MIT Libraries | DSpace@MIT

MIT Open Access Articles

Iterative user and expert feedback in the design of an educational virtual reality biology game

The MIT Faculty has made this article openly available. ***Please share*** how this access benefits you. Your story matters.

As Published: 10.1080/10494820.2019.1678489

Publisher: Informa UK Limited

Persistent URL: https://hdl.handle.net/1721.1/136492

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

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike

Massachusetts Institute of Technology

Iterative user and expert feedback in the design of an educational virtual reality biology game

This study focuses on an educational game titled *Cellverse*, a two-player cross-platform VR project intended to teach high school biology students about cell structure and function. In *Cellverse*, players work in pairs to explore a human lung cell and diagnose and treat a dangerous genetic disorder. *Cellverse* is being designed by the Collaborative Learning Environments in Virtual Reality (CLEVR) team, an interdisciplinary team consisting of game designers, educational researchers, and graduate and undergraduate students. Using a design-based research approach, we have enlisted the help of both subject matter experts and user testers to iteratively design and improve *Cellverse*. The objective of this paper is to share how user and expert feedback can inform and enhance the development of learning games. We describe how we gather and synthesize information from in-game observations, semi-structured interviews, and video data to review and revise our game. We discuss the input of subject matter experts, present feedback from our user testers, and describe how input from both parties influenced the design of *Cellverse*. Our results suggest that including feedback from both experts and users has provided information that can clarify gameplay, instruction, subject portrayal, narrative, and in-game goals.

Keywords

Virtual reality; educational games; serious games; K-12 education; embodied learning; collaborative learning

Introduction

As the technology becomes more affordable, virtual reality (VR) is becoming increasingly viable in K-12 and undergraduate circles (Castaneda, Cechony, Bautista, 2017). Many of the educational experiences currently available in VR address learning in science, technology, engineering, and mathematics (STEM) domains. Biology VR games, animations, and simulations have been developed for both educational and for entertainment purposes. Applications of VR in biology have been shown to improve high school students' understandings in abstract concepts in microbiology (Tan and Waugh, 2017; Minogue et al., 2006) to improve college students' understanding of molecular structures (University of Stavanger website, 2017), to enhance learning of anatomy through interaction compared to traditional methods (Jang, Vitale & Jyung, 2017), and to enable virtual laboratories for students (Potkonjak, 2016; Roschelle et al., 2017). Science, technology, engineering, and mathematics (STEM) topics have been the focus of a few games (Bonde, Makransky, Wandall, et al., 2014). Games like *InCell* and *Cellscape,* which are accessible via commercial VR headsets, take users on virtual "tours" of cells and cell functions (InCell Website, 2017; Cellscape website, 2017). Although these games are immersive, most of them are neither interactive nor collaborative, and allow only for single-player experiences. Furthermore, many of them do not depict cells in an authentic manner, taking liberties with size and scale, density, number of organelles, and other vital factors without regard for how these depictions may affect learners' conceptions.

School curricula are beginning to incorporate 21st century skills such as critical thinking and collaboration (Fiore et al., 2017); consequently, collaboration within VR has also become relevant to educators (Merchant, Goetz, Cifuentes, Keeney-Kennicutt, & Davis, 2014). Cross-platform collaborative games blend together VR and 2D platforms (Gugenheimer, Stemasov, Sareen, & Rukzio, 2017). Games like *Black Hat Cooperative* and *Keep Talking and Nobody Explodes* entertain players by providing different user and environmental interfaces to in- and out-of-VR players and challenging them to cooperatively solve puzzles or other fast-paced challenges. However, there are few viable examples of educational games that involve two or more players in and out of VR.

The goal of the Collaborative Learning Environments in Virtual Reality (CLEVR) project is to develop an educational game, *Cellverse*, to help students learn about cell structure, the process of transcribing DNA to RNA, and translating RNA to proteins (the central dogma). We have used a design-based research methodology (Collins, Joseph, Bielacyzc, 2004; DBRC, 2003), conducting tests and interviews with users and experts throughout the design process. The data we collected have helped us iteratively create an experience that offers learners rich, immersive opportunities to collaboratively investigate and explore the cell from the inside out.

Theoretical Background: STEM learning theories and collaborative games

Well-designed game-based learning experiences have the potential to engage students in the learning environment through different educational pathways. For example, in one segment within the simulation *The Body VR: Journey inside a cell*, viewers enter the bloodstream and experience the journey of a red blood cell spreading oxygen into the body. Being in the environment enables experiential learning, where viewers can learn through placing themselves within a virtual space (Kolb, 2014; Kolb 1984). Research suggests another potential added value for enabling learners to interact directly with the environment, enabling embodied cognition. In turn, the embodiment theory of learning proposes that knowledge is best imparted when learning events relate to physical actions (Kiefer & Trumpp, 2012). Weisberg & Newcombe (2017) further develop this theory by linking embodied cognition to the development of spatial awareness, which is identified as an important part of understanding in STEM domains (Stieff and Uttal, 2015). Thus, incorporating physical movement and gesture can increase an individual's capacity to learn complex domains by creating multiple ways of interacting with material. For example, students who learned about biomolecular models with haptics were better able to understand how molecules fit together than students using models without haptics (Schönborn, Bivall, Tibell, 2011). Studies of embodied learning environments suggest that direct manipulation of physical or virtual representations of materials can enhance learning as compared to traditional methods (Webster, 2016; Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004) guides the design of the CLEVR project *Cellverse* will also incorporate embodied learning by requiring students to use gestures and movement to navigate through the cell. Research connecting embodied cognition and virtual reality is still nascent (Weisberg and Newcombe, 2017); however, the research base on embodied cognition began long before virtual reality's rising popularity and includes research on manipulatives in mathematics (Martin & Schwartz, 2005; Schwartz & Martin, 2006), and in 2D video games such as Tetris (Kirsch & Maglio, 1994; Pouw, van Gog & Pass, 2014).

The game-based format of Cellverse provides a high degree of interaction between students and the concepts included in VR environment (Lindgren et al., 2016; Jang et al. 2017). A significant piece of this interaction is movement around the cellular environment, which introduces a spatial dimension to a cell that is often missed in 2D textbook representations. The two players involved develop the spatial understanding of the virtual environment over time to locate clues that will help them figure out what is wrong with the cell. Strong spatial awareness is linked to improved understanding of STEM topics (Uttal and Cohen, 2012). The game provides ongoing textual and visual feedback to the players, which may also assist the learning process (Merchant et al, 2014). In creating Cellverse, we intend to present a topic that is often passive and vocabulary-based into an experience that builds on our current understanding of embodied learning by incorporating action to cognition and development of spatial understanding of cells. By incorporating ongoing user testing, we can document how embodied cognition and spatial understanding might be developed through game design and adjust the design to further enhance embodied and spatial cognition. and understand the role of embodied cognition in their understandings. This paper documents ongoing user testing and how it has helped us gauge the effectiveness of our design in meeting those aims.

Collaboration in Embodied Learning

Education has expanded beyond conceptual and content knowledge to build students' 21st century skills such as critical thinking and collaboration to help students prepare for the workforce (OECD, 2017). One area of interest is in collaborative problem solving. Collaborative problem-solving links individuals in "sharing the understanding and effort required to come to a solution and pooling their knowledge, skills and efforts to reach that solution (Fiore et al., 2017).” As research in blended reality collaborative environments continues to develop (Bower, Lee & Dalgarno, 2016), broader research on collaboration becomes more applicable and useful.

Supporting collaborative problem solving requires individuals to actively communicate, interact, negotiate, and establish social relationships among the group. Johnson & Johnson’s (2005) theory of positive interdependence suggests that optimal collaboration originates from a situation where all individuals need to contribute for the group to succeed in its mission. Laal (2013) further divides interdependence into nine subcategories; for this study, we focus on their category of role and resource interdependence. Roles are an important construct in helping team members negotiate interactions in virtual environments (VE) and VR (Koutsabasis, Vosinakis, Malisoba, Paparounas, 2012), and in determining who should complete tasks (Li & Zhou, 2016; Lui & Slotta, 2014). Resource interdependence can be fostered by giving individuals access to different technology such as a head mounted display (HMD) and a tablet (Gugenheimer et al., 2017), or giving each participant varied different tools and capabilities (Jensen, 2016). We are interested in understanding how our game can establish role and resource interdependence among learners.

Another construct that is useful in exploring the opportunities for collaborative problem-solving is the concept of productive failure. Productive failure, studied by Kapur (2015), suggests that deeper learning can occur in situations when participants’ learning generate solutions to an unfamiliar concept with limited or no support. In Kapur’s experiments, students were given educational activities to complete without educator “cognitive guidance or support”; it was discovered that allowing these students to leverage their previous knowledge and natural intuition, thus forcing them to experiment with solutions and potentially fail before succeeding, enabled greater learning and retention than fully guided experiences. Some modern video games are a particularly insightful source of the power of productive failure - levels or challenges may require multiple attempts and strategies to complete. Failure is integral to gameplay; a well-designed game enables players to embrace process as part of learning to succeed (Gee, 2014). In designing Cellverse, we consider the role of productive failure as a careful balance between challenge and support.

Designing *Cellverse*

Within the framework of design-based research, we iteratively investigate and improve our game throughout the duration of its development (Collins, Joseph, & Bielaczyc, 2004; DRBC, 2003). Kurilovas (2016) recommends using both a bottom-up (user-based) and a top-down (expert-based) review to evaluate the quality of VR experiences. User-based testing recognizes the importance of understanding how the users interact with the virtual environment, while experts provide insight into the technical quality of the system (Kurilovas, Serikoviene, Viorikari, 2014). While Kurliovas (2016) focused on experts in VR, we explore how experts in science and science education have informed game design. We propose three design conjectures (Sandoval, 2014) that link our theoretical and conceptual starting points to our inquiry into the process of creating the game.

(1) Incorporating domain experts (top-down) in the vision and creation of a serious game will enable the design of an authentic representation of the cell.

(2) Iterative user testing (bottom-up) will reveal the learning affordances (e.g. embodied learning) in the context of a biological cell.

(3) Ongoing user testing will inform the creation of a cross-platform collaborative game.

The CLEVR project is funded by Oculus VR and was built to support the Oculus Rift system. Our team consists of education researchers, game designers, and undergraduate and graduate students skilled in programming, front-end development, UI design, 3D modeling, and/or digital art. *Cellverse*, the game, was built in Unity and programmed in C#, with cellular and protein models created in Blender and Maya or downloaded and modified from open-source sites like the Protein Data Bank. When developing *Cellverse* or conducting user tests, we used a PC with an AMD Ryzen 5 1600X Six-Core Processor and a Nvidia GeForce GTX 1080 graphics card, running Windows 10 64-bit. The navigator tablet was also built in Unity and tested on several tablet laptops running Windows 10 64-bit.

Methods and Analysis

Subject Matter Expert Interviews

In the context of designing *Cellverse*, subject matter experts (SMEs) are individuals with knowledge of science and science education, particularly cellular and molecular biology. We identified SMEs through online research, networking, and suggestions by the researchers themselves. Upon meeting, we demonstrated the game and gathered their feedback. We prepared broad questions about their field of expertise and specific questions about the game with possible applications within *Cellverse*. During the interviews, members from our research, design, and implementation teams took notes and shared them with the experts after the meeting to confirm accuracy (Thomas, 2017).

User Testing Process

INSERT TABLE 1 HERE

For formal playtests, we created a "Multiplayer User Testing Feedback Form," which records user reactions and interactions with major facets of the game: tutorial, environment, gameplay mechanics, and multiplayer function. Depending on whether the user tested the explorer (in-VR) or navigator (tablet) role, they participated in a slightly different pre-testing procedure. In each test, a team member explained the premise of the game, assisted users with the headset if applicable, and asked them a series of questions before, during, and after the game using a structured protocol. These questions were used to document and evaluate user interaction with *Cellverse* and to gauge our own ability to guide users through the experience. They were also designed to be open-ended, to allow users to be as honest and critical as possible. Some questions have been altered, revised, or removed depending on our needs.

The feedback notes contain observations on user behavior and information about users' backgrounds (such as student, teacher, researcher). User tests were recorded on video and stored in a private server where team members reviewed and took notes on significant findings. All user names have been replaced with pseudonyms in this analysis; while the team knows the first names of participants, they are otherwise anonymous.

The study has been reviewed by the Institutional Review Board (1709095354), and both users and subject matter experts gave consent to include their responses in the study.

Results

The results below are presented chronologically, from spring 2017 to early summer 2018. We will first discuss the input of SMEs (top-down), present feedback from our user testers (bottom-up), and finally explain how input from both parties influenced the design of *Cellverse* at that respective stage. We will also discuss the decisions that we made regarding SME and user suggestions, which include changes we added to the game as well as suggestions that we ultimately chose not to implement. We will finally discuss features of the most current version of *Cellverse* (as of summer 2018).

Summer 2017 - Developing the *Cellverse* Environment

SME Feedback

Between May and July 2017, we interviewed SMEs to further inform the overall design and direction of the game.[1] The primary topics of discussion included depictions of size and scale in VR; educational games within a classroom setting; misconceptions about cells, viruses, and other bodily microstructures; and the “life cycle” of viruses.

Experts consulted during this period were primarily experienced in various fields of microbiology, particularly virology. Initially, we explored viruses as a context for the game, and sought out experts in virology to discuss ways of integrating viruses into the game.[2] After learning about the diversity in how viruses infect and multiply in cells, we decided to focus on genetics, which is a more relevant topic in current high school curricula. Many of the virus-based suggestions regarding *Cellverse* gameplay thus became obsolete and were not implemented. We also started using the Next Generation Science Standards (NGSS, 2013) as a baseline for the educational content of *Cellverse*.

[1] Interview with Matthew Schnepps; Interview with Lourdes Aleman; Interview with Shane Tutweiler; Interview with David Walt.

[2] Interview with Gail Jones; Interview with Connie Cepko; Interview with Daniel Kurtizkes; Interview with Tyler Krause.

Feedback integration into design

Our initial focus in the beginning of Summer 2017 was on size and scale in biology, which we eventually narrowed down to size and scale in cell biology. The first iterations of *Cellverse* investigated various forms of movement and scaling and were created in simplistic VR environments that paid little heed to true scientific accuracy. The navigation method that we chose to continue working on, coined the “Spider-Man” style (Fig. 1), allows users to move their virtual forms around the cell environment using a pointer attached to the right-hand touch controller. “Spider-Man” allows for a significant amount of free movement and has been the movement style that we have used for subsequent iterations.

INSERT FIGURE 1 HERE

Fall 2017

SME Feedback

Interviews during the fall focused on single-player interaction and exploration within the VR space. Most of the SMEs invited to test and discuss the game were experienced with cell biology, specifically cell structure and function, and as such provided helpful input on the accuracy of our visual depictions and in-game science.

Input from SMEs was influential in constructing *Cellverse* as a scientifically accurate and visually interesting human epithelial lung cell.[3] We implemented suggestions from David Goodsell, a molecular biology professor and artist including creating a more polarized cell (as would happen in most lung cells), building a larger and more accurate nucleus model among

[3] Interview with David Goodsell.

other major organelles, and creating visual “markers” for genetic disease, and increasing the overall density of *Cellverse* to make it more realistic. Professor Iain Cheeseman and his graduate students suggested several ideas for cellular animations, which add a “living” quality to *Cellverse*.[4]

We interviewed Dr. Hongmei Mou, a specialist in personalized medicine for lung and airway diseases. Dr. Mou’s input solidified our decision to focus on the genetic lung disease cystic fibrosis (CF). We also learned about the numerous classes of CF, and used them to compose a list of symptoms that players must consider during the game.[5]

We did not incorporate all of the SME’s suggestions. For example, technological limitations and potential nausea among players prevent us from fully recreating the density or viscosity of the cell to the level that Dr. Goodsell suggested. Several of the other decisions were compromises between learning objectives, the principles of game design, and the technical capabilities of VR. We continually balance creating an accurate, and physically comfortable experience.

User testing feedback

User testing, which began in earnest during the fall of 2017, focused on comfort, in-game movement, and environmental feedback, and users’ understandings and expectations regarding cell structure. At this point we had not yet added goals or collaboration to the *Cellverse* interface; it was an explorable environment with minimal interaction.

[4] Interview with Ian Cheeseman et al.
[5] Interview with Hongmei Mou.

Reflections

User testing suggested that the design should include an engaging environment that prompted individuals to link their biology understandings with the virtual experience. Nine of the 25 testers specifically remarked on the dense nature of the environment, focusing on prolific organelles like mitochondria and ribosomes. Six users also tried to connect their observations with their preexisting biology knowledge, calling out specific organelles such as mitochondria, ribosomes, and “lysosomes - I forgot about those.” Finally, three users described how the environment was fun. They gave affective feedback such as “It's kind of amazing - it's just fun. I don't think my high school biology did cells justice.”

Users’ suggestions for the game fell under three major themes: direction, interaction, and aesthetic (nine mentioned at least one of these themes). Users requested more gameplay elements that would provide direction such as a “find-it list,” a “minimap,” or “given objectives” such as tutorials or quests. Many of these features are part of our ongoing design process and will be added over time. Users wanted more interaction with organelles in the cell and functionality such as “interactive elements on the left hand,” and toggleable “layers to the visuals.” For aesthetics, players wondered about the “choices made behind the color palette.” Five players wanted greater dynamism and to be able to see animated cell processes.

Improvements

Authenticity is one of our major goals and is something that is not always focused on in biology visualizations. Our SMEs recommended realistically dense environments that could support deeper thinking about cell processes and anomalies, and our users were surprised and engaged by those environments. Several players mentioned that they wanted to see a “task list,” a “map,” or simply more information, which were not in the game build at that time but were slated for future development. User requests were later reflected in some of our next steps, which included developing a game goal, increasing the methods of interacting with the environment, and making *Cellverse* more visually engaging.

Feedback integration into design

We remodeled our generic-looking cell as an epithelial lung cell in anticipation of possible diseases that could be used in-game. We increased the model count of several major organelles, including mitochondria and lysosomes; we also added a network of microtubules to better illustrate the cell's shape (Figure 2). By the end of the fall semester, we had rebuilt *Cellverse* into a larger, denser, and more visually interesting environment.

INSERT FIGURE 2 HERE

Winter 2018

We focused primarily on implementing suggestions and concentrating on user testing from January 2018, maintaining contact with a smaller set of SMEs to confirm that our game remained scientifically accurate.

January 2018

A total of ten users tested the Winter 2018 version of *Cellverse* and completed the feedback form. Collaboration was first introduced in the January playtest, where the explorer (in-VR) had an internal view of the cell and the tablet user, or navigator, had a birds-eye view of the cell and a "tasklist" of disease symptoms they were expected to search for within *Cellverse*

(Figures 3 and 4). The task list was influenced by discussion with genetics and disease experts from the fall and contained a list of possible genetic disorders that could be diagnosed by the absence, presence, and/or malformation of certain cell structures. The navigator also had access to a "beacon mode" that allowed them to place beams of light within *Cellverse* for the explorer to see. (This function was replaced by a different function, which we will describe below in "Current version of *Cellverse*".) The explorer and navigator roles were designed to require users to work together to solve the problem.

Our user feedback, particularly from navigators, suggested that having a partner was useful (2 users), but ultimately unnecessary (3 users). Although explorers observed that "it was nice to have a second opinion," navigators commented that "I could have done it all by myself" because "I had the description of the disease and my partner didn't." Users gave suggestions on how to better involve partners, including understanding the different views of *Cellverse* and accessing interactive tools that can affect each partner's view.

INSERT FIGURE 3 HERE

INSERT FIGURE 4 HERE

Improvements

This playtest enabled us to focus on balancing the information between players during the game, and we drew upon the users' suggestions to do so. We made sure that players know that there are two views and have a sense of what their partner might see. We made the navigator's view even more distinct from the explorer's view by reducing the detail in the navigator view, requiring more interaction between the navigator and explorer. We adjusted the task list so that explorer observations were needed to progress in the game. To make the roles more distinct and interdependent, we began considering implementing a “staining and tagging” function that mimicked real-world scientific data collection. We used scientific research techniques that the SMEs mentioned in the interviews as a way of maintaining authenticity in the game.

Spring 2018

User Testing Feedback

As the length and complexity of the game increased, we conducted more focused and longer user testing sessions with fewer participants. For March, we tested the game with four people, or two explorer/navigator pairs. In the March playtest, we realized that we needed to continue to balance information by clarifying labels in the tutorials. Additionally, the navigator was still leading the collaboration, which suggested there needed to be more balance of power between the partners.

Feedback from the May playtest suggested that we made progress on our goal of collaboration, as all four playtesters worked together during the activity as opposed to one partner taking over. One player explained she and her partner worked together because “I had examples and definitions, [and] I had to ask what he was seeing and comparing”, while another mentioned that they had to ask the partner who was inside the cell for “specific names and colors” to progress in the game. Having a partner also helped users identify and understand the objective of the game.

Players still requested the ability to change viewpoints to get different perspectives on the cell, and to be able to go inside the nucleus. These requests foreshadowed our work on another upcoming aspect of the game, a “nanoscope" view, to allow the explorer to view organelles and cell functions at a smaller scale and thus with much greater detail.

Summer 2018

Summer 2018 includes our most recent version of *Cellverse*. We spoke to several medical researchers to further improve our understanding of the effect of cystic fibrosis within lung cells.[6] We have also maintained contact with preexisting SMEs and have continued to develop CLEVR and *Cellverse* in a scientifically accurate manner.

User Testing Feedback

We had been testing paper prototypes of both explorer and navigator tutorials since January and received some constructive feedback on both. June’s playtest focused on the digital version of the game tutorial. In addition to digitizing the tutorials, the navigator view had been streamlined and improved. Four users tested the game during this period.

As surmised from user feedback, there are a few technical issues that we will need to address to further streamline the tutorial. Researchers observed that technical glitches within *Cellverse* resulted in confusion among users, including lack of guidance for some in-game functions and easily skippable tutorial instructions.

Some of our attempts to enable communication caused frustration between the players. For example, the tablet *Cellverse* view (depicting a “healthy” cell) was slightly different from the in-VR *Cellverse* view in terms of color and structure, which players did not always notice or

[6] Interview with Brian Lin and Jaimee Elizabeth Hoefert, July 26th, 2018.

communicate to each other. This resulted in confusion when players were attempting to describe certain organelles to each other in terms of color - for example, the centrosome, which was green in navigator view but is blue (a symptom of a fake placeholder disease) in VR view. Color confusion will likely be mitigated by the implementation of a microscopy technique as described below, which means that the organelles in the *Cellverse* environment will initially start out in greyscale.

Current version of Cellverse

As of summer of 2018, *Cellverse* design and gameplay allows two players to collaborate and communicate while traveling around the virtual “world” of *Cellverse*, a human epithelial lung cell. One player wears a VR head-mounted display (the “explorer”), and the second player uses an interactive tablet interface connected to the same virtual space (the “navigator”). Each player is granted a different perspective of the cell; the explorer has a three-dimensional internal view, where they are placed “inside” the cell and can travel around freely (Figure 7), while the navigator is given a “bird’s-eye” view of the same cell that can be rotated horizontally, vertically, and along the z-axis (Figures 5 and 6). The explorer can view descriptions of various cell organelles and components by selecting them and viewing a description on an adjoining clipboard. The navigator has a less detailed view of the cell and is provided with a reference of cell and disease information that the explorer cannot access. The navigator can communicate with the explorer, either verbally or by using simultaneous label-free autofluorescence-multiharmonic (SLAM) microscopy as a selection technique. SLAM uses light-based imaging to stain living tissue as opposed to traditionally used dyes.

We focused on balancing player roles, roles, and resources to foster collaboration (Thompson et al., 2018). The explorer's view is more detailed and includes cell structures and gameplay functions not shown on the navigator's user interface and lets them peruse nanoscopic cellular functions through the "nanobot" function. This affords the explorer the ability to view parts of the cell at an even smaller scale and see DNA, RNA, actin, proteins, ribosomes, and other structures too small to see at the "main" *Cellverse* view. The navigator has textual information pertaining to cellular disorders and diseases that are unavailable to the explorer, and access to a SLAMming mode that can select different organelles to make them visible to the explorer.

Some functions have not yet been implemented into playable builds of *Cellverse* but are currently under construction. After diagnosing the disease, players will rearrange DNA sequences to build a cure, much like how real-world scientists use gene therapy as a medical treatment for cystic fibrosis. We have also created a narrative that gives players context for what players are doing within the game, placing them as interns inside a hospital trying to help diagnose a sick patient.

INSERT FIGURE 5 HERE

INSERT FIGURE 6 HERE

INSERT FIGURE 7 HERE

Discussion

The top-down and bottom-up approach recommended by Kurilovas (2016) has been very useful in our design process. Consulting domain experts (top-down) during the process helped us design an authentic cell environment and focus the narrative on diagnosing and curing the cell. Experts have informed how we represent cystic fibrosis at the cellular level, have critically evaluated the accuracy of the portrayal of cell structures and processes, and have provided ideas about how players can interact with the cell environment. Our cellular depictions have changed directions based on information experts have shared that are at the forefront of scientific understanding of the causes of cystic fibrosis. In seeking expert feedback, we are continually reminded of the dynamic nature of science knowledge. The virtual cellular environment is designed to reflect a realistic population of organelles – far more mitochondria and lysosomes than are regularly included in 2D representations. To prevent players from being overwhelmed by the sheer number of interactive items, they can physically select one organelle, move towards it, and turn on a virtual clipboard to review its function. The ability to select and move towards one organelle at a time scaffolds the experience so that players can focus on problem-solving even as they are navigating through the densely packed cell.

User feedback helped us choose locomotion and selection techniques for the explorer, and the “pinch and zoom” and rotation functions the navigator has on the tablet. Interacting with a shared environment through a tablet and headset enabled us to establish possible roles and tasks. We learned that dividing resources such as information, tools, and viewpoints of the cell between the two players was critical in involving both players in the problem-solving process and developing a positive interdependence that involved both players in the problem-solving process (Johnson and Johnson, 2005; Laal, 2013). While roles and resources may be viewed as distinct aspects of interdependence, the two are tightly linked. Playing the game activated participants’ prior knowledge of cells; elaborating upon those ideas with a partner enabled participants to experience failure in a productive way (Kapur, 2015). Productive failure also played an important role in encouraging players to explore, discuss, hypothesize, and test potential solutions to the game. User feedback regarding game difficulty allowed us to shape a gameplay experience that provided a challenge without making the solution too easy or infuriating to solve.

Our next set of SMEs are science educators and high school students. We will be shifting focus from teacher feedback on the game to teacher and student input regarding using *Cellverse* in the classroom with the goal of involving our target audience of high school students, the concept of productive failure will become more salient. One of our ongoing research questions regards the amount of support we need to give in order to ensure a fruitful learning experience.

Conclusion

As a cross-platform multiplayer VR and tablet game designed for the high school biology classroom, CLEVR is an ambitious project. We have taken into consideration that a game played in a collaborative classroom environment requires different design decisions than a single-player or multiplayer game in a purely virtual environment. The optimal way to discover and test our design decisions has been through iterative user testing, and we have developed CLEVR and Cellverse with that understanding in mind. This method will become more relevant as VR moves from a novel, limited experience to a more accessible and widespread educational tool. It is important to ensure that these games are accurate, entertaining, and collaborative.

Designing in virtual reality is resource-intensive and ensuring that these resources are well-allocated is a serious challenge. Our aim in sharing this narrative is to help other creators consider top-down and bottom-up approaches by describing the data collection process and the resulting benefits to provide inspiration for future researchers seeking to design and create educational games for academic subjects. As educational games are infused into curricula and formal education, it will become increasingly important to create games that both portray the subject matter in an authentic way and are also easy to understand and play in a classroom or other large group setting. Including multiple stakeholders as informants is time intensive; however, we believe this will result in a more effective and engaging experience.

Acknowledgements

CLEVR is supported in part by Oculus Education. We also acknowledge the MIT students, staff, scientists, and individuals who gave us feedback on the game.

Declaration of Interest

The authors of this manuscript have no conflict of interest, as we do not benefit financially from this project and have no stock or other personal financial involvement with any organization or entity discussed in this manuscript.

References

Bower, M., Lee, M. J. W., & Dalgarno, B. (2017). Collaborative learning across physical and virtual worlds: Factors supporting and constraining learners in a blended reality environment. *British Journal of Educational Technology*, 48(2), 407–430. https://doi.org/10.1111/bjet.12435.

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General, 136*(4), 539-550.

Castaneda, L., Cechony, A., & Swanson, T. (2016). Implications of Virtual Reality in Applied Educational Settings. Retrieved from http://foundry10.org/research/the-implications-of-vr-in-applied-educational-settings/.

Cellscape: A Virtual Reality Journey Inside a Cell. (n.d.). Retrieved July 9, 2018, from https://www.cellscape.net/about.

Chan, M. S., & Black, J. B. (2006). Direct-manipulation animation: incorporating the haptic channel in the learning process to support middle school students in science learning and mental model acquisition. In Proceedings of international conference of the learning sciences. Mahwah, NJ: LEA.

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 & Education*, 85, 59-73.

Christopoulos, A., Conrad, M. & Shukla, M. (2018). Increasing student engagement through virtual interactions: How? *Virtual Reality*, 22(4), 353-369.

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

Day, S. B., & Goldstone, R. L. (2011). Analogical Transfer from a Simulated Physical System. *Journal of Experimental Psychology*: Learning Memory and Cognition, 37(3), 551–567. https://doi.org/10.1037/a0022333.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5-8.

Gee, J. P. (2014). What video games have to teach us about learning and literacy. Macmillan.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and Imagined Activity Can Enhance Young Children's Reading Comprehension. Journal of Educational Psychology, 96(3), 424–436. https://doi.org/10.1037/0022-0663.96.3.424.

Gugenheimer, J., Stemasov, E., Sareen, H., & Rukzio, E. (2017, May). FaceDisplay: Enabling Multi-User Interaction for Mobile Virtual Reality. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (pp. 369-372). ACM.

Fiore, S. M., Graesser, A., Greiff, S., Griffin, P., Gong, B., Kyllonen, P., ... & Soulé, H. (2017). Collaborative problem solving: considerations for the National Assessment of Educational Progress.

Han, I., Black, J. B., & Hallman Jr., G. (2009, April). Are simulation and physical manipulation different in improving conceptual learning and mechanical reasoning? Paper presented at the 2009 AERA Annual Meeting, San Diego.

Herrington, J., Reeves, T.C., & Oliver, R. (2010). Researching authentic e-learning. *A guide to authentic e-learning.* London and New York: Routledge.

InCell VR. (n.d.). Retrieved July 9, 2018, from https://luden.io/incell/.

Jang, S., Vitale, J. M., Jyung, R. W., & Black, J. B. (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Computers and Education*, *106*. http://doi.org/10.1016/j.compedu.2016.12.009

Jensen, C. G. (2017). Collaboration and Dialogue in Virtual Reality. *Journal of Problem Based Learning in Higher Education*, 5(1), 85-110.

Johnson, D. W.; Johnson, R. (2005). "New developments in social interdependence theory". *Psychology Monographs.* 131 (4): 285–358. doi:10.3200/mono.131.4.285-358.

Kapur, M. (2015). Learning from productive failure. *Learning: Research and Practice*, 1(1), 51-65.

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

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. Cognitive Science, 18, 513–549.

Koutsabasis, P., Vosinakis, S., Malisova, K., & Paparounas, N. (2012). On the value of virtual worlds for collaborative design. *Design Studies*, *33*(4), 357-390.

Kurilovas, E. (2016). Evaluation of quality and personalisation of VR/AR/MR learning systems. *Behaviour & Information Technology*, *35*(11), 998-1007.

Kurilovas, E., Serikoviene, S., & Vuorikari, R. (2014). Expert centred vs learner centred approach for evaluating quality and reusability of learning objects. *Computers in Human Behavior*, *30*, 526-534.

Laal, M. (2013). Positive interdependence in collaborative learning. *Procedia-Social and Behavioral Sciences*, *93*, 1433-1437.

Lave, J., & Wenger, E. (1998). Communities of practice. Retrieved June 9, 2008 ST--Communities of practice.

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

Li, L., & Zhou, J. (2016, December). Virtual reality technology based developmental designs of multiplayer-interaction-supporting exhibits of science museums: taking the exhibit of virtual experience on an aircraft carrier in China science and technology museum as an example. In Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1 (pp. 409-412). ACM.

Liu, D., Dede, C., Huang, R., & Richards, J. (Eds.). (2017). *Virtual, Augmented, and Mixed Realities in Education*. Springer.

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.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: adapting and reinterpreting physical environments in the development of fraction concepts. Cognitive Science, 29(4), 587–625.

Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Kwok, O. M., 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.

Minogue, J., Jones, M. G., Broadwell, B., & Oppewall, T. (2006). The impact of haptic augmentation on middle school students' conceptions of the animal cell. *Virtual Reality*, *10*(3-4), 293-305.

NGSS Lead States, (2013). *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press. HS-LS1 From Molecules to Organisms: Structures and Processes. Retrieved from http://www.nextgenscience.org/sites/default/files/dci-arrangement/HS-LS-1062117withfooter.pdf.

OECD (2017), PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science, Preliminary Version, OECD Publishing, Paris.

Potkonjak, V., Gardner, M., Callaghan, V., Mattila, P., Guetl, C., Petrović, V. M., & Jovanović, K. (2016). Virtual laboratories for education in science, technology, and engineering: A review. *Computers & Education*, *95*, 309-327.

Pouw, W. T., Van Gog, T., & Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives. Educational Psychology Review, 26(1), 51-72.

Prabhat, A. F., Katzourin, M., Wharton, K., & Slater, M. (2008). A comparative study of desktop, fishtank, and cave systems for the exploration of volume rendered confocal data sets. *IEEE Transactions on Visualization and Computer Graphics, 14*(3), 551-563.

Roschelle, J., Martin, W., Ahn, J., & Schank, P. (2017). Cyberlearning community report: The state of cyberlearning and the future of learning with technology.

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the learning sciences, 23*(1), 18-36.

Schwartz, D. L., & Martin, T. (2006). Distributed learning and mutual adaptation. Pragmatics & Cognition, 14(2), 313–332.

Schönborn, K. J., Bivall, P., & Tibell, L. A. (2011). Exploring relationships between students' interaction and learning with a haptic virtual biomolecular model. *Computers & Education*, *57*(3), 2095-2105.

Siegler, R. S., & Ramani, G. B. (2009). Playing Linear Number Board Games-But Not Circular Ones-Improves Low-Income Preschoolers' Numerical Understanding. *Journal of Educational Psychology*, 101(3), 545–560. https://doi.org/10.1037/a0014239

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality. *Frontiers in Robotics and AI*, *3* (December), 1–47. http://doi.org/10.3389/frobt.2016.00074.

Stieff, M., & Uttal, D. (2015). How much can spatial training improve STEM achievement? *Educational Psychology Review*, 27(4), 607-615.

Stolz, S. A. (2015). Embodied Learning. *Educational Philosophy and Theory*, 47(5), 474–487. https://doi.org/10.1080/00131857.2013.879694.

Tan, S., & Waugh, R. (2014). Use of virtual-reality in teaching and learning molecular biology. *3D Immersive and Interactive Learning* (pp. 17–43). https://doi.org/10.1007/978-981-4021-90-6_2.

Thomas, D. R. (2017). Feedback from research participants: are member checks useful in qualitative research? *Qualitative Research in Psychology*, *14*(1), 23-41.

University of Stavanger Website (2017). Teaching Molecular Biology in Virtual Reality. Accessed 7/7/2017. https://www.uis.no/news/teaching-molecular-biology-in-virtual-reality-article120293-8865.html.

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why, and how?. *Psychology of Learning and Motivation*, 57, 147-181.

Van der Land, S., Shouten, A. P., Van den Hoof, B., & Feldberg, F. (2011). Modeling the Metaverse: A Theoretical Model of Effective Team Collaboration in 3D Virtual Environments. *Virtual Worlds Research*, 4(3).

Vincini, P. (2002). The nature of situated learning. *Innovations in Learning.* Retrieved from https://sites.tufts.edu/ets/files/2012/12/newsletter_feb_20031.pdf.

Vosinakis, S. & Koutsabasis, P. (2013) Interaction design studio learning in virtual worlds. *Virtual Reality*. 17: 59.

Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, 24(6), 1319-1333.

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.

You, S., Tu, H., Chaney, E. J., Sun, Y., Zhao, Y., Bower, A. J., 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.

Appendix

User Testing Feedback Form (June 2018 Version)

Pre-testing Questions *(These questions were asked before the playtesting experienced, and involved a simple sit-down interview with the user tester.)*

Name(s):

You (VR user) are a(n):

Middle/High School Teacher

Middle/High School Student

High School Student

Undergraduate student

Graduate student (Masters/PhD)

Researcher/Scientist

Other:

You (tablet user) are a(n):

Undergraduate student

Graduate student (Masters/PhD)

High School Student

Researcher/Scientist

Middle/High School Teacher

Middle/High School Student

Other:

Have you used VR before? If so, what did you do/what did you play?

(If used VR before) Have you used collaborative or multiplayer VR before? If so, what did you do/what did you play?

What is the relationship between you and your partner? How well do you know each other?

Don't know each other

Know each other somewhat (acquaintances)

Know each other very well

Other:

[Internal] Did they make any comments about the narrative? *(Note: 'Internal' henceforth refers to inquiries or observations that were not asked of the user, but were recorded by team members to analyze user behavior.)*

VR questions - during gameplay

[Internal] About how long did the player use to finish the tutorial?

[Internal] What is the player doing during the time between tutorial instructions?

[Internal] Are there any signs of impatience or enjoyment during the tutorial? Ex. sigh, "this is so long," trying to skip ahead or "Wow," "cool," "I know how to do ___ now!"

Once the user has had some time in the headset, ask the following questions. How does the headset feel? How do you feel using the hand controllers?

Now that you’re in the VR environment, is it different from what you expected? If so, how?

How do you feel about moving through the environment? Do you feel nauseous at all?

[Internal] What features did the player not use, or have trouble with during the gameplay? (Choices: Right hand to select, Left hand to see clipboard, Point and press button to move)

[Internal] How do they act when they communicate? Are they calm, frustrated, ...

After Tutorial: Ask the player what they think the goal of the game is. Also ask them how they can complete that goal. Write their response below.

Tablet questions - during gameplay

[Internal] About how long did the player use to reach the midway ("Explore around!" or "Your partner has joined ..." instruction) of the tutorial? The end of the tutorial? (Please time the player.)

[Internal] Write down some observations while the player is doing the tutorial. Are they trying out the action after each tooltip?

[Internal] Are there any signs of impatience or enjoyment during the tutorial? Ex. sigh, trying to skip ahead, "this is so long," or "Wow," "cool," "I know how to do ___ now!"

After Tutorial: Ask the player what they think the goal of the game is. Also ask them how they can complete that goal. Write their response below.

After some time, ask the following questions. How do you feel about the controls?

How intuitive do you find the UI? Do things respond as expected?

Is there a tool or function not available that you wished you had?

[Internal] What features did the player not use, or have trouble with during the gameplay? (Choices: info mode, Beacon mode, Pinch to zoom, Rotate, Wheels, Beacon, Undo beacon, Reset, Clear, Disease information)

[Internal] Did the navigator finish the entire tutorial before the explorer or after the explorer?

Before:

After:

Other:

[Internal] How do they act when they communicate? Are they calm, frustrated, ...

Post-Testing

Did you find the experience engaging? Why or why not?

Was the tutorial engaging? Why or why not?

Was the tutorial informative? Why or why not?

How do you feel about the tutorial instructions in general? What was confusing, if any? What was clear, if any? (Please include whether you were a VR or tablet player.)

Did you understand how to move and operate the touch controllers or tablet from the tutorial?

Did you understand what was the objective of the game from the tutorial? Any comments?

Did you understand that you had to work with your partner from the tutorial? Any comments?

Did you feel that you had too short, enough time, or too long to learn how to navigate in VR or on the tablet? (Please include whether you were a VR or tablet player.)

Did you feel that the narrative helped you understand the experience?

Did you feel that you learned anything new about cells from this experience? How does this differ from images of cells that you may have seen in school or in the media?

Was having a partner useful, or do you think that could you have done everything by yourself?

What kinds of information did you share with your partner? What cooperation was necessary to meet your objectives?

Is there anything you would have changed or added to the interaction? What could have enhanced the multiplayer experience? (Tools, features, ...)