

Theory-Driven Design in Metaverse Virtual Reality Learning Environments: Two Illustrative Cases

Taehyun Kim¹, James Planey¹, and Robb Lindgren¹

Abstract—The metaverse entails modes of interactivity that hold enormous potential for education in various contexts and domains. The metaverse represents a convergence of technologies, such as immersive virtual reality (VR) and augmented reality, that allow for multimodal engagements with digital objects, virtual environments, and people. In this article, we focus specifically on the ways by which VR interactions with the metaverse can enhance learning and education. In our view, two primary issues have so far hindered the development of successful metaverse educational applications: first, the lack of theory-driven designs that use technologies such as VR in ways that are consistent with what we know about how people learn, and second, insufficient methods of evaluating metaverse technologies that go beyond usability and instead capture their efficacy for improving learning outcomes. To address these issues, this article aims to explore how three learning theories—experiential learning theory, distributed cognition theory, and embodied learning theory—can be applied to the design of educational VR, and how these theories can be leveraged to better support educational applications of the metaverse. We also introduce two science education VR environments developed in our lab that employ these theories to demonstrate how the design and development of the educational metaverse can be guided by research and evaluated on its ability to generate new learning.

Index Terms—Distributed cognition theory, embodied learning theory, experiential learning theory, metaverse learning design, virtual reality (VR).

I. INTRODUCTION

INTERACTIONS with the metaverse possess a wealth of educational potential [1]. The metaverse represents a convergence of technologies, such as virtual reality (VR) and augmented reality (AR), that allow for multimodal engagements with digital objects, virtual environments, and people [2]. Specifically, VR has recently attracted significant attention from educational researchers and designers [3], [4], [5] because it provides the potential for students to engage in a variety of learning environments in which they can directly interact with virtual objects, providing them with first-hand experience and is aligned with a learner-centered approach [6]. The metaverse,

Manuscript received 15 December 2022; revised 8 June 2023 and 14 August 2023; accepted 17 August 2023. Date of publication 22 August 2023; date of current version 14 December 2023. (Corresponding author: Taehyun Kim.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by University of Illinois at Urbana-Champaign Institutional Review Board under Application Nos. 22628 and 20163.

The authors are with the Department of Curriculum and Instruction University of Illinois Urbana-Champaign, Champaign, IL 61820 USA (e-mail: thk4@illinois.edu; planey@illinois.edu; robbld@illinois.edu).

Digital Object Identifier 10.1109/TLT.2023.3307211

catalyzed by emerging technologies such as extended reality (XR), is a vast, cohesive, enduring, and communal domain. It is seen as cyberspace's digital “big bang” and serves as an umbrella term for the immersive technologies that enable access to it [7]. However, research on these immersive technologies is still in its early stages, and due to the differences in affordances between VR and AR in educational settings, it is necessary to consider the distinctions between these two technologies from the design phase itself [8]. Thus, while we acknowledge that the educational metaverse will likely entail the full spectrum of XR technologies [7], we focus in this article specifically on the kinds of interactivity afforded in immersive VR because currently VR is the most available and the most researched metaverse component that has been applied in educational settings.

In considering whether and how new technologies are adopted in education, the perceived benefits of technology for learning are extremely important. If the affordances of digital learning environments and the technology that supports them are not communicated to learners, educators, and policymakers effectively, then even the existence of empirical studies supporting these enactments will have a muted impact. In an attempt to examine the issue of perceived affordance, Steffen et al. [9] introduced both AR and VR to 263 university students through both descriptions and short hands-on interactive demos. After experiencing both technologies, they were asked to rate a list of eight identified potential affordances that were most productive for each technology platform in relation to “physical reality.” The students’ ratings of VR were significantly higher (when compared to AR or physical reality) for dimensions related to reducing physical, emotional, and mental risks. VR was also rated significantly higher for the ability to recreate existing aspects of the physical world (reducing costs or allowing for more participation) and creating situations or representations that do not exist in the real world (the ability to depict the abstract and overcome limitations of space and time). It is important to note that the students were not asked about the affordances of the technology for a specific application and were simply rating their general perceptions. However, it is reasonable to assume that efforts to design VR applications in education will be more successful if they are more aligned with students’ perceptions of what kinds of learning is afforded by VR.

Approaching from an educator’s perspective, Minocha et al. [10] worked with 22 science and geography teachers to design a series of lessons and field trips augmented through the use of smartphone-based Google Cardboard VR viewers. While lower in the degree of body input and flexibility of

implementation when compared to higher fidelity VR systems, smartphone viewing platforms still provide some degree of bodily input through the ability to use the head and body to turn and attend to specific content displayed in a 360° sphere presented around the user. After enacting the lesson sequence with 521 students, Minocha et al. [10] asked the educators to reflect on the technology in a series of semistructured interviews. Out of these interviews, several major perceived affordances of VR were identified. Educators felt that the “visual authenticity” (the ability to observe an environment naturally), perceived spatial relationships, student agency, and immersion in the content were all enhanced through the VR lesson implementations. In addition, educators valued the ability to visualize content that was not accessible outside of a virtual world (visiting historical or dangerous locations, manipulating time and space, and visualizing the theoretical). This study shows that teachers, in addition to students, are already primed to leverage the affordances of VR in education even with relatively little experience with the technology. Studies such as these provide strong motivation to design, conduct, and communicate quality research on the specific ways VR experiences in the metaverse can serve as an effective learning tool.

Despite perceived benefits, educational VR research is still in its infancy, and most of the existing educational VR applications are not designed based on established learning theories [11], [12], [13], [14]. The lack of learning theory in the design of educational VR has led to inconsistent and difficult-to-interpret results in empirical studies. For example, some studies have found that the use of VR enhances learning effectiveness [15], [16], [17], [18], fosters learning engagement [19], [20], and promotes a positive attitude toward learning [21], [22], [23], [24]. However, the measures of learning and engagement tend to vary dramatically across these studies. Further, some studies have revealed a negative impact on learning, such as Rasheed et al. [25] who found that traditional teaching methods were more effective than VR for helping students retain factual information.

Fowler [11] points out that many VR implementation studies take a technocentric approach to design and analysis, focusing on increasing the fidelity of the environment and input systems over more pedagogical design approaches. This point is highlighted by Makransky et al. [26], who show that simply increasing the fidelity of a science lab simulation, without attempting to leverage the interactive affordances of VR, increases ratings of perceived presence, but decreases learning measured via pre-and postassessment. Makransky et al. [26] urge a more learner-centered approach to the design of VR science learning environments that prioritize pedagogical affordances over visual and technical fidelity. Furthermore, the evaluation of educational VR applications has primarily focused on their usability instead of learning outcomes [4]. The lack of concrete guidance for how VR can be used to enhance understanding and conceptual development means that the educational metaverse may unfold in ways that are not optimal for supporting students’ learning. We begin this article by examining the ways that prominent learning theories can be applied to the design and evaluation of educational VR.

II. THEORIES THAT HIGHLIGHT THE BENEFITS OF EDUCATIONAL METAVERSE ENVIRONMENTS

A. Experiential Learning Theory

Experiential learning theory emphasizes the role of experience in the learning process. In this theory, learning is defined as “the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience” [27]. Experiential learning encourages students to experience environments directly through exploration and discovery. The core idea of experiential learning is that the students play an active role in constructing knowledge through behavioral changes by interacting with the environment [28]. However, real experience sometimes can be expensive (e.g., practicing surgery skills), dangerous (e.g., experimenting with hazardous chemicals), or even impossible (e.g., diving into a human skin cell). Consequently, experiential learning theory supports the use of virtual experiences [29].

To apply experiential learning theory to the educational metaverse in applications such as VR environments, students should be able to perceive their experiences as authentic [27] while they are in the VR environment. Many scholars have agreed that one of the distinctive characteristics of VR is the “feeling of presence” [30], [31], [32]. Presence refers to the psychological state in which students experience the feeling of actually “being there” [31], [33]. This is a natural advantage of VR due to the 360 immersive virtual environments within which it situates its users. However, it is still important to consider how VR environments can be rendered more authentic. According to Kwon [34], when designing VR, designers need to allow the mediated environment to reproduce the real environment as it is and be able to interact with the mediated environment in real time. The more authentic a VR environment is designed, the more sense of “being there” the students feel. This has become increasingly possible since 2015 when commercial VR hardware with high resolution, high frame rate, lightweight, and low price was released [35].

However, VR is not always effective in creating authentic experiences that support cognition and lead to new learning. Kalantari et al. [36] evaluated participants’ cognitive test results across identical real and VR environments. They used a rendering tool to precisely replicate a real-world classroom setting in VR. Participants’ accuracy on the cognitive tests did not differ significantly between the VR environments and identical physical classrooms. Therefore, the value of VR environments likely does not come from digital replication of real-world experiences, rather it comes from the augmentation of the real world or creating perspectives that cannot be adopted in the real world because they are dangerous, expensive, or require multiple trials and errors [37]. For example, Carruth [38] conducted a study to improve workers’ understanding of a hazardous workplace by applying a VR application based on experiential learning, suggesting it was a safer and more cost-effective form of worker education than on-site training. Boedecker et al. [39] developed a VR environment that teaches surgeons about liver surgery procedures. Five highly experienced surgeons concluded that

VR was easier and more effective than conventional surgical practices using 3-D printed models.

Experiential learning theory gives us two promising design insights for the educational metaverse. First, immersive virtual learning environments such as VR should focus on augmenting or extending real world context or give them an experience that is not feasible in the real world. Because there is still a significant expense to developing VR content [40], designers need to be able to justify developing a virtual experience rather than simply having learners experience a phenomenon or activity in the real world. Second, VR learning environments should make every effort to give students a sense of “being there.” Giving students the option to actively participate in the simulation rather than simply observe, it is one method to make this achievable. One noteworthy aspect is that VR environments contain a variety of authentic 3-D models connected to the target learning content and enhanced with spatial information. This allows students to reduce their cognitive efforts while interacting with a learning experience, extending their learning from the mind to the environment. This extension is a core process of another prominent learning theory, known as distributed cognition.

B. Distributed Cognition Theory

Distributed cognition theory adds another layer to experiential learning by attempting to break down where and how cognition is occurring during a learning experience. When a learner interacts with their environment, cognition is not solely occurring in the brain but is distributed across an individual between the brain and their environment [41], [42]. The forms that distributed cognition can take are diverse, from social interactions with other learners, the creation of external knowledge representations (sketches or notes), to physically manipulating resources in the environment around a learner [43]. Broadly, when taking a distributed cognition perspective to the learning process, one must recognize the environment and the tools within it as critical components of a more complex learning model, moving beyond the assessment of a single learner’s mental model progression and acknowledging the role, a learner plays in the larger distributed learning ecosystem [44].

When examined from the perspective of digital environments for learning, distributed cognition theory can play a key role in shaping the design of environments by focusing on how digital interfaces and interactions can reduce cognitive load on the learner and support extended engagement [45], [46]. Hollan et al. [45] highlight distributed cognition’s role in shaping technology interactions via their development of an integrated framework in which distributed cognition is the foundation of repeated cycles of observation, theory, and design. This cycle relies on four core principles drawn from distributed cognition:

- 1) people establish and coordinate structure in their environment;
- 2) effort is spent in maintaining this structure;
- 3) people off-load cognitive effort to the environment whenever possible;
- 4) cognitive loads can better be managed via social organization.

Mcveigh and Isbister [47] refresh this approach as they address the unique affordances for enhancing distributed cognition via VR social interaction spaces. They envision VR spaces that augment not only the individual’s interaction with the environment (e.g., manipulating parts of the environment, embedding media, and creating content via sketching or modeling) but also a user’s social interactions with others (e.g., nonverbal body cues, gestural communication, and facilitating feedback).

When considering distributed cognition’s contribution to the design of metaverse learning environments, designers must critically examine the dimensions of interaction agency and interaction visibility as they develop experiential and distributed learning experiences. Designing for interactivity that facilitates distributed cognition means the careful consideration of not only how a learner interacts with structured and scripted elements of the environment (e.g., selecting a highlighted plant in a virtual forest to add it to an interface-based inventory) but also how the added agency of interaction can productively support how a learner can use the environment to store, modify, and express their learning (e.g., allowing users to manually lay out their collected plants to organize them by shared characteristics). Here, the challenge is designing to find a balance between high interactive agency (with a potentially deleterious lack of learning support) and tightly structured and scripted linear interaction progressions that overly restrict agency. Fully leveraging distributed cognition also means that the visibility of both the learner and their interactions with the environment are critical for multiuser environments. This can take the form of avatars that still connect directly to the body of the user (such as mapping hands to controller positions to facilitate their use in pointing and gesturing), systems that allow for the creation or annotation of elements within an environment, or dynamic changes to an avatar that are reflective of the current state of a user (e.g., visible tools held by an avatar changing to inform other users what task is being performed). In both of these contexts, the interactions are inherently mediated by the body of the user when explored via VR, and many examples of design influenced by distributed cognition also leverage embodied theories of cognition and learning.

C. Embodied Learning Theory

There is a wealth of research on the positive relationships between learning and embodied actions, such as gestures. Piaget [48] was an early proponent of the idea that sensorimotor behavior helps build knowledge and that body actions are neither independent from the brain nor purely unidirectional. More recent psychologists and learning scientists have revealed that gestures play a critical role in thinking and reasoning (e.g., Goldin-Meadow [49], McNeill [50], and Roth [51]). Research has since corroborated this stance, showing that cognition is deeply rooted in the relations between gestures with the world and our sensorimotor perception [52], [53].

Adding a motoric modality to the learning signal can activate more neural paths, which can make learning signals or memory tracking stronger. Accordingly, some researchers have argued that integrating gestures into learning behavior should

strengthen memory traces [54], [55]. There is also a direct connection between individuals' embodied interactions and their development of mental models. Nathan et al. [56] demonstrated how the use of embodied gestures helped students develop mental models around science reasoning, and, perhaps more critically, the restriction of gestures impaired their ability to make inferences about the content they were learning. As such, the activation of motor control systems and embodied interactions can be seen as a critical component for the development of mental models and sensemaking.

In addition, Stieff et al. [57] showed that students achieved better learning outcomes when they actually executed the relevant gestures rather than simply observing them. This study conducted an experiment with organic chemistry undergraduate students attempting translations between organic chemistry molecular representations. Here, the researchers randomly assigned 70 undergraduate students to one of three conditions: 1) control text-only group, 2) observed gestures group, and 3) observed and executed gestures group. Among the three groups, the group that observed the experimenter's gestures and then imitated them showed the most effective learning outcomes. This may indicate that students exhibit better learning outcomes when they can execute their gestures compared to simply seeing others' (e.g., teacher, peer) gestures.

This finding shows us the possibilities of metaverse VR from the perspective of embodied learning theory. With more natural user interfaces emerging, learners can move their whole bodies without restriction to their interactions with the VR learning environment. In particular, recently released VR devices (e.g., Oculus Quest 2) allow users to perform gestures that are similar to those in the real world through internal tracking sensors and controllers. Embodiment in VR can be particularly compelling compared to less immersive media because students can control their avatars with body movements, and actual body views are replaced by virtual bodies [58]. Therefore, although VR is in the early stages, some researchers have indicated its effectiveness, which is designed based on embodied learning theory.

Price et al. [59] designed a VR environment in which elementary school students could learn the concept of the Cartesian coordinate system. The aim of this simulation was for two students to collaborate to reach target coordinates to collect flowers. One student, who uses VR, moves their body to go to the right coordinates. The other student, who uses a 2-D monitor, helps the VR user by stating the coordinate values while observing the movements of the VR user on a 2-D monitor. In this environment, the VR student's body becomes a tangible resource for thinking, learning, and joint activities through embodied experience and where body movement, position, and orientation are made visible to collaborators. The results showed the students used their bodies as metaphors to connect different expressions (between VR user and 2-D monitor user) and to successfully master the concept of the Cartesian coordinate system.

Embodied learning theory gives us two design insights for VR learning environments. First, VR learning environments should facilitate users' natural bodily movements within a simulation. It is important to provide an opportunity for learners to move

their bodies and interact with virtual learning objects in VR environments as a means of forging connections between their motor system and new knowledge. Second, designers should ensure that a learner's bodily movements in VR environments can be meaningfully connected to target learning outcomes. Not all gestures will necessarily help students' learning (e.g., moving one's hand back and forth from their body likely will not help students understand the mechanics of gears), and thus, designers should give careful consideration to how a particular body action in VR will help or hinder student understanding of novel learning content. In many cases, this may entail guiding the learner's body to act as if it has become part of the system they are learning about, or helping a learner express an important process or mechanism through gestures.

D. Summary of Design Insights From Theories

We explored three learning theories that are worth considering when designing educational metaverse VR environments. Experiential learning theory focuses on how learners can perceive a designed VR environment as an authentic experience and how this might help them to learn specific content through that experience. To that end, it is critical to understand and leverage the affordances of VR interactions that lead to an experience that feels authentic and meaningful. Furthermore, designers should consider how a designed VR environment can extend beyond the limitations of a real experience. Distributed cognition theory focuses on how various resources existing within a VR environment provide a medium to facilitate learners' cognitive processes. Distributed cognition assumes that how learners interact with resources (e.g., knowledge presentations, virtual objects) in the VR environment affects learners' cognitive processes. To this end, it is necessary to find a productive balance between giving users the agency to interact in ways meaningful to their own learning and structuring high-quality learning experiences. In a multiuser environment, it is important to design learners' interactions to be recognizable to other learners, enabling the distribution of learning not just with the environment but between users in the environment. Finally, embodied learning theory focuses on how learners' bodily movements affect their learning by embedding knowledge within the sensorimotor system. When designing interactive systems, such as metaverse VR, elicited body movements, and gestures should be aligned with the learning content. Table I summarizes the design insights derived from the theories discussed above.

III. THEORY-BASED DEVELOPED METAVERSE VR LEARNING ENVIRONMENTS

The following projects will be used to highlight the importance of specific theoretical approaches to the design of metaverse learning environments. Each project takes a different approach to communicating learning and user interaction. Comparing and contrasting these projects will provide insight into some of the ways that specific theoretical foundations can inform design and assessment.

TABLE I
THEORY-DRIVEN EDUCATIONAL METAVERSE VR DESIGN INSIGHTS

Learning Theory	Design Insights
Experiential Learning Theory	<ul style="list-style-type: none"> - Leverage VR affordances to create virtual experiences that are authentic and meaningful - Position learners as active participants in the simulation
Distributed Cognition Theory	<ul style="list-style-type: none"> - Balance interaction agency (allowing users more control over how they interact and interpret content) - Make interactions perceivable to a learner as well as other users (facilitate social dimensions of distributed cognition)
Embodied Learning Theory	<ul style="list-style-type: none"> - Allow learners to freely use their bodies in the simulation - Align learners' bodily movements in the simulation with learning goals of the simulation

A. Metaverse VR Learning Environment 1: ChromosoME

ChromosoME was developed after investigating two existing studies on similar biology VR learning environments. First, Parong and Mayer [23] investigated students' learning outcomes through a media comparison study using a predeveloped VR simulation and PowerPoint Slides with the same learning contents. They used a biology VR simulation, called The Body VR: Journey Inside a Cell [60], which contained narration and immersive animations of the circulatory system and components of cells. In this simulation, students mostly played the role of observers as the simulation unfolded, and their interactions with the virtual objects within the simulation were limited. For example, students could touch virtual objects, but the touched objects only moved in the direction they were pushed and then returned to their original place. In other words, interaction with the virtual objects in the simulation did not greatly impact students' understanding of the learning content. In addition, the simulation automatically proceeded along with the included narration, regardless of the student's actions. According to the results, students who learned through the VR lesson performed significantly worse on transfer tests, expressed higher emotional arousal, showed more extraneous cognitive load, and reported less engagement based on electroencephalogram measures than those who learned through the PPT slides, with or without practice questions. Therefore, they argued that VR simulations created high emotional arousal and cognitive distractions, leading to poorer learning outcomes than traditional instruction methods.

The other study referenced, by Nasharuddin et al. [61], conducted a pre- and post-test analysis to compare students' knowledge about the human cell division process before and after using a mobile-based VR learning environment. This learning environment had three modes: "Open Note," "InCell VR," and "Mini Game." In "Open Note" mode, students could add or edit notes on their mobile devices, while in "Mini Game" mode, quizzes, including jigsaw puzzles, were used to measure students' learning outcomes. These two modes did not use VR devices but instead used mobile phone screens. In "InCell VR" mode, students could examine 3-D images in a mobile-based

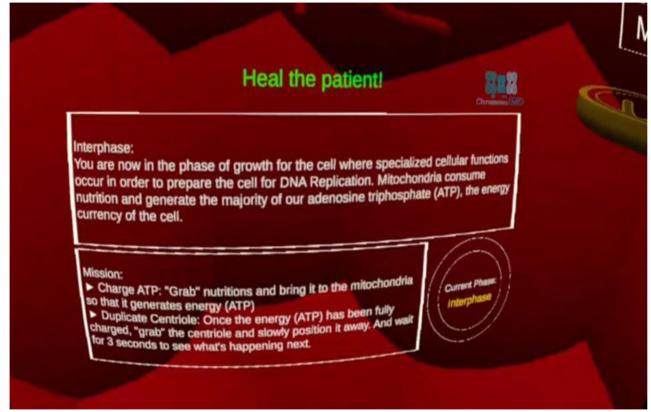


Fig. 1. Information panel providing phase descriptions and missions for the current stage.

immersive VR environment. Students were not allowed to interact with the simulation in this VR learning environment; instead, they could look around in 360° views within the 3-D-modeled cell. Their pre- and post-test results showed that there is a significant learning gain before and after using all the modes of learning environment.

Despite the mixed findings, these two studies indicate that both VR learning environments have room for improvement by applying the theory-driven design insights presented in Table I. Learners in both VR environments primarily assumed the role of observers, rather than having an active role in the simulations. They had limited options for interacting with virtual objects, and there is room to enhance the relationship between these interactions and the learning contents. Therefore, ChromosoME deals with learning content similar to the above two VR environments, but it was designed and developed in consideration of the VR learning environment design insights we described above.

ChromosoME is an interactive simulation of the cell division process powered by Oculus Quest, a VR hardware that allows full hand and head tracking without any other external sensors. The simulation presents a formalized representation of the cell division process. The aim is to make the learner's hands become a part of the actual cell division process in order to participate in what they are trying to understand. While users are interacting with the simulation, an accompanying interface offers a description of the current phase and the associated missions that students need to complete (see Fig. 1). The content displayed in the information panel is automatically updated whenever the phase changes.

Table II lists the series of missions and gestures that the learners need to perform in the simulation. The simulation begins at a site of cell division in a human patient. The primary objective of a user is to "heal the patient" that has sustained a cut to their arm. To achieve this goal, the user needs to use gestures that map to the cell division process so that chromosomes in a single cell can successfully divide into two identical daughter cells through the cell division process of mitosis. Users are required to use their hands according to tasks and can guess how to move by looking at the movements of other scripted objects (objects that have predefined animations and interactions) or reading mission



Fig. 2. (Left) Grab and detach the chromatids with both hands. (Right) Spread the arms and place chromatids at each node.

TABLE II
LIST OF ACTIVITIES IN THE CHROMOSOME SIMULATION

Phase	Mission	Gesture
Interphase	- Grab nutrients and give them to the mitochondrion so that it generates energy. - Duplicate centriole.	- Grab nutrients (e.g., protein) and bring them to the mitochondrion. - Grab the centriole with both hands and slowly move your hands apart.
Prophase	- Convert your hand into a pair of chromosomes.	- Grab and hold the DNA for about 3 seconds.
Metaphase	- Move the chromosomes to the equator after the spindle fibers are automatically attached to each chromosome.	- Move both hands to position the chromosomes in the center of the cell.
Anaphase	- Separate the paired chromosomes into anodes.	- Move both hands apart and spread your arms as wide as possible (see Fig. 2).
Telophase	- Separate the mother cell into two identical daughter cells.	- No gestures are required (simulation will play automatically).

and cell status in the interactive user interface. Therefore, in this simulation, there is a set of correct gestures that users need to perform to complete a process of cell division. In addition, students must complete the cell division process three times to completely heal the wound. This is done for two reasons: 1) to give students time to adjust to the simulation and VR device because they may be unfamiliar with both and 2) to reflect the need for numerous cell division processes to treat wounds because cell division is a persistent phenomenon in our bodies. We developed the simulation based on a design-based approach [62] while collaborating with a local high school teacher.

We carefully considered theory-based design insights in the simulation development process and Table III details how they were applied to the simulation.

B. Metaverse VR Learning Environment 2: Immersive Production of Representations Exploring Science Sketching (ImPRESS)

The ImPRESS project aims to combine the learning benefits of drawing representations of science knowledge with the

TABLE III
HOW CHROMOSOME FITS INTO OUR DESIGN INSIGHTS

DESIGN INSIGHTS	How the Insights Applied to Our Simulation
Leverage VR affordances to create virtual experiences that are authentic and meaningful	- Learning with authentic tasks. - Let learners feel they are really located in the human skin cell.
Position learners as active participants in the simulation	- Let learners decide their learning pace. - Give them a specific role to perform.
Balance interaction agency (allowing users more control over how they interact and interpret content)	- The simulation does not interfere with when and how the learners perform the gestures as long as they follow the core concepts. - Provide selective virtual objects that learners attend to in each phase.
Make interactions perceivable to a learner as well as other users (facilitate social dimensions of distributed cognition)	- Current version did not integrate social aspects. - In a future version, scripted object cues can be turned into other students' interactions.
Allow learners to freely use their bodies in the simulation	- Allowing learners' full-body interactions using wireless controllers.
Align learners' bodily movements in the simulation with learning goals of the simulation	- Let learners' bodies become a part of the learning contents (e.g., their hands converted into chromosomes). - Learners' gestures are closely related to learning concepts.

spatial affordances of a VR environment. Drawing as an act of representing one's knowledge has been a focus for learning researchers for some time [63]. Recently, there has been a resurgence of interest in drawing's role within science education, as well as the potential role technology has to play in facilitating the drawing process [63], [64]. A recent review of the drawing and technology-assisted drawing literature found that, while drawers do need support to facilitate the drawing process, the strongest evidence for learning benefits still originate from the process of hand drawing or sketching [65]. As an act of distributed cognition, drawing allows a learner to offload mental representations and conceptual relationships to the environment (paper, whiteboard, digital drawing). While this process has been shown to be valuable for helping learners assess their knowledge and synthesize new knowledge, the majority of research around the value of drawing has been done via 2-D mediums [66], [67]. When the learning content has a strong component of spatial complexity, the benefits of drawing representations can begin to fade as learners grapple with the challenges of how to represent spatial information within their 2-D drawings. It is at this intersection of sketched knowledge representations and spatially complex learning content that VR-based (tracked headset and controllers) drawing programs have the potential to allow a learner to fluidly engage with spatial dimensions as they create their drawings (see Fig. 3). Drawing in VR further unleashes the distributed cognition benefits of drawing to learn, as drawing elements can now be placed in space around the user creating a more embodied distributed model of their knowledge.



Fig. 3. Student leveraging the ImPRESS drawing platform in class with their work visible on the display in the background.

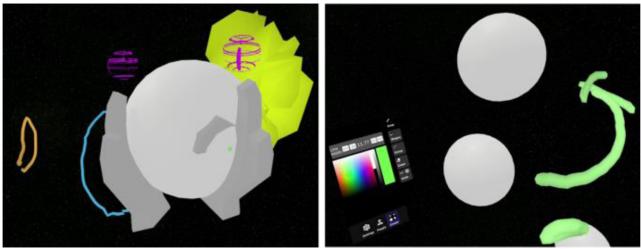


Fig. 4. Two methods of environment interaction in ImPRESS. (Left) Using button presses and gestures to manipulate drawn objects (here grabbing and pulling to resize a sphere). (Right) Using the interface to select tools and change the line color.

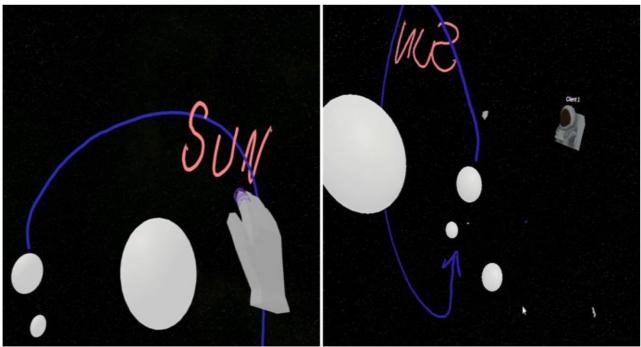


Fig. 5. Student working in ImPRESS viewed via their headset view (Left) and the simultaneous view of an observer (Right).

While there are currently numerous programs that support drawing and 3-D modeling in VR, ImPRESS has been designed and refined to address two main goals: 1) streamline a particular set of features most conducive to the creation and revision of spatially-complex science explanations and 2) facilitate the sharing and cocreation of VR knowledge representations. These goals were established after a series of pilot drawing sessions utilizing off-the-shelf VR drawing software and are described in the following section.

To address design goal 1) a priority was placed on maximizing the connection of embodied inputs to the creation and spatial manipulation of drawn content. Visually, this takes the form of the user's controller input represented as a pair of virtual hands (see Figs. 4 and 5, left). When drawing is initiated, the index finger

is pointed outward and a purple context icon appears above the finger at the point of interaction. When grabbing is initiated, the other four fingers on the hand curl and a small green dot on each hand can be used to help position the hands to manipulate small objects. Actions such as drawing, moving, and resizing elements are all completed via quick button and controller motion actions without the need to interact with the 2-D user interface (see Fig. 4, left). This is in part achieved via context-aware interaction with the controllers. For example, holding down one controller trigger button and moving the controller through VR space draws strokes in the environment. However, with both the left and right controller triggers pressed, drawing is automatically disabled and the user is able to rescale (push or pull controllers), rotate (move controllers around a midpoint), and move (slide both controllers left, right, up, or down) the whole environment. A return to drawing simply requires the release of one controller trigger. Similarly, for moving or resizing single objects in the environment, the learner leverages the VR representation of their hands to grab (reach into the object then hold a controller button) and resize (reach and with both controllers simultaneously, then move the controllers closer or further apart). This allows the user to quickly construct and begin to modify and manipulate a drawn explanation fluidly with their body and controller motions as the primary means of environment manipulation. When the UI must be accessed, it is presented as needed via holding down a button to allow a drawer to change line color and thickness, erase, undo, group, and place 3-D shapes and the menu is automatically minimized when the button is released (see Fig. 4, right).

While prior literature around the value of drawing has shown that some of the most notable learning effects are from primarily hand-sketched representations [65], during the pilot sessions some users reported difficulty generating 3-D solids that they felt were satisfactory representations of the concepts they were communicating (such as spheres to represent the Earth, Sun, and Moon). Here, while the VR interface was affording a more natural progression of creating, reaching, grabbing, and moving objects to communicate spatial information, the act of sketching applied in a 3-D environment adds a layer of complexity to the initial creation of conceptual representations. In order to be able to support drawers who encountered difficulty in the initial steps of the creation of a VR knowledge representation, ImPRESS also contains a "shapes" tool that allows a drawer to place spheres, cubes, cylinders, capsules, or planes as scaffolding for their hand-sketched content. All shapes are placed via the same input sequence as the draw tool, and can then be manipulated or resized with the same input actions as sketched content (the object being manipulated in Fig. 4 (left) is a sphere from the shape tool). These shapes have the potential to move a drawer past a communicative block of generating specific spatial representations while still allowing for the incorporation into a larger freehand sketch (Fig. 5 is an example of sketching in tandem with spheres to generate a representation of the Earth, Sun, and Moon).

To achieve design goal 2) ImPRESS has also been designed to leverage WebXR standards to deliver the VR experience. This means that multiple users can simultaneously engage with the construction of a VR knowledge representation, as well as enable non-VR users to observe and navigate the drawing space

TABLE IV
HOW IMPRESS FITS INTO OUR DESIGN INSIGHTS

DESIGN INSIGHTS	How the Insights Applied to Our Simulation
Leverage VR affordances to create virtual experiences that are authentic and meaningful	<ul style="list-style-type: none"> - Pilot work with existing VR drawing platforms identified areas for improvement. - Drawing as a productive process for engaging in learning has a strong body of literature. - Spatial affordances of the VR system go beyond 2D drawing or physical object manipulation.
Position learners as active participants in the simulation	<ul style="list-style-type: none"> - User interaction is directly linked to the value of the system, students create the environment content. - Learners can engage directly with others in the drawing space or verbally with observers.
Balance interaction agency (allowing users more control over how they interact and interpret content)	<ul style="list-style-type: none"> - Value of the system relies heavily on user agency as learners construct knowledge representations and communicate with others. - Combining the benefits of freehand sketching with the availability of simple 3D models as drawing supports.
Make interactions perceivable to a learner as well as other users (facilitate social dimensions of distributed cognition)	<ul style="list-style-type: none"> - Visual feedback via avatar finger positions and interaction icons. - Tracked avatar hands allow for gestural communication with other users. - All in-process actions (drawing, erasing, etc.) are visible to others.
Allow learners to freely use their bodies in the simulation	<ul style="list-style-type: none"> - Minimal restrictions on how users can navigate the drawing space with their bodies (walking/crouching) or the tracked hand controllers (object or whole world position and scaling).
Align learners' bodily movements in the simulation with learning goals of the simulation	<ul style="list-style-type: none"> - Structured gestural actions are linked to the process of scale and perspective changes at both the object and world scale.

via a web browser (see Fig. 5). All users' avatars and hand positions are updated in real time, as well as the finger positions communicating drawing or grabbing.

In its focus on user-generated objects to construct the VR experience, ImPRESS is one example of how immersive learning can take place using more general, creativity-focused tools. While other programs exist to support the creation of immersive art, the design decisions behind ImPRESS focus on the value of the system for sketching knowledge explanations (streamlining features and facilitating sharing and cocreation). Here, the focus on maximizing embodied input allows users to fluidly engage in drawing, object manipulation, and environment-level changes in scale and perspective. This then enables a variety of novel and productive interactions as students engage in the construction and exploration of their knowledge representations. While other mediums might address a subset of the affordances of immersive drawing (sketching and revising quickly via whiteboards or manipulating physical models), ImPRESS smoothly integrates the benefits of having student-generated drawings of their learning while unlocking the scale and perspective freedom of VR. Table IV lists how ImPRESS aligns with our established design insights for effective metaverse learning environments.

IV. PILOT STUDY ENACTMENT

In this section, we aim to present two pilot studies conducted for each of the metaverse VR simulations. For ChromosoME, we executed a pilot study to explore the potential effectiveness of a VR simulation in a learning context, developed based on our proposed theory-driven design insights. As for ImPRESS, we analyzed the learning effects facilitated by the 3-D drawing tools of pre-existing off-the-shelf software, Tilt Brush. Following this, we aim to detail the process of identifying participant approaches to VR drawing and the potential affordances of the technology as justifications for the ImPRESS design priorities. Although these simulations have been developed relatively recently, and despite the obstacles in conducting comprehensive in-classroom controlled studies due to the pandemic, we intend to demonstrate the potential of our simulations through lab-based pilot studies conducted with a small dataset.

A. ChromosoME Pilot Study

The pilot study was conducted with undergraduate students from a midwestern University who were enrolled in a non-STEM course. Participants were recruited and received one course credit as compensation. A total of seven students (five males, two females) participated in the pilot research. The research procedures were as follows. First, the students were given a paper-based material about cell division, which they studied independently for approximately 5 min. This paper-based material contained the exact same information presented in the ChromosoME simulation. After studying the paper material, the students were interviewed and asked to describe the process of cell division as detailed as possible based on what they learned (Pre-VR simulation interview). Following this, they engaged with the ChromosoME simulation, and after playing, they were again asked to describe the process of cell division as though they were explaining it for the first time (Post-VR simulation interview).

We evaluated these students' descriptions of cell division based on seven key knowledge elements (both the paper-based material and the ChromosoME simulation addressed these key knowledge elements, which were selected during the simulation development phase in collaboration with local high school teachers). We assessed how many of these key elements were explained by the students during both the pre- and post-VR simulation interviews. Table V summarizes the results, which showed that students' overall explanations of the key elements increased. For instance, for key element 1, only one student explained this aspect before playing the VR simulation, but all seven students mentioned it after engaging with the VR simulation. With the exception of element 3, more students were able to explain the key elements after playing the VR simulation compared to before. Therefore, we observe a trend indicating that students have learned about important elements of cell division in a metaverse VR learning environment designed based on our design insights.

However, as mentioned above, while ChromosoME simulation generally demonstrated a tendency to promote learning of key elements, the opposite result emerged for specific learning

TABLE V
COMPARISON OF STUDENTS' KEY ELEMENTS EXPLANATION COUNTS: PRE- AND POST-VR SIMULATION PLAY

Key Elements	Counts of Before VR Play	Counts of After VR Play
1. Mitochondria consume nutrition and generate energy to prepare cell division.	1	7
2. DNA condenses as a chromosome shape.	5	6
3. The spindle has captured all the chromosomes.	3	2
4. Line all the chromosomes up at the metaphase plate (middle of the cell).	5	7
5. Chromatids separate from each other.	3	5
6. Chromatids pulled towards opposite ends of the cell.	1	3
7. Cell divides into two identical daughter cells.	6	6

elements, particularly element 3. We intend to discuss this further in Section V.

B. ImPRESS Pilot Study

The pilot study that informed the design priorities of the ImPRESS environment was conducted with undergraduate students at a midwestern university and utilized the off-the-shelf VR drawing software Tilt Brush. While in-person data collection was not possible due to COVID-19 mitigation steps at the time of data collection, an attempt was made to preserve as much as possible the interactions between a drawer and facilitator as the participant engaged in a prompted exploration of their knowledge around the lunar phases. To facilitate this, VR equipment was dropped off at the residence of a participant while the facilitator remained outside in a vehicle and connected via a video meeting. After an introduction to the VR equipment, the facilitator then connected wirelessly to the participant's VR headset to launch and monitor the drawing session.

Each session was audio recorded and screen captured from the perspective of the VR user. Participants were first introduced to the VR equipment and drawing software. Once comfortable, participants then completed two nondomain training tasks to verify their use of the drawing controls. The first training task focused on basic checks of headset comfort and foundational drawing procedures (creating strokes, color change, erasing, and brush size). The second task introduced more complex manipulations such as grabbing and moving strokes, rescaling drawn items and environment, and navigating within the VR space without physical movement. After training was complete (approximately 15 min), the participants then received the lunar phases drawing prompt. For the prompt, participants were asked to explain the mechanism behind the change in lunar phases while drawing and thinking out loud.

In total, seven sessions were conducted, with four female and three male participants. The age range of all participants was 19–22, and of the seven, only three had previous experience with VR (two with some, and one a frequent VR user).

TABLE VI
COMPARISON OF STUDENTS' KEY ELEMENTS EXPLANATION COUNTS: PRE- AND POST-VR SIMULATION PLAY

Observed affordances	Description	Count of Participants
Moving drawn elements during construction	Grabbing and moving a drawn element	6
Changing the world scale and world position	Changing the world scale to increase or decrease the perceived size of the whole drawing environment and to move through the drawing space	6
Initially sketching in 2D with 2D elements arranged in 3D space	Sketching along a single plane as if on paper (often of a textbook-style representation)	5
Revising 2D elements into 3D	Initial constructions are fleshed out into 3D volumes or erased and replaced	3
Scaling and/or positioning drawn elements to align with a target perspective	A target perspective is designated (such as Earth-based) and is used as a reference to scale and move drawing elements	3
Resizing drawn elements	Selecting and rescaling an element after it has been drawn	3
Immediately sketching with 3D shapes	Using the drawing tools to sketch approximations of spherical volumes for representation of the earth, moon, and sun	2
Manipulating drawn elements to illustrate a dynamic process	Grabbing pieces of a drawing and moving them to illustrate motion, orbits, or other dynamic relationships	1

To begin a high-level analysis of the kinds of approaches and potential affordances of the VR drawing system, session recordings were reviewed holistically in order to generate a collection of observed actions and processes around the construction of a lunar phases knowledge representation. Table VI lists these actions and processes, and how many participants used them. After the initial prompt, participants took multiple approaches to initiate the construction of their lunar phases knowledge representation. A majority of the participants (five) engaged in a process of first creating 2-D sketches within the drawing environment (e.g., rings to represent the Sun, Earth, or Moon), often as acts of recall of prior knowledge and “textbook figure” representations of the lunar phases. Of the five participants that initially engaged in the 2-D representation construction, three participants went on to revise their 2-D elements into 3-D. The remaining two participants immediately began their drawings with 3-D shapes.

Broadly, the general spatial affordances of VR were used by a majority of the participants. Changing the “world scale” (uniformly shifting the scale of all items at the same time) was performed by six participants. This was frequently used to move in and out of the drawing to facilitate its construction but also to take in an overview of their work in preparation for revisions. World scale changes frequently occurred alongside rotation and translation within the environment (six participants). The majority of participants (six) grabbed and repositioned elements of

their drawing during construction; however, resizing of elements after they were drawn was only leveraged by three participants. While this was a productive process for these participants as they revised their construction, the fact that the ability to resize a drawn element required multiple steps to initiate (switching to “select” mode, activating “resize” mode, then using buttons and controller movements to resize) could have limited its use. Finally, one participant moved drawn elements not only to revise their work but also to use them as manipulatives in illustrating the orbital path of the Moon around the Earth.

V. DISCUSSION AND FUTURE DIRECTION

In this article, we argued that it is important to design metaverse VR learning environments informed by learning theories and following design insights grounded in those theories. To demonstrate how design insights can be derived from learning theories, and how these insights can then be applied to metaverse VR learning environments, we presented two examples of educational VR applications. ChromosoME requires learners to perform a specific set of gestures and students can learn about the cell division process while interacting with the simulation, whereas ImPRESS focuses on supporting students’ ability to explain their understanding of scientific phenomena through 3-D drawing. Despite the different aims and learning content of these two applications, theory-driven design insights are well integrated into our simulations, as described in Tables III and IV.

First, in the case of experiential learning theory, the focus is on how to derive effective learning outcomes when designing learners’ experiences in a VR learning environment. ChromosoME offers a unique experience to learners by positioning them within human skin cells and converting their hands into chromosomes to interact with other cell components. ImPRESS empowers learners with the experience of 3-D drawing, with spatial freedom that is difficult to replicate outside of an immersive virtual environment. Design insights derived from experiential learning theory encouraged us to articulate what experiences learners will encounter in the metaverse VR learning environments, and how those experiences will cultivate student learning.

Second, distributed cognition theory explains how the surrounding environment can support learners’ cognitive efforts. In applying this theory to the design of a metaverse VR learning environment, it is important to consider when and how to make virtual objects available within the VR environment to help the learner–environment interaction processes. For example, ChromosoME provided only selective interactable virtual objects around the learners at each stage of cell division, allowing the learners to determine when and how they interact. ImPRESS lets learners use simple 3-D models to support their drawing (knowledge construction), and all of their interactions are visually shared with others in real time. Design insights derived from distributed cognition theory encouraged us to consider the effects of when and how learners interact with (or be aware of) virtual objects in metaverse VR environments and the implications these effects have on the learning process.

Finally, embodied learning theory suggests that our bodily movements play an important role in shaping our cognitive

processes and that our thoughts and ideas do not exist independently of our physical experiences. Therefore, both of the metaverse VR learning environments described here allow learners to interact with virtual objects using whole body movements. In addition, to make learners’ bodily movements more meaningful, ChromosoME allows learners’ virtual hands to be converted into chromosomes, which then take on important roles in the cell division process. This allows learners to perceive their bodies actually becoming an important part of the learning concepts. ImPRESS also puts structured gestural actions which are linked to the process of scale and perspective changes at both the object and world scale. Design insights derived from embodied learning theory encouraged us to clarify how the learners’ bodily movements would affect their learning within metaverse VR environments.

However, we are not suggesting that our design insights should necessarily be considered in the development of VR simulations for all learning content and situations. As evidenced by the ChromosoME pilot results presented earlier, key element 3 appeared to be overlooked and not explained by students after learning through VR. This implies that learning outcomes may vary depending on the specifics of the learning content and what we prioritize in the simulation development process.

To elaborate, key element 3 in ChromosoME requires students to understand that spindle fibers connect all chromosomes, enabling them to move during the metaphase and anaphase stages. However, as shown in the metaphase task in Table II, students do not actively connect the spindle fibers to the chromosomes; instead, the simulation is designed to automatically establish this connection. This means students can learn about this process by reading the informational panels within the VR simulation, but the simulation automatically executes this action without requiring students to interact with virtual objects or pay specific attention.

This design decision was taken primarily for safety reasons, as it enables students to play while remaining stationary, given the significant distance between the starting point of the spindle fibers (the centrioles) and the position of the chromosomes within the cell. ChromosoME was developed in collaboration with teachers for use in local high school classrooms, and we designed it to minimize unnecessary movement within the VR environment for safety considerations. Consequently, we chose not to strictly adhere to our design insights for this key element, prioritizing student safety instead. In such instances, it might be necessary to provide additional supportive materials to ensure students have adequately learned these aspects following the VR simulation play.

In the future, discussions on how to assess learning processes in metaverse VR learning environments developed based on clear learning theory and design insights will be critical. The above-mentioned three theories have common aspects. First, they all highlight students’ role as active learners, and second, the way students experience the learning environments is not from a single modality (e.g., text, audio, video, gesture, etc.). Therefore, compared to traditional learning environments, VR-based metaverse learning environments have extended to include wider sources of student interaction data. Various researchers argued

that the synchronization and combining of data from several modalities enable a more thorough analysis of the student's learning cues in their associated research [68], [69], [70]. In addition, collecting multimodal data became feasible with the advent of recent sensor technologies and technological advances. However, how to analyze multimodal data remains a substantial challenge because of the complexity of multimodal interactions [71].

Multimodal learning analytics (MMLA) is one of the novel analysis methods that aims to understand the learning processes from collected multimodal data [68]. The purpose of MMLA is to harness the affordances of multimodal sensors and computational analysis to obtain a better sense of students' learning. MMLA includes principled but new methods for collecting, analyzing, coordinating, and presenting visual, aural, gestural, spatial, linguistic, and other data in online and offline learning environments when students participate in educational tasks [72], [73]. MMLA can provide more complex and rich insights than traditional learning analytics that are highly structured and have limited student–system interaction, especially when it comes to metaverse VR learning environments [68]. Therefore, in the future, research should be conducted on how MMLA methods can be applied to metaverse VR learning environments to assist our understanding of students' learning processes. Leveraging learning theories to justify and assess design elements of metaverse learning environments as demonstrated with ChromosoME and ImPRESS is a process that will be critical for the growth of meaningful and authentic learning within the metaverse.

ACKNOWLEDGMENT

The projects described in this article are the result of support provided by the Technology Innovation in Educational Research and Design (TIER-ED) at the University of Illinois at Urbana-Champaign. ChromosoME was supported via the TIER-ED Graduate Student Fellowship Program, while ImPRESS was supported by the TIER-ED Pilot Project. The authors would like to thank Dr. M. Tissenbaum for his invaluable support on both projects.

REFERENCES

- [1] J. Wu and G. Gao, "Edu-metaverse: Internet education form with fusion of virtual and reality," in *Proc. 8th Int. Conf. Humanities Social Sci. Res.*, 2022, pp. 1082–1085.
- [2] S. Mystakidis, "Metaverse," *Encyclopedia*, vol. 2, no. 1, pp. 486–497, 2022.
- [3] D. Checa and A. Bustillo, "A review of immersive virtual reality serious games to enhance learning and training," *Multimedia Tools Appl.*, vol. 79, no. 9/10, pp. 5501–5527, 2019.
- [4] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design Elements, lessons learned, and research agenda," *Comput. Educ.*, vol. 147, 2020, Art. no. 103778.
- [5] C. Snelson and Y.-C. Hsu, "Educational 360-degree videos in virtual reality: A scoping review of the emerging research," *TechTrends*, vol. 64, no. 3, pp. 404–412, 2019.
- [6] K. T. Henson, "Foundations for learner-centered education: A knowledge base," *Education*, vol. 124, no. 1, pp. 5–16, 2003.
- [7] M. Wang, H. Yu, Z. Bell, and X. Chu, "Constructing an edu-metaverse ecosystem: A new and innovative framework," *IEEE Trans. Learn. Technol.*, vol. 15, no. 6, pp. 685–696, Dec. 2022.
- [8] T. Kim, J. Planey, and R. Lindgren, "Comparing virtual and augmented reality: Learning and interaction effects in astronomy classrooms," in *Proc. 17th Int. Conf. Learn. Sci.*, 2023, pp. 295–301.
- [9] J. H. Steffen, J. E. Gaskin, T. O. Meservy, J. L. Jenkins, and I. Wolman, "Framework of affordances for virtual reality and augmented reality," *J. Manage. Inf. Syst.*, vol. 36, no. 3, pp. 683–729, 2019.
- [10] S. Minocha, A.-D. Tudor, and S. Tilling, "Affordances of mobile virtual reality and their role in learning and teaching," *Electron. Workshops Comput.*, vol. 2, 2017, Art. no. 100018.
- [11] C. Fowler, "Virtual reality and learning: Where is the pedagogy?," *Brit. J. Educ. Technol.*, vol. 46, no. 2, pp. 412–422, 2014.
- [12] E. Johnston, G. Olivas, P. Steele, C. Smith, and L. Bailey, "Exploring pedagogical foundations of existing virtual reality educational applications: A content analysis study," *J. Educ. Technol. Syst.*, vol. 46, no. 4, pp. 414–439, 2017.
- [13] T. A. Mikropoulos and A. Natsis, "Educational virtual environments: A ten-year review of empirical research (1999–2009)," *Comput. Educ.*, vol. 56, no. 3, pp. 769–780, 2011.
- [14] L. Stefan, F. Moldoveanu, and D. Gheorghiu, "Evaluating a mixed-reality 3D virtual campus with Big Data and learning analytics: A transversal study," *J. e-Learn. Knowl. Soc.*, vol. 12, May 21, 2016, Accessed: Dec. 15, 2022. [Online]. Available: <https://www.learntechlib.org/p/173466/>
- [15] Y. H. Cho and K. Y. Lim, "Effectiveness of collaborative learning with 3D virtual worlds," *Brit. J. Educ. Technol.*, vol. 48, no. 1, pp. 202–211, 2015.
- [16] C. Dede, T. A. Grotzer, A. Kamarainen, and S. Metcalf, "EcoXPT: Designing for deeper learning through experimentation in an immersive virtual ecosystem," *J. Educ. Technol. Soc.*, vol. 20, no. 4, pp. 166–178, Oct. 2017.
- [17] J. A. Frank and V. Kapila, "Mixed-reality learning environments: Integrating mobile interfaces with laboratory test-beds," *Comput. Educ.*, vol. 110, pp. 88–104, 2017.
- [18] E. Loup-Escande, R. Frenoy, G. Poplimont, I. Thouvenin, O. Gapenne, and O. Megalakaki, "Contributions of mixed reality in a calligraphy learning task: Effects of supplementary visual feedback and expertise on cognitive load, user experience and gestural performance," *Comput. Hum. Behav.*, vol. 75, pp. 42–49, 2017.
- [19] K.-E. Chang, C.-T. Chang, H.-T. Hou, Y.-T. Sung, H.-L. Chao, and C.-M. Lee, "Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum," *Comput. Educ.*, vol. 71, pp. 185–197, 2014.
- [20] F. Ke, S. Lee, and X. Xu, "Teaching training in a mixed-reality integrated learning environment," *Comput. Hum. Behav.*, vol. 62, pp. 212–220, 2016.
- [21] A. Grabowski and J. Jankowski, "Virtual reality-based pilot training for underground coal miners," *Saf. Sci.*, vol. 72, pp. 310–314, 2015.
- [22] G.-J. Hwang, P.-H. Wu, C.-C. Chen, and N.-T. Tu, "Effects of an augmented reality-based educational game on students' learning achievements and attitudes in real-world observations," *Interactive Learn. Environments*, vol. 24, no. 8, pp. 1895–1906, 2015.
- [23] J. Parong and R. E. Mayer, "Learning science in immersive virtual reality," *J. Educ. Psychol.*, vol. 110, no. 6, pp. 785–797, 2018.
- [24] E. Ucar, H. Ustunel, T. Civelek, and I. Umut, "Effects of using a force feedback haptic augmented simulation on the attitudes of the gifted students towards studying chemical bonds in virtual reality environment," *Behav. Inf. Technol.*, vol. 36, no. 5, pp. 540–547, 2016.
- [25] F. Rasheed, P. Onkar, and M. Narula, "Immersive virtual reality to enhance the spatial awareness of students," in *Proc. 7th Int. Conf. HCI*, 2015, pp. 154–160.
- [26] G. Makransky, T. S. Terkildsen, and R. E. Mayer, "Adding immersive virtual reality to a science lab simulation causes more presence but less learning," *Learn. Instruct.*, vol. 60, pp. 225–236, 2019.
- [27] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*. Upper Saddle River, NJ, USA: Prentice-Hall, 1984.
- [28] C. Vaz de Carvalho, "Improving experiential learning with haptic experimentation," *Int. J. Online Biomed. Eng.*, vol. 9, no. S8, 2013, Art. no. 7.
- [29] I. Duncan, A. Miller, and S. Jiang, "A taxonomy of virtual worlds usage in education," *Brit. J. Educ. Technol.*, vol. 43, no. 6, pp. 949–964, 2012.
- [30] W. Barfield and C. Hendrix, "The effect of update rate on the sense of presence within virtual environments," *Virtual Reality*, vol. 1, no. 1, pp. 3–15, 1995.
- [31] C. Heeter, "Being there: The subjective experience of presence," *Presence, Teleoperators Virtual Environments*, vol. 1, no. 2, pp. 262–271, 1992.

- [32] M. Slater and M. Usoh, "Representations systems, perceptual position, and presence in immersive virtual environments," *Presence, Teleoperators Virtual Environments*, vol. 2, no. 3, pp. 221–233, 1993.
- [33] M. Slater and S. Wilbur, "A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments," *Presence, Teleoperators Virtual Environments*, vol. 6, no. 6, pp. 603–616, 1997.
- [34] C. Kwon, "Verification of the possibility and effectiveness of experiential learning using HMD-based immersive VR Technologies," *Virtual Reality*, vol. 23, no. 1, pp. 101–118, 2018.
- [35] E. Johnston, G. Olivas, P. Steele, C. Smith, and L. Bailey, "Exploring pedagogical foundations of existing virtual reality educational applications: A content analysis study," *J. Educ. Technol. Syst.*, vol. 46, no. 4, pp. 414–439, 2017.
- [36] S. Kalantari, J. D. Rounds, J. Kan, V. Tripathi, and J. G. Cruz-Garza, "Comparing physiological responses during cognitive tests in virtual environments vs. in identical real-world environments," *Sci. Rep.*, vol. 11, no. 1, 2021, Art. no. 10227.
- [37] D. M. Markowitz, R. Laha, B. P. Perone, R. D. Pea, and J. N. Bailenson, "Immersive virtual reality field trips facilitate learning about climate change," *Front. Psychol.*, vol. 9, 2018, Art. no. 2364.
- [38] D. W. Carruth, "Virtual reality for education and workforce training," in *Proc. 15th Int. Conf. eLearning Technol. Appl.*, 2017, pp. 1–6.
- [39] C. Boedecker et al., "Using virtual 3D-models in surgical planning: Workflow of an immersive virtual reality application in liver surgery," *Langenbeck's Arch. Surg.*, vol. 406, no. 3, pp. 911–915, 2021.
- [40] S. Bouchard, S. Côté, J. St-Jacques, G. Robillard, and P. Renaud, "Effectiveness of virtual reality exposure in the treatment of arachnophobia using 3D games," *Technol. Health Care*, vol. 14, no. 1, pp. 19–27, 2006.
- [41] E. Hutchins, "How a cockpit remembers its speeds," *Cogn. Sci.*, vol. 19, no. 3, pp. 265–288, 1995.
- [42] J. van Dijk, R. van der Lugt, and C. Hummels, "Beyond distributed representation," in *Proc. 8th Int. Conf. Tangible, Embedded Embodied Interaction*, 2013, pp. 181–188.
- [43] E. Hutchins, "The distributed cognition perspective on human interaction," *Roots Hum. Sociality*, vol. 9, pp. 375–398, 2020.
- [44] E. Deirick, R. B. Shapiro, M. P. Ahrens, R. Fiebrink, P. D. Lehrman, and S. Farooq, "Using distributed cognition theory to analyze collaborative computer science learning," in *Proc. 11th Annu. Int. Conf. Int. Comput. Educ. Res.*, 2015, pp. 51–60.
- [45] J. Hollan, E. Hutchins, and D. Kirsh, "Distributed cognition," *ACM Trans. Comput.-Hum. Interaction*, vol. 7, no. 2, pp. 174–196, 2000.
- [46] D. Kirsh, "Embodying cognition and the magical future of interaction design," *ACM Trans. Comput.-Hum. Interaction*, vol. 20, no. 1, pp. 1–30, 2013.
- [47] J. McVeigh-Schultz and K. Isbister, "A 'beyond being there' for VR meetings: Envisioning the future of remote work," *Hum.–Comput. Interaction*, vol. 37, no. 5, pp. 433–453, 2021.
- [48] J. Piaget, *The Origins of Intelligence in Children*. England, U.K.: Routledge, 1952.
- [49] S. Goldin-Meadow, *Hearing Gesture: How Our Hands Help Us Think*. Cambridge, MA, USA: Harvard Univ. Press, 2005.
- [50] D. McNeill, *Hand and Mind: What Gestures Reveal About Thought*. Chicago, IL, USA: Univ. of Chicago Press, 1992.
- [51] W.-M. Roth, "Gestures: Their role in teaching and learning," *Rev. Educ. Res.*, vol. 71, no. 3, pp. 365–392, 2001.
- [52] A. M. Glenberg, J. K. Witt, and J. Metcalfe, "From the revolution to embodiment," *Perspectives Psychol. Sci.*, vol. 8, no. 5, pp. 573–585, 2013.
- [53] M. Wilson, "Six views of embodied cognition," *Psychon. Bull. Rev.*, vol. 9, no. 4, pp. 625–636, 2002.
- [54] S. C. Broaders, S. W. Cook, Z. Mitchell, and S. Goldin-Meadow, "Making children gesture brings out implicit knowledge and leads to learning," *J. Exp. Psychol., Gen.*, vol. 136, no. 4, pp. 539–550, 2007.
- [55] S. Goldin-Meadow, "Learning through gesture," *Wiley Interdiscipl. Rev., Cogn. Sci.*, vol. 2, no. 6, pp. 595–607, 2011.
- [56] M. J. Nathan and C. V. J. Martinez, "Gesture as model enactment: The role of gesture in mental model construction and inference making when learning from text," *Learn., Res. Pract.*, vol. 1, no. 1, pp. 4–37, 2015.
- [57] M. Stieff, M. E. Lira, and S. A. Scopelitis, "Gesture supports spatial thinking in STEM," *Cogn. Instruct.*, vol. 34, no. 2, pp. 80–99, 2016.
- [58] J. O. Bailey and J. N. Bailenson, "Considering virtual reality in children's lives," *J. Child. Media*, vol. 11, no. 1, pp. 107–113, 2017.
- [59] S. Price, N. Yiannoutsou, and Y. Vezzoli, "Making the body tangible: Elementary geometry learning through VR," *Digit. Experiences Math. Educ.*, vol. 6, no. 2, pp. 213–232, 2020.
- [60] The Body VR LCC, "The body VR: Journey inside a cell on steam," Oct. 3, 2016, Accessed: Dec. 15, 2022. [Online]. Available: https://store.steampowered.com/app/451980/The_Body_VR_Journey_Inside_a_Cell/
- [61] N. A. Nasharuddin, N. A. Khalid, and M. Hussin, "InCell VR: A virtual reality-based application on human cell division for mobile learning," *Int. J. Interactive Mobile Technol.*, vol. 15, no. 2, 2021, Art. no. 55.
- [62] S. Barab and K. Squire, "Design-based research: Putting a stake in the ground," *J. Learn. Sci.*, vol. 13, no. 1, pp. 1–14, 2004.
- [63] S. Ainsworth and K. Scheiter, "Learning by drawing visual representations: Potential, purposes, and practical implications," *Curr. Directions Psychol. Sci.*, vol. 30, no. 1, pp. 61–67, Feb. 2021, doi: [10.1177/0963721420979582](https://doi.org/10.1177/0963721420979582).
- [64] S. P. Wu and M. A. Rau, "How students learn content in science, technology, engineering, and mathematics (STEM) through drawing activities," *Educ. Psychol. Rev.*, vol. 9, pp. 1–34, 2019.
- [65] J. G. Cromley, Y. Du, and A. P. Dane, "Drawing-to-learn: Does meta-analysis show differences between technology-based drawing and paper-and-pencil drawing?," *J. Sci. Educ. Technol.*, vol. 29, no. 2, pp. 216–229, 2019.
- [66] L. Fiorella and Q. Zhang, "Drawing boundary conditions for learning by drawing," *Educ. Psychol. Rev.*, vol. 30, no. 3, pp. 1115–1137, 2018.
- [67] P. Van Meter and J. Garner, "The promise and practice of learner-generated drawing: Literature review and synthesis," *Educ. Psychol. Rev.*, vol. 17, no. 4, pp. 285–325, 2005.
- [68] P. Blikstein and M. Worsley, "Multimodal Learning Analytics and education data mining: Using computational technologies to measure complex learning tasks," *J. Learn. Analytics*, vol. 3, no. 2, pp. 220–238, 2016.
- [69] R. Luckin and M. Cukurova, "Designing educational technologies in the age of AI: A learning sciences-driven approach," *Brit. J. Educ. Technol.*, vol. 50, no. 6, pp. 2824–2838, 2019.
- [70] C. P. Rosé, E. A. McLaughlin, R. Liu, and K. R. Koedinger, "Explanatory learner models: Why machine learning (alone) is not the answer," *Brit. J. Educ. Technol.*, vol. 50, no. 6, pp. 2943–2958, 2019.
- [71] H. Sung, M. Swart, and M. J. Nathan, "Methods for analyzing temporally entangled multimodal data," in *Proc. 16th Annu. Conf. Int. Soc. Learn. Sci.*, 2022, pp. 242–249.
- [72] M. Worsley, D. Abrahamsen, P. Blikstein, S. Grover, B. Schneider, and M. Tissenbaum, *Situating Multimodal Learning Analytics*. Berkeley, CA, USA: UC Berkeley, 2016.
- [73] M. Worsley and P. Blikstein, "Using multimodal learning analytics to study learning mechanisms," in *Proc. 7th Int. Conf. Educ. Data Mining*, 2014, pp. 431–432.



Taehyun Kim received the B.A. degree in educational technology from Andong National University, Andong, South Korea, in 2017 and the M.Ed. degree in education from Sungkyunkwan University, Seoul, South Korea, in 2019. He is currently working toward the Ph.D. degree in department of curriculum and instruction, focusing on how immersive biology simulations facilitate students complex science understanding, University of Illinois Urbana-Champaign, Champaign, IL, USA.

He is working as a Research Assistant with Embodied and Immersive Technologies Lab and Pixel Playground Lab. He is interested in designing emerging media platforms, such as virtual and mixed reality, simulations, or video games with the framework of embodied cognition and design-based research. He seeks to understand how these technologies can be used to help learning and teaching situations, especially in STEM domain. Specifically, his research focuses on how the body-based interactions with the designed immersive media can facilitate complex understanding of learning contents, and how the media can be adequately designed to include these types of interactions.

Mr. Kim is a member of the International Society of the Learning Sciences, the American Educational Research Association, and the International Learning Analytics and Knowledge Conference.



James Planey received the B.A. degree in psychology from Illinois Wesleyan University, Bloomington, IL, USA, in 2006, and the M.S. degree in science education and the Ph.D. degree in curriculum and instruction, exploring immersive VR drawing for science concept representation from the University of Illinois Urbana-Champaign, Champaign, IL, USA, in 2012 and 2023, respectively.

He is currently a Postdoctoral Researcher and a Science Education Design Lead with the Molecule Maker Lab Institute, University of Illinois Urbana-Champaign. He has previous experiences as a high school Science Teacher, Technology Integration Coach, Curriculum Designer, and informal Nature Educator. He is interested in how immersive technologies, such as VR and AR can be meaningfully integrated into science classroom instruction to support student learning as they engage with spatially complex content.

Dr. Planey is a member of the International Society of the Learning Sciences, the National Science Teaching Association, the American Educational Research Association, and the National Association for Research in Science Teaching.



Robb Lindgren received the B.S. degree in computer science from Northwestern University, Evanston, IL, USA, in 2000, and the master's degree in psychology and the doctorate degree in learning sciences and technology design from Stanford University, Stanford, CA, USA, in 2009, where he conducted research on perspective-taking in virtual environments.

He is currently an Associate Professor of educational psychology with the University of Illinois Urbana-Champaign (UIUC), Champaign, IL, USA, where he leads the Embodied and Immersive Technologies Lab. He is also currently the Director of the Technology Innovations in Educational Research and Design initiative at UIUC that seeks to build a cross-campus community focused on the use of emerging technologies to address critical issues in education. His research pertains to the design of digital learning environments, theories of learning and interactivity, and methods for conducting research on how people learn through their interactions with emerging technologies such as virtual reality and augmented reality. He has authored or coauthored previously on the design and learning effects of extended reality technologies in journals, such as the *Journal of Educational Psychology* and *Educational Technology Research and Development*. He is particularly interested in theories and designs that acknowledge the embodied nature of human cognition and learning, and ways to seed new discourse and understanding through embodied activity such as gesturing and role-play.

Dr. Lindgren is an active member of the International Society of the Learning Sciences, and he was the Division C Jan Hawkins Early Career Award Winner for the American Educational Research Association in 2017.