Towards Collaborative Learning in Virtual Reality: A Comparison of Co-Located Symmetric and Asymmetric Pair-Learning

Tobias Drey
Institute of Media Informatics, Ulm University
Ulm, Germany
tobias.drey@uni-ulm.de

Patrick Albus
Institute of Psychology and Education, Ulm University
Ulm, Germany
patrick.albus@uni-ulm.de

Simon der Kinderen
Institute of Media Informatics, Ulm University
Ulm, Germany
simon.der@uni-ulm.de

Maximilian Milo
Institute of Media Informatics, Ulm University
Ulm, Germany
maximilian.milo@uni-ulm.de

Thilo Segschneider
Institute of Media Informatics, Ulm University
Ulm, Germany
thilo.segschneider@uni-ulm.de

Linda Chanzab
Institute of Media Informatics, Ulm University
Ulm, Germany
linda.schroedl@uni-ulm.de

Michael Rietzler
Institute of Media Informatics, Ulm University
Ulm, Germany
michael.rietzler@uni-ulm.de

Tina Seufert
Institute of Psychology and Education, Ulm University
Ulm, Germany
tina.seufert@uni-ulm.de

Enrico Rukzio
Institute of Media Informatics, Ulm University
Ulm, Germany
enrico.rukzio@uni-ulm.de

ABSTRACT

Pair-learning is beneficial for learning outcome, motivation, and social presence, and so is virtual reality (VR) by increasing immersion, engagement, motivation, and interest of students. Nevertheless, there is a research gap if the benefits of pair-learning and VR can be combined. Furthermore, it is not clear which influence it has if only one or both peers use VR. To investigate these aspects, we implemented two types of VR pair-learning systems, a symmetric system with both peers using VR and an asymmetric system with one using a tablet. In a user study (N=46), the symmetric system statistically significantly provided higher presence, immersion, player experience, and lower intrinsic cognitive load, which are all important for learning. Symmetric and asymmetric systems performed equally well regarding learning outcome, highlighting that both are valuable learning systems. We used these findings to define guidelines on how to design co-located VR pair-learning applications, including characteristics for symmetric and asymmetric systems.

CCS CONCEPTS

• **Human-centered computing → User studies**; **Virtual reality**; *Collaborative interaction*; *Tablet computers*; • **Applied computing → Collaborative learning**; *Interactive learning environments.*

KEYWORDS

pair-learning; collaborative learning; symmetric and asymmetric system; virtual reality; signaling

ACM Reference Format:
Tobias Drey, Patrick Albus, Simon der Kinderen, Maximilian Milo, Thilo Segschneider, Linda Chanzab, Michael Rietzler, Tina Seufert, and Enrico Rukzio. 2022. Towards Collaborative Learning in Virtual Reality: A Comparison of Co-Located Symmetric and Asymmetric Pair-Learning. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3491102.3517641

1 INTRODUCTION

Learning is an essential part of everybody's life. It is the base for our economic success and accompanies us from childhood on. Two examples that foster learning are virtual reality (VR) and pair-learning. VR was frequently used by previous research for learning applications due to its benefits regarding active engagement, motivation, and interest of students [8, 13, 48, 49, 61], as well as higher learning outcome [4, 19, 42, 55]. VR further offers the possibility to create low-cost and high-quality laboratories and to have expeditions to dangerous places [24, 64]. Pair-learning, which means collaborative learning between two persons, is another VR independent approach to increase learning outcome [40, 72, 75]. It is known to enhance motivation, social presence, and individual self-esteem among students [59, 75, 76].

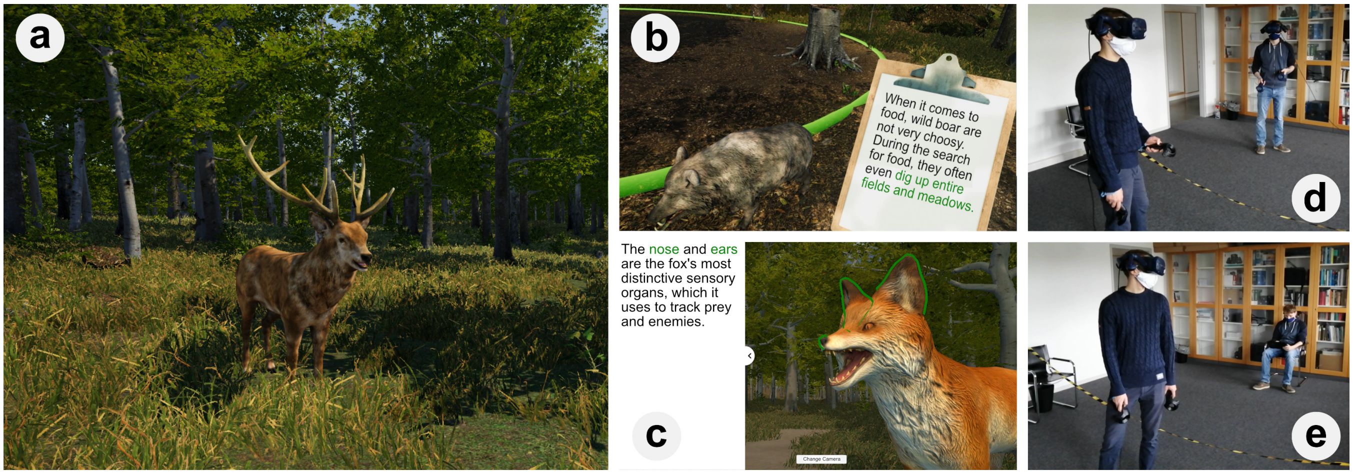
Unfortunately, very little is known about a combination of VR and pair-learning [61], but due to the previously stated benefits, we assume that with VR pair-learning, it is possible to create environments beneficial for learning.

Previous works regarding collaborative VR systems showed differences in communication and working together depending on how these systems are implemented. One approach is a symmetric

---

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04.
https://doi.org/10.1145/3491102.3517641



When it comes to food, wild boar are not very choosy. During the search for food, they often even dig up entire fields and meadows.

The nose and ears are the fox's most distinctive sensory organs, which it uses to track prey and enemies.

Figure 1: We developed a VR pair-learning application about forest animals (a) and investigated in a user study (N=46) if a symmetric (b, d) or asymmetric (c, e) co-located system impacts presence, immersion, player experience, cognitive load, and learning outcome. The effect of signaling important details (green markings, b, c) on these measures was furthermore investigated.

system where all users use the same type of device, e.g., two VR head-mounted displays (HMDs) [16, 20, 53, 60, 92], the other is an asymmetric system where the users have different types of devices, e.g., one VR HMD and one tablet/smartphone/laptop [15, 30, 57, 58] to join the virtual environment (VE). Whereas symmetric systems allow interaction as equal peers [37, 91, 92], which could foster presence, immersion, and communication, asymmetric systems can be used to guide and supervise the immersive peer [15, 56, 58], which perfectly fits the roles of students and teachers in the context of education. They also require fewer safety precautions than symmetric systems where both users can theoretically hit each other when in the same physical space and are occluded by the HMD. Furthermore, they are cheaper as only one HMD is needed, and laptops or tablets are often already available. This allows easy asymmetric co-located or remote collaboration out-of-the-box (see Oculus spectator view [33]), e.g., at home. These characteristics of symmetric and asymmetric systems have prompted our research.

In a pair-learning system, learners may be distracted from learning not only by the many visual objects in a VE but also by the additional peer. This additional distraction could have a negative impact on learning outcome because cognitive capacity is strained by unnecessary stimuli (e.g., unimportant body movement of peer or animated background not relevant for learning). Therefore, it may be especially important for virtual pair-learning environments that learners can distinguish relevant stimuli from irrelevant [65]. Guiding the attention of learners can be an appropriate solution. Different ways exist for combinations of different roles or symmetric and asymmetric systems to guide the attention of peers in non-learning contexts [15, 45, 52, 56]. Signaling, which means guiding the attention of learners by visual signs, can be such a mean and is counted to the classic multimedia design principles. Signaling highlights, for example, important objects or regions in the VE [3, 38, 50] and can influence the cognitive load of learners tested in single-user systems [3, 14]. Nevertheless, studies exist that such classic multimedia design principles (e.g., signaling or redundancy effect), are differently than one would expect not generally valid for all kinds of multimedia-based learning and cannot simply, without validation, be transferred to other learning systems, e.g., VR learning [46]. This work will, therefore, investigate if signaling can be transferred to VR pair-learning.

As no work exists that directly compares symmetric and asymmetric systems for VR pair-learning, including the influence of signaling on these [61, 72] and previous works exist that classic multimedia design principles cannot be transferred to other learning systems, we investigated this in an extensive and holistic approach. With our approach we investigated not only learning outcome but also factors supportive for learning (see [3, 8, 13, 48, 49, 51, 61]; used in the research questions (RQs)), and defined the following RQs:

RQ 1: *"Does symmetric collaboration compared to asymmetric collaboration in VR pair-learning environments impact presence, immersion, player experience (PX), motivation, cognitive load, and learning outcome of students and teachers?"*

RQ 2: *"Does signaling in symmetric and asymmetric collaborative VR pair-learning environments impact presence, immersion, PX, motivation, cognitive load, and learning outcome of students and teachers?"*

We implemented a pair-learning environment where a teacher teaches students typical characteristics of forest animals (see Figure 1). This topic was perfect for our VR learning unit, as it allowed us to use high-fidelity animal models in true size to teach their characteristics and allowed the teachers and students to observe them in their direct vicinity without risk and inside their natural habitat, an experience no zoo with real animals or a book can provide. We further chose this topic, as nature awareness is very important due to growing environmental pollution and climate change, but it is relatively low in the general population. Therefore, it was a topic in which most potential participants would have low prior knowledge, which made it possible to measure our learning unit's learning outcome. Due to the lack of forest animal experts in the general population, which could have mimed a teacher, we created teaching sheets for our learning unit, which contained all information about the animals as explanatory text. The teaching sheets were visible to only one participant during pair-learning so that this participant could act in the role of a teacher, whereas the other participant acted as a student. As both participants would potentially have low prior knowledge, this setup created a learning situation for both participants.

To answer our RQs, we developed two different pair-learning prototypes and compared them against each other. In a symmetric prototype, the student and teacher used VR HMDs and explored the virtual forest together as equal peers (see Figure 1b/d). In the asymmetric prototype, only the student explored the forest with a VR HMD while the teacher was guiding using a spectator view on a tablet seeing the VE through the student's eyes (see Figure 1c/e). This allowed the teacher to follow the student accurately and mimicked state-of-the-art HMD spectator views (e.g., Oculus Quest [33]). Both prototypes had the possibility to enable signaling for highlighting essential information (see Figure 1b/c). This was the base for our $2x2$ study design that further considers the two roles student and teacher.

Based on our assumptions, we formulate the following hypotheses:

- **H1:** *"Symmetric collaboration leads to higher presence, immersion, PX, and motivation and positively influences cognitive load compared to asymmetric collaboration."*
- **H2:** *"Asymmetric pair-learning systems can provide excellent and similar learning outcome for both learners."*
- **H3:** *"Signaling in VEs independently of the setup enhances presence, immersion, PX, motivation, and learning outcome and positively influences cognitive load compared to non-signaling."*

Our user study's findings (N=46) show that presence, immersion, and PX were statistically significantly higher, and cognitive load was statistically significantly lower in the symmetric condition with moderate and high effect sizes. This was also the case for the students who used an immersive HMD in all conditions and shows that not only the own device but also the device of the peer influenced these measures. The symmetric and asymmetric conditions performed great and equally well regarding learning outcome. According to the qualitative feedback, communication was easier for participants in the symmetric system showing the importance of an avatar for both peers. Signaling showed statistically significantly higher presence, PX, and better cognitive load for students, which supports learning and indicates that it can be transferred to VR pair-learning. Based on these findings, we define six guidelines on how to design co-located VR pair-learning applications to support researchers, educators, and software developers in their work.

The threefold contribution of our work is:

1. The comparison of symmetric and asymmetric as well as signaling and non-signaling VR pair-learning environments in a user study (N=46) with six dependent variables (presence, immersion, PX, motivation, cognitive load, learning outcome), eight control variables (age, gender, VR experience, participants relationship, spatial abilities (spatial visualization, figure difference recognition), working memory capacity, task duration), and a qualitative thematic analysis based on semi-structured interviews.
2. Design insights for a complex symmetric and asymmetric VR pair-learning application that optionally enables signaling to highlight essential information.
3. The definition of six novel guidelines on how to design symmetric as well as asymmetric co-located VR pair-learning applications, based on our qualitative and quantitative results.

2 RELATED WORK

Our work is based on previous work regarding VR in education, pair-learning, the signaling principle, and symmetric and asymmetric systems in VR. All these topics have benefits by themselves, but we will show that we are the first who brought them all together.

2.1 VR in Education

Various previous works have used VR for education and reported positive effects. Mantovani et al. [49] discussed VR in relation to educational theory and pedagogical practice in order to establish a possible theoretical basis for VR learning. They showed that VR could foster active engagement due to increased motivation and interest of students. According to a systematic literature review (SLR) by Pirker et al. [61], VR has advantages regarding interaction and immersion, visualization and metaphors, playful design, and social experiences relevant for educational systems.

Using GEtiT, Oberdörfer et al. [55] compared a VR with a desktop system. They had the result that participants had higher enjoyment and a more intuitive knowledge demonstration with VR, which indicates that VR is beneficial for learning. Krokos et al. [42] compared a VR HMD condition with a desktop condition for virtual memory remembering tasks. Their results show that the HMD users were able to recall the memory more accurately. As a reason for their results, they named the spatial awareness of the HMD users as an important success factor.

A similar positive effect for learning was found by Allcoat and von Mühlenen [4], comparing VR with a desktop and a video condition. While these positive effects support learning in VR, they do not come at no cost. Compared to traditional learning environments, learners have to orient themselves spatially and select the relevant visual stimuli from a multitude of visual stimuli. This may result in an increased cognitive load [3]. However, Radianti et al. [65] looked at several studies in a meta-analysis of learning in VR and found that previous studies that measured learning outcome in VR lacked a foundation of learning theory in many cases. Therefore, they propose that the effectiveness of virtual learning environments and the underlying learning processes should be investigated in more detail. Their findings are the reason why we based our work on learning theories.

One learning theory that is already established in conventional learning environments is the cognitive theory of multimedia learning (CTML [50]). The CTML describes relevant cognitive processes that take place during learning, such as the process of selection in which information that is subjectively relevant to the learner is selected and stored in the limited working memory (see Baddeley [7]) for further processing, leading to a bottleneck for learning when irrelevant information is selected [50].

Another theory, the Cognitive Load Theory (CLT), can be used to look more closely at the processes in working memory and the three involved cognitive load types. According to Chandler & Sweller's [14], these are the intrinsic cognitive load (ICL) (the complexity of the learning unit itself), the extraneous cognitive load (ECL) (caused by suboptimal design of learning material), and the germane cognitive load (GCL) (the engagement of the learner required to understand the learning unit). Optimally, the ECL should be reduced, the GCL maximized, and the ICL controlled. For example, too large ECL can cause cognitive overload in working memory [51].

We used the benefits of VR in education for our pair-learning system and considered CTML and CLT during the design and evaluation of it.

2.2 Pair-Learning

One of our research's main aspects is that we enable two users to learn collaboratively, which we will refer to as *pair-learning*. While VR was frequently used for education, social collaborative experiences are underrepresented, as reported by a SLR of Pirker et al. [61]. This shows our work's importance, as cooperative learning methods can increase student achievements, mutual concern among students, and student self-esteem [75]. It further enhances motivation and social presence [59, 76] and was, according to Johnson et al. [40], the most effective learning approach comparing computer-assisted cooperative, competitive, and individual learning.

Mantovani et al. [49] stated that ideally, a teacher should support the student's learning process. When comparing a single-user with system-based teaching with a two-user version, in which one user was a teacher, Simeone et al. [72] showed that the two-user version scored higher regarding overall preference and clarity compared to the single-user version that used animation sequences for teaching. Thompson et al. [83] and Uz-Bilgin et al. [86] explored how to foster collaboration between students in a VR/tablet-based educational game and defined a guiding role provided with all necessary information on the tablet, similar to a teacher. Their results show that such roles can shape collaboration between users. With RoleVR, Lee et al. [44] showed that roles can be an important factor in asymmetric systems and that the non-HMD user feels as immersed as the HMD user.

Based on these works, we decided to use a cooperative learning approach. They brought us, furthermore, to the decision to add a teacher role to our learning environment and to assign that role to a human user rather than just mimicking a teacher through the system.

2.3 The Signaling Principle

When creating virtual learning environments, the question arises of how to design virtual learning units so that they can promote the learning processes. As mentioned above, one challenge when learning in VR can be distinguishing the stimuli that are relevant to learning from those that are irrelevant. Learners also have to orient themselves more spatially in a VE, so the selection of relevant information can be more straining on the cognitive capacity of the learners'. To facilitate the selection process in VR based on the design principles of CTML, the signaling principle can be used. As Albus et al. [3] showed, signaling can improve learning outcome in single-user VR compared to a non-signaling group. Furthermore, they found that using the signaling principle in VR can also increase the GCL, i.e., the cognitive load that promotes learning. While there are different variants of signaling, such as highlighting [38], motion [2, 29], or annotations [3], we have chosen highlighting for this learning environment because it can focus the learner's attention on relevant things in a visually loaded VE [38] (see also the theory of Witmer and Singer [90] regarding presence) and works with non-moving animal models, allowing participants to focus the highlighted regions statically. We refrained from a more complex highlighting, which for example, is additionally animated or spatially separated, in order to avoid confounding of the presumed effects and to allow a more precise interpretation, if the existing findings and theory for signaling can be transferred to pair-learning in VR.

2.4 Symmetric and Asymmetric Systems in VR

Besides typical single-user VR systems, multi-user systems exist as well, which were implemented in a symmetrical way (all users use the same devices, e.g., two VR HMDs [16, 20, 60, 92]), or in an asymmetrical way (users have different devices, e.g., VR HMD and tablet/smartphone/laptop [15, 30, 57, 58, 67]).

2.4.1 Symmetric Systems. The benefits of symmetric VR systems for education, such as higher presence and an enjoyable experience, were investigated for several years, as the work of Jackson et al. [37] shows. Collaborative systems can improve users' efficiency, as shown by Pinho et al. [60], who found the efficiency of cooperative object manipulation with a VR HMD based system to be higher compared to a single-user manipulation. Zaman et al. [92] created a symmetric system for VR HMDs to design spatial room layouts collaboratively and showed that it is important for communication to have a shared perspective with all partners. Similar results were found by Nguyen et al. [53] with CollaVR, a collaborative VR video editing tool. Another scene editing authoring approach was presented with Spacetime by Xia et al. [91], combining single-user interaction with fluid multi-user collaboration of the VR users.

Dey et al. [20] investigated the user experience (UX) of symmetric VR systems by sharing the emotional states of two collaborators based on their heart rates, which the participants preferred compared to no emotional feedback. Collaboratively training surgeons in VR was investigated by Chheang et al. [16]. They had positive reviews regarding usability and usefulness, which shows that symmetric collaborative VR systems have the potential for a learning application.

These previous works have in common that the symmetric approach was never explicitly the topic of the study nor compared to an alternative approach, e.g., an asymmetric system. Our work instead directly investigates the influence of a symmetric and asymmetric system on collaborative learning in VR.

2.4.2 Asymmetric Systems. In contrast to many symmetric systems, asymmetric systems often directly investigate the asymmetric setup, e.g., they investigate the interaction between VR and desktop or tabletop users [26, 35, 78] or investigate the asymmetric system for tasks like authoring of VEs [34].

Grandi et al. [30] created a system for co-located object interaction in mixed reality (MR) and compared both a symmetric VR/VR and mobile augmented reality (AR)/AR setup with an asymmetric VR/mobile AR setup in a user study (N=36). The users' performance using the VR/VR and the VR/mobile AR system were similar and statistically significantly better compared to the mobile AR/AR system. This shows that peering an HMD user with an asymmetric partner can have similar systems quality than a symmetric system.

A good UX of all asymmetric users is important and was investigated by Olin et al. [57] for a system combining a smartphone and an HMD user. Both their users felt present in the VE, and their participants preferred to face each other during conversation. Gugenheimer et al. [31] presented ShareVR, a co-located asymmetric system for HMD and non-HMD users, using a tracked display and a floor projection. Their gamified experiences showed that their collaborative system increased presence, enjoyment, and social interaction for all users. Similar results were presented by Brondi et al. [13] when comparing a keyboard- and mouse-based with an HMD condition. They showed that both users were very satisfied with their system and aware of the others' actions. Schott et al. [67] created a multi-user liver anatomy education system with a VR, AR, and spectator mode. In a first qualitative usability walkthrough, they received positive feedback, which shows that such symmetric and asymmetric learning systems provide a promising base for further research.

Besides these co-located setups, asymmetric remote collaboration is possible as well. It can be used to guide an AR/VR user by a remote advisor. This was used for remote maintenance [15, 56] or guiding a VR user through a VE [58]. The remote user can use an AR/VR HMD [56, 63] or other devices such as a tablet or laptop [15, 58].

These examples show that non-HMD users can feel present as well in asymmetric setups and that concepts exist that can even improve the presence, enjoyment, and social interaction of all users. The flexibility of asymmetric systems to support different roles with different tasks and competencies perfectly fits our student and teacher approach. As facing each other was preferred [57], we used a co-located setup for our system and not a remote one.

With our work, we are the first who combined the benefits and potentialities of VR in education, pair-learning, signaling, and symmetric and asymmetric system to create a role-based (student and teacher) learning environment.

3 PAIR-LEARNING PROTOTYPE

To answer our RQs, we implemented a pair-learning prototype and developed a learning unit.

3.1 Learning Unit

To enable pair-learning, two students were situated together in the same VE. We decided to teach them about the typical characteristics of forest animals. For the development of our learning unit, we included experts for the domain forest animals and experts from teaching and learning research. The content of the learning unit was shown as teaching sheets to one of the two students. We will further call them teachers. Teachers should teach the content to the other student but also learn it themselves. The teaching sheets ensured that all teachers taught their student the same content to be able to compare the study results. We aimed to develop a conversation between the two roles by providing individual questions to student and teacher.

The VE, in which the participants found themselves, is a forest with small clearings and different animals (see Figure 1a/b/c). We chose the three forest natives, deer, fox, and boar, for our application. Each of the animals was standing on a different clearing. The learning unit taught the students about the species of deer that it is a mammal and cloven-hoofed, how to determine its age, details about the antler, the eating behavior, and facts about their teeth. It further introduced the fox as a canine and provided details about its tail and scent gland. The last clearing showed a boar, which is also a cloven-hoofed animal that stirs up entire meadows during foraging and likes wallowing in mud. Boars are omnivores, and the size of their tusks can distinguish female and male boars. By using VR, it was possible to teach the characteristics of the animals with high-fidelity 3D models in true size. It was possible to approach and observe them in their vicinity in their natural habitat without risk, an experience that no zoo, reserve, book, or standard learning application can provide. Besides the animal models, dedicated models for deer antlers and teeth, as well as stirred-up meadows, were created to visually support all aspects of the learning unit. Signaling was used to guide the attention of the users to the specific body parts and bring them into focus. Before the learning unit started, all participants went through a tutorial scene to get familiar with the VE, the devices, the interaction, and the communication between peers. We provide the teaching sheets as supplementary material.

3.2 Pair-Learning Prototype Design

In this section, we introduce all features that were implemented for our pair-learning application, beginning with the symmetric setup, where both participants were in VR, and will afterward explain the changes for the asymmetric setup in which the teachers used a tablet.

3.2.1 Symmetric Pair-Learning Prototype. Since there were always two participants in the same VE, we used avatars to visualize their position and movements as well as hand gestures. The avatars further made it possible to see the direction of the other person's gaze by using the head tracking of the HMD (see Figure 2a). The avatars consisted of a head and two hands which were tracked by the HMD and the controllers. Student and teacher could be distinguished by their head-wear. We chose a minimalist human-like appearance to not distract the participants. We modeled no body and legs, which should have no effect on immersion and body ownership [47, 70]. A similar avatar design was used by *Spacetime* [91] for collaborative VEs before. It also should prevent an uncanny valley effect [43].

The participants could walk freely inside the VE within the tracking space of the HMD, and teleportation was possible to travel longer distances [10] (see Figure 2e). We included features to help the participants keep orientated in the VE, such as a path connecting

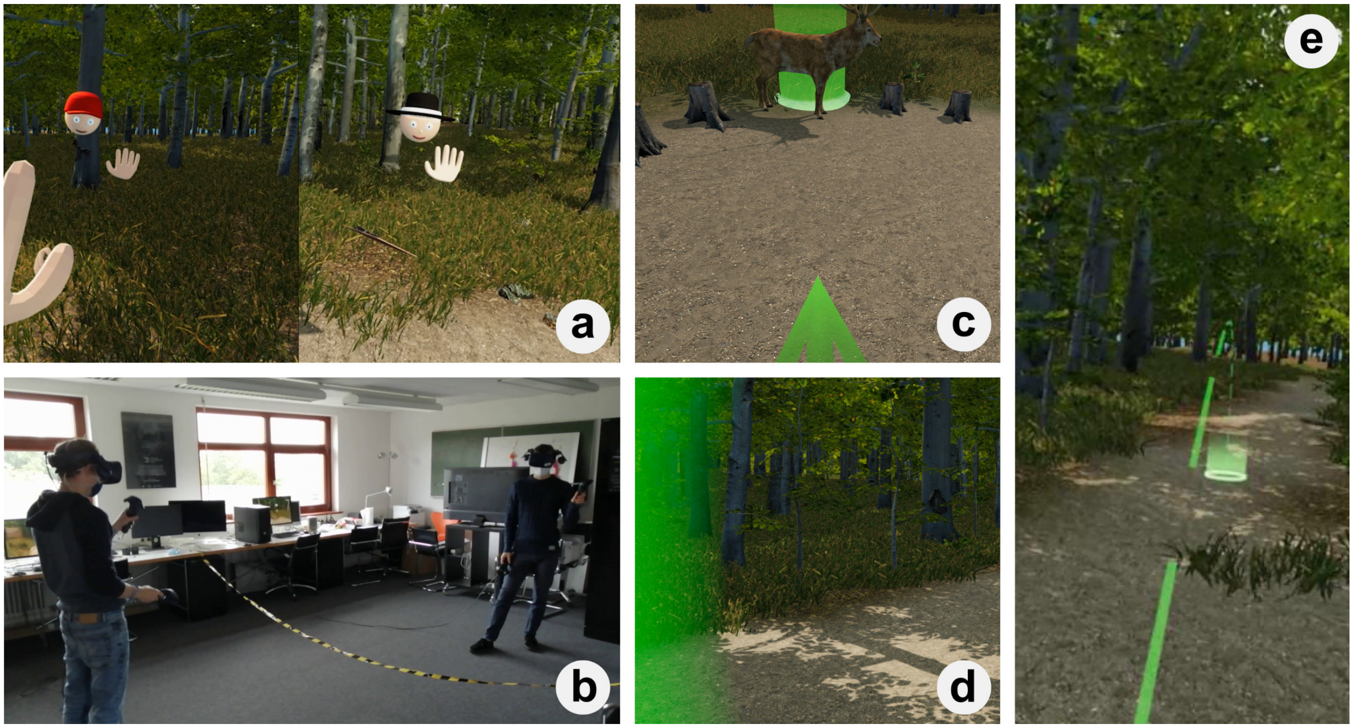


Figure 2: In the symmetric VR/VR setup, both participants were visualized with an avatar (a) mimicking the participants' head and hand movements (b). To keep the orientation in the VE, an arrow showed the direction to the next way-point (c). If the arrow was not in the field of view, the left or right edge of the field of view lighted up green and showed the direction (d). The participants could walk freely within the tracking space of the HMD inside the VE or teleport (e) to travel long distances.

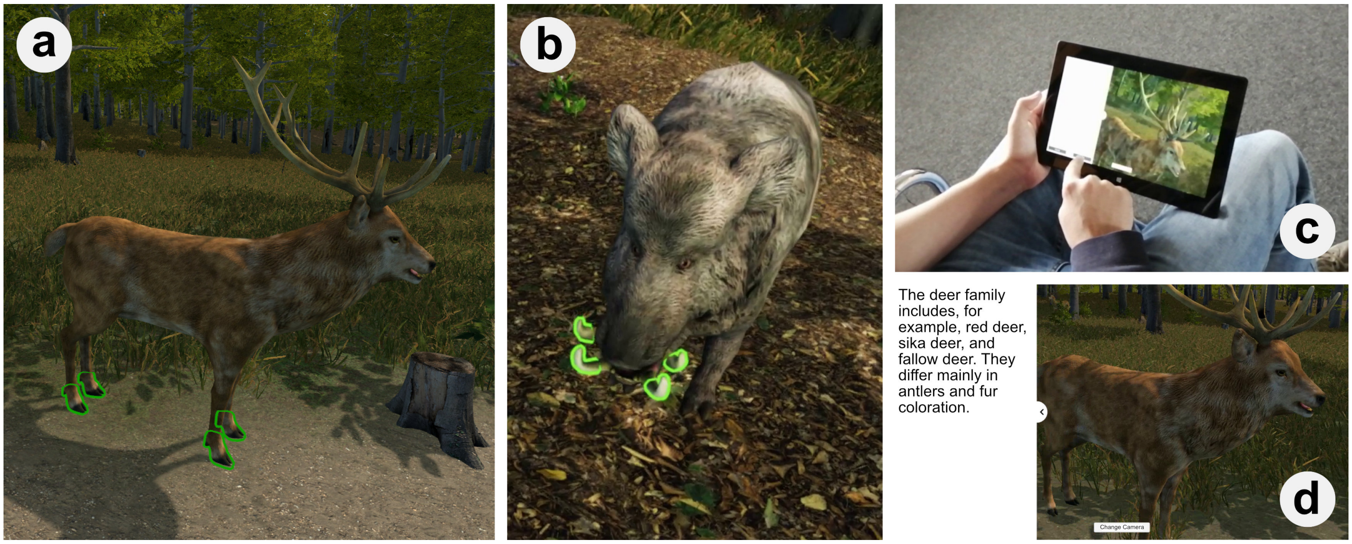
the clearings as well as a navigation system indicating the direction of the next way-point with an arrow (see Figure 2c). If the arrow was not in their field of view, the left or right edge of the display lighted up green (see Figure 2d).

It was possible to grab certain objects (e.g., antler and teeth of a deer) to allow participants to take a closer look. When an object was grabbed, the avatar's hand disappeared so that the object was not covered.

To answer RQ 2, *"Does signaling in symmetric and asymmetric collaborative VR pair-learning environments impact presence, immersion, PX, motivation, cognitive load, and learning outcome of students and teachers?"*, we implemented a signaling feature. We chose colored outlining for highlighting body parts as this was proven to be effective for guiding attention in learning applications [38] (see Figure 1c, Figure 3a/b). Signaling was automatically activated and deactivated according to the learning unit's progress, and it was always placed in front of all other objects so that it was visible even if the view on the highlighted part was obstructed. If the highlighted object was not in the participants' field of view, the edges of the display were highlighted similarly as for navigation (see Figure 2d).

Since clipboards are less cognitively demanding than an overlay interface [6], one was implemented to show the teaching sheets to the teachers as proposed by Thompson et al. [83] and Uz-Bilgin et al. [86] (see Figure 1b). Currently highlighted animal parts were also highlighted in the teaching sheets by appearing in a green font so that the teacher always knew what to talk about. For better comfort and readability, the clipboard could be fixated in mid-air.

3.2.2 Asymmetric Pair-Learning Prototype. For RQ 1, *"Does symmetric collaboration compared to asymmetric collaboration in VR pair-learning environments impact presence, immersion, PX, motivation, cognitive load, and learning outcome of students and teachers?"*, we implemented a prototype where the teacher did not symmetrically join the student in the VE but accompanied the student asymmetrically with a tablet. Therefore, we made some adjustments to the teacher application. Considering the design goals 1 and 2 of Kumaravel et al. [84] for asymmetric VR systems, we implemented a spectator view where teachers could see the environment through the student's eyes (see Figure 3c/d). Teachers also could detach themselves from the student's camera at any time and look around and rotate the camera at the current location freely. Teachers could not move through the forest by themselves to force them to always spectate the student and thereby observe everything the student does. This mimics an advanced state-of-the-art HMD spectator views (e.g., Oculus Quest [33]). We think this simplified the interaction for teachers as no own avatar has to be controlled so that they can be fully concentrated on teaching. It further reduces potentially distracting cues in the VE as no avatar is visible, which could improve learning [50]. Due to our goal to simplify the interaction and the interface of the teacher to reduce cognitive load and



The deer family includes, for example, red deer, sika deer, and fallow deer. They differ mainly in antlers and fur coloration.

Figure 3: For signaling, we highlighted body parts with green colored outlines (a, b). In the asymmetric tablet condition, the teacher observed the student through a spectator view. The teaching sheets were displayed on the left side of the screen (c, d).

to implement no features which are not available in the symmetric version, we refrained from implementing multi-modal context sharing as suggested by design goal 4 of Kumaravel et al. [84].

Since the teachers had no virtual hands or any avatar, the presentation of the teaching sheets had to be adapted as well. The teaching sheets could be displayed via a button on the left side of the tablet's screen (see Figure 1c, Figure 3c/d). Care was taken to ensure that not too much of the screen was used by this while simultaneously guaranteeing that the font remained legible. Similar to the symmetric version, signaling was activated and deactivated automatically depending on the learning unit and provided, therefore, spatial referencing as suggested by design goal 3 of Kumaravel et al. [84].

3.3 Implementation and Setup

Our collaborative pair-learning application was based on Unity [82]. For high graphical fidelity, we used the assets "Forest Environment - Dynamic Nature" [81] and "Forest Animals Pack" [80]. Non-user specific and non-directional ambient forest sounds were played by external speakers to increase immersion. We used two workstation computers with an Intel Core i5 processor, one with an NVIDIA GTX 1080 and one with an NVIDIA GTX 970 graphics card. One HTC Vive Pro HMD was attached to each computer. The used tablet was a Microsoft Surface 2 Pro. Graphic settings were reduced for the tablet version. For synchronizing the instances via the network, we used the asset "Dark Rift Networking 2" [79] for our client-server architecture. The setup featured a $4x4$ meters co-located VR tracking space with two SteamVR 2.0 base stations. For both the symmetric VR/VR and the asymmetric VR/tablet conditions, the tracking space was evenly split in half with barrier tape (see Figure 4). For the VR/tablet conditions, a chair was provided, facing towards the student's VR space (see Figure 4b). Due to the co-located setup, both users were able to communicate verbally with each other without the need for telecommunication devices.

4 USER STUDY

We conducted a user study to answer RQ 1, *"Does symmetric collaboration compared to asymmetric collaboration in VR pair-learning environments impact presence, immersion, PX, motivation, cognitive load, and learning outcome of students and teachers?"*, and RQ 2, *"Does signaling in symmetric and asymmetric collaborative VR pair-learning environments impact presence, immersion, PX, motivation, cognitive load, and learning outcome of students and teachers?"*.

4.1 Study Design

Our prototype had a symmetric as well as an asymmetric condition with one student and one teacher role for the participants. Signaling could be further enabled or disabled in both conditions by us to highlight important objects in the learning unit. As participants were tested regarding the learning outcome, they could only participate in one condition. Therefore, we chose a $2x2$ between-subject study design with the two roles student and teacher (see Figure 5).

A preliminary study was conducted to test the functionality of the prototype and to improve its usability. As a result, we shortened the learning unit and improved the navigation arrow, how to grab objects, and the readability of the teaching sheets.

4.1.1 Dependent Variables.

The study was designed to assess the participants' presence, immersion, PX, motivation, and cognitive load, which positively influence learning (see [3, 8, 13, 48, 49, 51, 61]), and learning outcome. For presence in VEs multiple definitions exist as well as ways to measure it by, e.g., questionnaires or physiological measures [68, 73, 77]. We refer to the definition of presence as the feeling of "being there", which means presence reflects the feeling of actually being in a different environment than the physical one (see Skarbez et al. [73], Slater [74], and Witmer et al. [89, 90]). For our understanding of presence, it is not important that users have the feeling of being non-mediated, and we think of presence as a non-binary state [73]. We chose to use a questionnaire

Figure 4: The study setup for the symmetric VR/VR (a) as well as the asymmetric VR/tablet (b) conditions consisted of a $4x4$ meters VR tracking space evenly split in half with barrier tape. For the VR/tablet conditions, a chair was provided, facing towards the student's VR space.

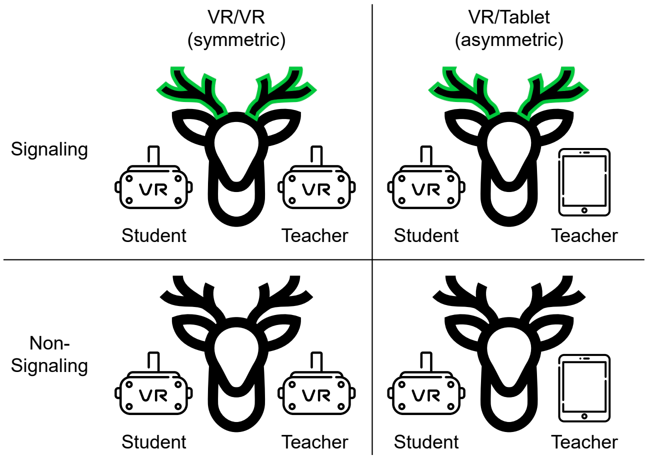


Figure 5: The study design consisted of a symmetric VR/VR condition where both participants, student and teacher, used an HMD and an asymmetric VR/tablet condition where the participant with the teacher role used a tablet instead. It further contained a signaling and non-signaling condition resulting in a $2x2$ design with the two roles student and teacher. As we measured the learning outcome of each condition, we used a between-subject design. Icons made by Freepik from www.flaticon.com

to measure presence, as this is still the most used way of measurement [77]. Different popular ones exist [25, 69, 77], such as the Presence Questionnaire (PQ) by Witmer and Singer [90] (Google Scholar citations: 6047, 12/04/21), the Slater-Usoh-Steed Presence Questionnaire (SUS PQ) by Usoh et al. [85] (Google Scholar citations: 837, 12/04/21), and the Igroup Presence Questionnaire (IPQ) by Schubert et al. [68] (Google Scholar citations: 1534, 12/04/21). We chose the PQ by Witmer and Singer [90] as it provides questions regarding distraction factors related to learning in a VE (see Jamet et al. [38], see section 2.3) and subscales for a more nuanced analysis as well as an analysis of more different components [69]. It is further validated, highly used, and recommended (see Souza et al. [77]) and provides the lowest variance of the three, which is a sign of its reliability [69]. For immersion in VEs also multiple definitions exist. There is the more technical definition of Slater [74] which states that a fully immersive system can fake all physical stimuli for a user in a way that this user is not able to distinguish it from reality. Technical immersion is, according to Witmer and Singer [90] as well as Slater [74], very important that one can become present. Nevertheless, Witmer and Singer [90] as well as Jennet et al. [39] define immersion not only as technical but see a link to player states such as flow, engagement, or others [18, 22] which complement presence and are beneficial for learning (see section 1). Investigating the player state is important for our system, as it is an educational game/serious game, which is why we chose the game-related questionnaire Immersive Experience Questionnaire (IEQ) by Jennet et al. [39] (Google Scholar citations: 1844, 12/04/21). Its subscales *challenge, control, real world dissociation, emotional involvement, and cognitive involvement* further perfectly complement the subscales *involvement/control, natural, auditory, haptic, resolution, and interface quality* of the PQ, which makes this a proper combination of immersion and presence questionnaires for our RQs.

To investigate the player state further, we chose for PX the subscales *audiovisual appeal, goals and rules* (asks if the goals and rules of the learning unit were clear), and *ease of control* of the Player Experience Inventory (PXI) questionnaire [1]. We did not use the entire PXI because relevant subscales were already covered by other questionnaires, or we rated them as not relevant for our RQs. To measure participants' motivation in (experimental) learning and performance situations and how the symmetric or asymmetric setup influenced this, we used the Questionnaire to assess Current Motivation (QCM) [66]. For cognitive load, we used the cognitive load questionnaire of Klepsch et al. [41], which measures intrinsic cognitive load (ICL), extraneous cognitive load (ECL), and germane cognitive load (GCL).

We also measured the learning outcome as follows. The relevant knowledge before the learning unit was measured with a pre-test and after it with a post-test. Our forest animal domain experts and the teaching and learning experts were consulted again for the development of the tests. The tests were developed according to Bloom's taxonomy [9]. The pre-test dealt with domain-specific knowledge questions on the topic not part of the learning unit so that the learners were not primed for the questions relevant in the post-test. The post-test specifically covered only questions which were answered during the learning unit. We included knowledge-, comprehension-, and transfer-based questions according to Bloom [9]. The previously introduced teaching sheets allowed us to compare the results of the conditions. Without it, differences could be caused by different teachers (e.g., by missed information). Pre- and post-test were both rated following a sample solution to ensure objectivity. We provide the pre- and post-test, including the sample solution, as supplementary material.

4.1.2 *Control Variables.*

Besides the prior knowledge, we further collected control variables that may have effects on the dependent variables (see Albus et al. [3]). With demographic questions, we collected age, gender, and experience with VR of our participants. We also asked the Inclusion of Other in the Self (IOS) scale [5] to measure how well the participants knew each other. To measure spatial ability or, more accurately, spatial visualization of participants, we used a paper folding test [27] as a learning characteristic that could influence learning outcome in VR. This also applies to the figure rotation test [27], which measures participants' spatial ability to recognize differences in figures. With a working memory test, we measured working memory capacity as the third measure for learner characteristics. We developed for this purpose an online version of the Numerical Memory Updating Test [54] with the same algorithms as in the original version. In addition, we logged the duration participants spent in the learning application.

4.1.3 *Semi-Structured Interviews.*

To further get qualitative feedback from the participants (see Drey et al. [23]), we conducted semi-structured interviews at the very end of the study. We asked questions regarding the participants' UX, their satisfaction with the peer communication in the symmetric and asymmetric conditions, how aware they were of their peer and their peer's actions (see workspace awareness [32]), the learning unit's difficulty, and if they encountered any issues. The questions were individually tailored to the role (student/teacher) and the condition (VR/VR / VR/tablet, signaling / non-signaling). It were seven general questions and additional five to eight questions depending on the study condition and participants' role.

4.2 Procedure

We recruited our participants through convenience sampling and assigned them to the conditions based on a randomization list. Participants initially received a short introduction and were then separated to conduct the pre-test, answer the QCM and the paper folding test, and fill out demographics. The tests and questionnaires were filled out on individual laptops in separate rooms. This ensured that participants were alone while conducting the tests and questionnaires and could not influence each other. After this, the participants were paired in the study room, received the study hardware, and started the tutorial, followed by the learning unit. Returning from the learning unit, participants conducted the post-test and answered the cognitive load questionnaire, figure rotation test, PQ, IEQ, PXI, IOS, and working memory test. This was followed by semi-structured interviews. Questionnaires and interviews were conducted once again in separate rooms.

We defined cybersickness as a no-go criteria for our study and instructed our participants that they should inform us if they experience cybersickness. As no participant stated cybersickness, we assume it did not occur during our study runs. Each participant was compensated with 10 €. Due to the COVID-19 pandemic, the study was conducted based on a hygiene concept approved by the university, which caused no limitations regarding our study design.

4.3 Participants

We collected data from 23 pairs (N=46 participants). The participants were between 20 and 29 years old ($M = 24.23$, $SD = 2.67$), and 16 participants were female and 30 male, with 34 of them studying at the moment. Ten of them had never used VR HMDs, nine answered that they had used them once, five participants each used them two times and three times, one used them four times, and the remaining 16 participants used them more than four times. 44 participants stated to know each other, and the IOS questionnaire (asks regarding the relationship between the peers) showed that participants had a closer relationship (the exact values can be found in the appendix, Figure A1). The vision of the participants was normal or corrected to normal.

5 RESULTS

In the following, we present the quantitative and qualitative results of the user study to answer our RQs regarding symmetric and asymmetric pair-learning and the effects of signaling.

5.1 Quantitative Results

The first step of our analysis was to check the control variables for statistically significant correlations, which we then considered when analyzing our study conditions. We then looked for statistically significant differences of our $2x2$ study design conditions (e.g., VR/VR compared to VR/tablet) as well as when considering the roles (e.g., teachers in VR/VR compared to teachers in VR/tablet) depending on our RQs and hypotheses. We report the results relevant for the RQs and hypotheses in the following, which are shown in Figure 6 and Figure 7. Tables providing these and all other results/data in the highest level of detail can be found in the appendix (see Figure A1 and Figure A2).

5.1.1 *Statistical Methods and Control Variables.*

To analyze if there are potential covariates, we checked the following variables: gender, age, prior knowledge, spatial ability, working memory capacity, Inclusion of Other in the Self (IOS). Correlations with dependant variables were conducted. For the score of the post-test statistically significant correlations could be found with spatial ability ($r = .33, p = .026$) and working memory capacity ($r = .45, p = .002$). For extraneous cognitive load (ECL) statistically significant correlations could be found with the IOS ($r = -.38, p = .011$). The germane cognitive load (GCL) showed statistically significant correlations with gender ($r = .31, p = .034$). We also found that the Immersive Experience Questionnaire (IEQ) is statistically significantly correlated with age ($r = .30, p = .048$). No other statistically significant correlations for the covariates could be found (see Figure A3)). The average completion time the participants spent in the learning unit was $M = 13.97$ minutes ($SD = 3.77$), with no differences between conditions (VR/VR to VR/tablet, $F(1, 21) = 1.98, p = .174, \eta^2 = .086$, and signaling to non-signaling, $F(1,21) = .28, p = .604, \eta^2 = .013$, see also Figure A1).

All covariates were included in the following calculations concerning the variable they were correlated with. The statistical analyses were performed properly, and the appropriate requirements and assumptions of the test procedures were ensured (see Field [28]). For the ANCOVA, for example, the regression slopes were checked for homogeneity, and care was taken to ensure that the residuals were approximately normally distributed for each category of the independent variable. We also deliberately performed only the analyses relevant to the RQs and, therefore, do not report further p-values. This way, we avoided conducting p-hacking or p-fishing (see Cockburn et al. [37]). To counteract biases for multiple comparison problems such as the alpha error accumulation, the conservative Bonferroni correction was applied. The ANCOVA results for the dependent variables are reported in the following including always the effect sizes $partial\ \eta^2$ or η^2 (see Dragicevic [21]).

5.1.2 *VR/VR vs. VR/Tablet.*

Regarding RQ 1, *"Does symmetric collaboration compared to asymmetric collaboration in VR pair-learning environments impact presence, immersion, PX, motivation, cognitive load, and learning outcome of students and teachers?"*, we found statistically significant differences for presence, immersion, PX, motivation, and cognitive load (see Figure 6).

Presence was role independent statistically significantly higher within the VR/VR condition compared to the VR/tablet condition ($F(1,43) = 9.89, p = .003, \eta^2 = .187$), as well as role dependent for students ($F(1,21) = 6.81, p = .016, partial\ \eta^2 = .245$) and teachers ($F(1,20) = 5.35, p = .031, partial\ \eta^2 = .211$). The same applied for immersion which was also role independent statistically significantly higher in the VR/VR condition compared to the VR/tablet condition ($F(1,43) = 8.86, p = .005, \eta^2 = .171$). Similar the PX was statistically significantly higher in the VR/VR condition for the PXI subscales *audiovisual appeal* ($F(1,43) = 5.05, p = .030, \eta^2 = .105$) and *goals and rules* ($F(1,43) = 4.94, p = .031, \eta^2 = .103$). Regarding cognitive load, the tablet version showed a statistically significantly higher intrinsic cognitive load (ICL) for teachers ($F(1,21) = 5.11, p = .034, partial\ \eta^2 = .196$) which means that learning was more complex with the tablet.

Interestingly the results were different for motivation, which was statistically significantly higher in the VR/tablet condition compared to the VR/VR condition ($F(1,44) = 8.22, p = .006, \eta^2 = .157$).

The learning outcome was equally well for the conditions VR/VR and VR/tablet for both the students and teachers, and the results were not statistically significantly different ($F(1,20) = .26, p = .613, partial\ \eta^2 = .013$) (see Figure 7). This is supported by the descriptive data, which show no descriptive trend, but only slightly and neglectable higher values for VR/VR.

5.1.3 *Signaling vs. Non-Signaling.*

Regarding RQ 2, *"Does signaling in symmetric and asymmetric collaborative VR pair-learning environments impact presence, immersion, PX, motivation, cognitive load, and learning outcome of students and teachers?"*, we found statistically significant differences for presence, PX, and cognitive load (see Figure 6).

Presence was statistically significantly higher for the students in the signaling condition compared against the students in the non-signaling condition ($F(1,21) = 4.20, p = .027, \eta^2 = .167$). The same applied for the PX where the PXI subscale *audiovisual appeal* was statistically significantly higher for the signaling condition ($F(1,21) = 3.62, p = .036, \eta^2 = .147$). Regarding cognitive load, the students in the signaling condition showed statistically significantly more GCL compared to non-signaling ($F(1,19) = 4.80, p = .021, partial\ \eta^2 = .202$), which means that participants invested more resources in the learning process.

The post-test results showed no statistically significant differences for the signaling and non-signaling condition ($F(1,44) = 2.59, p = .057, \eta^2 = .056$; see Figure 7). The data shows a descriptive trend for the students in the signaling condition (higher mean, lower *SD*) but it was not statistically significant ($F(1,21) = 1.61, p = .109, \eta^2 = .071$) either.

5.2 Qualitative Results

After the post-test and questionnaires, we conducted an individual semi-structured interview with each participant. All interviews were conducted in the participants' native language and were on average 7:37 minutes long (*SD* = 2:48). One author conducted a reflexive inductive thematic analysis similar to Braun and Clarke [11, 12] and analyzed the answers based on the taken interview notes, which were completed afterward using the audio recordings. A coding framework is not required for this approach, and it can be conducted by a single person [12]. We used the six phases of thematic analysis by Braun and Clarke [11] and numbered the emerged themes which provide an overview about *(1) Symmetric and Asymmetric Pair-Learning*, *(2) Usability in Pair-Learning VEs*, and *(3) Learning Unit about Forest Animals*. We included multiple participant statements for good quality [11, 12]. Quoted participants are numerated according to their role with *S* for *student* and *T* for *teacher*. If nothing about the participants' role is stated, this finding applies to both roles.

5.2.1 *Theme 1: Symmetric and Asymmetric Pair-Learning.*

The concept of pair-learning in VR or with the VR/tablet condition was perceived very positively by 44 out of 46 participants. 21 out of 35 participants who used an HMD answered that the VR environment was the decisive factor that made it fun for them, and 13 out of 46 explicitly said they had fun specifically because they were able to interact with their partner. Both was independent of the condition. For example, *S1* said, *"I found it ... funny to walk through the forest, and the combination of the VE and the interaction [with the partner] was fun"*. All students felt that their teacher did a good job at presenting the learning material. Especially five out of eleven participants (*T9, T12, T16, T17, T21*) who had used the tablet in their role as a teacher had fun due to the pair-learning activity, but eight (*T9, T10, T16, T18, T19, T21, T22, T23*) stated that they would have preferred the VR HMD instead of the tablet because it seemed more interesting and fun.

The participants had nearly no communication issues between the two roles. Only *T4* and *T15* stated that their partner did not always seem to listen to them due to the distractions caused by the VR environment. Nevertheless, there could be improvements regarding communication. Nine participants of the VR/VR condition (*S1, S3, S13, S17, T1, T3, T5, T15, T20*, n=24) felt uncomfortable by the location offset between the position of the VR avatar of their

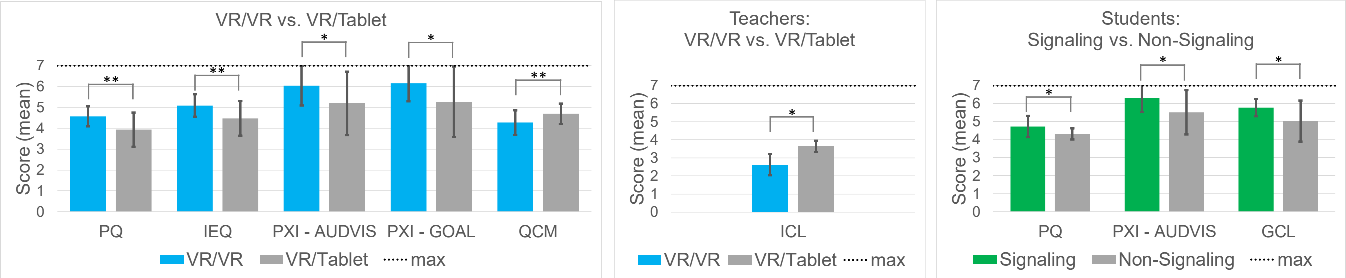


Figure 6: We found statistically significant results for presence (PQ), immersion (IEQ), PX (PXI subscales *audiovisual appeal* and *goals and rules*), and motivation (QCM) for the VR/VR vs. VR/tablet conditions independent of the roles. We further found statistically significant results for the teachers' cognitive load (ICL) in the VR/VR vs. VR/tablet conditions. In the signaling vs. non-signaling conditions, we found for students statistically significant results for presence (PQ), PX (PXI subscale *audiovisual appeal*), and cognitive load (GCL). Error bars show *SD*. Asterisk (*) indicates a statistically significant difference between conditions: p < .05 (*); p < .01 (**). Tables providing these results can be found in the appendix.

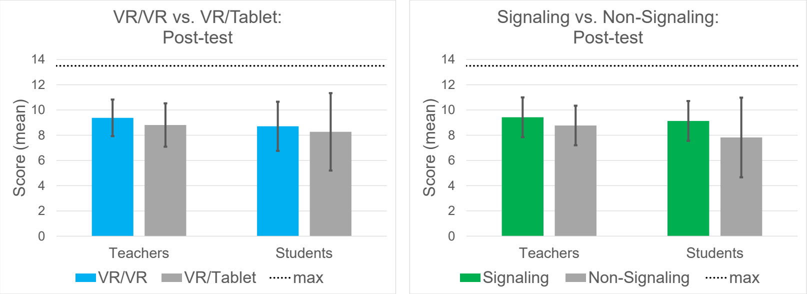


Figure 7: Our findings showed that the VR/VR and VR/tablet conditions performed equally well regarding the students' and the teachers' learning outcome. There were no statistically significant differences. Error bars show *SD*. A table providing these results can be found in the appendix.

partner and the voice direction of their partner in the co-located physical setup. *S13* said, *"I did sometimes teleport myself so that the [teacher's] voice came from the right direction because it was strange otherwise."*. No avatar for the teacher, on the other hand, as in the VR/tablet condition (n=22), also caused issues, and *S10, S11*, and *S18* stated that they would have liked visual feedback of their teacher. *S10* emphasized this by stating that, *"... it was sometimes very strange because you didn't see your partner ... but [still] heard him. And this [broke the immersion and] kept reminding you that you were in VR"*.

Three pairs (*S8, S15, S16, T10, T12, T19*) using the VR/tablet condition (n=22) mentioned moments of confusion regarding the teacher's view. Whenever teachers unlocked their camera view from the student to rotate the camera freely, they did not follow the student anymore, but the student was not automatically notified and aware of this. When the student then moved somewhere else, this could lead to misunderstandings, as they thought the teacher was aware of this movement. In such cases, workspace awareness, according to Gutwin and Greenberg [32], was lost.

5.2.2 *Theme 2: Usability in Pair-Learning VEs.*

Our pair-learning application was easy to use for 34 out of 46 participants, and only *S2* and *T1* stated difficulties with the controls. The other ten (*S1, S4, S7, S12, S19, T2, T3, T6, T7, T21*) managed to learn them after a short time. Especially the tutorial was helpful for all participants with low VR experience. For instance, *T2* mentioned, *"At the beginning I found it a little difficult, especially moving ... [and teleporting], but after I figured it out it was no longer a problem"*. Nevertheless, some participants forgot non-mandatory functionalities during the learning unit. *T5*, who used the VR application, said, *"I ... immediately forgot how to pick up [objects] and then I didn't pick up anything, but because I wasn't required to as a teacher, it was fine"*. Other teachers gave similar answers (*T1, T3, T6, T7, T8*, n=12) who used the VR application.

Most participants had no issues navigating in the VE. Especially the navigation arrow pointing towards the next way-point was perceived as very helpful by nine of the participants (*S1, S12, S15, S17, T2, T13, T14, T15, T20*, n=46). For example, *T13* said, *"You could always see the way-points very well, and if not you looked at the ground and then you had this big arrow which was helpful."*

The users of the tablet application (n=11) had no issues with the fewer functionalities compared to VR, but *T9, T10, T11*, and *T12* felt limited by them. *T9*, for instance, wished *"... to be able to do more because [I] ... was actually only reading the text aloud"*.

5.2.3 *Theme 3: Learning Unit about Forest Animals.*

30 out of 46 participants stated that they felt able to grasp the learning unit fully

and that the amount was just about right. No participant stated difficulties understanding our visual content of the high-fidelity 3D models. This shows that our learning unit was appropriate for our research topic.

The seven students (*S1, S2, S3, S9, S12, S14, S16*) who had difficulties following the learning unit named different reasons for this. For example, *S14* stated, *"No, [I was not able to follow the learning unit, because] I was distracted by all the objects [or rather] by the whole environment."*. A circumstance that could be targeted with signaling as it provides visual guidance, as we will discuss later. Nine teachers (*T1, T4, T7, T9, T10, T13, T15, T17, T18*) also answered that they could not completely grasp the learning unit, and *T9* explained, *"I could not really memorize the learning unit that well because I [was busy] ... with reading [the text] aloud and did, therefore ... not look [at the animals] in greater detail. Therefore, the learning effect was not so great for me"*. Further, some participants had difficulties memorizing the animal-specific terms.

6 DISCUSSION

To answer our RQ 1, *"Does symmetric collaboration compared to asymmetric collaboration in VR pair-learning environments impact presence, immersion, PX, motivation, cognitive load, and learning outcome of students and teachers?"*, and RQ 2, *"Does signaling in symmetric and asymmetric collaborative VR pair-learning environments impact presence, immersion, PX, motivation, cognitive load, and learning outcome of students and teachers?"*, we analyzed the quantitative results of our study conditions (see Figure 6 and Figure 7). We found statistically significant differences for presence, immersion, PX, motivation, and cognitive load, which we will discuss together with our qualitative results based on the thematic analysis. The effect sizes for all statistically significant results were, according to Cohen [17], moderate (>0.06) or high (>0.14).

6.1 Collaborative Learning

Students and teachers found the learning unit more present and immersive and had a better PX when the teacher was within the VE (see Figure 6). It was surprising how the peer's device influenced these measures even for the students, who used an immersive HMD in all conditions. These results show a benefit of the symmetric setup that was missing in the asymmetric VR/tablet setup. The lack of a teacher's avatar in the asymmetric setup meant that the students were less able to engage with the virtual world. A teacher who spoke to them, but was not in their world, pulled the students out of the illusion of the virtual world, reducing presence and immersion. It was also mentioned that it was easier if the teacher could show with its own avatar the student what to pay attention to in the virtual world. The teachers' limitation of only being able to express themselves verbally when using a tablet made it sometimes problematic to convey the learning unit effectively. This might explain why we could observe a statistically significantly higher cognitive load (ICL was significant) for the teachers using the tablet, meaning learning was more difficult for them.

Most participants' also explicitly stated that the VR/VR condition was preferred. For teachers, this was described as more fun, and almost all teachers in the tablet condition mentioned that they would have preferred to experience the application in VR, which is a justified wish looking at the better results of the PX in the VR/VR condition. Students also found it more motivating when the teacher was part of the VE. This is a surprising and very valuable result, as the motivation questionnaire (QCM) showed that participants in the asymmetric VR/tablet condition were statistically significantly more motivated before the learning unit based on the asymmetric setup, meaning that motivation has switched during the learning unit. We suspect that we had these results because in the symmetric VR/VR condition, the teachers were part of the virtual world, and thus a more personal teacher-student relationship, as well as easier communication, could be established (see Simeone et al. [72]). These findings support our hypothesis H1, *"Symmetric collaboration leads to higher presence, immersion, PX, and motivation and positively influences cognitive load compared to asymmetric collaboration."*.

Based on the stated related work, we already assumed that asymmetric systems could also be valuable learning systems. That is why we formulated our hypothesis H2, *"Asymmetric pair-learning systems can provide excellent and similar learning outcome for both learners."*, but it was surprising to see that our asymmetric prototype could even match the symmetric one regarding learning outcome (see Figure 7). As the effect size of the ANCOVA was small (see Cohen [17]), we assume that there unlikely will be a difference for larger participant groups and consider our result as solid. One explanation could be that the positive effect of pair-learning could compensate for the fact that the teacher was not in the VR environment and did not experience the true sized high-fidelity 3D models. Pair-learning had, therefore, maybe more influence on the learning outcome than VR, but as we had no pair-learning baseline condition, this is speculative at the moment but provides an interesting research direction for a future investigation.

Summing up, our developed asymmetric prototype can, therefore, very well be used for pair-learning, and H2 can be supported. Our findings also show the eligibility of commercial systems like the Oculus Quest [33] for pair-learning that support similar asymmetric implementations already out-of-the-box. As asymmetric systems only need one HMD, they are cheaper, require less conditions regarding the tracking space, and are, therefore, easier to set up than symmetric VR/VR systems. This means they are more likely to be used for pair-learning, especially at home and also at universities or schools. Summing up these findings, we conclude that asymmetric systems can be used for pair-learning, but symmetric systems should be preferred and used when possible.

6.2 Symmetric and Asymmetric Interaction

The tablet application allowed the teacher to stay close with the student and the possibility to see the world through the student's eyes as well as through a free rotatable camera. We assumed this as a simplification, as no own avatar had to be controlled, but we observed it differently. During the interviews, the teachers stated that they had difficulties orientating themselves in the virtual world, and they had to rely on the student's orientation. Managing their camera view while simultaneously following the student's actions and trying to present and explain the learning unit from the teaching sheets most likely overloaded them causing statistically significantly higher cognitive load (ICL was significant). Further, if the teacher stopped following the student's view, the student was not notified of this. According to the participants, this situation repeatedly led to miscommunications as both student and teacher lost workspace awareness. Summing up these findings, we suggest that attention has to be paid to facilitate interaction between student and teacher, especially in asymmetric teaching applications.

For pair-learning in general, symmetric and asymmetric, we recommend providing the student with visual feedback from the teacher, e.g., an avatar, additionally to the voice. Ideally, the perceived position of the voice should match the visual representation/avatar to prevent discomfort due to audio location mismatch, as stated by some participants. Automatic avatar regrouping to fit the physical with the virtual audio direction could be one solution as well as using telecommunication devices (e.g., headsets) even in co-located setups. Our guideline 4 takes this into account.

6.3 Signaling

The statistically significantly higher results for presence (PQ questionnaire; see Figure 6) indicate that signaling could focus the students' attention and involve them more (see the theory of Witmer and Singer [90]). The cognitive load questionnaire (GCL was statistically significant) showed that students invested more effort with signaling showing the importance of visual guiding. The statistically significantly higher scores of the PXI subscale *audiovisual appeal* further shows that signaling can improve graphical fidelity.

As these statistically significant differences were only measured between students, it indicates that signaling affected students more than teachers. This is likely explainable by the different tasks the roles were given. Students primarily focused on the animals and the environment itself and could, therefore, better spot the signaling. Teachers, on the other hand, had to focus on reading the teaching sheets.

The post-test results showed the descriptive trend that the students within the signaling condition had a higher mean score paired with a lower *SD* than those in the non-signaling condition (see Figure 7). This difference was not statistically significant but had a medium effect size indicating that it could become statistically significant with a larger number of participants. Furthermore, another form of signaling (e.g., motion [2, 29]) or a more prominent use of highlighting could provide higher effect sizes. It should be investigated in the future if such kinds of signaling would show not only descriptive but also statistically significant differences. A more noticeable form of signaling may also have an increased effect on teachers and make them look up from the teaching sheets. Therefore, future studies could vary different types of signaling (e.g., motion [2, 29], annotations [3]), or the complexity of the signals and even find out when visual signals in a VR environment become too complex to support the learner.

We conclude that the attention-directing function of signaling using highlighting can support learning applications in VR. It should be salient to draw attention, and it should be used both in symmetric and asymmetric setups. This confirms previous results of Albus et al. [3], Jamet et al. [38], and Mayer [50] also for VR pair-learning applications, which shows that the classic multimedia design principle signaling can be transferred to this learning setup. Therefore, we can support our hypothesis H3, *"Signaling in VEs independently of the setup enhances presence, immersion, PX, motivation, and learning outcome and positively influences cognitive load compared to non-signaling."* for presence, PX, and cognitive load for our participants in the student role.

7 GUIDELINES FOR VR PAIR-LEARNING APPLICATIONS

Based on our qualitative and quantitative findings, we define the following guidelines for VR pair-learning applications. The guidelines should stand for themselves, why we repeat deliberately some of our previously shown findings in the explanations of the guidelines. They are formulated for VR pair-learning but could be an inspiration for other collaborative symmetric and asymmetric VR systems as well.

1. *Symmetric VR pair-learning applications targeting two students should be preferred over asymmetric systems.*
   Our findings show that presence, immersion, and PX for all participants and cognitive load for the teachers were statistically significantly better in the symmetric setup. Participants further stated to have higher motivation and fun, and a symmetric setup provided easier communication and better out-of-the-box workspace awareness. Based on these supportive factors for learning (see [8, 13, 48, 49, 51, 61]), we advise using symmetric setups for VR pair-learning applications. Nevertheless, suppose a symmetric setup is impossible due to constraints such as a lack of HMDs, incompatible tracking space, or safety precautions for the complex setup, especially when used at home. In this case, the following guidelines should be adapted for an asymmetric setup, as our findings show that they can also provide valuable pair-learning regardless of lower values for presence, immersion, PX, and cognitive load.
2. *VR pair-learning applications should always allow both users to move and interact freely in the VE.*
   Our findings show that in the asymmetric system, the teachers' detachable spectator view made communication difficult for both participants as it was not obvious for them what exactly their peer was viewing at the moment. This resulted in a loss of workspace awareness. Therefore, we advise a similar implementation than in the symmetric system and allow all users to move around in the VR freely and to provide a visualization of the asymmetric (e.g., tablet) users' field of view for the VR users, which is not provided at the moment by state-of-the-art spectator views. This should be done by an avatar, as it has further benefits as well (see guideline 3), or if not possible (e.g., due to the learning scenario where a teacher has to supervise every action of the student) by other means such as, e.g., a light beam (see Peter et al. [58]), or an overlay display (see Kumaravel et al. [84]). Moving in the VE should be supported by a navigation feature that guides the direction to the next point of interest. Our findings show that a navigation arrow could be an appropriate implementation.
3. *Symmetric as well as asymmetric VR pair-learning applications should always provide an avatar for each user.*
   Our findings show that communication between participants was easier, more immersive, and more personal when

an avatar as a point of reference was available. We advise using avatars in symmetric as well as in asymmetric VR pair-learning applications (see Piumsomboon et al. [62]) and not just a spectator view (see Kumaravel et al. [84]). The avatars should have similar user representations, capabilities, and permissions to enhance communication and enable a more personal relationship compared to only voice chat (see Olin et al. [57]). Capabilities we could think of that provide possibly beneficial additional communication cues could be to show the gaze direction of a user (see guideline 2), gestures, as well as pointing with virtual hands, or a laser-pointer metaphor. An asymmetric user without an HMD should be able to perform the same actions and movements with the own avatar as the VR counterpart to enable as symmetric communication as possible.

4. *Co-located VR pair-learning applications should provide the possibility to correct the partner's voice and avatar's position offset.*
In co-located VR pair-learning applications, it can happen (e.g., due to locomotion techniques such as teleportation) that an offset between the partner's voice and the partner's avatar position emerges, which was stated as irritating by our participants. Therefore, we advise implementing a functionality so that users can correct this position offset easily. Different possibilities could be imagined to fix this, from just rotating the gaze directions of participants, suggesting a new position in the VE the user can move to, or to automatically regroup all avatars in the VE to match their virtual positions with the users' co-located positions. Another alternative to fix the voice offset is the mandatory use of telecommunication devices such as headsets even in co-located setups as it is common for remote setups.

5. *The information shown in VR pair-learning applications should guide the user to explore the VE and not solely catch the user's attention.*
We provided our participants in the role of a teacher with teaching sheets containing the learning unit's information. It described the forest animals shown in the VE. Whereas the students without the teaching sheets were concentrated on the virtual forest animal models, teachers stated that they were focusing on the presentation of the teaching sheets as well as the student's activity which made following the learning content and investigating the 3D animal models difficult. Therefore, we advise presenting information in a way that does not distract the user's attention from the explanatory VE and the presented 3D models. One solution could be to stationary locate information at the 3D models, for example, with annotations (see Albus et al. [3]). Therefore, users are guided to the 3D model when reading information.

6. *VR pair-learning applications should use salient signaling to focus the attention of users.*
Our findings show that presence, PX, and cognitive load were statistically significantly better with signaling for students. Our participants in the teacher role stated that signaling was not that salient for them due to other distractions such as the teaching sheets (see guideline 5). We advise to use signaling in VR pair-learning applications, but it should be salient for all users to guide their attention. Solutions could be to hide distractions when signaling is used or to use more prominent signaling methods such as motion (see [2, 29]) than just statically visualizing an object's outline.

8 LIMITATIONS

Although our sample size of N=46 participants provides limitations regarding statistical evaluations, we could still detect significant differences between the groups, most of all those with medium to high effect sizes (see Cohen [17]). For those with lower effect sizes, the results may become significant with a larger number of participants. Nevertheless, the found statistically significant effects were sufficient to answer our RQs. We had one technical limitation. For the tablet, we had to restrict the graphic settings and resolution due to performance reasons. Despite maybe not influencing the learning outcome, it is conceivable that it contributed to the lower scores of presence, immersion, and PX of the VR/tablet condition. Supposedly this effect was low or not present at all as the participants did not state issues with the tablet's performance during the interview.

9 FUTURE WORK

Our study setup was based on a student mimicking a teacher. When speaking about a well-versed teacher, our recommendations could change. The drawbacks that we found for the teachers using a tablet, such as lower motivation or higher cognitive load, are not relevant and might not apply to professional teachers. Therefore, we think that the VR/tablet setup should be tested with professional teachers in the future.

We further think that a VR/tablet setup could be a beneficial solution that scales for classroom-based collaborative learning, where one teacher supervises multiple students. While our results show that explaining and teaching was perceived better when the teacher was joining the students in VR, it should be investigated if the tablet could be used to supervise students when they conduct exercises in VR with a top-down view or multi-view tools best suited for a traditional screen as provided by a tablet. It would furthermore allow the teacher to do other tasks in the real world while supervising. Teachers outside VR have further benefits regarding health and safety, as falls or injuries of the users could be avoided by a supervising teacher, especially when multiple children are using VR HMDs simultaneously, e.g., in class. The possibility to seamlessly switch back from the tablet into VR with an HMD, for example, with approaches such as Slice of Light [87], would allow the teacher to support students when needed.

As we found no difference in learning outcome for our symmetric and asymmetric setup, this could be an indicator that pair-learning has more influence on the learning outcome than using VR or a tablet. With our current study, which had no baseline pair-learning condition, this assumption is speculative, but it could show potential future research directions. For example, this could be investigated in the future by, e.g., comparing a single-user VR learning environment with a traditional paper-based pair-learning setup.

Our findings further suggest to investigate the effects of a remote setup on immersion, presence, PX, communication, and learning outcome, which could prevent the perceived audio location offset between the avatar and the participant's position in the physical room, as telecommunication devices such as headset are mandatory.

Another research direction for pair-learning is AR both HMD as well as tablet-based, which provides further possibilities to consider and include the real environment as a context for the learning unit. AR education is still dominated by single-user systems [36]. It should be investigated how our guidelines would fit into a collaborative AR setup, for example to the ones presented in the survey of Sereno et al. [71]. Do our findings regarding symmetric and asymmetric collaborative learning also apply to AR? Based on our guidelines, we could assume that symmetric AR would also work best due to benefits regarding the possible symmetric and equal communication and interaction. However, we also assume that asymmetric AR would work as well as an effective learning system. Further research directions could be if avatars are still necessary and what are the differences between co-located and remote setups? If and what kind of gaze visualization is necessary? How should attention guiding and signaling be performed? How is AR pair-learning compared to VR pair-learning regarding learning outcome, player experience, or cognitive load? In which learning context do AR and VR work best, and how can they be combined? These are just some possible research questions. Piumsomboon et al. [62] already investigated how an MR remote collaboration could look like and how this affects avatars. Related work was also conducted by Kumaravel et al. [84] for asymmetric MR collaboration and Wells and Houben [88] for handheld group collaboration. Nevertheless, applying their results to pair-learning in AR as we did with the state-of-the-art for VR is a topic for future work.

10 CONCLUSION

To investigate the design of VR environments for pair-learning, we compared a symmetric and asymmetric setup based on HMDs and a tablet in a study with 46 participants regarding presence, immersion, PX, motivation, cognitive load, and learning outcome. We further investigated the impact of signaling on these measures. Our findings show that the symmetric VR condition had statistically significant benefits for presence, immersion, PX, and cognitive load. It further enhances motivation, fun, and the participants' communication. We further found statistically significant results for signaling regarding presence, PX, and cognitive load, showing benefits for learning. The symmetric and asymmetric conditions performed equally well regarding learning outcome, which shows that both setups are valuable learning tools. Summing up the benefits, we recommend to use symmetric systems for pair-learning and to use signaling to highlight important content. Based on these findings, we defined six guidelines on how to build VR pair-learning applications and discuss future research directions for asymmetric VR pair-learning applications as they are also relevant as cheaper, smaller, and state-of-the-art systems. Our findings are a profound base to help researchers, educators, and software developers to create valuable VR learning software.

ACKNOWLEDGMENTS

This work was conducted within the project ”AuCity 2 - AR in University Teaching Using the Example of Civil Engineering“ and within the project "AuCity 3 - Collaborative and Adaptive MR in Higher Education Teaching Using Civil Engineering as an Example“ both funded by the Federal Ministry of Education and Research (BMBF).

REFERENCES

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020), 102370. https://doi.org/10.1016/j.ijhcs.2019.102370

[2] Richard A. Abrams and Shawn E. Christ. 2003. Motion Onset Captures Attention. *Psychological Science* 14, 5 (2003), 427–432. https://doi.org/10.1111/1467-9280.01458 arXiv:https://doi.org/10.1111/1467-9280.01458 PMID: 12930472.

[3] Patrick Albus, Andrea Vogt, and Tina Seufert. 2021. Signaling in virtual reality influences learning outcome and cognitive load. *Computers & Education* 166 (2021), 104154. https://doi.org/10.1016/j.compedu.2021.104154

[4] Devon Allcoat and Adrian von Mühlenen. 2018. Learning in virtual reality: Effects on performance, emotion and engagement. *Research in Learning Technology* 26 (Nov. 2018). https://doi.org/10.25304/rlt.v26.2140

[5] Arthur Aron, Elaine N Aron, and Danny Smollan. 1992. Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology* 63, 4 (1992), 596.

[6] Sarune Baceviciute, Aske Mottelson, Thomas Terkildsen, and Guido Makransky. 2020. Investigating Representation of Text and Audio in Educational VR Using Learning Outcomes and EEG. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376872

[7] A Baddeley. 1992. Working memory. *Science* 255, 5044 (1992), 556–559. https://doi.org/10.1126/science.1736359 arXiv:https://science.sciencemag.org/content/255/5044/556.full.pdf

[8] Christopher Berns, Grace Chin, Joel Savitz, Jason Kiesling, and Fred Martin. 2019. MYR: A Web-Based Platform for Teaching Coding Using VR. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (Minneapolis, MN, USA) *(SIGCSE '19)*. Association for Computing Machinery, New York, NY, USA, 77–83. https://doi.org/10.1145/3287324.3287482

[9] Benjamin S Bloom et al. 1956. Taxonomy of educational objectives. Vol. 1: Cognitive domain. *New York: McKay* 20 (1956), 24.

[10] Evren Bozgeyikli, Andrew Raij, Srinivas Katkoori, and Rajiv Dubey. 2016. Point & Teleport Locomotion Technique for Virtual Reality. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) *(CHI PLAY '16)*. Association for Computing Machinery, New York, NY, USA, 205–216. https://doi.org/10.1145/2967934.2968105

[11] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[12] Virginia Braun and Victoria Clarke. 2020. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology* 0, 0 (2020), 1–25. https://doi.org/10.1080/14780887.2020.1769238

[13] Raffaello Brondi, Leila Alem, Giovanni Avveduto, Claudia Faita, Marcello Carrozzino, Franco Tecchia, and Massimo Bergamasco. 2015. Evaluating the Impact of Highly Immersive Technologies and Natural Interaction on Player Engagement and Flow Experience in Games. In *Entertainment Computing - ICEC 2015*, Konstantinos Chorianopoulos, Monica Divitini, Jannicke Baalsrud Hauge, Letizia Jaccheri, and Rainer Malaka (Eds.). Springer International Publishing, Cham, 169–181.

[14] Paul Chandler and John Sweller. 1991. Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction* 8, 4 (1991), 293–332. https://doi.org/10.1207/s1532690xci0804_2

[15] Henry Chen, Austin S. Lee, Mark Swift, and John C. Tang. 2015. 3D Collaboration Method over HoloLens™ and Skype™ End Points. In *Proceedings of the 3rd International Workshop on Immersive Media Experiences* (Brisbane, Australia) *(ImmersiveME '15)*. Association for Computing Machinery, New York, NY, USA, 27–30. https://doi.org/10.1145/2814347.2814350

[16] V. Chheang, P. Saalfeld, T. Huber, F. Huettl, W. Kneist, B. Preim, and C. Hansen. 2019. Collaborative Virtual Reality for Laparoscopic Liver Surgery Training. In *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. 1–17. https://doi.org/10.1109/AIVR46125.2019.00011

[17] J Cohen. 1988. Statistical power analysis for the behavioral sciences. *Hoboken: Taylor and Francis* (1988).

[18] Mihaly Csikszentmihalyi. 1975. Beyond boredom and anxiety: Experiencing flow in work and play. *San Francisco/Washington/London* (1975).

[19] A. Dengel and J. Mägdefrau. 2020. Immersive Learning Predicted: Presence, Prior Knowledge, and School Performance Influence Learning Outcomes in Immersive Educational Virtual Environments. In *2020 6th International Conference of the Immersive Learning Research Network (iLRN)*. 163–170. https://doi.org/10.23919/

iLRN47897.2020.9155084

[20] Arindam Dey, Thammathip Piumsomboon, Youngho Lee, and Mark Billinghurst. 2017. Effects of Sharing Physiological States of Players in a Collaborative Virtual Reality Gameplay. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4045–4056. https://doi.org/10.1145/3025453.3026028

[21] Pierre Dragicevic. 2016. *Fair Statistical Communication in HCI*. Springer International Publishing, Cham, 291–330. https://doi.org/10.1007/978-3-319-26633-6_13

[22] Tobias Drey, Fabian Fischbach, Pascal Jansen, Julian Frommel, Michael Rietzler, and Enrico Rukzio. 2021. To Be or Not to Be Stuck, or Is It a Continuum? A Systematic Literature Review on the Concept of Being Stuck in Games. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 229 (oct 2021), 35 pages. https://doi.org/10.1145/3474656

[23] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio. 2020. *VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376628

[24] Tobias Drey, Pascal Jansen, Fabian Fischbach, Julian Frommel, and Enrico Rukzio. 2020. Towards Progress Assessment for Adaptive Hints in Educational Virtual Reality Games. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3382789

[25] Tobias Drey, Michael Rietzler, and Enrico Rukzio. 2021. Questionnaires and Qualitative Feedback Methods to Measure User Experience in Mixed Reality. arXiv:2104.06221 [cs.HC]

[26] Thierry Duval and Cedric Fleury. 2009. An Asymmetric 2D Pointer/3D Ray for 3D Interaction within Collaborative Virtual Environments. In *Proceedings of the 14th International Conference on 3D Web Technology* (Darmstadt, Germany) *(Web3D '09)*. Association for Computing Machinery, New York, NY, USA, 33–41. https://doi.org/10.1145/1559764.1559769

[27] R Ekstrome, J French, H Harman, and D Dermen. 1976. Manual for kit of factor referenced cognitive tests. *Princeton, New Jersey: Educational Testing Service* 1976 (1976).

[28] Andy Field. 2013. *Discovering statistics using IBM SPSS statistics*. sage.

[29] Steven L Franconeri and Daniel J Simons. 2005. The dynamic events that capture visual attention: A reply to Abrams and Christ (2005). *Perception & psychophysics* 67, 6 (2005), 962–966.

[30] J. G. Grandi, H. G. Debarba, and A. Maciel. 2019. Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 127–135. https://doi.org/10.1109/VR.2019.8798080

[31] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4021–4033. https://doi.org/10.1145/3025453.3025683

[32] Carl Gutwin and Saul Greenberg. 2002. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)* 11, 3 (2002), 411–446.

[33] David Haeney. 2021. *Oculus Quest Can Now Cast To PC, App Getting 'Spectator Mode' Next Year.* https://uploadvr.com/quest-pc-casting-spectator-mode/ (Accessed on 03/30/2021).

[34] R. Holm, E. Stauder, R. Wagner, M. Priglinger, and J. Volkert. 2002. A combined immersive and desktop authoring tool for virtual environments. In *Proceedings IEEE Virtual Reality 2002*. 93–100. https://doi.org/10.1109/VR.2002.996511

[35] Hikaru Ibayashi, Yuta Sugiura, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Masaaki Mochimaru, and Takeo Igarashi. 2015. Dollhouse VR: A Multi-View, Multi-User Collaborative Design Workspace with VR Technology. In *SIGGRAPH Asia 2015 Emerging Technologies* (Kobe, Japan) *(SA '15)*. Association for Computing Machinery, New York, NY, USA, Article 8, 2 pages. https://doi.org/10.1145/2818466.2818480

[36] María-Blanca Ibáñez and Carlos Delgado-Kloos. 2018. Augmented reality for STEM learning: A systematic review. *Computers & Education* 123 (2018), 109–123. https://doi.org/10.1016/j.compedu.2018.05.002

[37] Randolph L. Jackson, Wayne Taylor, and William Winn. 1999. Peer Collaboration and Virtual Environments: A Preliminary Investigation of Multi-Participant Virtual Reality Applied in Science Education. In *Proceedings of the 1999 ACM Symposium on Applied Computing* (San Antonio, Texas, USA) *(SAC '99)*. Association for Computing Machinery, New York, NY, USA, 121–125. https://doi.org/10.1145/298151.298219

[38] Eric Jamet, Monica Gavota, and Christophe Quaireau. 2008. Attention guiding in multimedia learning. *Learning and Instruction* 18, 2 (2008), 135–145. https://doi.org/10.1016/j.learninstruc.2007.01.011

[39] Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies* 66, 9 (2008), 641–661. https://doi.org/10.1016/j.ijhcs.2008.04.004

[40] Roger T. Johnson, David W. Johnson, and Mary Beth Stanne. 1986. Comparison Of Computer-Assisted Cooperative, Competitive, And Individualistic Learning. *American Educational Research Journal* 23, 3 (1986), 382–392. https://doi.org/10.3102/00028312023003382 arXiv:https://doi.org/10.3102/00028312023003382

[41] Melina Klepsch, Florian Schmitz, and Tina Seufert. 2017. Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology* 8 (2017), 1997. https://doi.org/10.3389/fpsyg.2017.01997

[42] Eric Krokos, Catherine Plaisant, and Amitabh Varshney. 2019. Virtual memory palaces: immersion aids recall. *Virtual Reality* 23, 1 (2019), 1–15.

[43] Marc Erich Latoschik, Daniel Roth, Dominik Gall, Jascha Achenbach, Thomas Waltemate, and Mario Botsch. 2017. The Effect of Avatar Realism in Immersive Social Virtual Realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg, Sweden) *(VRST '17)*. Association for Computing Machinery, New York, NY, USA, Article 39, 10 pages. https://doi.org/10.1145/3139131.3139156

[44] Jiwon Lee, Mingyu Kim, and Jinmo Kim. 2020. RoleVR: Multi-experience in immersive virtual reality between co-located HMD and non-HMD users. *Multimedia Tools and Applications* 79, 1 (2020), 979–1005.

[45] Jie Li, Yiping Kong, Thomas Röggla, Francesca De Simone, Swamy Ananthanarayan, Huib de Ridder, Abdallah El Ali, and Pablo Cesar. 2019. Measuring and Understanding Photo Sharing Experiences in Social Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300897

[46] Tzu-Chien Liu, Yi-Chun Lin, Tzu-Ning Wang, Shih-Ching Yeh, and Slava Kalyuga. 2021. Studying the effect of redundancy in a virtual reality classroom. *Educational Technology Research and Development* 69, 2 (2021), 1183–1200.

[47] J. Lugrin, M. Ertl, P. Krop, R. Klüpfel, S. Stierstorfer, B. Weisz, M. Rück, J. Schmitt, N. Schmidt, and M. E. Latoschik. 2018. Any "Body" There? Avatar Visibility Effects in a Virtual Reality Game. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 17–24. https://doi.org/10.1109/VR.2018.8446229

[48] Jason Madar, Adina Goldberg, and Kim Lam. 2018. "Hour of Code" With Virtual Reality. In *Proceedings of the 23rd Western Canadian Conference on Computing Education* (Victoria, BC, Canada) *(WCCCE '18)*. Association for Computing Machinery, New York, NY, USA, Article 16, 7 pages. https://doi.org/10.1145/3209635.3209636

[49] Fabrizia Mantovani. 2001. 12 VR learning: Potential and challenges for the use of 3d environments in education and training. *Towards cyberpsychology: mind, cognition, and society in the Internet age* 2, 207 (2001).

[50] Richard Mayer. 2005. *The Cambridge handbook of multimedia learning*. Cambridge university press.

[51] Richard E. Mayer and Logan Fiorella. 2014. *Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles* (2 ed.). Cambridge University Press, 279–315. https://doi.org/10.1017/CBO9781139547369.015

[52] Fares Moustafa and Anthony Steed. 2018. A Longitudinal Study of Small Group Interaction in Social Virtual Reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 22, 10 pages. https://doi.org/10.1145/3281505.3281527

[53] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. 2017. CollaVR: Collaborative In-Headset Review for VR Video. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 267–277. https://doi.org/10.1145/3126594.3126659

[54] K. Oberauer, H.-M. Süß, R. Schulze, O. Wilhelm, and W.W. Wittmann. 2000. Working memory capacity — facets of a cognitive ability construct. *Personality and Individual Differences* 29, 6 (2000), 1017–1045. https://doi.org/10.1016/S0191-8869(99)00251-2

[55] Sebastian Oberdörfer, David Heidrich, and Marc Erich Latoschik. 2019. Usability of Gamified Knowledge Learning in VR and Desktop-3D. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300405

[56] Ohan Oda, Carmine Elvezio, Mengu Sukan, Steven Feiner, and Barbara Tversky. 2015. Virtual Replicas for Remote Assistance in Virtual and Augmented Reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 405–415. https://doi.org/10.1145/2807442.2807497

[57] Patrick Aggergaard Olin, Ahmad Mohammad Issa, Tiare Feuchtner, and Kaj Grønbæk. 2020. Designing for Heterogeneous Cross-Device Collaboration and Social Interaction in Virtual Reality. In *32nd Australian Conference on Human-Computer Interaction* (Sydney, NSW, Australia) *(OzCHI '20)*. Association for Computing Machinery, New York, NY, USA, 112–127. https://doi.org/10.1145/3441000.3441070

[58] Mark Peter, Robin Horst, and Ralf Dörner. 2018. Vr-guide: A specific user role for asymmetric virtual reality setups in distributed virtual reality applications. In

Proceedings of the 10th Workshop Virtual and Augmented Reality of the GI Group VR/AR.

[59] D. N. E. Phon, M. B. Ali, and N. D. A. Halim. 2014. Collaborative Augmented Reality in Education: A Review. In *2014 International Conference on Teaching and Learning in Computing and Engineering*. 78–83. https://doi.org/10.1109/LaTiCE.2014.23

[60] Marcio S Pinho, Doug A Bowman, and Carla M Dal Sasso Freitas. 2008. Cooperative object manipulation in collaborative virtual environments. *Journal of the Brazilian Computer Society* 14, 2 (2008), 53–67.

[61] Johanna Pirker, Andreas Dengel, Michael Holly, and Saeed Safikhani. 2020. Virtual Reality in Computer Science Education: A Systematic Review. In *26th ACM Symposium on Virtual Reality Software and Technology* (Virtual Event, Canada) *(VRST '20)*. Association for Computing Machinery, New York, NY, USA, Article 8, 8 pages. https://doi.org/10.1145/3385956.3418947

[62] Thammathip Piumsomboon, Gun A. Lee, Jonathon D. Hart, Barrett Ens, Robert W. Lindeman, Bruce H. Thomas, and Mark Billinghurst. 2018. Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173620

[63] Thammathip Piumsomboon, Gun A. Lee, Andrew Irlitti, Barrett Ens, Bruce H. Thomas, and Mark Billinghurst. 2019. On the Shoulder of the Giant: A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3290605.3300458

[64] Veljko Potkonjak, Michael Gardner, Victor Callaghan, Pasi Mattila, Christian Guetl, Vladimir M. Petrović, and Kosta Jovanović. 2016. Virtual laboratories for education in science, technology, and engineering: A review. *Computers & Education* 95 (2016), 309–327. https://doi.org/10.1016/j.compedu.2016.02.002

[65] Jaziar Radianti, Tim A. Majchrzak, Jennifer Fromm, and Isabell Wohlgenannt. 2020. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education* 147 (2020), 103778. https://doi.org/10.1016/j.compedu.2019.103778

[66] Falko Rheinberg, Regina Vollmeyer, and Bruce D Burns. 2001. FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern-und Leistungssituationen (Langversion, 2001). *Diagnostica* 2 (2001), 57–66.

[67] Danny Schott, Patrick Saalfeld, Gerd Schmidt, Fabian Joeres, Christian Boedecker, Florentine Huettl, Hauke Lang, Tobias Huber, Bernhard Preim, and Christian Hansen. 2021. A VR/AR Environment for Multi-User Liver Anatomy Education. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 296–305. https://doi.org/10.1109/VR50410.2021.00052

[68] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (06 2001), 266–281. https://doi.org/10.1162/105474601300343603 arXiv:https://direct.mit.edu/pvar/article-pdf/10/3/266/1623697/105474601300343603.pdf

[69] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. *Using Presence Questionnaires in Virtual Reality*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300590

[70] Sven Seele, Sebastian Misztal, Helmut Buhler, Rainer Herpers, and Jonas Schild. 2017. Here's Looking At You Anyway! How Important is Realistic Gaze Behavior in Co-Located Social Virtual Reality Games?. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17)*. Association for Computing Machinery, New York, NY, USA, 531–540. https://doi.org/10.1145/3116595.3116619

[71] Mickael Sereno, Xiyao Wang, Lonni Besancon, Michael J Mcguffin, and Tobias Isenberg. 2020. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics* (2020), 1–1. https://doi.org/10.1109/TVCG.2020.3032761

[72] Adalberto L Simeone, Marco Speicher, Andreea Molnar, Adriana Wilde, and Florian Daiber. 2019. LIVE: The Human Role in Learning in Immersive Virtual Environments. In *Symposium on Spatial User Interaction* (New Orleans, LA, USA) *(SUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 5, 11 pages. https://doi.org/10.1145/3357251.3357590

[73] Richard Skarbez, Frederick P. Brooks, Jr., and Mary C. Whitton. 2017. A Survey of Presence and Related Concepts. *ACM Comput. Surv.* 50, 6, Article 96 (nov 2017), 39 pages. https://doi.org/10.1145/3134301

[74] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557.

[75] Robert E. Slavin. 1980. Cooperative Learning. *Review of Educational Research* 50, 2 (1980), 315–342. https://doi.org/10.3102/00346543050002315 arXiv:https://doi.org/10.3102/00346543050002315

[76] Hyo-Jeong So and Thomas A. Brush. 2008. Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors. *Computers & Education* 51, 1 (2008), 318–336. https://doi.org/10.1016/j.compedu.2007.05.009

[77] Vinicius Souza, Anderson Maciel, Luciana Nedel, and Regis Kopper. 2021. Measuring Presence in Virtual Environments: A Survey. *ACM Comput. Surv.* 54, 8, Article 163 (oct 2021), 37 pages. https://doi.org/10.1145/3466817

[78] Yuta Sugiura, Hikaru Ibayashi, Toby Chong, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Takashi Shinmura, Masaaki Mochimaru, and Takeo Igarashi. 2018. An Asymmetric Collaborative System for Architectural-Scale Space Design. In *Proceedings of the 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry* (Tokyo, Japan) *(VRCAI '18)*. Association for Computing Machinery, New York, NY, USA, Article 21, 6 pages. https://doi.org/10.1145/3284398.3284416

[79] Unity Technologies. 2021. *Unity Asset Store: Dark Rift Networking 2.* https://assetstore.unity.com/packages/tools/network/darkrift-networking-2-95309 (Accessed on 04/02/2020).

[80] Unity Technologies. 2021. *Unity Asset Store: Forest Animals Pack.* https://assetstore.unity.com/packages/3d/characters/animals/forest-animals-pack-4990 (Accessed on 04/02/2020).

[81] Unity Technologies. 2021. *Unity Asset Store: Forest Environment - Dynamic Nature.* https://assetstore.unity.com/packages/3d/vegetation/forest-environment-dynamic-nature-150668 (Accessed on 04/02/2021).

[82] Unity Technologies. 2021. *Unity Homepage.* https://unity.com (Accessed on 04/02/2020).

[83] Meredith Thompson, Hidai Olivas-Holguin, Annie Wang, Jing Fan, Katharine Pan, David Vargas, and Joanna Gerr. 2018. Rules, roles, and resources: Strategies to promote collaboration in virtual reality contexts. In *Workshop Position Paper for CHI 2018.*

[84] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. 2020. *TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators.* Association for Computing Machinery, New York, NY, USA, 182–195. https://doi.org/10.1145/3379337.3415827

[85] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. 2000. Using Presence Questionnaires in Reality. *Presence* 9, 5 (2000), 497–503. https://doi.org/10.1162/105474600566989

[86] Cigdem Uz-Bilgin, Meredith Thompson, and Melat Anteneh. 2020. Exploring How Role and Background Influence Through Analysis of Spatial Dialogue in Collaborative Problem-Solving Games. *Journal of Science Education and Technology* 29, 6 (2020), 813–826.

[87] Chiu-Hsuan Wang, Chia-En Tsai, Seraphina Yong, and Liwei Chan. 2020. Slice of Light: Transparent and Integrative Transition Among Realities in a Multi-HMD-User Environment. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 805–817. https://doi.org/10.1145/3379337.3415868

[88] Thomas Wells and Steven Houben. 2020. CollabAR – Investigating the Mediating Role of Mobile AR Interfaces on Co-Located Group Collaboration. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376541

[89] Bob G. Witmer, Christian J. Jerome, and Michael J. Singer. 2005. The Factor Structure of the Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* 14, 3 (06 2005), 298–312. https://doi.org/10.1162/105474605323384654 arXiv:https://direct.mit.edu/pvar/article-pdf/14/3/298/1624394/105474605323384654.pdf

[90] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* 7, 3 (06 1998), 225–240. https://doi.org/10.1162/105474698565686 arXiv:https://direct.mit.edu/pvar/article-pdf/7/3/225/1836425/105474698565686.pdf

[91] Haijun Xia, Sebastian Herscher, Ken Perlin, and Daniel Wigdor. 2018. Spacetime: Enabling Fluid Individual and Collaborative Editing in Virtual Reality. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 853–866. https://doi.org/10.1145/3242587.3242597

[92] Cagri Hakan Zaman, Asiya Yakhina, and Federico Casalegno. 2015. NRoom: An Immersive Virtual Environment for Collaborative Spatial Design. In *Proceedings of the International HCI and UX Conference in Indonesia* (Bandung, Indonesia) *(CHIuXiD '15)*. Association for Computing Machinery, New York, NY, USA, 10–17. https://doi.org/10.1145/2742032.2742034

APPENDIX

Students

	VR/VR (*N* = 12)	VR/Tablet (*N* = 11)	Signaling (*N* = 12)	Non-Signaling (*N* = 11)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Dependent Variables				
PQ (max=7)	4.78 (.43)	4.28 (.48)	4.73 (.59)	4.32 (.31)
IEQ (max=7)	5.10 (.53)	4.78 (.79)	4.95 (.73)	4.94 (.63)
PXI - EOC (max=7)	6.33 (1.09)	5.58 (1.72)	6.19 (1.30)	5.73 (1.62)
PXI - AUDVIS (max=7)	6.14 (.89)	5.73(1.28)	6.33 (.80)	5.52 (1.23)
PXI - GOAL (max=7)	6.42 (.53)	5.52 (1.52)	6.14 (1.08)	5.82 (1.32)
QCM (max=7)	4.31 (.71)	4.70 (.47)	4.65 (.61)	4.33 (.63)
ICL (max=7)	3.21 (1.29)	4.05 (1.17)	3.67 (1.28)	3.55 (1.35)
ECL (max=7)	2.58 (1.51)	3.00 (1.48)	2.31 (1.03)	3.30 (1.75)
GCL (max=7)	5.28 (1.18)	5.58 (.54)	5.78 (.48)	5.03 (1.14)
Post-test (max=13.5)	8.71 (1.95)	8.27 (3.07)	9.13 (1.58)	7.82 (3.16)
Control Variables				
Gender (female) N (%)	3 (25.00)	6 (54.50)	5 (41.70)	4 (36.40)
Age	24.75 (2.86)	23.27 (2.45)	24.00 (3.0)	24.09 (2.07)
Duration (mm:ss)	15:00 (3:53)	12:50 (1:48)	14:22 (4:04)	13:31 (3:34)
Pre-test (max=18)	10.79 (1.88)	10.46 (2.44)	11.42 (1,29)	9.78 (2.55
Spatial Ability (%)	.77 (.12)	.62 (.15)	.72 (.14)	.68 (.16)
Working Memory (max=9)	3.33 (.99)	3.70 (1.25)	3.75 (1.14)	3.2 (1.03)
IOS (max=7)	4.83 (1.90)	5.50 (2.01)	5.17 (1.64)	5.10 (2.33)

Teachers

	VR/VR (*N* = 12)	VR/Tablet (*N* = 11)	Signaling (*N* = 12)	Non-Signaling (*N* = 11)
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Dependent Variables				
PQ (max=7)	4.34 (.44)	3.59 (.98)	3.92 (1.09)	4.01 (.51)
IEQ (max=7)	5.08 (.59)	4.16 (.78)	4.34 (.97)	4.90 (.56)
PXI - EOC (max=7)	4.67 (1.06)	5.24 (1.71)	5.36 (1.04)	4.55 (1.68)
PXI - AUDVIS (max=7)	5.94 (1.04)	4.67 (1.62)	5.58 (1.10)	5.03 (1.80)
PXI - GOAL (max=7)	5.88 (1.08)	5.03 (1.89)	5.55 (1.37)	5.36 (1.81)
QCM (max=7)	4.17 (.46)	4.69 (.53)	4.56 (.59)	4.26 (.48)
ICL (max=7)	2.63 (.64)	3.64 (1.40)	3.33 (1.39)	2.86 (.87)
ECL (max=7)	2.47 (.87)	2.21 (.78)	2.41 (.93)	2.27 (.71)
GCL (max=7)	4.86 (1.18)	4.94 (1.17)	4.83 (1.18)	4.97 (1.16)
Post-test (max=13.5)	9.38 (1.45)	8.81 (1.72)	9.42 (1.58)	8.77 (1.57)
Control Variables				
Gender (female) N (%)	4 (33.30)	3 (27.3)	5 (41.70)	2 (18.20)
Age	24.83 (2.79)	24.00 (2.61)	24.00 (2.89)	24.91 (2.47)
Duration (mm:ss)	15:00 (3:53)	12:50 (1:48)	14:22 (4:04)	13:31 (3:34)
Pre-test (max=18)	11.67 (2.26)	9.59 (1.56)	10.38 (1.04)	11.00 (2.40)
Spatial Ability (%)	.76 (.11)	.75 (.16)	.73 (.15)	.78 (.12)
Working Memory (max=9)	4.36 (1.03)	4.27 (.65)	4.45 (.69)	4.18 (.98)
IOS (max=7)	5.00 (1.27)	5.73 (.79)	5.18 (1.40)	5.55 (.69)

Figure A1: Results (means *M* and standard deviations *SD*) split by teachers and students for the conditions VR/VR and VR/tablet and signaling and non-signaling. Variables are split into dependent and control variables. Abbreviations: Presence Questionnaire (PQ), Immersive Experience Questionnaire (IEQ), Player Experience Inventory (PXI), PXI - AUDVIS = PXI subscale audiovisual appeal, PXI - EOC = PXI subscale ease of control, PXI - GOAL = PXI subscale goals and rules, Questionnaire to assess Current Motivation (QCM), intrinsic cognitive load (ICL), extraneous cognitive load (ECL), germane cognitive load (GCL), Inclusion of Other in the Self (IOS)

	VR/VR (*N = 24*)	VR/Tablet (*N = 22*)
	M (SD)	*M (SD)*
Dependent Variables		
PQ (max=7)	4.57 (.48)	3.93 (.82)
IEQ (max=7)	5.09 (.54)	4.47 (.83)
PXI - AUDVIS (max=7)	6.04 (.95)	5.19 (1.52)
PXI - GOAL (max=7)	6.16 (.87)	5.27 (1.69)
QCM (max=7)	4.27 (.59)	4.69 (.49)

Figure A2: Statistically significant results (means *M* and standard deviations *SD*) split by the VR/VR and VR/tablet conditions. Abbreviations: Presence Questionnaire (PQ), Immersive Experience Questionnaire (IEQ), Player Experience Inventory (PXI), PXI - AUDVIS = PXI subscale audiovisual appeal, PXI - GOAL = PXI subscale goals and rules, Questionnaire to assess Current Motivation (QCM)

		Gender	Age	Pre-test	IOS	Working Memory Capacity	Spatial Ability	PQ	IEQ	PXI - EOC	PXI - AUDVIS	PXI - GOAL	QCM	ICL	GCL	ECL	Post-test
Gender	Corr. accord. to Pearson	1															
	Signifikanz (2-sided)																
	N	46															
Age	Corr. accord. to Pearson	0.187	1														
	Signifikanz (2-sided)	0.214															
	N	46	46														
Pre-test	Corr. accord. to Pearson	0.020	-0.127	1													
	Signifikanz (2-sided)	0.894	0.399														
	N	46	46	46													
IOS	Corr. accord. to Pearson	0.048	0.063	-0.002	1												
	Signifikanz (2-sided)	0.759	0.686	0.991													
	N	44	44	44	44												
Working Memory Capacity	Corr. accord. to Pearson	0.081	-0.063	0.225	-0.043	1											
	Signifikanz (2-sided)	0.601	0.685	0.142	0.784												
	N	44	44	44	44	44											
Spatial Ability	Corr. accord. to Pearson	0.133	-0.166	0.108	0.143	0.080	1										
	Signifikanz (2-sided)	0.383	0.277	0.480	0.356	0.606											
	N	45	45	45	44	44	45										
PQ	Corr. accord. to Pearson	-0.032	0.199	0.028	0.048	-0.135	0.084	1									
	Signifikanz (2-sided)	0.837	0.190	0.856	0.757	0.383	0.583										
	N	45	45	45	44	44	45	45									
IEQ	Corr. accord. to Pearson	-0.034	,297*	0.170	0.218	-0.159	0.076	,740**	1								
	Signifikanz (2-sided)	0.825	**0.048**	0.266	0.156	0.303	0.620	**0.000**									
	N	45	45	45	44	44	45	45	45								
PXI - EOC	Corr. accord. to Pearson	0.046	-0.130	0.025	0.225	-0.026	-0.113	0.210	0.142	1							
	Signifikanz (2-sided)	0.766	0.393	0.871	0.143	0.865	0.461	0.166	0.352								
	N	45	45	45	44	44	45	45	45	45							
PXI - AUDVIS	Corr. accord. to Pearson	-0.177	0.073	0.260	0.118	-0.130	-0.144	0.253	,325*	,368*	1						
	Signifikanz (2-sided)	0.246	0.634	0.084	0.447	0.400	0.347	0.094	**0.029**	**0.013**							
	N	45	45	45	44	44	45	45	45	45	45						
PXI - GOAL	Corr. accord. to Pearson	0.042	0.247	-0.044	0.096	0.110	-0.246	0.151	0.148	,519**	,452**	1					
	Signifikanz (2-sided)	0.785	0.102	0.774	0.536	0.477	0.104	0.323	0.333	**0.000**	**0.002**						
	N	45	45	45	44	44	45	45	45	45	45	45					
QCM	Corr. accord. to Pearson	-0.067	0.108	0.154	0.157	0.095	-0.138	-0.038	0.111	0.131	0.038	-0.160	1				
	Signifikanz (2-sided)	0.660	0.476	0.308	0.310	0.539	0.367	0.804	0.467	0.390	0.803	0.293					
	N	46	46	46	44	44	45	45	45	45	45	45	46				
ICL	Corr. accord. to Pearson	-0.103	-0.050	-0.200	-0.258	-0.099	-0.287	-0.014	-0.104	-0.067	-0.173	-0.021	0.207	1			
	Signifikanz (2-sided)	0.496	0.741	0.182	0.091	0.524	0.056	0.930	0.497	0.661	0.256	0.891	0.167				
	N	46	46	46	44	44	45	45	45	45	45	45	46	46			
GCL	Corr. accord. to Pearson	,313*	-0.006	0.110	0.274	0.139	0.131	0.288	,318*	,369*	0.079	0.291	0.209	0.133	1		
	Signifikanz (2-sided)	**0.034**	0.969	0.467	0.072	0.369	0.390	0.055	**0.033**	**0.013**	0.606	0.052	0.164	0.378			
	N	46	46	46	44	44	45	45	45	45	45	45	46	46	46		
ECL	Corr. accord. to Pearson	-0.113	-0.280	0.025	-,378*	-0.288	-0.121	-0.060	-0.103	-0.250	-0.290	-,323*	-0.071	,458**	-0.250	1	
	Signifikanz (2-sided)	0.453	0.060	0.869	**0.011**	0.058	0.427	0.693	0.499	0.098	0.053	**0.031**	0.639	**0.001**	0.094		
	N	46	46	46	44	44	45	45	45	45	45	45	46	46	46	46	
Post-test	Corr. accord. to Pearson	0.141	-0.200	0.268	0.060	,447**	,331*	0.088	0.163	0.110	0.045	0.085	-0.059	-0.224	0.178	-0.229	1
	Signifikanz (2-sided)	0.351	0.182	0.072	0.700	**0.002**	**0.026**	0.566	0.285	0.471	0.770	0.578	0.695	0.135	0.237	0.125	
	N	46	46	46	44	44	45	45	45	45	45	45	46	46	46	46	46

**. The correlation is significant at the 0.01 level (2-sided).

*. The correlation is significant at the 0.05 level (2-sided).

Figure A3: This table shows the correlations between the control variables and the dependent variables. Abbreviations: Presence Questionnaire (PQ), Immersive Experience Questionnaire (IEQ), Player Experience Inventory (PXI), PXI - AUDVIS = PXI subscale audiovisual appeal, PXI - EOC = PXI subscale ease of control, PXI - GOAL = PXI subscale goals and rules, Questionnaire to assess Current Motivation (QCM), intrinsic cognitive load (ICL), extraneous cognitive load (ECL), germane cognitive load (GCL), Inclusion of Other in the Self (IOS)