is to segment information, rather than provide all the information at once (Mayer & Moreno 2002; Rey et al, 2019). In *Cellverse*, players started in a more sparsely populated part of a cell called a "projection". That way users could become familiar with the game controls before being immersed in the center of the densely packed cell. While all players gained a sense of the dynamic cell environment, players with more background knowledge were able to make connections between their existing ideas and the objects in the game than players who had less background knowledge. A certain level of background knowledge in biology transformed an overwhelming environment prompting extraneous processing into an opportunity to connect ideas into a more authentic context, a form of essential processing. Furthermore, knowledgeable players were able to channel the ideas about cell environments and processes into a better idea of the process of translation, which could be a form of generative processing. Future research should explore the specific level of knowledge needed to leverage cognitive load in VR learning environments.

6.1.2 Authenticity: Best Practices

1. *Establish scope and focus authenticity directly on learning goals.* When designing learning experiences in VR, we recommend focusing authenticity on aspects of the game that are directly related to the learning objectives. VR is time intensive to develop, therefore it is helpful to have a clear vision of the learning goals and refine that vision as the project progresses.
2. *Consult subject matter experts to inform design and guide learning goals, as well as increase action-based authenticity of the experience.* We drew on many sources of knowledge in the game design but found insights and feedback from subject matter experts especially helpful. SMEs provided insights from the cutting edge of biology knowledge as well as foundational ideas about how to connect the game to student-appropriate learning objectives. They also allowed us to promote authenticity of action, allowing students to take on tasks that real-life scientists would do.
3. *Consider how levels of authenticity, particularly action-based and environmental authenticity, impact cognitive load.* At the time of publication of this article, VR remains a novelty for many people. First time users can be overwhelmed by a complex VR environment, and although authenticity lends itself to increased realism, it can also create high cognitive load. Designers should consider what aspects of their experience *should* be authentic, particularly in context to learning goals. There is the possibility of complexity itself being the learning goal. In our studies, VR enabled even novices to experience complex models through embodied learning, however the level of authenticity is linked to the level of cognitive load, which we discuss further below. Starting players in a less dense environment gave users time to learn the game controls and options. To support a range of learners, designers may consider scaffolding highly complex experiences through embedded cognition within the environment, allowing "layers" of complexity that can be turned on and off, and prompting learners' conceptual frameworks through pretraining (Makransky et al., 2019).

6.2 Interactivity

Cellverse has also provided insight into the role of interactivity in learning. Interactivity is closely linked to *presence*, or the feeling of being in the virtual environment. Presence enables learners to interact more deeply with the content being learned. Rather than passively viewing an experience, the learner actively navigates through and interacts with the virtual world. Interactivity thus builds upon presence when a user is in a well-designed immersive virtual environment (Makransky & Petersen, 2021).

Many cell biology lessons are similar to vocabulary lessons, where associating abstract shapes with anthropomorphic definitions (e.g. the nucleus being the "brain" of the cell or the mitochondria as the "powerhouse" of the cell). These types of lessons exemplify passive learning that can be gleaned from reading textbooks and watching videos. However, the format and types of information available to learners have expanded beyond passive learning. We were curious about how learners perceived interaction within *Cellverse* compared to other materials they use or have used in K-12 biology classes. In our preliminary surveys, we asked study participants how they preferred to learn new biology concepts. While some reported they would ask a teacher or parent or consult a textbook, participants overwhelmingly preferred turning to the internet and other virtual sources -- virtual resources mentioned by name included Khan Academy, Wikipedia, and YouTube. During the Fall 2020 studies, we interviewed participants (n = 113) and inquired as to how *Cellverse* compared to other ways they learned biology. Moreover, nearly all users felt that they were "present" in the VR environment (n = 111). Over 40% of the students (n = 47) mentioned that the game enabled them to interact directly with the material (Thompson et al, 2020). The level of interactivity in the virtual reality game resonated with the learners in part because it closely matched their own media-rich personal learning experiences. One student commented that "it was cool to look around the cell and be in there because you don't normally get the opportunity to visualize it". About a third (39/113) of the students described the experience as "visual". Some students explained how the visually rich VR experience was a better match for their preferred learning strategy as a "visual learner." Of course, not all of the students viewed the added interactivity as a benefit. Some students described feeling lost or disoriented. One individual mentioned that "the movement was a little weird because you had to point everywhere".

The interactive elements integrated into *Cellverse* are designed to give players both structure and agency. Players have the agency to explore the environment, select organelles, learn more about them by opening up the clipboard, and collecting samples of possible evidence for the type of cystic fibrosis in the game. Even as the player is crafting their own tour, *Cellverse* also has features that focus players' attention on specific parts of the game. For example, the NPC FR3ND guides players through the initial tutorial. If the player hesitates for an extended time period, FR3ND provides hints for them to "look for the organelle with translating bound ribosomes" (the rough endoplasmic reticulum), to "press B to launch a nanobot" (to see the process of translation at a nanoscopic level), and so on. The game also includes built-in checklists that automatically collect the evidence the player gathers during gameplay. These scaffolds combine the embodied learning enabled by learning through the game and are supported by embedded cognition within the environment (Pouw et al, 2014). This embodied "hands-on" experience with the cell made the complex and abstract environment more understandable, even for introductory biology students (Uz Bilgin et al, 2020).

6.2.1 Interactivity: Lessons Learned

One very common theme throughout our user interviews was that *Cellverse* was more "hands-on" than other cell biology learning experiences that the participants had experienced. Indeed, experiencing the game helped students engage with an abstract concept by directly interacting with or manipulating the virtual environment. (As mentioned earlier, a high level of presence may have also helped to contribute to this "hands-on" feeling ubiquitous among users.) The ability to interact with the microscopic VE through a microbot and nanobot may have enabled users to make more authentic mental models of cells. This was evidenced by how organelle frequencies and the level of complexity increased in VR users' cell drawings after playing *Cellverse* (Fig. 2). These drawings provide visual evidence of users' shifting mental models, particularly as many users came to understand that their previous schematic-type image of cells was inaccurate in terms of scale and density. Users were not simply passively observing their surroundings while in *Cellverse*, but understanding their relative positions within the virtual environment.

Navigating the cell as a 3D environment enabled users to gain a sense of placement and space. While navigation could be difficult at first, particularly for users who were new to VR or 3D video games, users' knowledge of how to navigate their virtual space gradually improved as they spent more time playing *Cellverse*. We investigated this phenomenon in the "route knowledge task" of the data collection process, where users were asked by presiding researchers to navigate to specific areas or organelles in the cell as quickly as possible. The route knowledge task was performed twice -- once near the beginning of the user's session, and one near the end. Regardless of their levels of experience with VR or with biology, users were consistently faster at completing the route knowledge task near the end. This suggests to us that educational topics that require a strong sense of spatiality -- understanding how objects in a 3D space relate to one another -- can be well expressed within VR. Spatial abilities are increasingly important for K-12 learners to develop, as they can play a crucial role in development of professional skills, for instance in learning surgical techniques (Abe et al., 2018). In terms of VR experiences, spatial abilities may also be closely associated with the development of presence (Coxon et al., 2016). In this respect, designers should note that different levels of spatial abilities might result in different levels of spatial presence among different users. This is the reason why all learners do not equally benefit from the same VR technology. In our single-player game study, we noted an association between attention and spatial presence (Uz Bilgin & Thompson, 2021), which is an important consideration for designers who aim for a strong sense of spatial presence in their VE.

Interestingly, we found an association between attention and visual spatial imagery ability. Designers need to take into consideration that spatial abilities might have an effect on how people pay attention to the stimulus in VR. Although triggering learners' attention might be accomplished with highly-immersive technologies, designers should recall that different levels of spatial abilities take part in attention allocation. Educational VR designers should give users with low spatial ability enough support to engage effectively in the game and provide users with high spatial ability enough challenge to sustain a high level of engagement in the experience. VR training can have a significant positive effect on visuospatial orientation ability of people with disabilities, both in VR and real-life environments (de la Torre-Luque et al, 2017). Enhancing spatial ability using VR environments may help learners transfer these abilities into non virtual situations.

We also noted that players' heightened interest in a learning topic is directly associated with increased attention to the VE. This association does not seem to be impacted by other factors -- in this single player version of *Cellverse*, results showed that prior content knowledge, experience with VR, and gaming experience did not impact learners' formation of spatial presence in VR (Uz-Bilgin & Thompson 2021, under review). Domain interest and spatial abilities led to higher levels of attention, which resulted in a stronger feeling of presence in the game. Designers should focus on how to trigger learners' attention in VR using spatial and interactive elements, particularly elements that directly correlate with the relevant learning topic.

6.2.2 Interactivity: Best Practices

1. *Consider how VR allows for new engagement methods with learning topics.* Biology has changed, as has modern media -- teaching methods, conversely, have remained stagnant. VR's affordances may allow students to engage with core academic subjects from new perspectives.

2. *Interactive modalities are useful, as they allow direct manipulation of the learning at hand.* Designers should aim to create a rich environment that actively engages the learner in exploration and critical thinking. Interactive elements can also streamline knowledge gain by embedding non-essential knowledge into the virtual environment. Embedded learning is at play when users can efficiently access such information in order to refresh their memories or apply previous knowledge to the task (Pouw et al, 2014).
3. *Effective immersion within the VR environment requires linking interaction to learning goals.* In other words, the interactive elements in the game or simulation should not be extraneous, but directly relevant to the learning goal.
4. *Certain topics that require a strong contemplation of spatiality -- or where objects on a 3D plane are in relation to each other -- can be effectively expressed within VR.*
5. *The perception of VR environments as 3-dimensional enables learners to practice and develop spatial abilities, regardless of previous ability.* Designers should consider how to support and challenge learners with different levels of spatial ability in 3D space, and can thus be used to leverage a stronger understanding of spatiality.

6.3 Collaboration

One of the goals of *Cellverse* was to help players learn and practice collaborative problem solving. Collaborative problem solving is defined as four stages: (1) Exploring and understanding, (2) Representing and formulating, (3) Planning and executing, and (4) Monitoring and reflecting (Fiore et al., 2017). The collaborative version of *Cellverse* includes two users playing at once: the Explorer, who wears the head-mounted display and is immersed in the VR environment, and the Navigator, who observes the same cellular environment via a “bird’s-eye” view on a touchscreen tablet, as shown in Figure 3. We designed the game with cross-platform advantages in mind; the Explorer has a deeper, more detailed view of their surroundings and the Navigator has extensive reference materials about the game. We tailored the information for each role in order to establish positive interdependence, a concept describing situations where collaboration is necessary to complete a task. Data collection and analysis for collaborative *Cellverse* differed from our procedures in the single-player experience. As we were designing the VE and gameplay, we collected data from video recordings, transcripts, observation notes, and interviews of participants. We found that player-to-player dialogue during the game was an excellent resource for tracking collaboration. We analyzed the data both qualitatively, looking for themes in what the partners discussed, and quantitatively, using epistemic network analysis to identify patterns in how the partners’ discussion progressed. In addition to collaboration, we also conducted joint studies of the change in players’ biology knowledge and in the players’ development of spatial presence as a result of playing the game, and looked at a range of ages and biology backgrounds, including middle school, high school, students in a workforce development program, university students, and adults.

Midway through the project (Fall 2019), the development trajectory of *Cellverse* changed from cross-platform multiplayer to VR-exclusive single-player. This section discusses *Cellverse* as it existed between Summer 2018 - Summer 2019, as well as that period’s corresponding studies (see Table 1 for details).

INSERT FIGURE 3 HERE

Figure 3: Two players using the collaborative version of *Cellverse*

6.3.1 Establishing Collaboration

A central question in our inquiries around collaborative *Cellverse* was "are the partners working together"? Our early design stages revealed limited collaboration between partners. Navigator users reported that they possessed all of the information they needed to solve the game without the Explorer's input (Wang et al., 2019). In addition to the rules and the roles we had built into the design, we reallocated resources so that both players had information that was both unique to their role and was critical to game play. In our most recent collaborative studies, we noticed that players moved through a pattern of interactions that mirrored collaborative problem solving. Players began by orienting themselves with the environment, establishing a shared language of the environment, finding clues and determining whether those clues were relevant, and finally making a decision about the diagnosis and recommended treatment (Thompson et al., 2020). We found that teams went through many cycles of finding and examining clues that could be grouped into these four stages. The initial stages featured two-way communication (stages 1-4), then included orientation (stages 5-9), then moved towards orienting and discussing (stages 10-17), and ended with discussion (stages 18-21). Furthermore, partners continuously used biology terms throughout their conversations. Patterns of collaborative problem solving were similar across groups of different ages and levels of biology knowledge. Furthermore, Navigator and Explorer dialogue was continuous throughout the game, suggesting that the information exchange between the two players was useful in progressing through the game (Thompson et al, 2021).

6.3.2 Influence of Roles on Spatial Awareness

Partner dialogues also offered us clues regarding their mutual understandings of their environment. This mutual understanding reveals how Navigators and Explorers developed a sense of spatial presence in the game (Uz-Bilgin et al., 2020). "Spatial knowledge" in the context of *Cellverse* includes players' knowledge of the location of organelles, their ability to find different ways to navigate through the cell, and their ability to find and recognize clues to diagnose the cell and finish the game. Our studies suggest that the player's role and corresponding viewpoint affect how the player communicates their ideas about the virtual environment to their partner. For example, within the two-player cross-platform experience of *Cellverse*, the Navigator's global view allowed them to understand the perspective of the Explorer ("Where are you?"), and enabled the Navigator to direct the Explorer to different areas of the cell by sharing spatial information with the Explorer. This capability offloads the Explorer's task of where to search next to the Navigator, effectively reducing the Explorer's mental workload. ("Move toward the yellow round nucleus."). The HMD gives the Explorer a close-up view of the environment and a strong sense of presence from a first-person perspective. This perspective prompts players to use ego-centered references ("I'm by the Golgi Body, where do I go next?") as they describe the environment. The way the Explorers described themselves as "in" the environment through language indicates that the user feels that they are "there", an indicator of presence. We also noted that prior knowledge of cell biology affects spatial ability. Learners with high prior knowledge describe fewer instances of "spatial unawareness" ("I don't know where I am", "I'm lost") while collaborating with their partners in *Cellverse*. Mental awareness of location and surroundings were all affected by users' level of background knowledge about cell biology.

6.3.3 Collaboration: Lessons Learned

One core goal of collaborative *Cellverse* was creating positive interdependence between the Explorer and the Navigator, ensuring that both parties contribute equally to the problem-solving process. Earlier versions of collaborative *Cellverse* were problematic in that the Navigator had enough information to complete the entire game by themselves. In one of these studies, the Navigators in a small playtest session (N=4) both stated in post-interviews that they "did not need the Explorer to solve the challenge" (Thompson et al., 2018). This was corroborated by data from other researchers on the team, who noted that the Navigators they observed took the lead in each session and appeared to be in control of gameplay. In other words, creating balance of information between partners was not a straightforward task and required careful design and redesign.

We addressed the lack of collaboration by reducing the amount of information available to the Navigator, requiring additional interaction between the partners. Once positive interdependence was established, we began studying interactions between users in greater detail. We examined dialogue between the Navigator (on tablet) and the Explorer (in HMD) between 8 pairs of players, four pairs from a middle/high school and four pairs from a biotechnology workforce development program. Although background knowledge did affect game experience, we also found that the collaborative problem solving process was similar even between groups that had different levels of cell biology knowledge (Thompson & Uz-Bilgin, 2021). Despite the discrepancy of knowledge, pairs' processes of approaching the problem were very similar, suggesting to us that collaboration could be developed through educational VR regardless of a users' previous level of experience with a topic.

Partners' similarities in approaching the collaborative version of *Cellverse* may have been intrinsic to the game's design, as *Cellverse* has a narrative that might encourage a very specific approach to gameplay. However, because comparisons between different iterations of *Cellverse* are needed to confirm such a claim, we plan to explore this possibility in future studies.

6.3.4 Collaboration: Best Practices

1. *When creating a collaborative VR experience, balance of information is critical.* Allowing players to have equal footing in sharing and contributing not only makes gameplay more interactive, but also more enjoyable for all participants. Thus, designers developing collaborative VR must be careful when dividing information among roles, and focus on promoting interdependence among players so that they must depend on each other's knowledge to produce the best results.
2. *Learning through collaborative problem-solving can be useful for learners of all backgrounds and levels of knowledge.* Our observations of players of varying backgrounds suggest that diverse learners can learn and practice collaborative problem solving through a single game.
3. *Dialogue between partners makes thinking "visible,"* or audible through dialogue. Single-player VR games do not instinctively lend themselves to communication, but involving multiple players naturally encourages users to voice their ongoing thoughts, as players discuss how they want to approach the game. This is useful for researchers interested in studying users' perceptions of the game.
4. *Splitting roles can distribute cognition between players and thus lower cognitive load for each individual player.* Although we did not study this systematically, we noticed that users with low levels of biology knowledge in the collaborative game were less likely to report feeling "overwhelmed" than users with low levels of biology knowledge in the single-player game. In the single-player game, the user had to assimilate the information about the environment and formulate their next step. Splitting roles allowed players to tackle challenging problems together - because of this, pair play required less external guidance than the single-player game.

7 Summary of Best Practices

For authenticity:

1. *Establish scope and focus authenticity directly on learning goals.*
2. *Bring in subject matter experts to inform design and guide learning goals.*

3. *The level of complexity should be directly linked to learning objectives to manage players' cognitive load.*
4. *Striving for authenticity of environment and authenticity of action within the XR environment can leverage the affordances XR provides in presence and agency.*
5. *Authenticity of narrative* can both motivate users to try the game and provide an opportunity to learn the topic in the game.

For interactivity:

1. *Designing within VR provides learners with a high level of interactivity with the topic, enabling embodied learning.*
2. *Interactive modalities are useful, as they allow direct manipulation of the learning.*
3. *Effective immersion within the VR environment requires linking interaction to learning goals.*
4. *Certain topics that require a strong contemplation of spatiality -- or where objects on a 3D plane are in relation to each other -- can be effectively expressed within VR.*
5. *The perception of VR environments as 3-dimensional enables learners to practice and develop spatial abilities, regardless of previous ability.*

For collaboration:

1. *When creating a collaborative VR experience, balance of information is critical.*
2. *Learning through collaborative problem-solving can be useful for learners of all backgrounds and levels of knowledge.*
3. *Collaboration makes thinking "visible", enabling the study of and reflection upon collaborative problem solving.*
4. *Splitting roles, particularly in a graphically intense experience like Cellverse, appears to distribute cognition between players and thus lower cognitive load.*

8 Conclusion

By summarizing the last few years of *Cellverse*'s development through the lenses of authenticity, interactivity, and collaboration, we have been able to reflect upon the trajectory of a long-term project and its numerous implications for designing and developing VR for learning. We have also gained a more well-rounded understanding of the affordances and drawbacks of VR as a technology that can benefit the future of learning. Through different studies and physical settings, we note that a clear understanding of the subject matter, particularly critical frameworks or models, allowed users to gain the most benefit from a VR experience. Authenticity allows for a more accurate mental model of the learning, but comes at a cost of increased cognitive load. As a result, the level of complexity in the experience should be directly linked to the learning goals. Interactivity enables users to apply their knowledge and utilize their virtual environment through learning. Finally, collaboration in VR offers opportunities for users to connect, interact, and disseminate information with each other in a shared VE. The opportunity to build a shared understanding of a situation and work together to solve problems are critical skills in a workforce that continues to become more interdisciplinary and virtual.

Researchers must consider how VR can bolster learning and how VR tools can be used within educational contexts (Dalgarno et al, 2011). Designing effective VR-based learning experiences lies at the nexus of theories and frameworks within the domains of education, game design, and cognitive science. More than anything, such design requires balance. Designers should also be aware that creating VR requires constant rebalancing of game design and information, particularly when supplemented with feedback from users and SMEs. In sharing these findings, we aim to offer a thoughtful insight into the best practices of educational VR in both harnessing and tempering its affordances. We hope that the future educators, researchers, and designers interested in or already working with immersive VR will find our summaries useful.

As of 2021, VR remains costly to create and implement. Streamlining the development process is critical for any educational project with limited time or resources. To address this concern, contemporary and future designers and educators may find value in reviewing and implementing our "lessons learned" and "best practices." We must also keep in mind that access to technologies like VR remains inequitable across regions and school systems. In order to develop thoughtful and inclusive VR experiences that appeal to a wide audience, we emphasize the importance of testing with users from diverse backgrounds. Embedding the feedback of diverse voices within the initial design promotes a more inclusive experience by the end of the development process. There are many challenges remaining for VR in the near future, but our experiences suggest that VR is a useful research tool that can allow for increased learner engagement and collaboration within an immersive virtual environment.

9 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

10 Ethics Statement

The studies involving human participants were reviewed and approved by the Committee on the Use of Human Experimental Subjects (COUHES) at MIT. The participants provided their written informed consent to participate in this study.

11 Author Contributions

Annie Wang and Meredith Thompson collaborated to conceptualize the paper, create the outline. They divided up the sections of the paper between and each wrote different sections of the paper. Cigdem Uz-Bilgin contributed the sections on spatial presence and collaboration and spatial skills. Eric Klopfer edited and provided comments on paper drafts before final submission.

12 Funding

CLEVR was sponsored by the Oculus Foundation, the Jamil World Education Lab (J-WEL) Education Grants, and the MIT Integrated Learning Initiative (MITili) Grant.

13 Acknowledgments

We acknowledge the contributions of the CLEVR game designers Dan Roy, lead programmer Philip Tan, program coordinators Rik Eberhardt and Judy Perry, the subject matter experts we spoke to, the science teachers who implemented the game in their classrooms, the MIT faculty, undergraduates, and graduate students who worked on the project, and our study participants.

13.1 References

Ahn, J., Asbell-Clarke, J., Berland, M., Chase, C., Enyedy, N., Fusco, J., Gardner, S., Grover, S., Halverson, E., Jona, K., Chad Lane, H., Martin, W., Mercier, E., Moher, T., Ogan, A., Pinkard, N., Polman, J., Roschelle, J., Schank, P., … Worsley Editors, M. (2017). *Cyberlearning Community Report | October 2017 a Cyberlearning Community Report: The State of Cyberlearning and the Future of Learning With Technology* (pp. 12–16). SRI International. http://circlcenter.org/wp-content/uploads/2017/07/CyberlearningCommunityReport2017.pdf

Amiel, T., & Reeves, T. (2008). Design-Based Research and Educational Technology. *Educational Technology & Society*. https://doi.org/10.1590/S0325-00752011000100012

Bailenson, J. (2018). *Experience on Demand: What Virtual Reality Is, How It Works, and What It Can Do*. W. W. Norton & Company.

Castaneda, L., Cechony, A., & Swanson, T. (2017). *Implications of Virtual Reality in Applied Educational Settings*. https://www.foundry10.org/research/technology-and-learning

Chiu, J. L., Dejaegher, C. J., & Chao, J. (2015). The effects of augmented virtual science laboratories on middle school students' understanding of gas properties. *Computers and Education*, *85*, 59–73. https://doi.org/10.1016/j.compedu.2015.02.007

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15–42.

Coxon, M., Kelly, N., & Page, S. (2016). Individual differences in virtual reality: Are spatial presence and spatial ability linked? *Virtual Reality*, *20*(4), 203-212. https://doi.org/10.1007/s10055-016-0292-x

Cuseo, J. (1992). Cooperative Learning vs. Small-Group Discussions and Group Projects: The Critical Differences. *Cooperative Learning and College Teaching*, *2*(3), 5–10.

Cystic Fibrosis Foundation, & Md 20814301-951-4422 800-344-4823. (n.d.). *Basics of the CFTR Protein*. Retrieved May 23, 2021, from /Research/Research-Into-the-Disease/Restore-CFTR-Function/Basics-of-the-CFTR-Protein/

Daher, S., Kim, K., Lee, M., Schubert, R., Bruder, G., Bailenson, J., & Welch, G. (2017). Effects of Social Priming on Social Presence with Intelligent Virtual Agents. In J. Beskow, C. Peters, G. Castellano, C. O'Sullivan, I. Leite, & S. Kopp (Eds.), *Intelligent Virtual Agents* (pp. 87–100). Springer International Publishing. https://doi.org/10.1007/978-3-319-67401-8_10

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

Deniş Çeliker, H. (2015). Prospective science teachers' levels of understanding and explanation of animal and plant cells: Draw-write. *Journal of Baltic Science Education*, *14*, 501–512.

Fiore, S. M., Graesser, A., Greiff, S., Griffin, P., Gong, B., Kyllonen, P., Massey, C., O'Neil, H., Pellegrino, J., Rothman, R., Soulé, H., & von Davier, A. (2017). *Collaborative Problem Solving: Considerations for the National Assessment of Educational Progress*. National Center for Education Statistics.

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

Hew, K. F., & Cheung, W. S. (2010). Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: A review of the research. *British Journal of Educational Technology*, *41*(1), 33–55. https://doi.org/10.1111/j.1467-8535.2008.00900.x

Jacobson, J. (2017). Authenticity in Immersive Design for Education. In D. Liu, C. Dede, R. Huang, & J. Richards (Eds.), *Virtual, Augmented, and Mixed Realities in Education* (pp. 35–54). Springer. https://doi.org/10.1007/978-981-10-5490-7_3

Jang, S., Vitale, J., Jyung, R., Black, J. (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Computers and Education*, *106*, 150–165.

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

Johnson, D. W., & Johnson, R. T. (1999). Making Cooperative Learning Work. *Theory into Practice*, *38*(2), 67–73.

Johnson-Glenberg, M. C. (2017). Embodied Education in Mixed and Mediated Realties. In D. Liu, C. Dede, R. Huang, & J. Richards (Eds.), *Virtual, Augmented, and Mixed Realities in Education* (pp. 193–217). Springer. https://doi.org/10.1007/978-981-10-5490-7_11

Johnson-Glenberg, M. C. (2018). Immersive VR and Education: Embodied Design Principles That Include Gesture and Hand Controls. *Frontiers in Robotics and AI*, *5*. https://doi.org/10.3389/frobt.2018.00081

Johnson-Glenberg, M. C., & Megowan-Romanowicz, C. (2017). Embodied science and mixed reality: How gesture and motion capture affect physics education. *Cognitive Research: Principles and Implications*, *2*. https://doi.org/10.1186/s41235-017-0060-9

Jones, M. G., Andre, T., Superfine, R., & Taylor, R. (2003). Learning at the nanoscale: The impact of students' use of remote microscopy on concepts of viruses, scale, and microscopy. *Journal of Research in Science Teaching*. https://doi.org/10.1002/tea.10078

Kiefer, M., & Trumpp, M. (2012). Embodiment theory and education: The foundations of cognition in perception and action. *Trends in Neuroscience and Education*, *1*(1), 15–20. https://doi.org/10.1016/J.TINE.2012.07.002

Korbey, H. (2017). Will Virtual Reality Drive Deeper Learning? *Edutopia*. https://www.edutopia.org/article/virtual-reality-drive-deeper-learning-holly-korbey

Laal, M. (2013). Positive Interdependence in Collaborative Learning. *Procedia - Social and Behavioral Sciences*. https://doi.org/10.1016/j.sbspro.2013.10.058

Lee, M. (2009). How Can 3D Virtual Worlds be used to support collaborative learning? An analysis of cases from the literature. In *Journal of e-Learning and Knowledge Society* (Vol. 5). http://je-lks.org/ojs/index.php/Je-LKS_EN/article/view/300

Leinen, P., Green, M. F. B., Esat, T., Wagner, C., Tautz, F. S., & Temirov, R. (2015). Virtual reality visual feedback for hand-controlled scanning probe microscopy manipulation of single molecules. *Beilstein Journal of Nanotechnology*, *6*(1), 2148–2153. https://doi.org/10.3762/bjnano.6.220

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by Embodiment: Six Precepts for Research on Embodied Learning and Mixed Reality. *Educational Researcher*, *42*(8), 445–452. https://doi.org/10.3102/0013189X13511661

Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers and Education*, *95*, 174–187. https://doi.org/10.1016/j.compedu.2016.01.001

Liu, D., Dede, C., Huang, R., & Richards, J. (Eds.). (2017). *Virtual, Augmented, and Mixed Realities in Education*. Springer Singapore. https://doi.org/10.1007/978-981-10-5490-7

Lui, M., & Slotta, J. D. (2014). Immersive simulations for smart classrooms: Exploring evolutionary concepts in secondary science. *Technology, Pedagogy and Education*, *23*(1), 57–80. https://doi.org/10.1080/1475939X.2013.838452

Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): A Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educational Psychology Review*. https://doi.org/10.1007/s10648-020-09586-2

Mayer, R. E. (2020). Cognitive Foundations of Game-based Learning. In *Handbook of Game-Based Learning*. MIT Press.

Mayer, R. E., & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Mayer, R. E., Wells, A., Parong, J., & Howarth, J. T. (2019). Learner control of the pacing of an online slideshow lesson: Does segmenting help? *Applied Cognitive Psychology*, *33*(5), 930–935. https://doi.org/10.1002/acp.3560

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

Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Kwok, O., Cifuentes, L., & Davis, T. J. (2012). The learner characteristics, features of desktop 3D virtual reality environments, and college chemistry instruction: A structural equation modeling analysis. *Computers & Education*, *59*(2), 551–568. https://doi.org/10.1016/j.compedu.2012.02.004

Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, *140*, 103603. https://doi.org/10.1016/j.compedu.2019.103603

Montoro, D. T., Haber, A. L., Biton, M., Vinarsky, V., Lin, B., Birket, S. E., Yuan, F., Chen, S., Leung, H. M., Villoria, J., Rogel, N., Burgin, G., Tsankov, A. M., Waghray, A., Slyper, M., Waldman, J., Nguyen, L., Dionne, D., Rozenblatt-Rosen, O., … Rajagopal, J. (2018). A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature*, *560*(7718), 319–324. https://doi.org/10.1038/s41586-018-0393-7

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

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. http://www.nextgenscience.org/three-dimensions

Office of the Commissioner, O. of the. (2020, March 24). *FDA approves new breakthrough therapy for cystic fibrosis*. FDA; FDA. https://www.fda.gov/news-events/press-announcements/fda-approves-new-breakthrough-therapy-cystic-fibrosis

Paas, F., Tuovinen, J. E., Tabbers, H., & Gerven, P. W. M. V. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Pellas, N., Dengel, A., & Christopoulos, A. (2020). A Scoping Review of Immersive Virtual Reality in STEM Education. *IEEE Transactions on Learning Technologies*.

Pouw, W. T. J. L., van Gog, T., & Paas, F. (2014). An Embedded and Embodied Cognition Review of Instructional Manipulatives. *Educational Psychology Review*, *26*(1), 51–72. https://doi.org/10.1007/s10648-014-9255-5

Press, T. M. (n.d.). *Handbook of Game-Based Learning | The MIT Press*. The MIT Press. Retrieved May 21, 2021, from https://mitpress.mit.edu/books/handbook-game-based-learning

Rey, Günter Daniel, Maik Beege, Steve Nebel, Maria Wirzberger, Tobias H. Schmitt, and Sascha Schneider. "A Meta-Analysis of the Segmenting Effect." *Educational Psychology Review* 31, no. 2 (June 1, 2019): 389–419. https://doi.org/10.1007/s10648-018-9456-4.

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. https://doi.org/10.1371/journal.pone.0200724

Sandoval, W. A., & Bell, P. (2004). Design-Based Research Methods for Studying Learning in Context: Introduction. *Educational Psychologist*. https://doi.org/10.1207/s15326985ep3904_1

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality. *Frontiers in Robotics and AI*, *3*. https://doi.org/10.3389/frobt.2016.00074

Slater, Mel; Sadagic, Amela; Usoh, Martin; Schroeder, R. (2000). Small-group behavior in a virtual and real environment. *Presence*, *9*(1), 37–51.

Tan, S., & Waugh, R. (2014). Use of virtual-reality in teaching and learning molecular biology. In *3D Immersive and Interactive Learning*. https://doi.org/10.1007/978-981-4021-90-6_2

Thompson, M. M., Pastorino, L., Lee, S., & Lipton, P. (2016). Research and Teaching: Reenvisioning the Introductory Science Course as a Cognitive Apprenticeship. *Journal of College Science Teaching*, *46*(1).

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. In *Workshop Position Paper for CHI 2018*.

Thompson, M., Uz-Bilgin, C., Anteneh, M., Cho, L., Klopfer, E. (2021). Visualizing the Collaborative Problem Solving Process in an Immersive Cross Platform Game. *ILRN 2021 Conference Proceedings*. iRLN.

Thompson, M.; Wang, A.; Roy, D.; Klopfer, E. (2018). Authenticity, Interactivity, and Collaboration in VR learning games. *Frontiers in Robotics and AI, Section Virtual Environments Article*.

Thompson, M., Wang, A., Uz-Bilgin, C., Anteneh, M., Roy, D., Tan, P., Eberhart, R., Klopfer, E.. (2020). Influence of Virtual Reality on High School Students' Conceptions of Cells. *Journal of Universal Computer Science*, *6*(8), 929–946.

Torre-Luque, A. de la, Valero-Aguayo, L., & Rubia-Cuestas, E. J. de la. (2017). Visuospatial Orientation Learning through Virtual Reality for People with Severe Disability. *International Journal of Disability, Development and Education*, *64*(4), 420–435. https://doi.org/10.1080/1034912X.2016.1274022

Uz-Bilgin, C., & Thompson, M. (2021). Processing presence: How users develop spatial presence through an immersive virtual reality game. *Virtual Reality*. https://doi.org/10.1007/s10055-021-00528-z

Uz-Bilgin, C., Thompson, M., & Anteneh, M. (2020). Exploring How Role and Background Influence Gameplay Through Analysis of Spatial Dialogue in Collaborative Problem-Solving Games. *Journal of Science Education and Technology*, *29*(6), 813–826. https://doi.org/10.1007/s10956-020-09861-5

Venugopalan, P. L., Esteban-Fernández de Ávila, B., Pal, M., Ghosh, A., & Wang, J. (2020). Fantastic voyage of nanomotors into the cell. *ACS nano*, *14*(8), 9423-9439.

Wang, A. (2020). *Creators, classrooms, and cells: Designing for the benefits and limitations of learning in immersive virtual reality* [Thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/127661

Wang, A., Thompson, M., Roy, D., Pan, K., Perry, J., Tan, P., Eberhart, R., & Klopfer, E. (2019). Iterative user and expert feedback in the design of an educational virtual reality biology game. *Interactive Learning Environments*, *0*(0), 1–18. https://doi.org/10.1080/10494820.2019.1678489

Weisberg, S. M., & Newcombe, N. S. (2017). Embodied cognition and STEM learning: Overview of a topical collection in CR:PI. *Cognitive Research: Principles and Implications*, *2*(1), 38. https://doi.org/10.1186/s41235-017-0071-6

Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems thinker*, *9*(5), 2-3.

You, S., Tu, H., Chaney, E. J., Sun, Y., Zhao, Y., Bower, A. J., Liu, Y.-Z., Marjanovic, M., Sinha, S., Pu, Y., & Boppart, S. A. (2018). Intravital imaging by simultaneous label-free autofluorescence-multiharmonic microscopy. *Nature Communications*, *9*(1), 2125. https://doi.org/10.1038/s41467-018-04470-8

Yuan, S., Chan, H. C. S., & Hu, Z. (2017). Using PyMOL as a platform for computational drug design. *WIREs Computational Molecular Science*, *7*(2), e1298. https://doi.org/10.1002/wcms.1298

Zhou, Z., Hu, Z., & Li, K. (2016). Virtual Machine Placement Algorithm for Both Energy-Awareness and SLA Violation Reduction in Cloud Data Centers. *Scientific Programming*, *2016*(i). https://doi.org/10.1155/2016/5612039

(N.d.). Retrieved May 23, 2021, from https://scholar.google.com/scholar_url?url=https://link.springer.com/chapter/10.1007/978-981-10-5490-7_3&hl=en&sa=T&oi=gsb&ct=res&cd=2&d=15612111972567791367&ei=73uqYN2RB4-bmAGkqYPABA&scisig=AAGBfm2aNYRUTJTbB8P83RYAvG6sEXWOWA

14 Supplementary Material

Researchers who wish to view the dataset and instruments for the project may go to the website https://osf.io/bv89n/