



**HAL**  
open science

# Is It Real? – Learners’ Perceptions on Tele-Immersive 3D Video Technology and Its Further Use in K-12 Education

Erkki Rötönen, Calkin Suero Montero, Nicolas Pope, Erkki Sutinen

► **To cite this version:**

Erkki Rötönen, Calkin Suero Montero, Nicolas Pope, Erkki Sutinen. Is It Real? – Learners’ Perceptions on Tele-Immersive 3D Video Technology and Its Further Use in K-12 Education. Open Conference on Computers in Education (OCCE), Aug 2021, Tampere, Finland. pp.234-246, 10.1007/978-3-030-97986-7\_20 . hal-04161436

**HAL Id: hal-04161436**

**<https://inria.hal.science/hal-04161436v1>**

Submitted on 1 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

# Is It Real? – Learners’ Perceptions on Tele-immersive 3D Video Technology and its Further Use in K-12 Education

Erkki Rötönen<sup>1</sup>[0000-0003-0951-3390], Calkin Suero Montero<sup>2</sup>[0000-0001-5686-7285],  
Nicolas Pope<sup>1</sup>[0000-0002-5694-4970], and Erkki Sutinen<sup>1</sup>[0000-0002-1020-3325]

<sup>1</sup> Department of Computing, University of Turku, 20014 Turku, Finland

<sup>2</sup> School of Educational Sciences and Psychology, University of Eastern Finland,  
80101 Joensuu, Finland

erkki.rotkonen@utu.fi, calkins@uef.fi, nicolas.pope@utu.fi,  
derkki.sutinen@utu.fi

**Abstract.** Current immersive technology applications, such as virtual and augmented reality (VR/AR), are still limited in providing an immersive experience in which a user is perceived in 3 dimensions (3D) as their real self. Furthermore, research on such immersive environments applied to educational contexts is scarce. To investigate the unexplored opportunities of immersive technologies in education, we carried out a trial of our tele-immersive platform (TIP) with elementary school 6th grade learners (12–13-year-olds) working on an environmental study lesson. During our investigations, we applied qualitative analysis of the learners’ focus group interview, researchers’ field notes, and recordings of test sessions, alongside descriptive statistics of questionnaires. Learners expressed feelings of positive surprise and excitement, and felt the technology enabled their participation throughout the lesson. Learners recounted feelings of ‘being there’ inside the environment, and in the company of the teacher. Learners agreed that their sense of physical proximity and presence inside the immersive environment is similar to that of face-to-face interaction with the teacher. With our investigations, the opportunities of this emerging technology towards boosting online educational activities for learners to develop their 21C skills are highlighted. Our findings indicate that the introduced 3D TIP technology has the critical potential to overcome psychological strains due to physical distance in online education.

**Keywords:** Sense of presence, immersive learning environments, 21C skills

## 1 Introduction

The outbreak of the Covid-19 pandemic since 2020 has rapidly increased the demand for online communication in education. We see the use of online platforms ranging from asynchronous email, online discussion boards, and blogs, to synchronous video conferencing, text messaging and Voice-over-IP services (in which the learners and instructor are together online at the same time) as well as hybrid formats combining different forms of media (e.g., text, audio, video) and different timescales (e.g., asynchronous, synchronous) within the same meeting [1]. Technologies such as virtual (VR), augmented (AR) or extended reality (XR), might be used to create more realistic group encounters for distance learning and online collaborative activities, for instance. In this vein, immersive learning environments expanding traditional VR or telepresence

are emerging as a suitable solution for supporting novel forms of online collaboration and practices [2, 3]. Nevertheless, the fast demand for the development of immersive technologies in educational practices is ridden with limitations that can be addressed at the design stage, through the involvement of students and teachers throughout the design process, as advocated by co-design [4]. Still, there remains insufficient evidence about the effect of immersive 3D video technology on an overall sense of presence or on key factors related to learning. [3].

In previous work, in which primary school children collaborate across national, regional, political and cultural boundaries in the same immersive virtual learning space, Rötönen et al. (2019) [5] reported concrete suggestions that the learners provided for improvements, including the use of real faces rather than drawn avatars, and their wish to meet, get to know each other, talk, shake hands, play and be together in the virtual environment as normally as they act in real life.

To further investigate this matter, we have designed and implemented a live 3D tele-immersive platform (TIP) - an environment for K-12 education providing an immersive three-dimensional alternative to video. TIP integrates novel 3D video technologies, within which the teacher can be perceived in 3 dimensions, generating a sense of being with a real-life person, instead of an avatar. In our prototype, rather than using immersive virtual environments, real environments are digitally captured live to form a high-fidelity model of a physical space and its occupants.

Extending the work of Rötönen et al. (2019), our study is, therefore, a preliminary investigation of the experiences and sense of presence that TIP technologies can generate. We are seeking for signs of the potential positive effects of the TIP technology, as well as identifying the main challenges the technology presents in online educational contexts. Hence, this paper investigates further the opportunities and challenges in relation to the *sense of presence, technology* and *education* that is perceived within our TIP prototype during an online learning activity for elementary school children. We believe this kind of technology is very promising for creating and supporting novel forms of collaboration and practices for students and teachers during distance, as well as hybrid, learning activities. Our research questions, therefore, are:

RQ 1) *What are the perceptions of the learners towards the new 3D telepresence technology for education?*

RQ 2) *What impact does the new 3D telepresence technology currently have on the learners' sense of presence?*

## 2 Related Work

Information and communication technology (ICT) tools are now mainly used in schools in traditional ways to communicate, create, disseminate, store, and manage information, for instance, which has been reported to undermine the potential by which technology can reform the education realm [6]. However, as new pedagogical approaches emerge, aided by ICT tools, there is a change in education from the traditional teacher-centred classroom to a collaborative learner-centred classroom where learners can be actively engaged in the educational process, and knowledge, skills and attitudes are constructed

and distributed in a collaborative way. Furthermore, the trend of the 21C towards a globally connected world, facilitated by the access to online technologies, has also had a profound impact on developing online collaborative learning as a new paradigm, connecting learners globally beyond the classroom. When learners are geographically dispersed, online technologies can be used to create connections and communication for sharing ideas and to co-create new understandings with others, making the local connect with the global, and learning happens in interconnected situations, utilising online technologies [7]. The utilisation of ICT is a way for learners to develop 21C skills [8], to prepare them as global citizens. Thus, under the collaborative learning paradigm, a future school needs to be equipped with people, tools and partnerships that allow the school to support its learners and local communities to act in the world, to set problems and solve them, extend their curiosity and to act upon it from a global perspective [9].

In order to support online collaborative learning and connect globally, emerging immersive technologies may provide functional solutions, since they can support the user to perceive each other in a natural way inside an otherwise virtual, or digitally generated, environment. For instance, VR applications can enhance learning and advance social and creative skills, and the use of them in learning situations seems to improve digital-age interactive literacy, creative thinking, communication, collaboration and problem-solving ability, which constitute 21C skills [10]. However, currently available educational technologies, such as e-learning tools, or traditional VR platforms, such as Second Life and Active Worlds, do not offer realistic learning experience environments, as these technologies are not able to replicate the perception of presence, a sense of “being there” with another person as in real life, which is vital for providing close-to-real-life experience and interactions. Social interactions are particularly important in K-12 educational contexts [11], as peer learning and socially constructed knowledge have been demonstrated to be fundamental for the socio-cognitive development of the learners in that age.

However, the concept of *presence* is very broad and has a variety of definitions and meanings. In our study, we look at the *sense of presence* in three dimensions: *presence* can offer insights into the medium’s ability to provide the feeling that the user is “there” inside the media (*telepresence*), as well as in the “company of others” (*copresence*). At the same time, *presence* can measure the extent to which people feel that the interface is able to provide some sense of access to another’s mind and connect people (*social presence*) [12]. Presence has been recognised as a key performance goal for many systems e.g., social presence as direct contributor to the success of the educational experience [13]. However, it is still a complex task to evaluate presence in an objective way [14]. To our knowledge, there is yet no known work on measuring a sense of presence of 3D video technologies with children. Moreover, there is a challenge that the questions of existing data collection mechanisms are difficult for children to answer, due to them being targeted at adults (and in our case, also language translation being required).

In the research area of technology use to support computer mediated collaboration, there is now an increasing, but still limited, amount of work towards synchronous telepresence technologies in education. Asynchronous collaboration tools have been receiving more attention [15]. In terms of applications of live 3D capture technology in education, there has been little if any progress since a few early trials [16, 17]. There is

an abundance of work on immersive virtual learning environments or Collaborative Virtual Environments (CVEs), but not on 3D tele-immersive capture technology [15]. To our knowledge, there are currently three broad approaches being taken in industry and academia to achieve real time hyper-realistic 3D representation of people and spaces: a) Microsoft “Holoportation” or similar depth map fusion technologies [18]; b) Facebook “Codec Avatars” generated using machine learning; and c) dynamic reconstruction using only a single camera [19, 20]. Our own solutions at the University of Turku fit within category (a) currently.

The focus in the literature remains technical in nature, with little pedagogical evaluation and a very narrow view of the future potential. Moreover, there is little research on the pedagogy needed to make full use of the technology in K-12 education, and there is limited understanding of the affordances it may offer [3]. Our 3D TIP technology can be considered as a strong candidate to meliorate the shortcomings found in online technologies when applied to K-12 connective learning and education. In this paper, therefore, we present opportunities and challenges of a TIP prototype, towards its relevance in fostering a sense of presence among the learners that participated in our study.

### **3 3D Tele-immersive Platform Prototype**

In our study we use an in-house developed live 3D tele-immersive video platform (TIP). Our prototype setup consists of 6 stereo pairs of cameras, each pair connected to a desktop computer with a GPU (NVIDIA GeForce RTX 2080) for depth processing. The cameras are arranged around the edges of the room or group to passively capture the scene without intrusion. Once the depth data is captured from each camera, the camera views are merged together using camera pose information gained through a calibration step, to achieve a high-quality fusion with reduced noise distortion and errors in the different individual depth estimates from each camera.

Following fusion, the data are compressed as regular videos, using hardware for real-time performance. After transmission, the multiple video feeds for colour and depth are decoded, compression artefacts are filtered, and then arbitrary 3D views can be rendered locally using a straightforward point-cloud or mesh rendering combined with reprojection to the colour video feeds to add texture. Additionally, virtual objects, including live screen captures placed on walls or tables, can be mixed into the final scene. The overall latency of our prototype is around 200 ms, allowing a two-way experience where each physical space is captured and transmitted to the same web server. A viewing client can then obtain the data of multiple spaces and merge them together during the render step. The rendering is then viewed via a computer 2D display (using a web browser or an Open GL renderer), via an AR/VR headset or via a mobile phone.

### **4 Study Methodology**

Our work is framed within a pragmatic research paradigm and is using research through design as a methodology [21]. This methodology states that research may develop not only through increasing agreement, but also through discursiveness and elaboration.

Thus, rather than making statements about what is, the design aspect of the research is concerned with creating what might be [22]. In order to understand the applicability of 3D immersive video technology in real educational settings, we carried out testing sessions of an in-house prototype system together with learners, teachers and researchers. The prototype testing was designed to allow measuring the sense of presence that the technology fostered in the learners. As such, this study is framed within the following hypotheses:

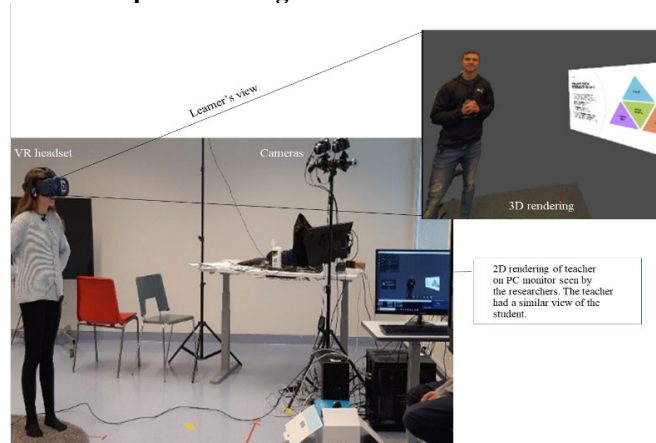
H1: *The perceptions of the learners on the use of 3D immersive video in education will be positive.* (RQ1)

H2: *Immersive 3D video technology will foster a high sense of presence in the learners.* (RQ2)

#### 4.1 Participants

Fourteen learners (6th grade, 12–13-year-olds) at one elementary school in Finland were chosen via random selection and given the opportunity to participate in testing our 3D tele-immersive video technology prototype. The elementary school involved is governed and financed by the municipality. Written permission was obtained from the parents or guardians of all participating learners, and the learners were also informed at the beginning about the nature of the research. Some learners had previously participated in research studies on unrelated VR technologies with our university. A single teacher at the same school was involved in the prototype testing and the learners knew the teacher well.

#### 4.2 Experimental Setup and Testing Sessions



**Fig. 1.** Experimental setup, learner's view, and 3D rendering

During testing, the learners were in the school whilst the teacher was in another physically separate space off site, at our university facilities. The system was set one-way so that the digital view of the captured physical space where the teacher was located was received by the learners using an AR/VR headset (see Fig. 1). The teacher viewed the

learners via a 2D display of a client computer using GeForce 1050+ and OpenGL renderer. This implementation was used in order to test a suitable trade-off between the system affordances, the costs of implementation, and the beneficial outcome of using this type of emerging technology in education from the learner's perspective (the learners are the ones that observe the teacher in 3 dimensions inside the immersive environment). Two researchers and a single video camera were located with the children to record the activity. The digital view of the VR headset viewpoint was recorded also for analysis. The 3D data was also recorded in addition to the video view of the participant. Between each participant, the equipment was disinfected due in part to the ongoing pandemic.

Each session with each learner lasted for approximately 20 minutes of which 10 minutes was the activity using the technology. After entering the testing room, the learner filled in a pre-questionnaire on the computer. The learner was then given brief instructions how to use the immersive 3D video prototype. After this, the learner was set to work on an educational activity together with the teacher within the virtual environment. The educational activity was a typical environmental study lesson related to the different kinds of fire, a subject that they had covered in their previous academic year and the instruction by the teacher was done by exploring the topic with the learner synchronously using a whiteboard. Each learner was assigned a random number which they then used to identify themselves to the observers and in their survey responses. Immediately following the learning activity, the learner answered the presence questionnaire using an online survey tool.

### 4.3 Data Collection

A presence questionnaire, researchers' observation, video analysis, and focus group interview were used to investigate the hypotheses. The questions in the questionnaire were translated into Finnish by an experienced teacher to ensure they were phrased appropriately for the learners in their native Finnish language. The measures used included *the Temple Presence Inventory* [23] *the Networked Minds Questionnaire* [24] and *the MEC Spatial Presence Questionnaire* [25]. The sessions as well as the digital view of the VR headset viewpoint, the 3D data and the video view of each participant were recorded for analysis. Two researchers were located with the learners to record the activity by taking field notes. The observations helped the researchers notice, identify, and understand different types of practical challenges that are faced during the sessions. After all the sessions were conducted a focus group interview was arranged with the learners. The purpose of the focus group interview was to deepen our understanding on the issues that were found from the observations. The interview protocol consisted of inquiring about the learners' feelings and experiences, feedback, and evaluation, as well as suggested improvements.

## 5 Results and Analysis

The recordings and observation notes thereof from the sessions, together with the interview transcript were content analysed. We used an inductive approach to thematic



analysis [26] which resulted in three themes: *sense of presence perceptions*, *technology perceptions* and *educational applications perceptions*. Here, we present these three key themes and discuss opportunities offered and further challenges to address in relation to the use and development of 3D tele-immersive video in educational contexts, also answering the hypothesis set in the study.

### 5.1 Sense of Presence Perceptions

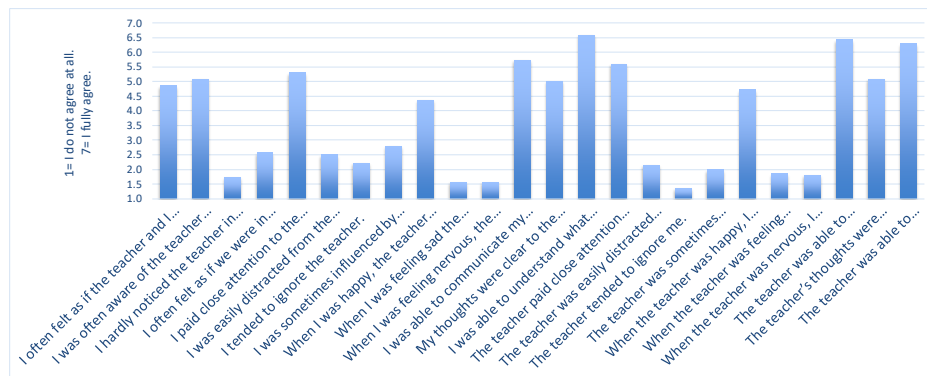
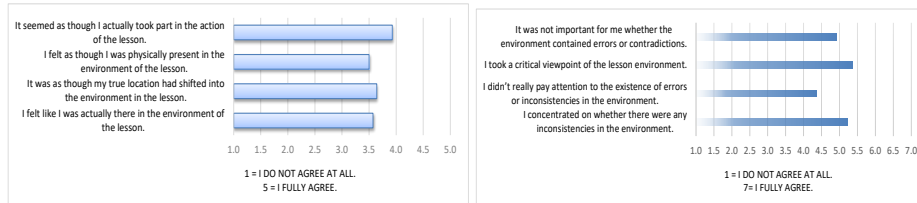


Fig. 2. Social presence questionnaire (descriptive statistics)

In relation to a sense of presence, from the social presence perspective, the learners reported that they highly perceived the teacher being immersed in the environment together with them (telepresence) also being easy to understand the teacher's expressions and ideas as if they were physically close (copresence), (see Fig. 2). For instance, item 2 (Q2), (*I often felt that the teacher was in the same space as me*) had an average response of over 5 on a 7-points Likert scale (1= completely disagree | 7= completely agree), while questions 3 (*I hardly notice the teacher in the space*) and 4 (*I felt as if we were in different places*) had an average score of 2.5 or below. Similarly, items 14, 22 and 24 related to the perception of understanding between the teacher and the learners had a high average response of 6 and above, indicating that it was easy for the students to understand the teacher's expressions and meanings within the virtual environment. This validates our H2 (also providing answers to RQ2).

The learners said in the focus group interview they had a feeling of being in the actual environment themselves and in the same space as the teacher, as indicated by the following quote: "... and it felt as if [I was] in the same place with the teacher. I didn't even notice that I was wearing the glasses." (Girl2). Furthermore, the learners reported that it was not really important for them whether the environment contained errors or contradictions in order to work on the lesson: "*I looked around in a way so how the place looked like, ..., looking.., like, where you are and all... like, What's in here?*" (Boy1 and Boy2). In general, the learners paid attention to the teacher, his voice and body expressions and gestures and they understood what the teacher meant during the lesson, which shows good potential for an educational application.



**Fig. 3.** (Left) Sense of self-location in the immersive environment. (Right) Perceived errors in the immersive environment

## 5.2 Technology Perceptions

We observed from the collected questionnaire that even though the learners enjoyed the immersive aspect of the technology (e.g., the feeling of ‘being there’) there are still several issues to address in terms of their self-location within the virtual environment. That is, for instance Fig. 3, Left, item Q4 “*I felt like I was actually there in the environment of the lesson*” had an average score of 3.6 from the respondents (Likert scale 1= not at all, 5= completely). This indicates that the learners, perhaps, could not perceive themselves physically in the immersive environment as fully as they could perceive their teacher. Furthermore, due to the preliminary nature of the technology the environment tested was very basic, without distracting backgrounds or clues as to where the learners could be located inside of it. Also, Fig. 3, Right, item Q3 “*I didn’t really pay attention to the existence of errors or inconsistencies in the environment*”, had average score of 4.4 (Likert scale 1= completely disagree, 7= completely agree), indicating that in fact the learners did pay attention to glitches that occurred in the TIP environment.

Nevertheless, the focus group discussion revealed that the learners were comfortable enough with the technology, even at this stage, to accept it despite the limitations of self-location and errors that it presented, having strong feelings of ‘being there’, in the company of their teacher. During our testing, even though some of the learners had previous experiences of VR environments this novel immersive technology was totally new and different from anything else they had experienced before: the first time they could see someone else as her/himself in a virtual environment. A sample quote in the following from the focus group interview highlights the issue:

Interviewer: “*How did you feel when you went to the VR environment and you saw the teacher...*”

Girl1: “*Interesting, cool*”

Girl2: “*It was quite odd. I didn’t expect to see him as a whole...*”

Boy1: “*It was kind of odd and interesting and exciting to see...*”

Boy2: “*It was kind of surprising, it was different than I expected.*”

These impressions of positive surprise and excitement also reveal a sense of novelty from the use of the technology, since all the students were experiencing a 3D immersive technology for the first time.

## 5.3 Educational Application Perceptions

The learners expressed in the focus group interview: *“It was quite odd, I didn’t expect to see him as a whole, when you could see the teacher moving and all.”* (Boy2). *“It was kind of odd and interesting and exciting to see all expressions and the movements and all the little things.”* (Girl1) This indicates support for the learners to perceive each other in a more realistic way inside a virtual environment enabling observation of subtle details of objects e.g. for recognising people’s faces and body language and (facial) expressions which can indicate the emotional state or level of comprehension of the others when communicating and working together.

In their answers, the learners also indicated they were able to communicate their intentions clearly to the teacher and that their thoughts were clear to the teacher, thus understanding each other meanings. The learners felt that the teacher paid close attention to them. Some of the learners mentioned they look curiously around the environment but that this did not actually disturb their concentration. They thought that when the novelty factor diminishes it would be even more like a normal learning environment for them. These reported experiences validate our H1, also offers a positive answer to RQ1.

## 6 Discussion

Based on results of preliminary experiments, we pose in Table 1 opportunities offered and further challenges to address use of immersive 3D video in educational contexts.

**Table 1.** Summary of challenges and opportunities offered by immersive 3D video

Perspective	Challenges	Opportunities
Sense of presence	HMD gadgets difficult usage. Lack of eye contact when using the headset. Lack of eyes and face gestures when using the headset.	Feeling of being inside the actual environment. Feeling of being in the same space as the other and in his/her company. Feeling of being able to understand the other’s intentions and feelings.
Technology	Need for technical assistance Laborious setup Cost of equipment HMD technology Environment background Technical glitches	Positive technology acceptance. Technology attractiveness. Support the learners to perceive each other in a more realistic way inside a virtual environment. Recognising and observing people’s faces and other objects in great detail. High-fidelity body language and facial expression. Ability to arbitrarily shift the focus of attention.
Education	Dependence on adults, with pre-established power dynamics, which may result in a power imbalance.	Learning and developing collaborative skills through online interaction. Learning and developing digital skills for learners and teachers. Learning to express and communicate ideas in an online environment.

<b>Perspec- tive</b>	<b>Challenges</b>	<b>Opportunities</b>
	Dependence on specialised technical skills.	Wide array of tools available for online and hybrid collaboration. Non-obstructiveness enabling learners to focus on the learning process. Shared interaction with real objects (e.g., using a real whiteboard, not a digital one). Unrestricted or informal positioning in the physical environment

Regarding a sense of presence, we observed from the collected questionnaire that even though learners enjoyed the immersive aspect of the technology there are still issues to address in terms of their self-location. The learners were comfortable enough with the technology, even at this stage, to accept it despite limitations of self-location that it presented, having strong feelings of ‘being there’, in the company of their teacher.

In terms of the technology implementation, the learners found the idea of being immersed in a 3D environment appealing, in spite of the issues reported with the use of an early prototype technology. Through the technology, the learners could perceive subtle body language and expressions that indicate the emotional state or level of their teacher’s understanding of what the learner say. The technology also could offer the teacher the ability to direct their focus and look for cues on whether they and the task are being understood. A point of consideration, however, is the high costs of implementation of the technology at the present (about 30,000€ in this setup).

With regards to education, the technology could advance 21C skills in terms of collaborative online and hybrid learning. Furthermore, the technology is able to isolate or extract elements from the physical environment that could create distractions during interactive learning activities, or conversely it could augment the virtual environment with virtual elements. For online collaborative learning and connecting globally, beyond the classroom the technology could provide suitable solutions, since it can support the learners to perceive each other in a more realistic way inside a virtual environment enabling observation of subtle details of objects e.g. for recognising people's faces and body language and (facial) expressions which can indicate the emotional state or level of comprehension of the others when communicating and working together.

## 7 Conclusions

In this study, we investigated the experiences and sense of presence generated with a tele-immersive 3D video environment in K-12 education. We investigated what the main challenges are for the technology at the present as well as its potential positive effects. Based on the collected data, we were able to provide answers to our research questions regarding the preliminary investigation on the perceptions of the learners towards 3D telepresence technology for education and the impact the new medium currently has on the learners’ sense of presence. According to our assumptions and first

hypothesis, the perceptions of the learners on the use of immersive 3D video in education were positive. The learners expressed feelings of positive surprise and excitement and felt the technology enabled participation for them in the action of the lesson. Our second hypothesis was also confirmed, as immersive 3D video technology fostered a high sense of presence in the learners. The learners recounted the feeling that they were “there” inside the environment and in the company of the teacher being able to communicate their intentions and thoughts and to reciprocally understand the teacher.

Our study, however, presents limitations including the accessibility of the technology at present due to the high cost of the equipment, and persistent technical challenges, such as physical discomfort and software glitches that may hinder the system’s use during regular curriculum instruction and education. Also, the number of participants was small, making it difficult to draw conclusions that are statistically relevant. Nevertheless, it is important to highlight the novel nature of our TIP technology under development. It would be useful to further investigate the effectiveness of instruction within an immersive 3D video environment in multiple iterations and lengthy implementations, in order to counter the novelty effect in the new technology, and increase the credibility of research findings, particularly under the participatory design paradigm. More research and development work are also needed to ease user access with lighter technical setups and enhanced portability to, for instance, mobile devices.

## References

1. Nash, C.: Report on Digital Literacy in Academic Meetings during the 2020 COVID-19 Lockdown. *Challenges*, 11(2), (2020). <https://doi.org/10.3390/challe1102002>
2. Hacker, J., vom Brocke, J., Handali, J., Otto, M., Schneider, J.: Virtually in this together - how web-conferencing systems enabled a new virtual togetherness during the COVID-19 crisis. *European Journal of Information Systems* 29(5), 563-584 (2020). <https://doi.org/10.1080/0960085X.2020.1814680>
3. Luo H., Li G., Feng Q., Yang Y., Zuo M.: Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019. *J Comput Assist Learn* 1-15 (2021). <https://doi.org/10.1111/jcal.12538>
4. Johnson-Glenberg, M. C.: Immersive VR and Education: Embodied Design Principles That Include Gesture and Hand Controls. *Front. Robot. AI* 5:81 (2018). <https://doi.org/10.3389/frobt.2018.00081>
5. Rötönen, E., Islam, A.K.M.N., Sutinen, E.: Toward Pedagogy Driven Virtual Reality Learning Space Design. In: Tatnall, A., Mavengere, N. (eds) *Sustainable ICT, Education and Learning. SUZA 2019. IFIP Advances in Information and Communication Technology*, vol 564. Springer, Cham, Switzerland (2019). [https://doi.org/10.1007/978-3-030-28764-1\\_26](https://doi.org/10.1007/978-3-030-28764-1_26)
6. IIEP-UNESCO. <https://learningportal.iiep.unesco.org/en/issue-briefs/improve-learning/curriculum-and-materials/information-and-communication-technology-ict>, last accessed 2021/4/4
7. Lindsay, J.: Connecting beyond the classroom - Move from local to global learning modes. *Scan*. 36, 27-38 (2017).
8. Liesa-Orús, M., Latorre-Cosculluela, C., Vázquez-Toledo, S., Sierra-Sánchez, V.: The Technological Challenge Facing Higher Education Professors: Perceptions of ICT Tools for Developing 21C Skills. *Sustainability* 12(13), 5339 (2020). <https://doi.org/10.3390/su12135339>
9. Facer, K.: *Learning Futures: Education, Technology and Social Change*. 1st edn. Routledge, London (2011). <https://doi.org/10.4324/9780203817308>

10. Papanastasiou, G., Drigas, A., Skianis, C., Lytras, M., Papanastasiou, E.: Virtual and augmented reality effects on K-12, higher and tertiary education students' twenty-first century skills. *Virtual Reality* 23(4), 425–436 (2019). <https://doi.org/10.1007/s10055-018-0363-2>
11. Hurst, B., Wallace, R., Nixon, S. B.: The Impact of Social Interaction on Student Learning. *Reading Horizons: A Journal of Literacy and Language Arts*, 52(4) (2013). Retrieved from [https://scholarworks.wmich.edu/reading\\_horizons/vol52/iss4/5](https://scholarworks.wmich.edu/reading_horizons/vol52/iss4/5)
12. Nowak, K.L., Biocca, F.: The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence. Copresence, and Social Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 12(5), 481–494 (2003). <https://doi.org/10.1162/105474603322761289>
13. Garrison, D., Anderson, T., Archer, W.: Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education* 2, 87-105 (1999). [https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)
14. Herbelin, B., Vexo, F., Thalmann, D.: Sense of presence in virtual reality exposures therapy. (2003).
15. Ens, B., Lanir, J., Tang, A., Bateman, S., Lee, G., Piumsomboon, T., Billinghamurst, M.: Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies* 131, 81-98 (2019).
16. Kurillo, G., Bajcsy, R.: 3D teleimmersion for collaboration and interaction of geographically distributed users. *Virtual Reality* 17, 29–43 (2013). <https://doi.org/10.1007/s10055-012-0217-2>
17. Economou, D., Doumanis, I., Argyriou, L., et al.: User experience evaluation of human representation in collaborative virtual environments. *Pers Ubiquit Comput* 21, 989–1001 (2017). <https://doi.org/10.1007/s00779-017-1075-4>
18. Orts-Escolano, S., et al.: Virtual 3D Teleportation in Real-time. In: Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16), pp. 741–754. Association for Computing Machinery, New York, NY (2016). <https://doi.org/10.1145/2984511.2984517>
19. Zhang, H., Xu, F.: MixedFusion: Real-Time Reconstruction of an Indoor Scene with Dynamic Objects. *IEEE Transactions on Visualization and Computer Graphics* 24(12), 3137–3146 (2018). <https://doi.org/10.1109/TVCG.2017.2786233>
20. Marniok, N., Goldluecke, B.: Real-Time Variational Range Image Fusion and Visualization for Large-Scale Scenes Using GPU Hash Tables. In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) on 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 912–920. Piscataway, NJ (2018). <https://doi.org/10.1109/WACV.2018.00105>
21. Zimmerman, J., Forlizzi, J., Evenson, S.: Research through design as a method for interaction design research in HCI. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07), pp. 493–502. Association for Computing Machinery, New York, NY (2007). <https://doi.org/10.1145/1240624.1240704>
22. Gaver, W.: What should we expect from research through design? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12), 937–946. Association for Computing Machinery, New York, NY (2012). <https://doi.org/10.1145/2207676.2208538>
23. Lombard, M. et al.: Measuring Presence: The Temple Presence Inventory, *Proceeding Presence 2009*, 12th Int. Work. Presence, no. January, pp. 1–14, 2009.
24. Biocca, F., et al.: The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. 4th Annu. Int. Work. Presence, pp. 1–9. (2001).
25. Vorderer, P., et al.: MEC Spatial Presence Questionnaire (MEC-SPQ) Short Documentation and Instructions for Application. Rep. to Eur. Community, Proj. Presence MEC (2004).
26. Braun, V., Clarke, V.: Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 77–101 (2006). <https://doi.org/10.1191/1478088706qp063oa>