Using VR for Collaborative Learning: A Theoretical and Practical Lens

Jonathan Spike
Northern Illinois University, USA

Ying Xie
https://orcid.org/0000-0002-9189-6046
Northern Illinois University, USA

INTRODUCTION

The first known head-mounted, virtual reality display, nicknamed "the Sword of Damocles", was developed by Ivan Sutherland in 1968. Most notably, this early hardware allowed for the display to change perspective in the virtual world as the user moved his/her head (Mazuryk & Gervautz, 1999). Since then, the-head-mounted displays have gone through many iterations, including Thomas Furness' Visually Coupled Airborne Systems Simulator (VCASS), Virtual Visual Environment Display (VIVED), and recently Google Cardboard, and Oculus technologies. Virtual reality differs from other types of virtual environments such as virtual world. According to Xie (2010), virtual world is usually computer-based, and users interact with the world by controlling avatars whereas VR completely immerses users within a simulated realist environment by replacing the immediate surroundings. Gigante (1993) classified head-mounted virtual reality as "the illusion of participation in a synthetic environment rather than external observation of such an environment. VR relies on a three-dimensional, stereoscopic head-tracker display, hand/body tracking and binaural sound. VR is an immersive, multi-sensory experience" (p. 4).

BACKGROUND

Due to the unique features and increasing affordability of the head-mounted VR tools, practical applications and research studies sprouted in recent years. Besides studies leverage VR tools to promote academic learning in various disciplines (e,g., Ahn et al., 2016; Lee et al., 2017; Lisichenko, 2015; Xie et al., 2021), studies also examined the efficacy of features provided by the tool itself, for example, the ability of manipulating objects in the virtual space (Jang et al., 2017), the ability to see the world as the participant moved their head and body (Orman et al., 2017), Yet, most of these earlier applications leveraged VR tools in a non-social, user-alone manner. In other words, each user interacts with the environment alone without socializing with other users while immersed in the virtual reality, mainly due to technological limitations. Such a drawback, however, would be mitigated with the availability of newer platforms, such as Spatial, which allows for life-like human avatars to be embodied, controlled, and voiced by actual people rather than computer-generated avatars (Spatial Systems, 2021).

DOI: 10.4018/978-1-6684-7366-5.ch040

Focus of the Article

Due to the lack of a theoretical foundation to elucidate the learning processes within a collaborative VR environment, this chapter first establishes a working framework of how to facilitate VR-based collaborative learning through an aggregated lens including social translucence and embodied cognition. To help contextualize VR tools' ability to facilitate collaborative learning, three current virtual reality platforms are then evaluated in terms of collaborative potential. Based on analyses of a number of collaborative virtual reality studies, the chapter also provides recommended best practices for facilitating and designing collaborative virtual reality learning experiences.

Building a Theoretical Foundation for VR-Based Collaborative Learning

Since collaborative learning via VR tools is still a novel phenomenon, no established single theory or model could offer a thorough explanation of the learning process. However, many theories or conceptual frameworks referred to team collaboration, or technology-assisted collaborative activities, esp. in the field of organizational psychology (e.g., Colbry et al., 2014; Olson et al., 2013; Schulze & Krumm, 2017). In the effort to build a theoretical foundation toward a conglomerated view about how VR tools could facilitate team-based collaboration, in the section below, we explore a few theories and highlight the important theoretical concepts relevant to the phenomenon.

Root of Collaborative Learning Theory

The concept of collaborative learning originated from Vygotsky's (1978a, 1978b) sociocultural theory. Built on the assumption that learning is a constructive process, collaborative learning refers to an instructional strategy in which students of various abilities work in small teams toward one joint learning goal. The learning activities can vary, such as content comprehension, discourse advancement, problem solving, and product creation etc. Contrast to independent work, mutual understanding and interdependency are the fundamental key factors for any successful collaborative learning experience (Doolittle, 1997). As a result, finding effective communication channels and approaches to facilitate group interaction has been highlighted in the field of educational technology.

Technology-assisted Collaboration

The ever-changing and wide availability of novel technologies in recent years have inevitably brought about new opportunities and challenges to group interactions in collaborative learning situations (Karpova et al., 2009). Computer-mediated communication tools allow for a relatively easy and efficient way to exchange text, graphic, voice/audio, video, and most recently 3-D contexts or environments such as those in virtual reality. Since technology tools were not created equally, they provide different affordances for collaboration or collaborative learning.

Research found that in a text-based collaborative learning environment enabled by discussion boards and Google Docs, students agreed that there were a lot of quality interactions when actively engaged (Driver, 2002; Cundell & Sheepy, 2018). However, when there were few resources or technologies tools available, students experienced challenges in terms of communication, participation, accountability, and cohesive interaction (Bakir et al., 2020). Audio-based technology tools such as conference calls allow team members to converse collectively in real time. However, due to absence of circumstantial cues and nonverbal information, students had a hard time arriving at a productive, mutual understanding (Karpova et al., 2009).

In the article, "Communication and Collaboration in Distributed Cognition", Boland and Tenkasi (2001) recognized that in a distributed network, successful collaboration requires both message exchanges and a mechanism allowing members to "construct and reconstruct understandings of a situation", which they called "a boundary object" (p.63). The boundary object can be any form of representations that can be synchronously shared and revised as the focus of the conversation for meaning construction. Video-based communications seemed to be able to address many concerns with previous technology tools, especially when combined with many other functions. For example, Web conferencing allows for file sharing, chat, collaborative whiteboards, as well as synchronous video communication via webcam videos. (Bower et al., 2012). Other chief characteristics of synchronous video communication include its simultaneous occurrence with participants, full conveyance of facial expressions, partial conveyance of body language, and the transfer of speech (Blau et al., 2017). Recently, modern synchronous video platforms such as Microsoft Teams and Zoom have emerged as natural and easy-to-use alternatives to face-to-face interactions when participants cannot be in the same location (Blau et al., 2017; Srichanyachon, 2013). Yet, are these video platforms ready to replace or completely replicate face-to-face interactions in collaborative learning?

Bjørn and Ngwenyama (2009) put forward the need of "translucence" when teamwork is mediated by groupware. Translucence requires that the collaborative technology should be designed to "permit important but invisible social clues to be visible, thus enabling distributed collaborators to monitor and interpret each other's actions during collaboration" (p.232). According to the authors, the invisible social clues are not limited to the in-situ contextual cues immediately surrounding the collaborative activity. They encompass three contextual levels in the organization: the lifeworld, the organization structure and the work practice, all of which serve as "filters of the collective reality like a veil through which people observe and interpret the actions of others" (Ngwenyama & Klein, 1994, p. 133). In summary, to create a socially translucent system, the virtual technology must 1) transmit messages timely, 2) provide a mechanism for members to co-construct meaning and share understanding, 3) make collaborators' actions visible, 4) communicate nonverbal cues of team members, 5) sustain implicit social clues in a larger frame of reference. Although integrated video conferencing systems appeared to meet several of these requirements, they seemed inadequate to project a panoramic picture of the larger context for the collaborative work. The following table shows the capability of supporting social translucence by common technology tools as compared to VR.

Table 1. Social Translucence Affordances of Various Technologies

Social Translucence	Audio Only	Synchronous Video	Virtual Reality
Synchronous Audio	Yes	Yes	Yes
Co-construction tools	No – done separately	Yes– whiteboarding, image annotation, file/screen sharing	Yes - Add 3D objects, whiteboard notes, sticky notes, 2D objects, videos, Office 365/Google Drive files, Screenshare
Visible Actions	No	No	Yes – See participants manipulating objects in environment
Nonverbal Cues	No	Yes – see reactions from head up in typical web conferencing platform	Yes – movements in space, physical interactions, mouth movements
Authentic Environment	No	No	Yes – uploaded or LiDAR scanned virtual recreations of environments

Embodied Cognition

Since virtual reality creates a simulated environment where the body's functions are tracked to create an immersive experience for body and mind (Gigante, 1993), the theory of embodied cognation can also lend insights into the phenomenon of VR-based collaboration process (Ladendorf et al., 2019). Embodied cognition posits that the mind and body work cohesively in tandem with the context-rich environment (Clark, 1997; 1998; Wilson, 2002). The theory emphasizes the mind and body's reliance on a situated context for cognitive processes (Chiel & Beer, 1997; Clark, 1997), and learning depends on these authentic contexts to allow learners to apply their relevant skills (Brown et al., 1989, Lave & Wenger, 1990).

Another tenet of embodied cognition is that the physical world and space allows for offboarding cognitive load (Glenberg & Robertson, 1999; Brooks, 1991). At times, the environment can serve as a tool to store information and even manipulate said information for the user (Wilson, 2002). Beyond the physical affordances of the space, digital enhancements such as live visuals, responsive data collection, and representations help users process cognitive information in a more efficient manner (Danish et al., 2020).

A third key aspect of embodied cognition argues that the action of the cognitive activity is best served with both a visual and sense of motor control (Wilson, 2002). Prior studies noted that memory perceptions and performance improve when actions are done in a three-dimensional world or executed as modeled when done in the real world (Glenberg, 1997; Tucker & Ellis, 1998). Beyond the act of performing actions in the cognitive activity, motor control while collaborating with others has been shown to enhance learning opportunities (Danish et al., 2020).

Framework for Virtual Reality as Embodied Social Translucence

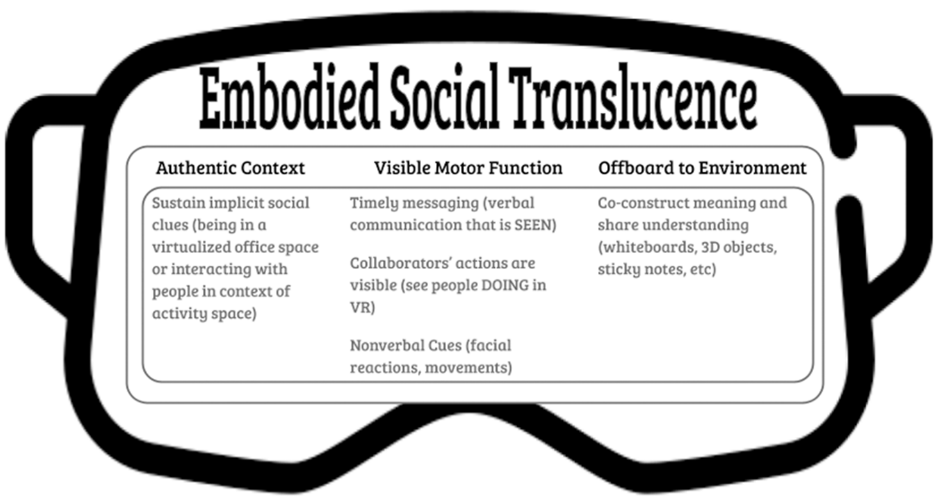
When considering whether a technology platform is appropriate for achieving the tenets of social translucence in group collaboration (Bjorn & Ngwenyama, 2009), the tool and its application will ideally align with many of the features of embodied cognition (Clark, 1997). Head-mounted virtual reality relies on the key components of embodied cognition such as the visible recreation of motor functions, situating users in authentic contexts, and affordances to offboard cognitive load to the digital environment (Parmar, 2017; Beck et al., 2013). Therefore, a proposed framework for viewing virtual reality as a vehicle for creating embodied, social translucence depends on the alignment of the three main tenets of embodied cognition addressing the five components of social translucence, as proposed in the previous sections.

Authentic contexts key to embodied cognition (Clark, 1997, 1998; Wilson, 2002; Brown et al., 1989) and presented in high fidelity by virtual reality environments connect to the social translucence tenet of witnessing "social clues," or visual representations of the culture and expectations of a contextualized space (Ngwenyama & Klein, 1994). When a digital avatar joins another collaborator in a realistic representation of the normal environment where group tasks would occur, the norms of the organization, social indicators, and other contextual clues shape the interactions (Bjorn & Ngwenyama, 2009).

Another characteristic of embodied cognition that aligns with virtual reality platforms is the ability to offboard cognitive load to the environment, as well as provide opportunities for storage and interaction with the information (Glenberg & Robertson, 1999; Brooks, 1991). Social translucence during collaborative activities relies on the co-construction of knowledge and skill development (Bjorn & Ngwenyama, 2009), and virtual reality often builds tools into the environment to facilitate co-construction of meaning (Spatial Systems, 2021; Engage XR Holdings, 2021; Virbela, 2021). Often, users import and manipulate 3D objects, digital whiteboards, videos, images, and annotated sticky notes within the 360-degree environments in order to share group member perspectives, ideas, and understanding (Spatial Systems, 2021; Engage XR Holdings, 2021; Virbela, 2021).

A quality that often separates embodied cognition in virtual reality from other forms of digital communication and collaboration is the act of seeing the visual and feeling the sensation of motor function during a cognitive activity (Glenberg, 1997; Tucker & Ellis, 1998; Wilson, 2002). Synchronous facial movements that coincide with verbal communication in virtual reality help to establish timely messaging between collaborators, a quality of effective social translucence (Bjorn & Ngwenyama, 2009). In addition, virtual reality often communicates nonverbal cues and attributes, such as collaborators turning their heads and bodies to face new speakers, moving to new locations based on conversations, and representing emotions via body language, such as clapping and high-fives (Spatial Systems, 2021; Engage XR Holdings, 2021), all aspects of team-based social translucence. Virtual reality provides accountability between team members through the visibility of collaborators' actions, often through their motor functions performed by virtual representations of arms, hands, and visuals of objects and documents they utilize in the virtual space, which provides a transparent look into their contributions to the group (Bjorn & Ngwenyama, 2009).

Figure 1. Embodied Social Translucence Model



Research Support for the Framework

Although collaborative VR tools are still novel, findings from recent research studies already provided empirical evidence for Embodied Social Translucence, the theoretical framework discussed above. For instance, Miller et al. (2021) performed a study investigating the level of nonverbal behavior synchronization between groups of three collaborators in virtual reality during a set of design tasks. In addition, the study investigated rates of synchrony in virtual reality when faced with formal and informal environments, as well as participants' sense of team effectiveness and overall team performance. During the study, participants engaged in a series of collaborative activities such as selecting an optimum product prototype or brainstorming ideas for an emerging industry. Results demonstrated that synchrony of nonverbal cues were achieved between collaborations of three people in virtual reality. Participants were able to achieve the sensation of seeing one another's head movements, gaze, and other nonverbal cues, which are key tenets of establishing social translucence in collaborative activities (Bjorn & Ngwenyama, 2009). The study also revealed that the more contextualized workspace of the conference room virtual environment had higher rates of nonverbal synchrony than the informal garage virtual environment (Miller et al., 2021). One potential explanation for this result traces back to the importance of social clues within the culture and environment of a collaborative space, as well as how the space allows for the co-construction of ideas. The conference room contained several indicators of an organizational workspace, with cultural indicators such as an office table and chairs, a professional whiteboard, and a city skyline outside of the window to provide social context and expectations of behavior (Miller et al., 2021; Bjorn & Ngwenyama, 2009).

Another study explored multi-user virtual reality using immersive screens and head-mounted viewers to bring two groups of participants in remote locations together (Beck et al., 2013). In the experiment, the two remote groups explored and discussed a virtual city model, and the technology allowed for the groups to see each others' movements, conversations, and facial expressions in the virtual space. The sensation of witnessing a collaborator's actions in the virtual world provide a key accountability factor in achieving social translucence (Bjorn & Ngwenyama, 2009) and motor function representation serves as a key tenet of embodied cognition (Wilson, 2002). The study participant dispositions revealed positive attitudes toward group coordination, understanding of teammate body language, and comprehension of gaze communication, or the social cue of making eye contact (Beck et al., 2013). Many of these nonverbal actions further establish collaborative social translucence, helping to aid in group cohesion (Bjorn & Ngwenyama, 2009).

A third study by Jensen (2017) explored how participants from backgrounds such as architecture, engineering, hospitality, and energy technology were able to collaborate in virtual reality on a design-based task. The participants utilized head-mounted virtual reality to gather and discuss how to restore a museum by navigating a virtual representation of the space, with each collaborator represented by a floating robot avatar (Jensen, 2017). Qualitative observations of the participants revealed that the virtual environment provided unique visual context that spurred discussions and idea sharing among the collaborators in the space (Jensen, 2017), which aligns with the social translucence concept of being able to co-contruct meaning (Bjorn & Ngwenyama, 2009), as well as the embodied cognition trait of the environment helping the user process information (Danish et al., 2020; Glenberg & Robertson, 1999; Brooks, 1991). While the environment did help in the processing of cognitive load, the virtualized space was not found to help establish social and cultural rules and norms that would come with the expectations of engineers, architects, and other professional roles on a job site (Jensen, 2017), which fails to meet the expectation of social clues common to achieving social translucence (Bjorn & Ngwenyama, 2009).

VR for Collaboration Practice

Given the close fidelity of a VR world to real life, the technology seems to offer many advantages that no other groupware could offer for collaborative learning. In the following section, we will further illustrate the capabilities of a number of existing VR tools for collaborative learning and discuss the related research findings.

The following table compares and contrasts three distinct VR tools for collaborative learning based on the theory of Embodied Social Translucence.

Table 2. Comparison of Three VR Platforms for Collaborative Learning

Embodied Social Translucence	Spatial	Engage VR	Frame VR
1. Timely messaging	Synchronous spatial audio, live translation	Synchronous spatial audio	Synchronous spatial audio
2. Co-construct meaning and share understanding	Add 3D objects, whiteboard notes, sticky notes, 2D objects, videos, Office 365/Google Drive files, Screenshare	Add 3D objects and animated 3D objects, whiteboard notes, sticky notes, 2D objects, videos, Office 365/Google Drive files, Screenshare	Add 3D objects, whiteboard notes, 2D objects, videos, Screenshare
3. Collaborators' actions are visible	See torso & above with accurate hand/arm movements, photo-realistic faces	See full body with accurate hand/arm, and leg movements, avatar faces	Only basic avatar with no visible motor function, can pin webcam
4. Nonverbal Cues	Can clap, high five, photo-realistic face has moving mouth during microphone use, pin webcam above avatar	Walking motion, controller vibration on touch, avatar faces (not photo-realistic)	Unrealistic "floating" and teleporting movement
5. Sustain implicit social clues	Ability to LiDAR map a room with a mobile device or design custom environment using modeling, skyboxes & more	Limited ability to upload own environments and contextualized spaces	Can upload 360-degree images and 3D environments, but not LiDAR scanning

Spatial

Spatial takes the novel approach of establishing social translucence through a photo-realistic representation of individuals superimposed on a three-dimensional body controlled by the user (Spatial Systems, 2021). Using a realistic avatar complete with motor functions visible from the torso up through the arms and hands, Spatial allows collaborator actions to be visible so that others may monitor each other's progress and have accountability (Bjorn & Ngwenyama, 2009). Like most head-mounted, virtual reality environments, Spatial uses 360-degree, verbal audio via microphones to provide timely communication among collaborators in a shared space. Synchronous verbal and textual communication provide a barrier-free and contextualized communication method key to social translucence. In terms of nonverbal cues, Spatial translates movement either through avatar "teleportation," joystick-operated locomotion, or virtual reality users walking in their physical space. Another nonverbal cue, touch, manifests in Spatial via a digital confetti animation when a user claps their digital hands together or "high fives" another collaborator's hand. In addition, when photo-realistic avatars in Spatial communicate verbally, their mouth moves in tandem with the audio (Spatial Systems, 2021). A major drawback to Spatial and other collaborative, virtual reality platforms is the inability to express emotion through facial reactions, emoticons, and other strategies (Spatial Systems, 2021). The availability of nonverbal communication tools in a collaboration space aids in clarifying meaning and avoiding misunderstanding (Bjorn & Ngwenyama, 2009), and the uneven access to such cues in Spatial is an area for growth.

One feature unique to Spatial is its ability to scan in both entire rooms and objects using Apple's LiDAR technology for customizing the virtual learning environment (Spatial Systems, 2021). By importing in those contextualized spaces, artifacts, and visuals that reinforce social norms, participants regain social clues from the organizational structures and work practices key to social translucence (Bjorn & Ngwenyama, 2009). Beyond contextualized and familiar learning and collaboration spaces, social translucence requires the ability to easily co-construct meaning with collaborators (Bjorn & Ngwenyama, 2009). Spatial addresses such needs by providing the ability to collaboratively whiteboard, annotate with digital sticky notes, import and manipulate 2D and 3D objects, screenshare, access Microsoft Office 365 and Google Drive files, and much more (Spatial Systems, 2021).

Engage VR

Much like Spatial, Engage VR (Engage XR Holdings, 2021) utilizes synchronous spatial audio to provide clear messaging instantaneously, a key aspect of social translucence and good collaborative practice (Bjorn & Ngwenyama, 2009). Engage VR does lack effective nonverbal cues found in other virtual platforms, with avatars that are only generated based on uploaded user images and not photorealistic, a lack of facial expressions or emotions, and no noticeable acknowledgement of user interactions, such as an animation or noise. Such a lack of nonverbal cues negatively affects users' sense of connectedness and social translucence in the collaborative space. Additionally, Engage VR typically does not support users designing or uploading their own custom environment or contextualized space (Engage XR Holdings, 2021), a feature that often helps establish the expectations, experiences, and cultures of the collaborative space, such as an office area, workspace, or classroom (Bjorn & Ngwenyama, 2009).

The act of physical representation in collaboration is a strength for Engage VR due to its full-motion representation of locomotion and library of animated 3D content that other platforms cannot match. When establishing a sense of social translucence in collaborative activities, seeing fellow teammates performing actions is crucial (Bjorn & Ngwenyama, 2009). Engage VR not only displays moving avatar mouths, arms, bodies, and hands, but also represents the leg locomotion of other users (Engage XR Holdings, 2021). The visual of fellow collaborators and their actions adds a layer of accountability to the social interactions. In addition to seeing others' actions, Engage VR effectively facilitates the ability to co-design and create in the virtual environment. Each user can upload documents, files, 3D objects, videos, and screenshares into the three-dimensional space, and Engage VR also provides an extensive library of not only 3D objects, but 3D animated objects and pre-designed learning simulations (Engage XR Holdings, 2021). Collaborators can utilize these immersive objects to gain a shared understanding of a concept, as well as work cohesively to construct new ideas, both key tenets of social translucence (Bjorn & Ngwenyama, 2009).

Frame VR

Frame VR is a browser-based platform first and foremost, focusing on access instead of high-fidelity experience. Frame VR relies on preset interaction spaces and 360-degree photo uploads, which provide fewer social clues present in the LiDAR-scanned and photo-realistic environments of other platforms (Virbela, 2021). With a lack of cultural markers, customs, and visual social landmarks in the virtual environment, users lose one aspect of social translucence (Bjorn & Ngwenyama, 2009). Frame VR also lacks nonverbal cues such as facial expressions, expression of arm, leg, and torso movements, featuring only cartoonish avatars with heads and bodies. While Frame VR does allow for the use of emoticons to float out from the avatars to express feelings such as happiness, sadness, laughter, and concern, these do not reflect social signals necessary to establish the social translucence necessary to collaborate effectively (Bjorn & Ngwenyama, 2009). One way Frame VR attempts to rectify this limitation is via a "floating" webcam of the user, which allows for full facial expressions and nonverbal cues, though it is detached from the avatar performing the collaborative actions (Virbela, 2021).

Frame VR facilitates the co-construction of concepts effectively through users' ability to import and place 3D objects, images, whiteboard notes, videos, screenshares, 360 images, and more (Virbela, 2021). In addition, FrameVR integrates spatialized audio that creates louder audio when collaborators are nearby, and quieter audio when communicating via distance (Virbela, 2021). While such collaborative activities and features are a strength of the platform, Frame VR does not communicate these actions clearly through the avatars or webcam, making it hard to interpret who is contributing what objects, information, and content to the space. Social translucence in cooperative work depends on transparent contributions through physical displays of actions performed (Bjorn & Ngwenyama, 2009), and Frame VR does not meet such a criterion.

SOLUTIONS AND RECOMMENDATIONS

Collaborative activities in virtual reality environments allow for high-fidelity experiences via their environments and realistic representations of end users in such environments. When designing a collaborative experience in virtual reality, facilitators should prioritize tools that create a more genuine and photorealistic representation of the person via the avatar. In addition, experience designers must make every effort to select platforms that allow for uploading of accurate digital representations of contextualized collaboration spaces relevant to the task at hand, such as a company building for corporate co-use, classrooms for educational teamwork, and laboratories for research cooperation. Technology such as LiDAR scanning of real environments and 3D modeling design software (Spatial Systems, 2021) make the prospect of high-fidelity and authentic spaces a very inexpensive and practical reality. Also, virtual reality platforms only need a webcam image to create a 3D model of a 360-degree avatar that reflects the end user's appearance, complete with moving mouth and blinking eyes (Spatial Systems, 2021). These realistic details serve to add immersion and social norms to collaboration scenarios, enhancing opportunities for efficient group work. In addition, simply providing VR devices with realist avatars or environments is never enough for achieving successful collaborative learning. Besides pedagogical strategies commonly used for collaborative learning such as setting up a joint goal or task, encouraging productive co-construction, and facilitating effective group dynamics and positive interdependency (Jeong & Hmelo-Silver, 2016), practitioners and instructional designers should also take the following techniques into consideration.

Allow Familiarization with the Platform and Virtual Environment

Individuals or groups designing collaborative experiences within virtual reality environments should provide participants with an opportunity to explore the digital space, as well as a chance to practice with the co-use tools in a low-stakes scenario. Users who were not afforded such introduction activities reported frustrations with collaboration later in their use of the virtual reality environment (Jensen, 2017). Experience designers might offer a challenge for participants to complete, such as a scavenger hunt, a design challenge, or a checklist of essential skills to help familiarize users with the platform and tools. Any authentic practice within the environment can help users not only make virtual reality more accessible, but also more conducive to collaborative activities.

Prioritize and Encourage Nonverbal Communication Options in Virtual Reality

Virtual reality presents the unique opportunity to engage in group activities in a 360-degree space with range-of-motion avatars, which opens up new possibilities to communicate via nonverbal cues and movement. Experience designers must consider both a platform that permits a large suite of nonverbal cues, such as gaze, emotional expressions, body language, as well as spatial movements like touching, walking, grabbing, and more actions. Research demonstrates that virtual reality can give users the sensation of eye contact and gaze direction (Miller et al., 2021; Beck et al., 2013), and these crucial social techniques help foster effective collaboration. When planning for co-use, designers should also encourage participants to utilize virtual reality's physical actions and movement capabilities. Many tools acknowledge actions such as high fives and pats on the back with haptic feedback to both users (Spatial Systems, 2021; Engage XR Holdings, 2021), and showing participants in collaborative activities these forms of nonverbal affirmation can strengthen group work and cohesiveness.

FUTURE RESEARCH DIRECTIONS

As head-mounted and high-fidelity virtual reality continues to add new features to enhance a sense of shared presence and communication, new and under-examined research opportunities will emerge. Future research might investigate the significance of more sophisticated body tracking, such as eye contact, facial expressions, and full range of extremities and its impact on ability to collaborate in a digital space. Similarly, future studies may explore how more interactive and realistic virtual environments may or may not impact synchronous users' sense of social translucence with each other as they collaborate or engage in simulations together. In addition, researchers may investigate the impact of feeling connected and socially present with others in a virtual space when using photorealistic and lifelike avatars compared to more generic-looking avatars designed by end users. Such studies could help shed light on how the technology, user experiences, and fidelity of the avatars might enhance the ability to use head-mounted virtual reality and virtual environments to accomplish collaborative tasks in authentic spaces.

CONCLUSION

Although virtual reality is becoming more accessible, affordable, and powerful, few studies have explored high-fidelity, head-mounted virtual reality with motion tracking in a collaborative environment. Future research in the field should clarify how to best leverage the technology to impact the lives of others. With the rise of online education, virtual reality collaboration and design carries major implications relating to how individuals can learn from a distance and still experience a sense of presence and connection to their fellow classmates. Policy makers should investigate the efficacy of relevant VR platforms that could support more flexible learning options for K-12, higher education, and even businesses and organizations to better serve the needs of their students, customers, or workers.

REFERENCES

Ahn, S. J., Bostick, J., Ogle, E., Nowak, K. L., McGillicuddy, K. T., & Bailenson, J. N. (2016). Experiencing nature: Embodying animals in immersive virtual environments increases inclusion of nature in self and involvement with nature. *Journal of Computer-Mediated Communication*, *21*(6), 399–419. doi:10.1111/jcc4.12173

Andrews-Todd, J., Jackson, G. T., & Kurzum, C. (2019). Collaborative problem-solving assessment in an online mathematics task. *ETS Research Report Series*, *2019*(1), 1–7. doi:10.1002/ets2.12260

Bakir, N., Humpherys, S., & Dana, K. (2020). Students' perceptions of challenges and solutions to face-to-face and online group work. *Information Systems Education Journal*, *18*(5), 75–88.

Beck, S., Kunert, A., Kulik, A., & Froehlich, B. (2013). Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics*, *19*(4), 616–625. doi:10.1109/TVCG.2013.33 PMID:23428446

Bjørn, P., & Ngwenyama, O. (2009). Virtual team collaboration: Building shared meaning, resolving breakdowns and creating translucence. *Information Systems Journal*, *19*(3), 227–253. doi:10.1111/j.1365-2575.2007.00281.x

Blau, I., Weiser, O., & Eshet-Alkalai, Y. (2017). How do medium naturalness and personality traits shape academic achievement and perceived learning? An experimental study of face-to-face and synchronous e-learning. *Research in Learning Technology*, *25*(0). Advance online publication. doi:10.25304/rlt.v25.1974

Boland, R. J., & Tenkasi, R. V. (2001). Communication and collaboration in distributed cognition. *Co-ordination theory and collaboration technology*, 51-66.

Bower, M., Kennedy, G. E., Dalgarno, B., Lee, M. J. W., Kenney, J., & de Barba, P. (2012). Use of media-rich real-time collaboration tools for learning and teaching in Australian and New Zealand universities. In M. Brown, M. Hartnett, & T. Stewart (Eds.), Future challenges, sustainable futures. In Proceedings of ascilite Wellington 2012 (pp. 133–144). Academic Press.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence Journal*, *47*(1-3), 139–160. doi:10.1016/0004-3702(91)90053-M

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42. doi:10.3102/0013189X018001032

Can, T. (2020). Training pre-service English language teachers with 3-d machinima. *The Turkish Online Journal of Educational Technology*, *19*(1), 53–65.

Chiel, H., & Beer, R. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body, and environment. *Trends in Neurosciences*, *20*(12), 553–557. doi:10.1016/S0166-2236(97)01149-1 PMID:9416664

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. MIT Press.

Clark, A. (1998). Embodied, situated, and distributed cognition. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science* (pp. 506–517). Blackwell.

Colbry, S., Hurwitz, M., & Adair, R. (2014). Collaboration Theory. *Journal of Leadership Education*, *13*(4). Advance online publication. doi:10.12806/V13/I4/C8

Cornell, H., Sayman, D., & Herron, J. (2019). Sense of community in an online graduate program. *Journal of Effective Teaching in Higher Education*, *2*(2), 117–132. doi:10.36021/jethe.v2i2.52

Cundell, A., & Sheepy, E. (2018). Student perceptions of the most effective and engaging online learning activities in a blended graduate seminar. *Online Learning*, *22*(3), 87–102. doi:10.24059/olj.v22i3.1467

Danish, J. A., Enyedy, N., Saleh, A., & Humburg, M. (2020). Learning in embodied activity framework: A sociocultural framework for embodied cognition. *International Journal of Computer-Supported Collaborative Learning*, *15*(1), 49–87. doi:10.100711412-020-09317-3

Doolittle, P. E. (1997). Vygotsky's Zone of Proximal Development as a Theoretical Foundation for Cooperative Learning. *Journal on Excellence in College Teaching*, *8*(1), 83–103.

Driver, M. (2002). Exploring student perceptions of group interaction and class satisfaction in the web-enhanced classroom. *The Internet and Higher Education*, *5*(1), 35–45. doi:10.1016/S1096-7516(01)00076-8

Duncan, J., & West, R. (2018). Conceptualizing group flow: A framework. *Educational Research Review*, *13*(1), 1–11. doi:10.5897/ERR2017.3313

Engage X. R. Holdings. (2021). *Engage VR* (Version 2.0) [Oculus Quest Software]. https://engagevr.io/

Gigante, M. A. (1993). Virtual reality: definitions, history and applications. In R. A. Earnshaw (Ed.), *Virtual Reality Systems* (pp. 3–14). Academic Press. doi:10.1016/B978-0-12-227748-1.50009-3

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*(1), 1–55. doi:10.1017/S0140525X97000010 PMID:10096994

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*(1), 1–26. doi:10.1080/01638539909545067

Gündüz, A. Y., Alemdag, E., Yasar, S., & Erdem, M. (2016). Design of a problem-based online learning environment and evaluation of its effectiveness. *The Turkish Online Journal of Educational Technology*, *15*(3), 49–57.

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 & Education*, *106*, 150–165. doi:10.1016/j.compedu.2016.12.009

Jensen, C. (2017). Collaboration and dialogue in virtual reality. *Journal of Problem-based Learning*, *5*(1), 85–110.

Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported collaborative learning: How to support collaborative learning? How can technologies help? *Educational Psychologist*, *51*(2), 247–265. doi:10.1080/00461520.2016.1158654

Karpova, E., Correia, A. P., & Baran, E. (2009). Learn to use and use to learn: Technology in virtual collaboration experience. *The Internet and Higher Education*, *12*(1), 45–52. doi:10.1016/j.iheduc.2008.10.006

Ladendorf, K., Schneider, D., & Xie, Y. (2019). Mobile-based virtual reality: Why and how does it support learning? In A. Zhang & D. Cristol (Eds.), *Handbook of Mobile Teaching and Learning* (2nd ed., pp. 1353–1371). Springer. doi:10.1007/978-981-13-2766-7_133

Lave, J., & Wenger, E. (1990). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.

Lee, S. H., Sergueeva, K., Catangui, M., & Kandaurova, M. (2017). Assessing Google Cardboard virtual reality as a content delivery system in business classrooms. *Journal of Education for Business*, *92*(4), 153–160. doi:10.1080/08832323.2017.1308308

Lisichenko, R. (2015). Issues surrounding the use of virtual reality in geographic education. *Geography Teacher*, *12*(4), 159–166. doi:10.1080/19338341.2015.1133441

Mazuryk, T., & Gervautz, M. (1999). *Virtual reality: History, applications, technology and future*. Vienna University of Technology.

Miller, M. R., Sonalkar, N., Mabogunje, A., Leifer, L., & Bailenson, J. (2021). Synchrony within triads using virtual reality. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW2), 1–27. 10.1145/3479544

Ngwenyama, O. K., & Klein, H. K. (1994). An exploration of expertise of knowledge workers: Towards a definition of the universe of discourse for knowledge acquisition. *Information Systems Journal*, *4*(2), 129–140. doi:10.1111/j.1365-2575.1994.tb00047.x

Olson, G. M., Malone, T. W., & Smith, J. B. (2013). *Coordination theory and collaboration technology*. Psychology Press. doi:10.4324/9781410605863

Orman, E., Price, H., & Russell, C. (2017). Feasibility of using an augmented immersive virtual reality learning environment to enhance music conducting skills. *Journal of Music Teacher Education*, *27*(1), 24–35. doi:10.1177/1057083717697962

Parmar, D. (2017). Evaluating the effects of immersive embodied interaction on cognition in virtual reality. *All Dissertations*, 1987. https://tigerprints.clemson.edu/all_dissertations/1987

Robbins, S., Gilbert, K., Chumney, F., & Green, K. (2019). The effects of immersive simulation on targeted collaboration skills among undergraduates in special education. *Teaching & Learning Inquiry*, *7*(2), 168–185. doi:10.20343/teachlearninqu.7.2.11

Sarker, S., Lau, F., & Sahay, S. (2000, January). Building an inductive theory of collaboration in virtual teams: An adapted grounded theory approach. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. IEEE. 10.1109/HICSS.2000.926934

Schulze, J., & Krumm, S. (2017). The "virtual team player" A review and initial model of knowledge, skills, abilities, and other characteristics for virtual collaboration. *Organizational Psychology Review*, *7*(1), 66–95. doi:10.1177/2041386616675522

Shaari, I., Chang, S., & Shanks, G. (2008). Virtual Teams: Information Types for Effective Functioning. *PACIS 2008 Proceedings*, 64.

Spatial Systems. (2021). *Spatial* (Version 3.0) [Oculus Quest software]. https://spatial.io

Srichanyachon, A. N. (2013). Attitudes of undergraduate students towards an online English class. *Turkish Online Journal of Distance Education*, *14*(2), 225–232.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(3), 830–846. doi:10.1037/0096-1523.24.3.830 PMID:9627419

Virbela. (2021). *Frame VR* [Web Virtual Reality Software]. https://framevr.io

Vygotsky, L. (1978a). Interaction between learning and development. *Readings on the Development of Children, 23*(3), 34-41.

Vygotsky, L. S. (1978b). Socio-cultural theory. *Mind & Society*, *6*, 52–58.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. doi:10.3758/BF03196322 PMID:12613670

Wood, D. J., & Gray, B. (1991). Toward a comprehensive theory of collaboration. *The Journal of Applied Behavioral Science*, *27*(2), 139–162. doi:10.1177/0021886391272001

Xie, T. (2010). Tools for teaching Chinese in the virtual world. *Journal of Technology and Chinese Language Teaching*, *1*(1), 59–70.

Xie, Y., Chen, Y., & Ryder, L. (2021). Effects of mobile-based VR on Chinese L2 students' oral proficiency. *Computer Assisted Language Learning*, *34*(3), 225–245. doi:10.1080/09588221.2019.1604551

ADDITIONAL READING

Erickson-Davis, C., Luhrmann, T. M., Kurina, L. M., Weisman, K., Cornman, A., & Bailenson, J. (2021). The sense of presence: Lessons from virtual reality. *Religion, Brain & Behavior*, *11*(3), 335–351. doi:10.1080/2153599X.2021.1953573

Han, E., Nowak, K. L., & Bailenson, J. N. (2022). Prerequisites for Learning in Networked Immersive Virtual Reality. *Technology, Mind, and Behavior, 3*(4).

Lowyck, J., & Pöysä, J. (2001). Design of collaborative learning environments. *Computers in Human Behavior*, *17*(5-6), 507–516. doi:10.1016/S0747-5632(01)00017-6

Palincsar, A. S., & Herrenkohl, L. R. (2002). Designing collaborative learning contexts. *Theory into Practice*, *41*(1), 26–32. doi:10.120715430421tip4101_5

Resta, P., & Laferrière, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review*, *19*(1), 65–83. doi:10.100710648-007-9042-7

Strickland, J., & Xie, Y. (2012). Cooperating or collaborating: Design considerations of employing Wikis to engage college-level learners. In C. Wankel & P. Blessinger (Eds.), *Increasing learner engagement through cutting-edge technologies* (pp. 17–45). Emerald Group Publishing Limited.

Tüzün, H., Bilgiç, H. G., & Elçic, S. Y. (2019). The effects of 3D multi-user virtual environments on collaborative learning and social presence. *International Electronic Journal of Elementary Education*, *11*(3), 221–231. doi:10.26822/iejee.2019349247

Xie, Y., Ryder, L., & Chen, Y. (2019). Using interactive virtual reality tools in an advanced Chinese L2 class: A case study. *TechTrend*, *63*(3), 251–259. doi:10.100711528-019-00389-z

Zhang, H., Yu, L., Ji, M., Cui, Y., Liu, D., Li, Y., Liu, H., & Wang, Y. (2020). Investigating high school students' perceptions and presences under VR learning environment. *Interactive Learning Environments*, *28*(5), 635–655. doi:10.1080/10494820.2019.1709211

KEY TERMS AND DEFINITIONS

Cognitive Load: Refers to the various tasks that demand learners' attention and mental processes when developing new skills.

Collaborative Learning: Originated from Vygotsky's (1978a, 1978b) sociocultural theory, refers to an instructional strategy in which students of various abilities work in small teams toward one joint learning goal.

Embodied Cognition: Refers to the learning theory emphasizing that the mind and body work cohesively in tandem with the context-rich environment.

Embodied Social Translucence: Is defined as the theoretical framework or model in the attempt to explain how to facilitate VR-based collaborative learning through an aggregated lens of social translucence and embodied cognition. It consists of three major components: 1) authentic contexts encompassing "social clues," or visual representations of the culture and expectations of a contextualized space; 2) the ability to offboard cognitive load to the environment, as well as opportunities for storage and interaction with the information; 3) synchronous communication of nonverbal cues such as facial movements that coincide with verbal communication.

Sociocultural Theory: Suggests that individuals develop as the result of social interactions, dialogue, and experiences in society, often learning from those with more expertise. (Vygotsky, 1987a, 1978b).

Translucence: A concept put forward by Bjørn and Ngwenyama (2009). It is defined as a characteristic of an environment that allows for the invisible social sues to be transparent to all the collaborators in that environment.

Virtual Reality (VR): Technology refers to those technology tools that completely immerses users within a simulated, realist, 3-D environment by replacing the immediate surroundings.