



The Theory of Immersive Collaborative Learning (TICOL)

Guido Makransky¹  · Gustav Bøg Petersen¹

Accepted: 6 October 2023 / Published online: 27 October 2023
© The Author(s) 2023

Abstract

The need to virtually collaborate across distributed locations has drastically increased. Developments such as the COVID-19 pandemic and new IT platforms like the metaverse have spurred a host of new immersive social applications that are accessed through head-mounted displays. This is expected to stimulate a surge in research on extended reality–supported collaborative learning (XRCL) which refers to distributed collaboration situations where immersive technology such as head-mounted displays are used as a medium for collaborative learning. The primary aim of this article is to critically examine the potential pedagogical benefits and limitations of using XRCL with the objective of developing a theoretical framework that describes the fundamental factors that make immersive collaborative learning unique: the theory of immersive collaborative learning (TICOL). In TICOL, we propose that technological features, social affordances, and pedagogical techniques can foster four psychological factors that we define as fundamentally different in XRCL compared to collaboration that occurs through traditional systems (e.g., laptops): social presence, physical presence, body ownership, and agency. These are central factors that we hypothesize can transform the processes and contexts of collaboration through their influence on the quality of cognitive and socio-emotional social interaction, the social space, and ultimately learning outcomes. Since XRCL research is in its infancy, we hope that TICOL can provide a theoretical basis for developing the field by motivating researchers to empirically challenge and build on our hypotheses and ultimately develop a deeper understanding of if and how immersive media influences collaborative learning.

Keywords Immersive learning · Extended reality · Computer-supported collaborative learning · Collaborative learning · Virtual reality

✉ Guido Makransky
gm@psy.ku.dk

¹ Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Introduction

Take a minute to stop what you are doing, and have a look around your environment. If you are in the presence of others, be that in a café, your living room, an office, in public transportation, or a park, there is a high probability that there are people around you looking into a device. Our lives have rapidly evolved over the last decade, and technology has infiltrated how we communicate, play, obtain information, and learn. In the next decade, new communication platforms like the metaverse (which is predicted to be the next iteration of the Internet; [Mystakidis, 2022](#)) are expected to stimulate a shift in communication, moving from information-based communication accessed through 2D screens to experience-based communication accessed through head-mounted displays (HMDs: [Plechata et al., 2022, b](#)). The timing of these developments matches a need to increase the quality of distributed collaboration in a globalized world where there is increased focus on factors such as unnecessary travel ([Gössling et al., 2019](#)). This need was accelerated during the COVID-19 pandemic when people relied on technology to connect and solve social, educational, and organizational challenges ([Almeida et al., 2020](#); [Bygstad et al., 2022](#); [European Commission, 2020](#)). Limitations to using standard media for distributed collaboration during the COVID-19 pandemic (e.g., [Bailenson, 2021](#)) have spurred a host of new extended reality (XR) social applications (e.g., VRChat, Mozilla Hubs, Horizon Worlds, Spatial, Engage VR). The implication of these developments is that when we look around our environment in 10 years, people may still be using a device, but rather than reading or passively watching a video on a 2D screen, they may be using XR where they are embodied as an avatar of their choice, actively engaging with other avatars in virtual worlds that feel real but are too expensive, dangerous, or impossible to experience in the physical world.

This possibility is both intriguing and daunting for fields like education wherein existing methods have long been criticized for not adapting to the opportunities and challenges of the twenty-first century ([Scott, 2015](#)). Research is therefore needed to harness the possibilities and understand the limitations of these technologies on human cognitive processing, communication, and learning. Reviews have specifically highlighted the need for theories and best practices to guide research and application development in the field of immersive learning (e.g., [Jensen & Konradsen, 2018](#); [Radianti et al., 2020](#)). With an expected surge of XR-supported collaborative learning (XRCL) content, it is important to develop empirical and theoretical knowledge that builds on the vast body of research on collaborative learning. The primary aim of this article is to critically examine the potential pedagogical benefits and limitations of using XRCL environments with the objective of developing a theoretical framework that describes the fundamental factors that make immersive collaborative learning unique. To this aim, we present the theory of immersive collaborative learning (TICOL), which provides an overview of the fundamental factors that are distinctive for collaborative learning in immersive virtual environments (see [Fig. 1](#)). The aim of TICOL is to build on the literature from computer-supported collaborative learning (CSCL)

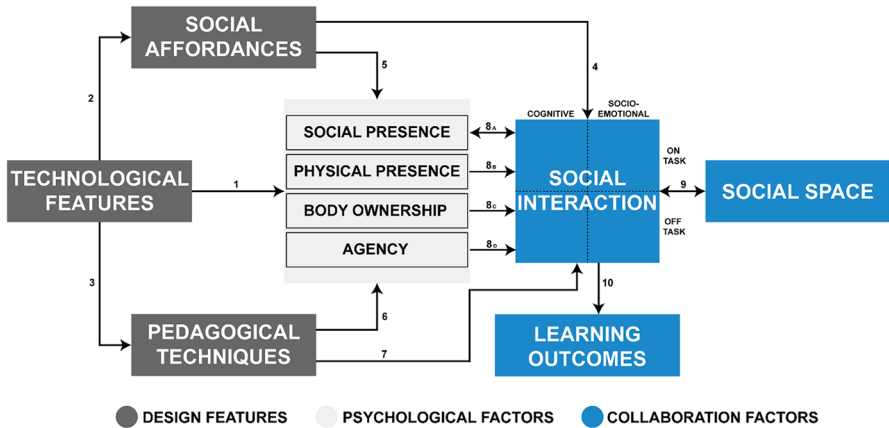


Fig. 1 An illustration of the theory of immersive collaborative learning (TICOL), including its central variables and the proposed relations between them

and immersive learning and develop a theoretical model that describes how the *fundamental psychological factors* of immersive media can facilitate collaborative learning. By fundamental psychological factors, we mean factors that result from using immersive technology, are considerably stronger than those with less immersive technology such as computers or mobile devices, and ultimately influence learning. We hope that this framework can guide a range of future lines of research and applications that can empirically test and further investigate the benefits and boundary conditions of using XRCL.

In TICOL, we highlight four psychological factors associated with XRCL that make collaborative learning fundamentally different compared to collaboration that occurs through traditional systems (e.g., laptops): *social presence*, *physical presence*, *body ownership*, and *agency*. Furthermore, we hypothesize that *technological features*, *social affordances*, and *pedagogical techniques* can foster these four psychological factors. Finally, we propose specific hypotheses about how each of these factors can influence different learning outcomes through their influence on the quality of *cognitive* and *socio-emotional social interaction*, the *social space*, and ultimately *learning outcomes* as illustrated in Fig. 1.

Defining Immersive Learning Environments

Extended reality (XR) is an umbrella term encapsulating augmented reality (AR), virtual reality (VR), and other technologies that enable an extension of reality. These technologies differ in how much of the outside world is included in the virtual experience. *Immersion* refers to the sensory fidelity offered by the system while shutting out the physical world (Bowman & McMahan, 2007; Cummings & Bailenson, 2016). Slater (2009) describes immersive systems in terms of the sensorimotor contingencies they support. Sensorimotor contingencies refer to actions that we are already familiar with in our interactions, such as turning our heads to change our

gaze direction (Slater, 2009). Based on this definition, an immersive VR system using an HMD is located on the far end of the immersive spectrum (see right part of Fig. 2), providing a vivid, multi-sensory experience through visual, auditory, kinaesthetic, and tactile displays (Slater, 2009).

Figure 2 illustrates how simulations accessed through a desktop computer or tablet are considered to provide low immersion because they have a limited field of view, and the user can see the physical world, whereas simulations accessed through HMDs are regarded as providing high immersion (Di Natale et al., 2020), because this equipment partially or completely shuts out the physical world, psychologically isolating the user in the virtual environment (Loomis et al., 1999). Augmented reality (AR) and augmented virtuality are located in the middle of Fig. 2, representing mixed reality applications which combine elements from the virtual and physical world and differ based on the extent to which the physical or virtual world is dominant in the field of view. AR applications (where virtual objects are overlaid in a real-world environment) are typically accessed through a 2D interface such as a phone or tablet and are considered to have lower immersion because the physical world dominates the field of view. Augmented virtuality applications are considered highly immersive because the virtual world dominates the field of view while simultaneously enabling elements from the real world such as a computer, mouse, and keyboard into the virtual experience.

TICOL focuses on describing the potential benefits and limitations of collaborating in immersive environments mediated by HMDs, and we refer to such collaborative learning experiences as extended reality–supported collaborative learning (XRCL). XRCL also encompasses modern HMDs that enable an immersive environment where users can integrate their computer, mouse, and keyboard into the virtual experience. Hence, we use the term XR rather than VR. The popularity of these systems is expected to surge in the near future as they provide the practicality of integrating elements from the real world while simultaneously allowing users to



Fig. 2 An immersion continuum ranging from low immersion systems such as desktop computers to high immersion systems such as head-mounted displays

manipulate who they are in terms of physical appearance, where they are in terms of virtual location, and what they can do in a highly immersive environment.

Defining Extended Reality–Supported Collaborative Learning (XRCL)

Collaborative learning is defined as “a situation in which two or more people learn or attempt to learn something together” (Dillenbourg, 1999, p. 1). Computer-supported collaborative learning (CSCL) investigates how information and communication technologies can facilitate group learning processes, knowledge sharing, and co-construction (Karel Kreijns et al., 2003; Resta & Laferrière, 2007; Stahl et al., 2006). A meta-analysis of research on CSCL by Chen et al. (2018) found that collaboration has a significant positive effect on knowledge gain, skill acquisition, and student perceptions and that computer use leads to positive effects on knowledge gain, skill acquisition, student perceptions, group task performance, and social interaction in collaborative learning contexts.

The literature on CSCL differentiates between collaborative learning *with* computers and *through* computers (Jeong & Hmelo-Silver, 2016). Collaboration with computers refers to collaboration in face-to-face settings around computers where computers become the focus of interaction. In another meta-analysis of CSCL in science, technology, engineering, and mathematics (STEM) education, Jeong et al. (2019) found a moderate but notable effect size advantage when using CSCL. In an analysis of the technology that was most effective, they found that simulations produced the greatest effects. In this context, simulations are usually used to help groups in face-to-face collaborative settings (Jeong et al., 2019), where learners use simulations to construct and test models like an understanding of electrical wiring (Liu & Su, 2011) or probability (Gürbüz & Birgin, 2012).

Immersive technology accessed through HMDs has also been used to create experiences that can facilitate discussions in face-to-face collaboration. For instance, Makransky and Mayer (2022) found that an immersive virtual field trip to Greenland was more effective than a video field trip when it was used in the exploration phase of a collaborative inquiry-based learning intervention. Similarly, Klingenberg et al. (2020) found that an immersive simulation led to significantly more learning than the same simulation presented through a desktop computer when it was followed by the collaborative activity of peer teaching. These examples revolve around face-to-face collaboration where students take on similar rules, roles, and tasks as in traditional CSCL with desktop computers but where simulations are used to elicit increased interest and engagement thereby facilitating a desired interaction among learners (Makransky & Petersen, 2021; Zurita & Nussbaum, 2004).

XRCL is not only useful as a means of supporting face-to-face interactions. In contrast, we argue below that immersive media has specific psychological factors that make collaboration through immersive media fundamentally different from traditional CSCL media such as desktop computers and mobile devices.

Collaboration through computers refers to distributed collaboration situations in which computers are used as a medium for social interaction (Jeong & Hmelo-Silver, 2016, p. 249). Similarly, TICOL deals with distributed collaboration situations

where XR technology (accessed through HMDs) is used as a medium for two or more people to interact via avatars, with the objective of learning (i.e., XRCL). We argue below that immersive media has unique characteristics and operationalize four factors that make immersive collaboration fundamentally different from collaboration via traditional media such as desktop computers and mobile devices. Furthermore, we argue that an understanding of these differences is important for optimizing XRCL and hypothesize how these four factors can influence learning outcomes.

Existing Research Related to XRCL

In general, the field of XRCL is still in its infancy. Han et al. (2022) recently reviewed the literature on collaborative virtual environments using HMDs or in some cases stereoscopic projection systems. They identified a total of 37 studies where at least 10 groups of people had been networked in VR and only a subset of those dealt specifically with learning. The main focus of these studies was the characteristics of the interaction between the participants such as presence, trust, body ownership, task performance, collaboration, or behaviour. An earlier meta-analysis performed by Zheng et al. (2018) examined ten articles that evaluated the effect of a collaborative learning prototype on student performance in comparison to a non-immersive-based approach. Although the mean effect size advantage of 0.41 for XRCL (labelled VRCL in their article) was encouraging, the authors did not provide an overview of the included articles or information about sample sizes or the quality of the methods.

The aforementioned work of Han et al. (2022) investigated collaborative learning during a higher education 10-week-long online course where 81 students met eight times using networked VR. The results showed that all measures including self-presence, social presence, physical presence, enjoyment, entitativity (degree to which a collection of people is perceived as a single, unified entity; Campbell, 1958), and realism (perceived photorealism of the VR environment and people) increased over time, suggesting that the advantages of XRCL may increase as participants adapt to the medium and are no longer affected by the novelty of the technology (Han et al., 2022).

Other studies have found that perceived presence and co-presence in a collaborative setting turned out to be higher in XRCL than in computer-based collaborations (Bayro et al., 2022; Sun et al., 2022). Many studies have investigated avatar appearance as a main driver for the experience of social presence during XRCL. Studies have found that behavioural realism (Herrera et al., 2018), locomotion, and continuous movements influence physical presence and social presence more than the style or look of the avatar (i.e., whether it is stylized, has a full body, etc.; Freiwald et al., 2021). In addition, research shows that when inferring the emotional state of a collaborator, participants prefer to focus on the voice instead of the properties of the avatar (Khojasteh & Won, 2021). Yoon et al. (2019) emphasized the context of collaboration as being more influential on social presence than avatar specifics (i.e., whether it is embodied, customized, etc.). A survey study by van Brakel et al. (2023) among anonymous social VR users concluded

that social presence and self-presence were predictors of perceived social support and that this perception was positively associated with users' well-being. The authors conclude that the affordances of social VR make it a particularly well-suited medium for facilitating beneficial interactions among users.

Sedlák et al. (2022) studied the effects of learning geography collaboratively in immersive VR and found that the collaborative learning group experienced significantly higher use of cognitive resources than the individual learning group. Akselrad et al. (2023) studied instances of miscommunication in social VR and found that technological failures such as "body crumple" (i.e., contorted virtual bodies), sound intrusion (i.e., unintended audio entering the microphone), and embodiment violations (i.e., getting "tangled up" in another's virtual body) inhibited effective communication. This highlights the importance of studying not only how XRCL can enhance collaborative learning but also how technological issues can inhibit collaborative learning.

Given the dearth of literature related to XRCL, TICOL also builds on a vast amount of research from the fields of CSCL and immersive learning. Meta-analyses have found a small effect size advantage for immersive lessons compared to less immersive technology (Coban et al., 2022; Wu et al., 2020). However, using XR without considering instructional design (Makransky et al., 2021; Makransky & Mayer, 2022), the affordances of the technology (Petersen et al., 2022), and the desired outcomes (Makransky, Borre-Gude, & Mayer, 2019) can lead to distraction and less learning (Makransky et al., 2021; Makransky, Terkildsen, & Mayer, 2019; Parong & Mayer, 2018) compared to traditional means of instruction. Similarly, in their meta-analysis, Jeong et al. (2019) found that CSCL outcomes were moderated by the educational level of learners, domains of learning, technology, and the pedagogy in which the technology was embedded. The key takeaway is that simply putting learners in an immersive environment will not lead to better learning and simply placing learners in a group and assigning them a task does not ensure that they will work together (Hughes & Hewson, 1998), coordinate their activities (Erkens et al., 2006), engage in effective collaboration (Hallett & Cummings, 1997), participate in beneficial discussions (Weinberger & Fischer, 2006), or learn more (Kirschner & Erkens, 2013, p. 1).

Each time a new medium enters the educational sphere, it generates over-expectations with respect to its intrinsic effects on learning (Clark, 1994; Dillenbourg et al., 2009). There is abundant evidence that the richness of a medium alone does not predict its effectiveness (Dillenbourg et al., 2009). Within the field of immersive learning, there is also evidence that immersive environments can increase presence and that this does not necessarily lead to more learning (Makransky et al., 2021; Makransky, Terkildsen, & Mayer, 2019; Parong & Mayer, 2018). The consequence is not only that exaggerated claims about the learning benefits of new media generate unfounded expectations, but also that they lead to the neglect of technological benefits (Dillenbourg et al., 2009). Therefore, our goal is to identify the specific psychological factors that are relevant to learning with immersive technology and use existing research to develop hypotheses about how they can influence collaborative learning.

The Building Blocks of TICOL

In addition to the literature on XRCL presented above, TICOL draws on CSCL literature, including Kreijns et al. (2013) *social aspects of CSCL research framework* and *collaborative cognitive load theory* (Kirschner et al., 2018). TICOL also builds on theories and principles of individual immersive learning, including the *immersion principle in multimedia learning* (Makransky & Mayer, 2022) and the *cognitive-affective model of immersive learning* (CAMIL; Makransky & Petersen, 2021), which describe the distinctive characteristics of learning in immersive environments. Below, we present TICOL and start by providing a description of what makes collaborative learning through immersive media fundamentally different from collaboration through traditional media. We then use existing research to describe the central factors and hypothesized relations between these factors.

The Theory of Immersive Collaborative Learning

Why Collaboration Through XR is Fundamentally Different from Collaboration Through Computers?

TICOL identifies four fundamental psychological factors that are central to XRCL: social presence, physical presence, body ownership, and agency. These four factors are identified as the main psychological factors in TICOL because they have theoretical and empirical value with regard to collaborative learning and are experienced more strongly in XR compared to less immersive media (Johnson-Glenberg, 2019; Makransky & Petersen, 2021). For instance, CAMIL (Makransky & Petersen, 2021) highlights *presence* and *agency* as the fundamental affordances of using immersive technology in learning. The immersion principle in multimedia learning describes how people learn better with immersive media than with standard media when immersive lessons are designed according to instructional design principles and the affordances of the technology, which include physical and social presence (Makransky, 2022; Makransky & Mayer, 2022). The sense of embodiment (sensation that arises in conjunction with being inside, having, and controlling a virtual body) has also been described as an important psychological factor for learning with immersive VR in a number of models and frameworks (e.g., Dincelli & Yayla, 2022; Shin, 2017; Slater, 2017). This construct consists of three dimensions, including body ownership (“this is my body”), agency (“it is me who is acting”), and self-location (“it is me who is here”; Kilteni et al., 2012; Mottelson et al., 2023), with the first two being expected to play a significant role in XRCL. In summary, social presence, physical presence, body ownership, and agency have been identified as central factors that differentiate learning in immersive environments from learning in less immersive environments in previous empirical research and theoretical models. There is a vast foundation of research about the operationalization and measurement of each of these factors which goes beyond the scope of the current article. Below, we will summarize a selection of this research and describe how each factor can influence XRCL.

Social Presence

Social presence has played an important role in CSCL and online learning as well as immersive technology research, which has led to a number of different definitions and conceptualizations of the construct. Kreijns et al. (2022) recently reformulated the original Short et al.'s (1976) definition of social presence as “the psychological phenomenon in which, to a certain extent, the others are perceived as physical “real” persons in technology-mediated communication enabled by computer-mediated communication tools and electronic platforms” (Kreijns et al., 2022, p. 156). As demonstrated by a recent review, experiences of social presence are commonly solicited by means of self-report measures (Oh et al., 2018). XRCL systems allow researchers to collect objective data, such as proxemics and mutual gaze, which are not measures of social presence but rather social responses that can happen as a function of social presence (Lee, 2004). A high level of social presence leads to behaviour that could be expected from similar situations in reality. For instance, people experiencing a high level of presence would maintain a farther distance from someone they did not know, or if someone were communicating a message, basic human rules of communication would prompt us to listen actively (Mayer, 2014). What is unique about immersive technology is that it affords a significantly higher level of social presence compared to less immersive media (Oh et al., 2018) while simultaneously allowing users to manipulate their physical appearance. For instance, imagine collaborating with someone who, in the blink of an eye, can shift their appearance from embodying an elderly Caucasian female avatar to a young Asian male avatar. The fact that a collaborator can instantaneously change who they are in terms of their physical characteristics in a learning setting where the social interaction is perceived as real fundamentally challenges the central assumptions of many educational psychological theories related to social interaction and stimulates many new research questions related to XRCL.

Physical Presence

Physical presence is defined as the psychological feeling of “being there” in the virtual environment (Ijsselstein & Riva, 2003). Currently, one of the most used conceptualizations of presence was coined by Slater (2009). According to Slater, presence consists of two dimensions: place illusion and plausibility illusion. Place illusion is, as the name indicates, the illusion of being in the place that is depicted in VR in spite of the knowledge that this is not the case (Slater et al., 2022). The cause of place illusion can be traced to the fact that perception in VR occurs through sensorimotor contingencies (i.e., immersion: e.g., rotating our heads to change our gaze direction; Slater et al., 2022). The other constituent of physical presence, plausibility illusion, refers to the illusion that events in VR are actually occurring despite the knowledge that these are digitally generated (Slater et al., 2022). The plausibility illusions are generated when the virtual environment is responsive to the user's actions, when events in the environments are happening with reference to the user, and when the virtual events happen in accordance with one's expectations (Slater et al., 2022). When both illusions (place and plausibility) are at play, people tend to

respond realistically to virtual situations and events despite the fact that they know these are not real (Slater et al., 2022). Experiences of physical presence are often collected via questionnaires administered after the learner has been immersed (e.g., Makransky et al., 2017). However, psychophysiological measurement has also been used (e.g., Schöne et al., 2023; Terkildsen & Makransky, 2019).

A meta-analysis by Cummings and Bailenson (2016) found that more immersion is linked to higher physical presence. Therefore, XRCL allows learners to change the physical context they are in instantaneously (Bailenson, 2018; Markowitz & Bailenson, 2021) while affording the feeling of “being there” to a greater extent than standard media. This allows learners to experience the illusion of being in and exploring different worlds, such as can be achieved via XRCL platforms like Engage VR (Han et al., 2022) where users can travel to an underwater coral reef, to a museum, or to Mars, all within seconds.

Body Ownership

Body ownership is the illusion that a virtual body belongs to oneself (Slater et al., 2022). The experience of body ownership is often measured via subjective questionnaires (e.g., Gonzalez-Franco & Peck, 2018). However, objective measures, such as galvanic skin response, have been used to infer body ownership in experimental setups involving threats to a virtual body (Yuan & Steed, 2010). Recall the prior example wherein we described how XRCL allows users to embody multiform avatars of different ethnicity or age. There is evidence that people align their behaviour and attitudes with their avatar’s characteristics (Yee & Bailenson, 2007). Although social psychological theories attempt to understand how people create and define their place in a social group based on characteristics they have had their entire lives (Tajfel & Turner, 2004), current learning theories do not account for what happens when these characteristics are transient. On a related note, body ownership illusions make it possible to experience the world from other people’s perspectives (Slater et al., 2010). This has led some to label VR as the ultimate “empathy machine” (Herrera et al., 2018). Maister and colleagues (2015) reviewed the evidence on the role of body ownership in implicit social bias and found that ownership of an out-group individual (i.e., a person with a different gender, age, or race from oneself) can reduce implicit biases against that group. However, little is known about the consequences of body ownership in collaborative learning scenarios. On the other end of the spectrum, XRCL may also provide a neutral environment where learners are able to interact and engage in productive collaboration free of stereotypes and biases that they would normally be unable to escape in the physical world since it affords the option of controlling neutral avatars that do not connote gender, race, etc. (Han et al., 2023).

Agency

Agency is a central aspect of human existence and refers to the experience of being the director of one’s own actions (David et al., 2008). Conceptually, agency (i.e., “it is me who is acting”) is different from body ownership (i.e., “this is my body”).

Agency is typically measured through judgements or attributions using self-reports or errors through misidentification (e.g., David et al., 2007; Frith et al., 1999; Motelson et al., 2023). Multivariate approaches that include implicit measures such as kinematics, eye movements, or brain activity have also been used to infer agency (e.g., Jeunet et al., 2018; Sperduti et al., 2011).

XRCL environments include visual, audio, kinesthetic, and tactile feedback, which makes it possible to interact with objects and others in a similar fashion to how these interactions take place in the physical world, but without the limitations of the physical world. XRCL environments allow learners to be and do whatever they want within the limits of the technological features, social affordances, and pedagogical techniques as illustrated in TICOL (see Fig. 1). For instance, a learner can tap a collaborator on the shoulder, hold their hand, and together jump off the Navajo Bridge in Arizona after visiting a national heritage site and then fly hand in hand to the Great Wall of China. Although these technological features in themselves do not ensure that learners will have a high sense of agency, it is more likely that learners will feel agency in the aforementioned example compared to a lesson where these actions are governed by the system (Petersen et al., 2022).

In the next section, we will provide a definition of the other central factors in TICOL prior to describing the specific predictions and central relationships proposed by the theory. This is followed by a description of the implications of using and designing XRCL environments, as well as a future research agenda.

Other Central Factors in TICOL

In addition to defining the fundamental psychological factors that make collaboration through immersive media fundamentally different from collaboration through traditional media, TICOL also proposes how these factors are related to design features and how they influence learning outcomes. Figure 1 illustrates the central variables and hypothesized relationships in TICOL. Prior to describing these relationships, we will describe each factor in TICOL in more detail starting with the design features on the left side of Fig. 1.

Technological Features

Technological features include hardware and software design specifications including, but not limited to, the level of immersion, the amount of control factors, and the representational fidelity of the environment (Makransky & Petersen, 2021). The level of immersion is determined by the visual, auditory, kinesthetic, and tactile elements of an XRCL environment and includes the tracking level, image quality, sound quality, the size of the field of view, and update rate, among other factors (Cummings & Bailenson, 2016; Slater, 2009). Control factors include the degree of control, immediacy of control, and mode of control (Witmer & Singer, 1998). Finally, representational fidelity includes the realism of avatars and virtual environments, as well as consistency of object, avatar, and agent behaviour including novel, sophisticated approaches based on large language models that enable agents

to respond to the user in natural language (Dalgarno & Lee, 2010; Park et al., 2023). The technical features of an XRCL environment influence the fundamental psychological factors in TICOL directly and dictate what social affordances and pedagogical techniques are possible. However, as we will describe below, technical features are distinct from social affordances or pedagogical techniques because the latter stimulate cognitive or socio-emotional interaction whereas technological features are simply design specifications of the XRCL.

Social Affordances

Kreijns et al. (2013) refer to social affordances in their framework as the social-contextual facilitators relevant to the learners' socio-emotional interactions. In other words, these serve non-cognitive aspects such as group culture and can include tangible tools that facilitate avatar-based interactions or intangible factors such as policies and rules governing the XRCL groups.

Pedagogical Techniques

Pedagogical techniques refer to instructional design factors such as the use of scripts (Vogel et al., 2017; Petersen et al., 2023), methods for peer feedback (Xiao & Lucking, 2008), role assignment (Strijbos et al., 2004), and the use of representational tools (Janssen et al., 2007). For instance, scripts are computational instructional supports that instruct learners about how to interact in ways that trigger the elaboration of the learning material (Vogel et al., 2017). An example of peer feedback is learning by teaching, which involves explaining the learning material to another (Fiorella & Mayer, 2016). Collaborative mapping is an example of a representational tool that involves jointly creating a diagram that represents the learning material (Adesope et al., 2022). According to collaborative cognitive load theory (CCLT), collaborative learning tasks need to be sufficiently complex to make up for the extra resources spent on communicating and coordinating (Kirschner et al., 2018). CCLT also highlights the fact that such tasks need to be sufficiently guided in order to not overload the learners (often achieved via scripts; Kirschner et al., 2018).

Social Interaction

In their framework, Kreijns et al. (2013) describe how social interaction may occur in the cognitive or socio-emotional dimension as well as in task or non-task contexts. Social interaction in the cognitive dimension serves the acquisition of knowledge and skills and includes activities such as knowledge sharing and monitoring. Conversely, interaction in the socio-emotional dimension involves processes related to group development and dynamics including encouragement and positive appraisal (Isohätälä et al., 2020). Social interaction that is on-task refers to the social interaction that takes place in the strictly defined task setting, which can be thought of as the virtual classroom, while interaction in non-task contexts refers to everything that can be considered outside of the learning setting, such as casual "getting to know one another" areas or a virtual recreational area which serves the purpose of

informal social interaction. It is important to note that task and non-task contexts may include cognitive and socio-emotional social interaction (and vice versa; Kreijns et al., 2013). Consistent with Kreijns et al. (2013), TICOL distinguishes between social interaction that occurs in the cognitive or socio-emotional dimension as well as in on-task or off-task contexts.

Social Space

Kreijns et al. (2003, p. 608) define social space as “the network of social relationships among the group members embedded in group structures of norms and values, rules and roles, beliefs and ideals.” Furthermore, Kreijns et al. (2022, p. 159) describe a sound social space as a group attribute “manifested by a sense of community, group climate, mutual trust, social identity, and group cohesion.” Social space also has a cultural structure because interpersonal relationships are embedded within norms and values, rules and roles, and beliefs and ideals (Blanchard & Lynne Markus, 2004; Brook & Oliver, 2002). Therefore, collaboration groups with a sound social space experience relational structures and a shared social identity that includes group cohesiveness, connectedness, mutual trust, a sense of belonging, a sense of community, a social climate, and an open atmosphere (Kreijns et al., 2022).

Learning Outcomes

Learning outcomes are defined as the particular knowledge, skills, attitudes, and behaviours that are acquired during collaborative learning. Based on Anderson et al.’s (Anderson et al., 2001) taxonomy of learning, teaching, and assessing, knowledge can include factual knowledge, conceptual knowledge, procedural knowledge, and meta-cognitive knowledge. Furthermore, we include the transfer of knowledge since it is highlighted as an ultimate goal of education (Mayer & Fiorella, 2022; Prawat, 1989).

Hypothesized Relationships Between Central Components of TICOL Based on Existing Research

The purpose of a collaborative learning environment is not simply to enable collaboration across distances, but to create conditions in which effective group interactions can occur (Dillenbourg et al., 2009). The crucial point in terms of building XRCL environments that are effective learning tools thus lies in understanding the complex interplay between design features, the fundamental psychological factors that make XRCL unique, and how these influence social interactions, social space, and learning outcomes. TICOL predicts many specific, hypothesized paths that are developed based on existing research from CSCL, online learning, immersive learning, and related literature. These hypothesized paths are described in the following and depicted in Fig. 1 as different paths labelled one through ten.

The Influence of Technological Features on the Four Fundamental Psychological Factors of XRCL (Hypothesized Path 1)

There is a vast amount of research investigating the technological features that influence the four fundamental psychological factors of XRCL. Summarizing the entire body of literature is beyond the scope of this article. Rather, we focus on presenting a selection of key findings regarding each of the four factors to support the hypothesized path 1. Regarding social presence, Oh et al. (2018) conducted a systematic review of 233 separate findings identified from 152 studies that investigate the factors that predict social presence. They found that technological features including depth cues, audio quality, haptic feedback, and interactivity most often had positive effects on social presence. Furthermore, there is meta-analytic evidence that more immersion is linked to higher physical presence (Cummings & Bailenson, 2016). More specifically, Cummings and Bailenson aggregated 115 effect sizes from 83 studies and found that immersion had a medium-sized effect on physical presence. Their results identified increased levels of user tracking, the use of stereoscopic visuals, and wider fields of view of visual displays as the most important technological features influencing physical presence. A recent review and meta-analysis by Mottelson et al. (2023) surveyed 111 papers including 4925 participants to investigate the technological features that influence body ownership. Mottelson and colleagues found that the manipulation of visuo-motor synchrony rendered the largest effects on body ownership and that congruence of appearance, perspective, visuo-tactile stimuli, and abstraction of the avatar were also effective manipulations. Makransky and Petersen (2021) described technological features such as being able to control a body representation (including factors such as degree of control, immediacy of control, and mode of control), as well as the ability to modify the environment and its objects, as the most important predictors of agency. Kilteni et al. (2012) described how the accordance between actual movements and corresponding visual feedback is important for creating agency, otherwise known as forward modelling of the central nervous system (Farrer et al., 2008). Mottelson et al. (2023) confirmed this association in their meta-analysis where they found that visual asynchrony was the most influential negative manipulation on agency.

The Association Between Technological Features and Social Affordances (Hypothesized Path 2)

The technological features of an XRCL system define how social affordances can be implemented into an XRCL environment. An example of a social affordance is a meet-and-greet area where learners can greet each other by giving a high five, which is a feature in the Engage VR platform (Han et al., 2022). Here, the technological features enable haptic feedback when learners high-five each other's avatars. Notice that haptic feedback is a technological feature of the XRCL environment, which intensifies the social affordance (the high five in the meet and greet area). The first thing a learner usually does when seeing a collaborator's avatar for the first time in an XRCL session in Engage VR is to approach them and give them a high five. This instigates immediate social interaction and

typically leads to informal conversation (socio-emotional interaction; hypothesized path 4, described below). Social affordances also increase social presence (e.g., due to the fact that the high five with another avatar increases the sense of interacting with another person; hypothesized path 5, described below) and body ownership (e.g., as the visual and tactile synchrony of seeing your virtual hand react to your actual movement from a first-person perspective increases the sense that it is your body; hypothesized path 5). The above example is a simple example of how the way in which social affordances are designed in XRCL platforms will depend on the technological features of the system.

The Association Between Technological Features and Pedagogical Techniques (Hypothesized Path 3)

Similar to the relation between technological features and social affordances, technological features also determine how pedagogical techniques can be implemented in an XRCL environment. Above, we referred to scripts, which have been shown to be effective in CSCL (Vogel et al., 2017). For instance, consider a co-construction activity in an XRCL environment where students are prompted to build an enlarged virtual cell as part of a virtual tour through the human bloodstream: Prompting students to collaboratively build the cell by combining the different parts can facilitate constructive social interaction in the cognitive dimension and lead to enhanced spatial knowledge for both learners (Petersen et al., unpublished manuscript). The technological features dictate the scripts that can be implemented in an XRCL environment. For instance, recent advances in artificial intelligence concerning large language models allow the integration of interactive pedagogical agents capable of supporting and guiding the collaborative learning process in a dynamic way (Park et al., 2023). Traditionally, however, VR scripts are modelled on real-life scaffolds such as an interactive virtual tablet that provides instructions (Makransky, Wismer, & Mayer, 2019).

Another example of a pedagogical technique is the signalling principle, which states that deeper learning is achieved when cues are incorporated that direct learners' attention toward critical elements of the multimedia learning material (van Gog, 2022). Notice that signalling is a pedagogical technique that can be incorporated into an XRCL environment differently depending on the technological features that are available. For instance, in the above example where students have to co-construct a cell, students are instructed to select the next component to build the cell. The student responsible for initiating the task could signal this by highlighting it from a list of components if the technological features are limited. Alternatively, if the technological features allow students represented as avatars to move freely around a 3D model of a cell and see the actions and gaze patterns of other collaborators, then simply looking at a particular component would provide a signal of the intention to use that as the next component to build the cell.

The Influence of Social Affordances on Social Interaction (Hypothesized Path 4)

While TICOL posits that social affordances influence social interaction indirectly through the four fundamental psychological factors (to be discussed below), these affordances can also influence social interaction directly. Kreijns and Kirschner et al. (2018) describe how social affordances influence social interaction through factors such as sociability (the degree to which the CSCL system supports socio-emotional interaction) and hedonicity (the extent to which a CSCL system is designed to incite pleasure and enjoyment during the interaction). Social media are good examples of platforms that have many types of social affordances that are designed to engage users to interact (Bucher & Helmond, 2018). Keenan and Shiri (2009) linked the sociability of four social media platforms including Facebook, Myspace, LinkedIn, and Twitter with the high degree of social interaction that these applications facilitate. Research on the adoption of social media has also found that hedonicity is a factor that plays a role in increasing social interaction and the adoption of these systems (van Koningsbruggen et al., 2017). A casual getting-to-know-one-another area, e.g., modelled on real-life conference facilities, is another example of a social affordance capable of sparking social interaction—particularly in the socio-emotional dimension.

The Influence of Social Affordances on the Four Fundamental Psychological Factors of XRCL (Hypothesized Path 5)

There is existing evidence that social affordances can influence social presence. For instance, in their review of the contextual factors that influence social presence, Oh et al. (2018) found that physical proximity, identity cues, and the personality/traits of virtual humans (social affordance factors) were often significant predictors of social presence. On a theoretical level, the *threshold model of social influence in digital environments* (TMSI) provides a framework for understanding how digital humans promote social presence and exert social influence (Blascovich, 2002; Ryan et al., 2019). According to TMSI, the extent to which digital humans elicit social presence, and thereby potentially social influence effects (i.e., changes in an individual's cognitions, attitudes, physiological responses, and behaviours), is a function of three independent factors: agent/avatar, communicative realism, and response system level. Communicative realism refers to the movement realism, anthropomorphic realism, and photographic realism of the virtual other and can be considered a type of social affordance that can influence social presence (Ryan et al., 2019).

There is also theoretical support for the link between social affordances and physical presence. As described above, when place and plausibility illusions are at play, people tend to respond realistically to virtual situations and events despite the fact that they know these are not real (Slater et al., 2022). Take for instance a group of students who are given the joint task of investigating the consequences of climate change on a coral reef. This could occur in an XRCL environment through exploring a vibrant coral reef together from both the perspective of the present time as well as the future to provide a different perspective that shows the consequences of climate change from a first-person perspective (Plechata, Morton, et al., 2022). Then,

students could choose to take a break by virtually visiting an exotic recreational environment where they can collaborate on a teambuilding activity. The social affordance of having an exotic recreational environment can prime the sense of “being in” and interacting with collaborators in the exotic recreational environment to the extent that place and plausibility illusions are incorporated into the XRCL.

There is empirical evidence supporting and describing how social affordances can also influence body ownership. A recent meta-analysis by Mottelson et al. (2023) found that appearance manipulations were effective in inducing body ownership. In XRCL settings, it is possible to alter avatars on many dimensions, including gender, age, race, height, and attractiveness, in addition to changing nonverbal behaviour such as smiling more or less. The social affordances of the XRCL can therefore influence body ownership to the extent that these potential manipulations can effectively induce body ownership.

Finally, social affordances are also hypothesized to influence agency. For instance, allowing learners to select their own avatar could increase their sense of agency compared to a lesson where they are randomly assigned one.

The Influence of Pedagogical Techniques on the Fundamental Psychological Factors of XRCL (Hypothesized Path 6)

The hypothesized path 6 predicts that pedagogical techniques can influence the fundamental psychological factors of TICOL, which in turn influence social interaction (hypothesized paths 8A–D described below). Dubosc et al. (2021) conclude that offering a task context that generates dependence between collaborators improves social presence in an immersive collaborative learning environment. This is consistent with the collaboration principle in multimedia learning, which predicts that the quality of interaction is enhanced when collaborators experience interdependence during learning (Janssen et al., 2022).

An example of a pedagogical technique which could influence cognitive social interaction through increasing the fundamental psychological factor of physical presence is that students could be given the task (script) of building an auto repair shop in the virtual environment by selecting and customizing available virtual assets (parts) prior to building a new motor. The pedagogical technique would be directly related to physical presence in the sense that students would feel like they are in an auto repair shop (compared to a lesson where they were not given that script and were presented with the motor without the environment). The task of building the auto repair shop would likely facilitate discussions about how to build the motor (on-task cognitive social interaction). Furthermore, providing the students the task of building the auto repair shop, rather than simply immersing them in the shop, would increase agency which would also likely stimulate more cognitive as well as socio-emotional social interaction. In general, pedagogical techniques can play a significant role in learners’ agency in XRCL. For instance, learners may experience low agency in an XRCL lesson where they passively sit through a direct instruction lecture in a virtual classroom with little control, compared to a discovery-learning lesson where they are able to gain the same knowledge through exploration (Makransky & Petersen, 2021).

The Influence of Pedagogical Techniques on Social Interaction (Hypothesized Path 7)

TICOL builds on the vast amount of CSCL research which highlights key processes that drive collaborative learning. Most pedagogical techniques in CSCL are designed to stimulate productive social interaction (e.g., Gunawardena, 1995; Rourke & Anderson, 2002). Kreijns et al. (2013 p. 237) describe how “without a specific pedagogy, social interaction among group members will not arise.” A great deal of CSCL research has focused on developing such CSCL pedagogy, and pedagogical techniques that have been successful at enhancing social interaction include the application of pedagogical scripts (e.g., Dillenbourg & Tchounikine, 2007; Fischer et al., 2013) or predefined roles (e.g., Strijbos & De Laat, 2010). We acknowledge this research and expect that these findings will generalize to XRCL. We therefore suggest that the design of XRCL applications build on the fundamental research that has been conducted in CSCL, viewed through the lens of TICOL.

While most pedagogical techniques are based on CSCL literature, XR environments also allow for novel pedagogical techniques that are not possible with less immersive media. For instance, the XRCL platform Engage VR makes it possible for students to record a session where their avatars are working on a task, then “pause” or “rewind” during a conversation, and be embodied into that scene at the same time as they are observing and recording the interaction, thereby allowing for a new dimension of generative learning activities where they can see their own avatar interact with others and reflect over their previous behaviour (Fiorella & Mayer, 2016).

The Influence of Each of the Four Fundamental Psychological Factors on Learning Outcomes Through Social Interaction and Social Space (Hypothesized Paths 8A–D, 9, and 10)

Since the objective of TICOL is to identify factors that differentiate XRCL from CSCL and to describe how these factors can influence learning outcomes through increased social interaction and improved social space, we will now describe these combined hypothesized paths for one fundamental psychological factor at a time in the following sections.

The Influence of Social Presence on Learning Outcomes Through Social Interaction and Social Space (Hypothesized Paths 8A, 9, and 10) There is abundant research that highlights how social presence influences social interaction in online learning groups (Tu & McIsaac, 2002; Zhao et al., 2014). For instance, Poth (2018) describes how the development of social presence can promote a more engaging and supportive educational experience, in which students become more motivated and can attain more success through social interaction. One of the primary goals of networked communication systems is to offer high levels of social presence (Biocca & Harms, 2002; Oh et al., 2018). Social presence is also a central component in many frameworks related to online learning, including Garrison’s (Garrison, 2000) community of inquiry framework, which includes social presence as a primary factor. Garrison and Arbaugh (2007) concluded from a literature review that social presence is one

important factor essential for creating a community of inquiry and for designing, facilitating, and directing higher-order learning. The research on the influence of student group cohesiveness and interaction on team effectiveness in online courses also suggests a strong relationship between social presence and learning outcomes (Hwang & Arbaugh, 2006; Williams et al., 2006). The paths from social presence to learning outcomes through social interaction (Paths 8A and 10) are also supported in previous research (Song & Yuan, 2015; Tu, 2000).

Social presence is a central component in enabling social interaction. What is unique about immersive technology is that it affords a significantly higher level of social presence compared to less immersive technology (Bailenson, 2018). The fundamental question is therefore how much social presence is needed to influence the quality, content, or intensity of the socio-emotional and cognitive social interaction. It is important to note that in their review, Oh et al. (2018) describe how social interaction also tended to “increase participants’ feelings of social presence” (p. 25). In this way, we propose a multidirectional path between social presence and social interaction.

In general, TICOL predicts that higher social presence can result in more social interaction but that this will only lead to improved social space and more learning to the extent that the social interaction is productive for the specific learning outcomes. Kirschner et al. (2009) refer to the transactional costs involved in collaboration and highlight how collaborative learning works best when the costs in terms of invested time and effort are exceeded by the benefits in terms of effectiveness and efficiency of learning. This is also formulated in the collaboration principle in multimedia learning, according to which the costs of collaboration (e.g., in terms of coordination) must be more than compensated for by the added value of being able to distribute information processing among group members in order for collaborative learning to be effective (Janssen et al., 2022). The authors highlight how learners in CSCL experience working in teams as positive to the extent that they feel certain that the extra time and effort that need to be invested to work together pays off. This means that the costs of communication with others and coordination of activities are compensated by the returns in terms of ease of learning. However, it is important to acknowledge that strong social space does not necessarily coincide with productive cognitive interactions. Social media platforms are an excellent example of platforms that use social affordances to prompt users to remain connected and engage in socio-emotional interaction, but this may not result in more cognitive interaction.

The Influence of Physical Presence on Learning Outcomes Through Social Interaction and Social Space (Hypothesized Paths 8B, 9, and 10) The affordances of 3D learning environments are well documented and include the facilitation of tasks that lead to enhanced spatial knowledge representations, greater opportunities for experiential learning, increased motivation/engagement, improved contextualization of learning, and richer/more effective collaborative learning as compared to tasks made possible by 2D alternatives (Dalgarno & Lee, 2010, p. 10). There is meta-analytic evidence that more immersion is linked to higher physical presence (Cummings & Bailenson, 2016). Therefore, one of the fundamental aims of TICOL is to identify the contexts where physical presence is essential for stimulating the different components of

social interaction to the extent that a better social space and ultimately better learning outcomes are produced.

Studies of immersive learning highlight application domains where the opportunity to virtually transport learners to any environment instantaneously is particularly useful. These include environments that would be dangerous (e.g., what to do if you get acid on your skin in a lab; Makransky, Borre-Gude, & Mayer, 2019), expensive (e.g., a virtual field trip to Greenland; Makransky & Mayer, 2022), or impossible (e.g., travelling through the body; Klingenberg et al., 2022) to experience in the real world (Bailenson, 2018; Makransky & Petersen, 2021), with the possibility of overcoming limitations of space and time (Plechata, Makransky, & Böhm, 2022). In collaborative learning settings, it can also be beneficial for learning to have group members who are physically separate work together in an environment that does not necessarily live up to these criteria but where the properties of the environment prompt physical presence which improves cognitive and socio-emotional interaction.

The role of physical presence in immersive learning has been studied extensively, and the CAMIL (Makransky & Petersen, 2021) and the immersion principle in multimedia learning (Makransky, 2022; Makransky & Mayer, 2022) describe how physical presence can be beneficial to leaning to the extent that the sense of “being there” in the virtual environment is important for the learning task and stimulates task involvement through variables such as situational interest. In collaborative learning, we hypothesize that the environmental context can play an even greater role in stimulating the cognitive and socio-emotional dimensions of social interaction and, through these, affecting the social space and the learning outcomes, as XRCL environments can provide a relevant shared contextualized experience. This can occur through establishing a joint context that is rich and engaging and allows for perspective-taking, which have all been shown to be important for successful collaborative learning. For example, the XRCL platform Engage VR (Han et al., 2022) offers a pedagogical technique that allows learners to change the physical context they are in instantaneously. As described above, this allows students to experience the illusion of being in and exploring different worlds, such as an underwater coral reef, then in an instant “be” in a museum, on Mars, or virtually travel to the same coral reef in the future where the habitat is destroyed. This could mean that a strong experience of physical presence could prime interaction in the socio-emotional domain (e.g., discussions about how the surroundings create a sense of fright in the group) and the cognitive domain (e.g., discussions about how the group can use the knowledge they have obtained to alleviate the consequences of climate change) while immersed in the underwater environment. Then, students could choose to take a break by virtually visiting an exotic recreational environment where they can collaborate on a teambuilding activity (non-task interaction in the socio-emotional dimension) and where students may even generate ideas and discuss how to solve the collaborative task (non-task cognitive dimension).

Co-construction has been highlighted as a central affordance of CSCL (Jeong & Hmelo-Silver, 2016). Optimal co-construction occurs when groups construct a shared frame of reference, or cognitive space, in which they obtain relevant knowledge, build on each other’s contribution, and co-construct potential solutions to build new knowledge (Matusov, 1996; Suthers, 2006). Representational tools have

been shown to be valuable for developing co-construction (LeBaron & Koschmann, 2003), and although the example described above would also be possible with mobile or desktop media, XRCL environments provide a contextualized experience that feels real (i.e., higher physical presence). This is specifically relevant for the transfer of learning where XRCL environments can be created to have high psychological and physical fidelity thereby decreasing the gap between the context where learning takes place and the context where the knowledge is intended to be used (Makransky, Borre-Gude, & Mayer, 2019).

The Influence of Body Ownership on Learning Outcomes Through Social Interaction and Social Space (Hypothesized Paths 8C, 9, and 10) XRCL environments make it possible to manipulate avatar body representations in terms of structure, morphology, and size (Kilteni et al., 2012), which can have an important psychological impact. According to the Proteus effect introduced by Yee and Bailenson (2007), people align their behaviour and attitudes with their avatar's characteristics. In a meta-analysis, Ratan et al. (2020) identified a small-to-medium-sized effect in favour of the Proteus effect across 46 experimental studies in which avatars with specific characteristics were randomly assigned to participants. The explanation for the effect can be found in the self-perception theory (Bem, 1972), which argues that people determine their attitudes by interpreting the meaning of their own behaviour in a situation, and from a stereotyping effect that posits that people behave in a situation according to how others would expect a person with such characteristics to behave (Yee & Bailenson, 2007). There is also evidence showing how the body ownership of specific avatars can influence social interaction. For instance, Yee and Bailenson (2007) found that people embodied as an attractive rather than an unattractive avatar in collaborative VR were willing to move closer to one another and disclose more information. In a recent study where researchers analyzed the linguistic patterns of group conversations in a collaborative VR course with 171 students, DeVaux et al. (2023) found that virtual representations of self were the most robust theme of conversations. In this way, body ownership was a source of socio-emotional social interaction.

The mechanisms through which body ownership can influence social interaction, social space, and ultimately learning outcomes are complicated and can include many other factors. Slater et al. (2010) showed that a body ownership illusion makes it possible to experience the world from other people's perspectives and that it provides a possibility to change roles and combat stereotypes. The science of implicit cognition suggests that many mental processes function implicitly (i.e., outside of the conscious attentional focus). Therefore, when collaborating, learners do not always have conscious, intentional control over such processes as social perception, impression formation, and judgement which motivate their actions (Greenwald & Krieger, 2006, p. 946). Processes that are relevant for XRCL include implicit attitudes (Greenwald et al., 1998), implicit stereotypes (Blair, 2001), implicit memory (Schacter, 1987), and implicit perception (Kihlstrom et al., 1992), which can influence the quality of the social interaction, social space, and ultimately learning outcomes. Prestige-biased social learning occurs when individuals predominantly choose to learn from a prestigious member of their group, i.e., someone who has

gained attention, respect, and admiration for their success in some domain (Brand et al., 2021). In XRCL, it is possible for collaborators to embody prestigious member of a group, and previous research has found that embodying an avatar who is perceived as very intelligent can lead to better performance on a cognitive task compared to embodying an avatar with normal intelligence (Banakou et al., 2018). This suggests that body ownership could potentially influence social interaction, social space, and learning outcomes as predicted by TICOL.

Many diversity-related educational initiatives have been suggested to promote positive intergroup relations (Denson, 2009), including implicit bias training (Girod et al., 2016). In XRCL, learners are able to embody an avatar of choice and collaborate in an environment where they are not met by the potential stereotypes and biases they would normally be unable to escape when interacting in the physical world which could influence learners' social interaction, social space, and ultimately learning outcomes according to TICOL.

The Influence of Agency on Learning Outcomes Through Social Interaction and Social Space (Hypothesized Paths 8D, 9, and 10) The literature on embodied cognition is one theoretical lens which helps understand how agency can influence social interaction. This literature describes how human cognition is deeply rooted in the body's interactions with the world and our systems of perception (Barsalou, 1999; Wilson, 2002). XRCL environments can be highly interactive platforms allowing for kinesthetic learning experiences, which has been shown to result in positive learning outcomes due to the high degree of collaboration (on-task cognitive interaction) engendered by the co-located environment (Johnson-Glenberg et al., 2014).

More broadly, social cognitive theory highlights the importance of collective agency which is a group of learners' belief in their collective capacity to achieve given attainments (Bandura, 2006). Meta-analyses have found that perceived collective efficacy accounts for a large portion of the variance in the quality of group functioning (Gully et al., 2002; Stajkovic & Lee, 2001). The CAMIL (Makransky & Petersen, 2021) describes how control factors which encompass variables such as degree of control, immediacy of control, and mode of control influence agency. Therefore, groups of learners would generally have lower agency in an XRCL where interaction is limited and where learners follow a fixed narrative compared to a highly interactive XRCL where groups are able to control many parameters and have the freedom to explore (Johnson-Glenberg, 2019). The more open and interactive an XRCL environment is, the more learners need to coordinate their activities which naturally leads to higher social interaction. Immersive learning research highlights how agency can influence cognitive and affective factors such as situational interest, intrinsic motivation, and embodied learning (Petersen et al., 2022) which are factors that can influence social interaction through higher engagement.

Although more freedom to explore can result in a higher sense of agency and social interaction, there is abundant evidence that unassisted discovery does not benefit novice learners, whereas different forms of scaffolding do (Alfieri et al., 2011). Furthermore, research has shown that inquiry-based learning can be more effective than other more expository instructional approaches as long as students are supported adequately (Lazonder & Harmsen, 2016). When students engage in

inquiry-based learning, their actions are often as important as their dialogue. Therefore, although increased freedom can result in more agency, it is important that pedagogical techniques are in place to ensure that students' exploration is appropriately scaffolded. This can be complicated and requires a certain balance, as a pedagogy that promotes efficient learning can cause learners to be less likely to discover novel information (Bonawitz et al., 2011).

The crucial point in terms of building XRCL environments that are effective learning tools according to TICOL thus lies in developing evidence of design factors that can increase the four fundamental psychological factors in a way that increases productive social interaction, can improve social space, and ultimately increase learning outcomes, that is, designing XRCL environments that optimize the benefits of XRCL while taking into account their limitations. There is abundant literature describing how VR environments can lead to hedonic participation where students use the learning experience for its entertainment qualities (Makransky, Terkildsen, & Mayer, 2019). This can lead to lower learning outcomes as the students' actions are not channelled toward the learning content. As an example, it may be tempting to visit Mars and play via an XRCL system rather than focusing on the on-task activity. However, this would be detrimental to learning outcomes.

The Association Between Social Interaction and Social Space (Path 9)

TICOL builds on CSCL literature and describes how the different kinds of social interaction are crucial for developing a positive social space which creates a feedback loop to social interaction and ultimately impacts learning outcomes (Kreijns et al., 2013). Many researchers highlight how a sound social space including positive interpersonal relationships and a sense of community can facilitate social interaction and vice versa, as high-quality collaboration can reinforce a sound social space (Haythornthwaite, 2002; Johnson & Johnson, 2009; Palloff & Pratt, 2005). Social interaction is thus a precursor of and a necessary element for the development of social space (Kreijns et al., 2022).

Essentially, providing group members with the possibility to embody an avatar of choice, with the possibility to interact with a high level of social presence in a relevant physical environment that feels real, while supporting agency, can stimulate social interaction. This can contribute toward a sound social space consisting of strong interpersonal relationships, trust, and a sense of cohesion, which reinforces social interaction and thereby learning according to the social aspects of CSCL framework (Kreijns et al., 2013).

The Influence of Social Interaction on Learning Outcomes (Path 10)

Successful collaborative learning depends on the extent to which groups engage in productive social interactions (Kreijns et al., 2013). Abundant CSCL research has investigated the conditions where specific interactions occur, as well as the interactions that are predictive of learning (Dillenbourg et al., 1996). Social interaction such as explanation, argumentation/negotiation, and mutual regulation is necessary

for group members to learn from each other in a CSCL environment, but this may not be sufficient: “Only when the group development results in a social space where trust, sense of community, and strong interpersonal relationships exist can CSCL pedagogy be successfully applied” (Kreijns et al., 2013, p. 229). The purpose of a CSCL environment is thus not simply to enable collaboration across distance, but to create conditions in which effective group interactions can occur (Dillenbourg et al., 2009). Above, we proposed how the four psychological factors that make XRCL unique can influence learning outcomes by increasing social interaction and improving social space. In the next section, we describe a research agenda for testing those assumptions and providing an evidence-based approach for developing the field of XRCL.

Future Research Agenda

Given the fact that XRCL research is in its infancy, we hope that TICOL can provide a theoretical basis for developing this field by motivating researchers to empirically challenge our assumptions and ultimately develop a deeper understanding of if, and how, immersive media influences collaborative learning. Rather than taking a technocentric approach to understanding XRCL, we hope that TICOL will pave the way for an evidence-based research agenda that will influence the future of collaboration, education, and work, in a world where we will increasingly rely on immersive technology to keep us connected and reach shared goals. This could take many forms. In a recent paper, Janssen and Kirschner (2020) proposed a research agenda to guide future CSCL research and contend that it is important to simultaneously study antecedents, processes, and consequences of collaboration. TICOL focuses on the process of learning in XRCL environments and the psychological factors that make XRCL unique. Future research could therefore focus on the following research questions that investigate how each of the fundamental psychological factors combined with the unique possibilities in XRCL influence learning outcomes through social interaction and social space as predicted by TICOL: (1) How does the potentially high level of physical presence in XRCL combined with the ability to instantaneously change the physical environment change the quality and quantity of social interaction, and does this influence social space and ultimately learning outcomes? (2) How does the potentially high level of social presence in XRCL combined with the possibility to change who one is (in terms of physical appearance) influence collaboration partners, and how does this influence social interaction, social space, and learning outcomes? (3) How do the characteristics and the possibility to customize one’s own avatar influence one’s behaviour, and how does this influence how people collaborate and learn? (4) What pedagogical design features can improve the quality of immersive collaborative learning, given that XRCL provides physically and socially unconstrained environments that allow learners to “be” anyone and “do” anything they want?

Future research could also focus on antecedents by systematically varying different design features described in TICOL and investigating how this influences the fundamental psychological factors, social interaction, social space, and ultimately

learning outcomes as predicted by TICOL. For instance, more research is needed to investigate how pedagogical techniques from CSCL literature generalize to XRCL, as well as research that investigates novel pedagogical techniques in XRCL. Other factors that are not currently included in TICOL can also influence learning outcomes in XRCL. Similar to how many people experienced a difficult transition with regard to using online meetings as a primary means of communication at the beginning of the COVID-19 pandemic due to technical issues (Raake et al., 2022), cognitive load (Bailenson, 2021), or usability issues (Fauville et al., 2021), the potentials of XRCL may be thwarted due to these factors as well as side effects such as simulation sickness (Biocca, 1992). Furthermore, a sound social space is ultimately the consequence of group members. Therefore, antecedents such as individual difference factors could influence how the psychological mediators in TICOL influence social interaction and through it social space and learning outcomes. In their meta-analysis, Jeong et al. (2019) found that the effects of a particular technology or pedagogy are likely to vary as a function of the collaboration mode, learner levels, and domains of learning. This suggests that the interplay between how the fundamental psychological factors of XRCL influence social interaction, social space, and learning outcomes will likely depend on a number of different factors (i.e., boundary conditions). Finally, the consequences of collaboration in XRCL need to be investigated by how the predictions in TICOL differ based on different types of learning outcomes as previous research has found that immersive learning is specifically effective for certain types of learning outcomes (Makransky & Petersen, 2021).

While it is possible to test the proposed hypotheses and assumptions of TICOL through theoretically driven research as proposed above, a bottom-up data-driven approach is also possible. XRCL makes it possible to measure central aspects of social interaction in a controlled environment that feels real (Han et al., 2022). That is, XRCL provides rich data about a group of learners including what they say, interpersonal distance, movement trajectory, gaze, gestures, and facial expressions. This represents a whole new level of social learning analytics (Kaliisa et al., 2022), which could be used to gain a deeper understanding of the fundamental assumptions and hypotheses in TICOL. To this end, data-driven machine learning approach including explainable AI methods such as SHAP (Lundberg & Lee, 2017) and permutation importance (Altmann et al., 2010) can be used to understand which are the most important predictors of learning (Deininger et al., 2023). While the amount of data that XRCL produces is extensive (Miller et al., 2023), it is important that future research is grounded in strong theory to guide how behavioural data is operationalized and used. Finally, ethical issues (e.g., Brown et al., 2023), data security (Chen et al., 2022), safety (Zallio & Clarkson, 2022), and the use of XRCL in diverse populations (e.g., Wang et al., 2023) are also important future XRCL research areas.

Conclusion

In this manuscript, we present the theory of immersive collaborative learning (TICOL). The theory draws on research from computer-supported collaborative learning (CSCL) including Kreijns et al.' (2013) social aspects of CSCL research

framework and theories and principles of individual immersive learning, including the immersion principle in multimedia learning (Makransky & Mayer, 2022) and the cognitive-affective model of immersive learning (CAMIL; Makransky & Petersen, 2021), to propose a model of extended reality-supported collaborative learning (XRCL). TICOL describes four fundamental psychological factors that make collaboration through immersive media fundamentally different from collaboration through traditional media such as laptops or tablets: social presence, physical presence, body ownership, and agency. The model describes how design factors including technological features, social affordances, and pedagogical techniques of an XRCL system influence the four fundamental psychological factors. Furthermore, the theory describes how these fundamental psychological factors influence social interaction, which can occur in the cognitive or socio-emotional dimension as well as in task or non-task contexts. Finally, the model describes how the different dimensions of social interaction can influence social space and ultimately learning outcomes. According to TICOL, the crucial point with regard to building XRCL environments that are effective learning tools lies in understanding the complex interplay between design factors, fundamental psychological factors, social interaction, social spaces, and learning outcomes.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adesope, O. O., Nesbit, J. C., & Sundararajan, N. (2022). The mapping principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 351–359). Cambridge University Press.
- Akselrad, D., DeVeaux, C., Han, E., Miller, M. R., & Bailenson, J. (2023). *Body crumple, sound intrusion, and embodiment violation: Toward a framework for miscommunication in VR*. Computer Supported Cooperative Work and Social Computing (CSCW '23 Companion) <https://stanfordvr.com/pubs/2023/body-crumple-in-vr/>
- Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, 103(1), 1–18. <https://doi.org/10.1037/a0021017>

- Almeida, F., Santos, J. D., & Monteiro, J. A. (2020). The challenges and opportunities in the digitalization of companies in a post-COVID-19 world. *IEEE Engineering Management Review*, 48(3), 97–103. <https://doi.org/10.1109/EMR.2020.3013206>
- Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics*, 26(10), 1340–1347.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy*. Longman.
- Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. Norton & Company.
- Bailenson, J. N. (2021). Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior*, 2(1). <https://doi.org/10.1037/tmb0000030>
- Banakou, D., Kishore, S., & Slater, M. (2018). Virtually being Einstein results in an improvement in cognitive task performance and a decrease in age bias. *Front. Psychol.*, 9, 917.
- Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, 1(2), 164–180.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 577–660.
- Bayro, A., Ghasemi, Y., & Jeong, H. (2022). Subjective and objective analyses of collaboration and co-presence in a virtual reality remote environment. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, VRW 2022* (pp. 485–487). <https://doi.org/10.1109/VRW55335.2022.00108>
- Bem, D. J. (1972). Self-perception theory. In *Advances in Experimental Social Psychology* (pp. 1–62). Academic Press.
- Bioocca, F. (1992). Will simulation sickness slow down the diffusion of virtual environment technology? *Presence: Teleoperators and Virtual Environments*, 1(3), 334–343. <https://doi.org/10.1162/pres.1992.1.3.334>
- Bioocca, F., & Harms, C. (2002). *Defining and measuring social presence: Contribution to the networked minds theory and measure*. International Workshop on Presence.
- Blair, I. V. (2001). Implicit stereotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 359–374). Psychology Press.
- Blanchard, A. L., & Lynne Markus, M. (2004). The experienced “sense” of a virtual community: Characteristics and processes. *Database for Advances in Information Systems*, 35(1), 64–79. <https://doi.org/10.1145/968464.968470>
- Blascovich, J. (2002). A theoretical model of social influence for increasing the utility of collaborative virtual environments. In *Proceedings of the 4th International Conference on Collaborative Virtual Environments* (pp. 25–30). <https://doi.org/10.1145/571878.571883>
- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, 120(3), 322–330. <https://doi.org/10.1016/j.cognition.2010.10.001>
- Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, 40(7), 36–43. <https://doi.org/10.1109/MC.2007.257>
- Brand, C. O., Mesoudi, A., & Morgan, T. (2021). *Trusting the experts: The domain-specificity of prestige-biased social learning*. <https://doi.org/10.31234/osf.io/28t5d>
- Brook, C., Oliver, R. (2002). Supporting the development of learning communities in online settings. In P. Barker, S. Rebelsky, Proceedings of the ED-MEDIA 2002 – World conference on educational multimedia, hypermedia & telecommunications (pp. 192–197).
- Brown, J., Bailenson, J., & Hancock, J. (2023). Misinformation in virtual reality. *Journal of Online Trust and Safety*, 1(5).
- Bucher, T., & Helmond, A. (2018). *The affordances of social media platforms* (pp. 233–254). The SAGE Handbook of Social Media.
- Bygstad, B., Øvrelid, E., Ludvigsen, S., & Dæhlen, M. (2022). From dual digitalization to digital learning space: Exploring the digital transformation of higher education. *Computers and Education*, 182. <https://doi.org/10.1016/j.compedu.2022.104463>
- Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, 3(1), 14–25. <https://doi.org/10.1002/bs.3830030103>

- Chen, J., Wang, M., Kirschner, P. A., & Tsai, C. C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research*, 88(6), 799–843. <https://doi.org/10.3102/0034654318791584>
- Chen, Z., Wu, J., Gan, W., & Qi, Z. (2022). Metaverse security and privacy: An overview. In *2022 IEEE International Conference on Big Data (Big Data)* (pp. 2950–2959). IEEE.
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42(2), 21–29. <https://doi.org/10.1007/BF02299088>
- Coban, M., Bolat, Y. I., & Goksu, I. (2022). The potential of immersive virtual reality to enhance learning: A meta-analysis. *Educational Research Review*, 36. <https://doi.org/10.1016/j.edurev.2022.100452>
- Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 19(2), 272–309. <https://doi.org/10.1080/15213269.2015.1015740>
- Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, 41(1), 10–32. <https://doi.org/10.1111/j.1467-8535.2009.01038.x>
- David, N., Cohen, M. X., Newen, A., Bewernick, B. H., Shah, N. J., Fink, G. R., et al. (2007). The extrastriate cortex distinguishes between the consequences of one's own and others' behavior. *NeuroImage*, 36(3), 1004–1014.
- David, N., Newen, A., & Vogeley, K. (2008). The “sense of agency” and its underlying cognitive and neural mechanisms. *Consciousness and Cognition*, 17(2), 523–534. <https://doi.org/10.1016/j.concog.2008.03.004>
- Deiningner, H., Lavelle-Hill, R., Parrisius, C., Pieronczyk, I., Colling, L., Meurers, D., et al. (2023). Can you solve this on the first try?—Understanding exercise field performance in an intelligent tutoring system. In *International Conference on Artificial Intelligence in Education* (pp. 565–576). Cham: Springer Nature Switzerland.
- Denson, N. (2009). Do curricular and cocurricular diversity activities influence racial bias? A meta-analysis. *Review of Educational Research*, 79(2), 805–838. <https://doi.org/10.3102/0034654309331551>
- DeVeaux, C., Markowitz, D. M., Han, E., Miller, M. R., Hancock, J. T., & Bailenson, J. N. (2023). Descriptive linguistic patterns of group conversations in VR. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* <https://stanfordvr.com/pubs/>
- Di Natale, A. F., Repetto, C., Riva, G., & Villani, D. (2020). Immersive virtual reality in K-12 and higher education: A 10-year systematic review of empirical research. *British Journal of Educational Technology*, 51(6), 2006–2033. <https://doi.org/10.1111/bjet.13030>
- Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning science* (pp. 189–211). Elsevier.
- Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp. 1–19). Elsevier.
- Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning: From design to orchestration. In N. Balacheff, S. Ludvigsen, T. Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-enhanced learning: Principles and products (Vol. 23)* (pp. 3–19). Springer.
- Dincelli, E., & Yayla, A. (2022). Immersive virtual reality in the age of the metaverse: A hybrid-narrative review based on the technology affordance perspective. *Journal of Strategic Information Systems*, 31(2). <https://doi.org/10.1016/j.jsis.2022.101717>
- Dubosc, C., Gorisse, G., Christmann, O., Fleury, S., Poinsot, K., & Richir, S. (2021). Impact of avatar anthropomorphism and task type on social presence in immersive collaborative virtual environments. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, VRW 2021* (pp. 438–439). <https://doi.org/10.1109/VRW52623.2021.00101>
- Erkens, G., Prangsa, M., & Jaspers, J. (2006). Planning and coordinating activities in collaborative learning. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative learning, reasoning, and technology* (pp. 233–264). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203826843-15>
- European Commission. (2020). Digital education action plan (2021–2027). <https://education.ec.europa.eu/focus-topics/digital-education/action-plan>
- Farrer, C., Bouchereau, M., Jeannerod, M., & Franck, N. (2008). Effect of distorted visual feedback on the sense of agency. *Behavioural Neurology*, 19(1–2), 53–57. <https://doi.org/10.1155/2008/425267>

- Fauville, G., Luo, M., Queiroz, A. C. M., Bailenson, J. N., & Hancock, J. (2021). Zoom exhaustion & fatigue scale. *Computers in Human Behavior Reports*, 4. <https://doi.org/10.1016/j.chbr.2021.100119>
- Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, 28(4), 717–741. <https://doi.org/10.1007/s10648-015-9348-9>
- Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, 48(1), 56–66.
- Freiwald, J. P., Schenke, J., Lehmann-Willenbrock, N., & Steinicke, F. (2021). Effects of avatar appearance and locomotion on co-presence in virtual reality collaborations. *Proceedings of Mensch Und Computer, 2021*, 393–401. <https://doi.org/10.1145/3473856.3473870>
- Frith, C. D., Perry, R., & Lumer, E. (1999). The neural correlates of conscious experience: An experimental framework. *Trends in Cognitive Sciences*, 3(3), 105–114.
- Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *Internet and Higher Education*, 10(3), 157–172. <https://doi.org/10.1016/j.iheduc.2007.04.001>
- Garrison, R. (2000). Theoretical challenges for distance education in the 21st century: A shift from structural to transactional issues. *International Review of Research in Open and Distance Learning*, 1(1), 1–17.
- Girod, S., Fassiotto, M., Grewal, D., Ku, M. C., Sriram, N., Nosek, B. A., & Valantine, H. (2016). Reducing implicit gender leadership bias in academic medicine with an educational intervention. *Academic Medicine*, 91(8), 1143–1150. <https://doi.org/10.1097/ACM.0000000000001099>
- Gonzalez-Franco, M., & Peck, T. C. (2018). Avatar embodiment. Towards a standardized questionnaire. *Frontiers in Robotics and AI*, 5. <https://doi.org/10.3389/frobt.2018.00074>
- Gössling, S., Hanna, P., Higham, J., Cohen, S., & Hopkins, D. (2019). Can we fly less? Evaluating the ‘necessity’ of air travel. *Journal of Air Transport Management*, 81, 101722.
- Greenwald, A. G., & Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law Review*, 94(4), 945–967. <https://doi.org/10.2307/20439056>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037//0022-3514.74.6.1464>
- Gully, S. M., Incalcaterra, K. A., Joshi, A., & Beaubien, J. M. (2002). A meta-analysis of team-efficacy, potency, and performance: Interdependence and level of analysis as moderators of observed relationships. *Journal of Applied Psychology*, 87(5), 819–832. <https://doi.org/10.1037/0021-9010.87.5.819>
- Gunawardena, C. N. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications*, 1, 147–166.
- Gürbüz, R., & Birgin, O. (2012). The effect of computer-assisted teaching on remedying misconceptions: The case of the subject “probability”. *Computers and Education*, 58(3), 931–941. <https://doi.org/10.1016/j.compedu.2011.11.005>
- Hallett, K., & Cummings, T. (1997). The virtual classroom as authentic experience: Collaborative, problem-eased learning in a WWW environment. In *13Th Annual Conference on Distance Teaching & Learning: Competition Connection Collaboration* (pp. 103–107).
- Han, E., Miller, M. R., DeVeaux, C., Jun, H., Nowak, K. L., Hancock, J. T., Ram, N., & Bailenson, J. N. (2023). People, places, and time: A large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. *Journal of Computer-Mediated Communication*, 28(2). <https://doi.org/10.1093/jcmc/zmac031>
- Han, E., Miller, M. R., Ram, N., Nowak, K. L., & Bailenson, J. N. (2022). Understanding group behavior in virtual reality: A large-scale, longitudinal study in the metaverse. In *72nd Annual International Communication Association Conference* <https://ssrn.com/abstract=4110154>
- Haythornthwaite, C. (2002). Building social networks via computer networks: Creating and sustaining distributed learning communities. In K. Renninger & W. Shumar (Eds.), *Building virtual communities: Learning and change in cyberspace* (pp. 159–190). Cambridge University Press.
- Herrera, F., Oh, S. Y., & Bailenson, J. N. (2018). Effect of behavioral realism on social interactions inside collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, 27(2), 163–182. https://doi.org/10.1162/PRES_a_00324
- Hughes, C., & Hewson, L. (1998). Online interactions: Developing a neglected aspect of the virtual classroom. *Educational Technology*, 38(4), 48–55 <https://www.jstor.org/stable/44429001>

- Hwang, A., & Arbaugh, J. B. (2006). Virtual and traditional feedback-seeking behaviors: Underlying competitive attitudes and consequent grade performance. *Decision Sciences Journal of Innovative Education*, 4(1), 1–28. <https://doi.org/10.1111/j.1540-4609.2006.00099.x>
- Ijsselstein, W., & Riva, G. (2003). Being there: The experience of presence in mediated environments. In G. Riva, F. Davide, & W. Ijsselstein (Eds.), *Being there: Concepts, effects and measurement of user presence in synthetic environments* (pp. 1–14). Ios Press.
- Isöhätälä, J., Näykki, P., & Järvelä, S. (2020). Cognitive and socio-emotional interaction in collaborative learning: Exploring fluctuations in students' participation. *Scandinavian Journal of Educational Research*, 64(6), 831–851. <https://doi.org/10.1080/00313831.2019.1623310>
- Janssen, J., & Kirschner, P. A. (2020). Applying collaborative cognitive load theory to computer-supported collaborative learning: Towards a research agenda. *Educational Technology Research and Development*, 68(2), 783–805.
- Janssen, J., Kirschner, F., & Kirschner, P. A. (2022). The collaboration principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 304–312). Cambridge University Press.
- Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior*, 23, 1105–1125. <https://doi.org/10.1016/j.chb.2006.10.005>
- Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, 23, 1515–1529. <https://doi.org/10.1007/s10639-017-9676-0>
- 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. <https://doi.org/10.1080/00461520.2016.1158654>
- Jeong, H., Hmelo-Silver, C. E., & Jo, K. (2019). Ten years of computer-supported collaborative learning: A meta-analysis of CSCL in STEM education during 2005–2014. *Educational Research Review*, 28. <https://doi.org/10.1016/j.edurev.2019.100284>
- Jeunet, C., Albert, L., Argelaguet, F., & Lécuyer, A. (2018). “Do you feel in control?": Towards novel approaches to characterise, manipulate and measure the sense of agency in virtual environments. *IEEE transactions on visualization and computer graphics*, 24(4), 1486–1495.
- Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, 38(5), 365–379. <https://doi.org/10.3102/0013189X09339057>
- Johnson-Glenberg, M. C. (2019). The necessary nine: Design principles for embodied VR and active stem education. In P. Díaz, A. Ioannou, K. K. Bhagat, & J. M. Spector (Eds.), *Learning in a Digital World* (pp. 83–112). Springer.
- Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, 106(1), 86–104. <https://doi.org/10.1037/a0034008>
- Kaliisa, R., Rienties, B., Mörch, A. I., & Kluge, A. (2022). Social learning analytics in computer-supported collaborative learning environments: A systematic review of empirical studies. *Computers and Education Open*, 3, 100073.
- Keenan, A., & Shiri, A. (2009). Sociability and social interaction on social networking websites. *Library Review*, 58(6), 438–450.
- Khojasteh, N., & Won, A. S. (2021). Working together on diverse tasks: A longitudinal study on individual workload, presence and emotional recognition in collaborative virtual environments. *Frontiers in Virtual Reality*, 2, 643331.
- Kihlstrom, J. F., Barnhardt, T. M., & Tataryn, D. J. (1992). Implicit perception. In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness: Cognitive, clinical, and social perspectives* (pp. 17–53). Guilford Press.
- Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4), 373–387. https://doi.org/10.1162/PRES_a_00124
- Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review*, 21(1), 31–42. <https://doi.org/10.1007/s10648-008-9095-2>
- Kirschner, P. A., & Erkens, G. (2013). Toward a framework for CSCL research. *Educational Psychologist*, 48(1), 1–8. <https://doi.org/10.1080/00461520.2012.750227>

- Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, J. R. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning*, 13, 213–233. <https://doi.org/10.1007/s11412-018-9277-y>
- Klingenberg, S., Fischer, R., Zettler, I., & Makransky, G. (2022). Facilitating learning in immersive virtual reality: Segmentation, summarizing, both or none? *Journal of Computer Assisted Learning*. <https://doi.org/10.1111/jcal.12741>
- Klingenberg, S., Jørgensen, M. L. M., Dandanell, G., Skriver, K., Mottelson, A., & Makransky, G. (2020). Investigating the effect of teaching as a generative learning strategy when learning through desktop and immersive VR: A media and methods experiment. *British Journal of Educational Technology*, 51(6), 2115–2138. <https://doi.org/10.1111/bjjet.13029>
- Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior*, 19, 335–353. [https://doi.org/10.1016/S0747-5632\(02\)00057-2](https://doi.org/10.1016/S0747-5632(02)00057-2)
- Kreijns, K., Kirschner, P. A., & Vermeulen, M. (2013). Social aspects of CSCLE environments: A research framework. *Educational Psychologist*, 48(4), 229–242. <https://doi.org/10.1080/00461520.2012.750225>
- Kreijns, K., Xu, K., & Weidlich, J. (2022). Social presence: Conceptualization and measurement. *Educational Psychology Review*, 34(1), 139–170.
- Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, 86(3), 681–718. <https://doi.org/10.3102/0034654315627366>
- LeBaron, C., & Koschmann, T. (2003). Gesture and the transparency of understanding. *Studies in language and social interaction*, 102–112.
- Lee, K. M. (2004). Presence, explicated. *Communication theory*, 14(1), 27–50.
- Liu, H. C., & Su, I. H. (2011). Learning residential electrical wiring through computer simulation: The impact of computer-based learning environments on student achievement and cognitive load. *British Journal of Educational Technology*, 42(4), 598–607. <https://doi.org/10.1111/j.1467-8535.2009.01047.x>
- Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, and Computers*, 31(4), 557–564. <https://doi.org/10.3758/BF03200735>
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Advances in neural information processing systems* (Vol. 30).
- Makransky, G. (2022). The immersion principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 296–303). Cambridge University Press.
- Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology*, 113(4), 719–735. <https://doi.org/10.1037/edu0000473>
- Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning*, 35(6), 691–707. <https://doi.org/10.1111/jcal.12375>
- Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior*, 72, 276–285. <https://doi.org/10.1016/j.chb.2017.02.066>
- Makransky, G., & Mayer, R. E. (2022). Benefits of taking a virtual field trip in immersive virtual reality: Evidence for the immersion principle in multimedia learning. *Educational Psychology Review*. <https://doi.org/10.1007/s10648-022-09675-4>
- Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*. <https://doi.org/10.1007/s10648-020-09586-2>
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225–236. <https://doi.org/10.1016/j.learninstruc.2017.12.007>
- Makransky, G., Wismer, P., & Mayer, R. E. (2019). A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *Journal of Computer Assisted Learning*, 35(3), 349–358. <https://doi.org/10.1111/jcal.12335>

- Matusov, E. (1996). Intersubjectivity without agreement. *Mind, Culture, and Activity*, 3(1), 25–45. https://doi.org/10.1207/s15327884mca0301_4
- Mayer, R. E. (2014). The cambridge handbook of multimedia learning, second edition. In *The Cambridge handbook of multimedia learning*, second edition (Vol. 1, Issue 3). <https://doi.org/10.1017/CBO97811139547369>
- Mayer, R. E., & Fiorella, L. (2022). *The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge University Press.
- Miller, M. R., DeVeaux, C., Han, E., Ram, N., & Bailenson, J. N. (2023). A large-scale study of proxemics and gaze in groups. In *IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. <https://doi.org/10.1109/VR55154.2023.00056>
- Mottelson, A., Muresan, A., Hornbæk, K., & Makransky, G. (2023). A systematic review and meta-analysis of the effectiveness of body ownership illusions in virtual reality. *ACM Transactions on Computer-Human Interaction*.
- Mystakidis, S. (2022). Metaverse. *Encyclopedia*, 2(1), 486–497. <https://doi.org/10.3390/encyclopedia2010031>
- Oh, C. S., Bailenson, J. N., & Welch, G. F. (2018). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI*, 5(114). <https://doi.org/10.3389/frobt.2018.00114>
- Palloff, R., & Pratt, K. (2005). *Collaborating online: Learning together in community*. Jossey Bass.
- Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. *ArXiv*. <https://doi.org/10.48550/arXiv.2304.03442>
- Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6), 785–797. <https://doi.org/10.1037/edu0000241>
- Petersen, G. B., Petkakis, G., & Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers & Education*, 179, 1–16. <https://doi.org/10.1016/j.compedu.2021.104429>
- Petersen, G. B., Stenberdt, V., Mayer, R. E., & Makransky, G. (2023). Collaborative generative learning activities in immersive virtual reality increase learning. *Computers & Education*.
- Plechátá, A., Makransky, G., & Böhm, R. (2022). Can extended reality in the metaverse revolutionise health communication? *NPJ Digital Medicine*. <https://doi.org/10.1038/s41746-022-00682-x>
- Plechátá, A., Morton, T., Perez-Cueto, F. J. A., & Makransky, G. (2022). Why just experience the future when you can change it: Virtual reality can increase pro-environmental food choices through self-efficacy. *Technology, Mind, and Behavior*, 3(4). <https://doi.org/10.1037/tmb0000080>
- Poth, R. D. (2018). Social presence in online learning. In M. Marmon (Ed.), *Enhancing social presence in online learning environments* (pp. 88–116). IGI Global. <https://doi.org/10.4018/978-1-5225-3229-3.ch005>
- Prawat, R. S. (1989). Promoting access to knowledge, strategy, and disposition in students. *A research synthesis*, 59(1), 1–41.
- Raake, A., Fiedler, M., Schoenberg, K., De Moor, K., & Döring, N. (2022). Technological factors influencing videoconferencing and zoom fatigue. *ArXiv Preprint ArXiv*. <http://arxiv.org/abs/2202.01740>
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147. <https://doi.org/10.1016/j.compedu.2019.103778>
- Ratan, R., Beyea, D., Li, B. J., & Graciano, L. (2020). Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: A meta-analysis of the proteus effect. *Media Psychology*, 23(5), 651–675. <https://doi.org/10.1080/15213269.2019.1623698>
- Resta, P., & Laferrrière, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review*, 19(1), 65–83. <https://doi.org/10.1007/s10648-007-9042-7>
- Rourke, L., & Anderson, T. (2002). Exploring social communication in asynchronous, text-based computer conferencing. *Journal of Interactive Learning Research*, 13, 259–275.
- Ryan, W. S., Cornick, J., Blascovich, J., & Bailenson, J. N. (2019). Virtual reality: Whence, how and what for. In A. Rizzo & S. Bouchard (Eds.), *Virtual reality for psychological and neurocognitive interventions* (pp. 15–46). Springer. https://doi.org/10.1007/978-1-4939-9482-3_2
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(3), 501–518. <https://doi.org/10.1037/0278-7393.13.3.501>

- Schöne, B., Kisker, J., Lange, L., Gruber, T., Sylvester, S., & Osinsky, R. (2023). The reality of virtual reality. *Frontiers in Psychology, 14*. <https://doi.org/10.3389/fpsyg.2023.1093014>
- Scott, C. L. (2015). *The futures of learning 3: What kind of pedagogies for the 21st century?* (pp. 1–21). Education Research and Foresight.
- Sedlák, M., Šašinka, Č., Stachoň, Z., Chmelfík, J., & Doležal, M. (2022). Collaborative and individual learning of geography in immersive virtual reality: An effectiveness study. *PLoS One, 17*(10). <https://doi.org/10.1371/journal.pone.0276267>
- Shin, D. H. (2017). The role of affordance in the experience of virtual reality learning: Technological and affective affordances in virtual reality. *Telematics and Informatics, 34*(8), 1826–1836.
- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3549–3557. <https://doi.org/10.1098/rstb.2009.0138>
- Slater, M. (2017). Implicit learning through embodiment in immersive virtual reality. *Virtual, Augmented, and Mixed Realities in Education*, 19–33.
- Slater, M., Banakou, D., Beacco, A., Gallego, J., Macia-Varela, F., & Oliva, R. (2022). A separate reality: An update on place illusion and plausibility in virtual reality. *Frontiers in Virtual Reality, 3*. <https://doi.org/10.3389/frvir.2022.914392>
- Slater, M., Spanlang, B., Sanchez-Vives, M. V., & Blanke, O. (2010). First person experience of body transfer in virtual reality. *PLoS One, 5*(5). <https://doi.org/10.1371/journal.pone.0010564>
- Song, M., & Yuan, R. (2015). Beyond social presence: Increasing cognitive presence through meaningful interaction. In *Proceedings of Global Learn Berlin 2015: Global Conference on Learning and Technology* (pp. 731–736).
- Sperduti, M., Delaveau, P., Fossati, P., & Nadel, J. (2011). Different brain structures related to self-and external-agency attribution: A brief review and meta-analysis. *Brain Structure and Function, 216*, 151–157.
- Stahl, G., Koschmann, T. D., & Suthers, D. D. (2006). CSCL: A historical perspective.
- Strijbos, J. W., & De Laat, M. F. (2010). Developing the role concept for computer-supported collaborative learning: An explorative synthesis. *Computers in Human Behavior, 26*(4), 495–505.
- Stajkovic, A. D., & Lee, D. S. (2001). *A meta-analysis of the relationship between collective efficacy and group performance*. Academy of Management.
- Strijbos, J. W., Martens, R. L., Jochems, W. M. G., & Broers, N. J. (2004). The effect of functional roles on group efficiency: Using multilevel modeling and content analysis to investigate computer-supported collaboration in small groups. *Small Group Research, 35*, 195–229. <https://doi.org/10.1177/1046496403260843>
- Sun, Y., Albeaino, G., Gheisari, M., & Eiris, R. (2022). Online site visits using virtual collaborative spaces: A plan-reading activity on a digital building site. *Advanced Engineering Informatics, 53*. <https://doi.org/10.1016/j.aei.2022.101667>
- Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning, 1*(3), 315–337. <https://doi.org/10.1007/s11412-006-9660-y>
- Tajfel, H., & Turner, J. C. (2004). The social identity theory of intergroup behavior. In *Political psychology* (pp. 276–293). Psychology Press.
- Terkildsen, T. S., & Makransky, G. (2019). Measuring presence in video games: An investigation of the potential use of physiological measures as indicators of presence. *International Journal of Human Computer Studies, 126*, 64–80. <https://doi.org/10.1016/j.ijhcs.2019.02.006>
- Tu, C.-H. (2000). On-line learning migration: From social learning theory to social presence theory in a CMC environment. *Journal of Network and Computer Applications, 23*(1), 27–37.
- Tu, C. H., & McIsaac, M. (2002). An examination of social presence to increase interaction in online classes. *The American Journal of Distance Education, 16*(2), 131–150.
- van Brakel, V., Barrera-Ángeles, M., & Hartmann, T. (2023). Feelings of presence and perceived social support in social virtual reality platforms. *Computers in Human Behavior, 139*. <https://doi.org/10.1016/j.chb.2022.107523>
- van Gog, T. (2022). The signaling (or cueing) principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (pp. 221–230). Cambridge University Press.
- van Koningsbruggen, G. M., Hartmann, T., Eden, A., & Veling, H. (2017). Spontaneous hedonic reactions to social media cues. *Cyberpsychology, Behavior, and Social Networking, 20*(5), 334–340. <https://doi.org/10.1089/cyber.2016.0530>

- Vogel, F., Wecker, C., Kollar, I., & Fischer, F. (2017). Socio-cognitive scaffolding with computer-supported collaboration scripts: A meta-analysis. *Educational Psychology Review*, 29, 477–511. <https://doi.org/10.1007/s10648-016-9361-7>
- Wang, X., Young, G. W., Plechatá, A., Mc Guckin, C., & Makransky, G. (2023). Utilizing virtual reality to assist social competence education and social support for children from under-represented backgrounds. *Computers & Education*, 201, 104815.
- Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers and Education*, 46(1), 71–95. <https://doi.org/10.1016/j.compedu.2005.04.003>
- Williams, E. A., Duray, R., & Reddy, V. (2006). Teamwork orientation, group cohesiveness, and student learning: A study of the use of teams in online distance education. *Journal of Management Education*, 30(4), 592–616. <https://doi.org/10.1177/1052562905276740>
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9(4), 625–636. <https://doi.org/10.3758/BF03196322>
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3), 225–240. <https://doi.org/10.1162/105474698565686>
- Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology*, 51(6), 1991–2005. <https://doi.org/10.1111/bjet.13023>
- Xiao, Y., & Lucking, R. (2008). The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment. *Internet and Higher Education*, 11, 186–193. <https://doi.org/10.1016/j.iheduc.2008.06.005>
- Yee, N., & Bailenson, J. (2007). The proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, 33(3), 271–290. <https://doi.org/10.1111/j.1468-2958.2007.00299.x>
- Yoon, B., Kim, H. I., Lee, G. A., Billinqhurst, M., & Woo, W. (2019). The effect of avatar appearance on social presence in an augmented reality remote collaboration. In *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings* (pp. 547–556). <https://doi.org/10.1109/VR.2019.8797719>
- Yuan, Y., & Steed, A. (2010). Is the rubber hand illusion induced by immersive virtual reality? *IEEE Virtual Reality Conference (VR)*. <https://doi.org/10.1109/VR.2010.5444807>
- Zallio, M., & Clarkson, P. J. (2022). Designing the metaverse: A study on inclusion, diversity, equity, accessibility and safety for digital immersive environments. *Telematics and Informatics*, 75, 101909.
- Zhao, H., Sullivan, K. P. H., & Mellenius, I. (2014). Participation, interaction and social presence: An exploratory study of collaboration in online peer review groups. *British Journal of Educational Technology*, 45(5), 807–819. <https://doi.org/10.1111/bjet.12094>
- Zheng, L., Xie, T., & Liu, G. (2018). Affordances of virtual reality for collaborative learning. In *2018 International Joint Conference on Information, Media and Engineering (ICIME)* (pp. 6–10). <https://doi.org/10.1109/ICIME.2018.00011>
- Zurita, G., & Nussbaum, M. (2004). Computer supported collaborative learning using wirelessly interconnected handheld computers. *Computers & Education*, 42(3), 289–314. <https://doi.org/10.1016/j.compedu.2003.08.005>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.