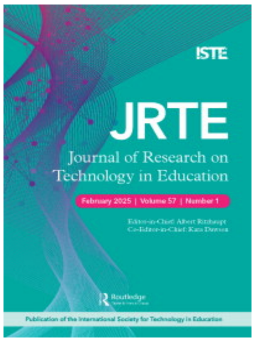

Journal of Research on Technology in Education

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ujrt20

Virtual reality and affective learning in commemorative history teaching: effects of immersive technology and generative learning activities

Miriam Mulders, Kristian H. Träg, Lilly Kaninski, Lara Kirner & Michael Kerres

To cite this article: Miriam Mulders, Kristian H. Träg, Lilly Kaninski, Lara Kirner & Michael Kerres (18 Feb 2025): Virtual reality and affective learning in commemorative history teaching: effects of immersive technology and generative learning activities, Journal of Research on Technology in Education, DOI: 10.1080/15391523.2025.2461524

To link to this article: https://doi.org/10.1080/15391523.2025.2461524

Published online: 18 Feb 2025.

JOURNAL OF RESEARCH ON TECHNOLOGY IN EDUCATION
https://doi.org/10.1080/15391523.2025.2461524

Virtual reality and affective learning in commemorative history teaching: effects of immersive technology and generative learning activities

Miriam Mulders, Kristian H. Träg, Lilly Kaninski, Lara Kirner and Michael Kerres

University of Duisburg-Essen, Essen, Germany

ABSTRACT

Learning with virtual reality (VR) can be highly motivating and conducive for cognitive and affective learning outcomes. For commemorative history education, VR enables novel experiences of testimonies co-created by survivors, witnesses, and museums. VR technologies allow learners to feel immediacy and an emotional connection to digitally reconstructed spaces and events of the past. While VR has proven to enhance affective learning outcomes by provoking emotions, interest, or motivation, its perceptual richness may also lead to distraction and cognitive overload. Generative learning activities can alleviate some of the limitations of learning with VR by helping learners to focus on the learning material. This study examines a highly engaging and historical immersive VR application and thereby investigates the effectiveness of the generative learning activities of *self-explaining* and *drawing*. Seventy-four undergraduate students explored a three-dimensional representation of the room where Anne Frank, a Jewish girl, was hiding during World War II. For the two experimental conditions, students had to either *create drawings* of Anne Frank's room and their own room or *verbally explain* how Anne Frank's living conditions would feel for them. For the control condition, students did not engage in a subsequent activity. Based on generative learning theory, we predicted that students engaging in generative learning activities would display higher posttest scores in knowledge and perspective taking than the control group. No such effects were found. Although the VR experience itself increased the ability to empathize with Anne Frank across all groups, it did not contribute to knowledge building. The study results indicate that even without any additional activity, VR can convey highly emotionally engaging testimonies and enables role taking, which suggests that VR is in particular suitable for affective learning.

ARTICLE HISTORY
Received 22 July 2024
Revised 29 October 2024
Accepted 28 January 2025

KEYWORDS
Virtual reality; educational media; educational technology; learning strategies; pedagogical issues

1. Introduction

Virtual Reality (VR) is a unique and promising learning tool that allows learners to immerse themselves in computer-generated and three-dimensional environments. It has the capability to enable highly engaging, interactive learning experiences since it can actively involve users in the learning process (e.g. Freina & Ott, 2015; Radianti et al., 2020). Previous research indicates that VR supports the experience of illusions such as presence and body ownership that can especially enhance affective learning outcomes by provoking emotions, interest, or motivation (Filter et al.,

CONTACT Miriam Mulders miriam.mulders@uni-due.de; Kristian H. Träg kristian.traeg@uni-due.de University of Duisburg-Essen, Essen, Germany.

2020; Makransky et al., 2021; Makransky & Petersen, 2021). Since VR experiences have the potential to engender a deep emotional engagement through spatial and temporal embeddedness, VR is seen to enhance a sense of relationality. Thus, various VR applications aimed at enhancing empathy-related skills have been developed (e.g. Bachen et al., 2012; Schutte & Stilinović, 2017; Sirkkunen et al., 2016). Especially VR applications where learners can inhabit the bodies of others allow role taking from another's perspective, support empathy and affective learning (Bertrand et al., 2018; Han et al., 2022).

However, while learning in VR applications can intellectually engage and emotionally involve learners, it may also lead to distraction and cognitive overload resulting in decreased performance (Mayer et al., 2023). Generative learning activities promise to alleviate some of these constraints by assisting learners to focus on learning materials and situational cues more intensively. They aim to foster a deeper processing of the material and improve understanding (Fiorella & Mayer, 2016). Accordingly, several studies have revealed that VR environments promote better learning when they incorporate multimedia design principles described as generative learning (Albus et al., 2021; Makransky, 2021). Activities such as summarizing and collaborating have been demonstrated to be effective when added in a typical VR training environment (e.g. Parong & Mayer, 2018; Petersen et al., 2023). Yet, less is known about the effect of generative learning activities for *affective* learning outcomes (e.g. taking the perspective of another person and training empathy in a VR setting).

The aim of the current study is to investigate an emotionally engaging VR application for commemorative history education. We are interested in differences regarding cognitive and affective learning outcomes in a highly engaging environment. Mainly, we assume that an immersive environment which depicts the reality or memory of a deceased person's life can facilitate empathy with the protagonist portrayed in the VR environment and can therefore address affective learning objectives effectively. Moreover, we examine the value added by two additional generative learning activities after the VR experience, namely *drawing* and *self-explaining*, and compare these two conditions to a control condition without an additional activity.

2. Theoretical background

Immersive learning, a subset of multimedia learning, focuses on adapted or simulated three-dimensional environments that are multisensory and that learners can explore and interact with (Makransky & Lilleholt, 2018; Sherman & Craig, 2018). Settings can be created that cannot be visited in real life due to various limitations (Freina & Ott, 2015; Mikropoulos & Natsis, 2011). Thus, VR is particularly relevant for learning experiences that are difficult to present adequately with the means of traditional media (Bailenson, 2018), like learning objects that consist of attitudes and affective learning. The idea that VR can trigger emotions and therefore evoke empathy more effectively than any other medium is linked to the brain's perception of VR experiences as real rather than mediated (Madary & Metzinger, 2016; Maister et al., 2015). As a result, learners may exhibit both emotional and behavioral responses to virtual experiences, which can have lasting effects. Learners feel as if they are either present in a different time at a different place or indeed experiencing the world through a different body, which may enhance both the urgency and the emotional impact of what learners are witnessing (Simine & Ch'ng, 2023).

Such virtual experiences may be suitable for teaching the history and life stories of previous and present generations with a high level of importance for our current socio-political debates. The VR environment *Clouds over Sidra*, for example, makes it possible to share the first-person perspective of the 12-year-old Syrian girl Sidra in a refugee camp in Jordan, home to thousands of Syrians fleeing war, experiencing her everyday life (Milk, 2015). Another VR application succeeded in reducing implicit racial bias against dark-skinned people when light-skinned participants embodied in a dark-skinned body (Peck et al., 2013). The *Kokoda VR* environment sends learners back to World War II in the Australian territory of Papua where they can pick up weapons, run through the battle fields and explore the historical grounds of the Kokoda Trail (Calvert & Abadia, 2020).

On the other hand, previous immersive multimedia endeavors that were not based in VR have produced mixed results. For example, enabling Instagram users to experience the actions of German anti-Nazi resistance activist Sophie Scholl in real time was perceived as somewhat controversial because the boundaries between fact and narrative or between fiction and reality seemed to blur (Thiele & Thomas, 2023). Whether parasocial relationships are an effective way of engaging learners with history education remains debatable, especially since that project seemed to have missed its target audience (Thiele, 2023), in addition to lacking embedding in wider institutional curricula, which would be recommended (Träg & Mulders, in press).

However, ethical challenges with immersive VR have also become apparent. While we acknowledge that VR is a handy technology to explore otherwise inaccessible places or points in time, it should always be kept in mind, especially in history education about sensitive topics, that researchers as well as programmers should not attempt to replace or supplant reality. Paradoxically, making voices of victims and witnesses of atrocities more accessible might also run the risk of making them appear more mundane (Knoch, 2021). Nash (2018), for example, warns against designing the virtual world as photo-realistically as possible because this could create an improper distance from victims and historical events. Moreover, Segovia and Bailenson (2009) found that VR can even create false memories. This is why on the one hand ethics are a necessary part of the computer science education of future programmers and developers (Horton et al., 2022), while on the other hand learning interventions should take care to emphasize the difference between reality and virtuality (Bunnenberg, 2020; Lewers, 2022).

In our study we take on the historical experience of the hiding place of people of Jewish origin during World War II. Through high-quality three-dimensional representations, VR allows learners to explore the hiding place room by room in a way that would be vastly more difficult in the real world.

2.1. Learning in immersive VR

Today, immersive VR is used in many areas, such as aviation, military training, gaming, engineering, simulating surgical procedures, psychological treatment, learning, and social skills (Baceviciute et al., 2021; Christopoulos & Mystakidis, 2023; Cipresso et al., 2018; Pellas et al., 2020; Wolf et al., 2021; Wolfartsberger, 2019; Wong & Lee, 2024). Based on promising examples, it is often predicted that immersive VR will be able to enhance traditional classroom learning (Calvert & Abadia, 2020; Snelson & Hsu, 2020). Still, successfully implementing immersive VR in the everyday classroom depends on briefing and debriefing activities as well as several factors on a micro- (i.e. characteristics of the learners), meso- (i.e. teacher- and classroom-specific influences) and macro-level (i.e. institutional and governmental factors) (Dengel et al., 2023; Träg & Mulders, in press).

VR technology allows easy access to points in time and space that would otherwise be difficult to reach (Janssen et al., 2016), meaning that it holds special value for history education, where it can enable learners to experience the lives of historical personalities (Frentzel-Beyme & Krämer, 2023; Mulders, 2023). Patterson et al. (2022) argue that relating to historical personalities can help facilitate learning the values of pluralistic democracy. Some studies show that history education aided by VR technology may increase empathy, presence, and academic performance (Abadia et al., 2019; Calvert & Abadia, 2020; Frentzel-Beyme & Krämer, 2023; Mulders et al., submitted; Zhang, 2019).

Tied to this, there is a growing body of evidence showing the benefits of learning academic content in immersive VR as compared to other media. Recent meta-analyses (Coban et al., 2022; Wu et al., 2020) and reviews (Parong, 2021; Radianti et al., 2020) found small effect size advantages of immersive VR when compared with non-immersive learning conditions. Immersive VR is said to increase learners' remembering of the concepts learned (Buttussi & Chittaro, 2023; Meyer et al., 2019), transfer of knowledge (Chittaro et al., 2018), practical skills (Barrett et al., 2024; Mulders et al., 2022), and their emotional performance (Cheng & Tsai, 2020; Mulders, 2023). Nonetheless, there are also studies suggesting that immersive VR does not make a significant difference in learning (Hassenfeldt et al., 2020; Leder et al., 2019).

Despite the postulated educational advantages linked to immersive VR, there are significant barriers impeding its implementation, such as the experience of physical discomfort (i.e. motion sickness) (Jensen & Konradsen, 2018), high arousal, or cognitive load triggered by information overload (Makransky et al., 2019).

2.2. Characteristics of immersive VR technologies

VR is an umbrella term that encompasses various technological systems (Rauschnabel et al., 2022; Wohlgenannt et al., 2020). In our study, we exclusively work with head-mounted display (HMD)-based VR. This VR technology monitors the learner's head movements and, in some cases, body movements (Sousa Santos et al., 2009). HMDs offer learners the opportunity to control the environment with a full stereoscopic view (Calvert & Abadia, 2020; Passig et al., 2016).

For the classification of VR, it is useful to distinguish between *immersion* and *presence* (Bowman & McMahan, 2007). Immersion is described as a set of technological features that provide a sense of reality to learners (Slater, 2018). In contrast, presence refers to the subjective experience of learners in terms of the degree to which they have the feeling of *being there* in a computer-generated environment (Slater et al., 2009). Some researchers argue that technologies with a high level of immersion create a higher perception of presence than low immersive VR technologies (Meyer et al., 2019).

Nevertheless, highly immersive media, such as HMD-based VR, possibly overwhelm learners with the many auditory and visual stimuli presented. Since human processing capacity is limited (e.g. Sweller, 2011), immersive VR in particular can be a hindrance to learning because the perceptual richness of the virtual world demands too many resources (Mayer et al., 2023). Several studies have already revealed that learning with VR can lead to cognitive overload (e.g. Albus et al., 2021; Makransky et al., 2019; Mayer et al., 2023; Meyer et al., 2019). According to Mayer's assumptions about multimedia learning (e.g. Mayer, 2005), cognitive overload appears because of an increase in intrinsic cognitive load due to the difficulty of the learning material combined with an increase in extraneous cognitive load due to the many stimuli afforded by the medium. One principle of multimedia learning, called the *immersion principle* (Makransky, 2021), specifies how VR training may benefit from added instructional guidance. In accordance with this principle, immersive VR trainings foster deeper learning when they are designed based on multimedia design principles. A recent meta-analysis investigating HMD-based VR revealed that more than one quarter of the examined studies demonstrated negative effects, underscoring the significance of considering the instructional design of VR-based learning environments (Wu et al., 2020). VR trainings that ignore these instructional principles may overtax working memory resources by being excessively intricate and complex, thereby competing with the processing of essential learning content (Chandler, 2009). Hence, instructional guidance is needed to help learners concentrate on the content presented in VR. Richard Mayer himself observed that the list of evidence-based principles for learning with different multimedia applications must continue to be developed, which applies in particular to technologies such as immersive VR (Mayer et al., 2020).

2.3. Generative learning activities when learning with immersive VR

Generative learning theory proposes that learners need to be actively involved in processing learning materials to integrate new information into the long-term memory (Wittrock, 2010). This involves three processes: (1) *selecting* the relevant information, (2) *organizing* it into a cohesive structure within working memory, and (3) *integrating* it with relevant knowledge retrieved from long-term memory (Fiorella & Mayer, 2016, 2021).

Fiorella and Mayer (2016) differentiate between eight generative learning activities (e.g. mapping, imagining) that should lead to active participation in the learning process. Such activities aim to reduce learners' cognitive load on working memory and improve learning outcomes by encouraging the learners to reflect on the learning material. For immersive VR, such activities are also considered to be conducive to learning (Makransky et al., 2021; Mulders et al., 2020; Yang et al., 2021). Some studies have already investigated the effectiveness of generative learning activities when learning with immersive VR, such as teaching (e.g. Klingenberg et al., 2020), summarizing (e.g. Parong & Mayer, 2018), and collaborating (e.g. Petersen et al., 2023).

In our current study, our emphasis is on employing prompts to encourage engagement in generative learning activities after the VR experience has been completed. In the Parong and Mayer (2018) study, students took off their HMDs and wrote a summary, whereas in the Petersen et al. (2023) study students created their summaries verbally in VR while wearing their HMDs. While these two studies analyzed the activity of summarizing, we asked learners to produce explanations and drawings. The generative learning activity of *self-explaining* involves a detailed statement by the learner to themselves and the generative learning activity of *drawing* encompasses creating a visual representation of the material (Fiorella & Mayer, 2016).

Furthermore, previous studies often focused on examining the effects of generative learning activities on cognitive learning (e.g. retention, transfer). We aim to investigate effects on an affective level as well, as VR has proven to be particularly effective in eliciting compassion and empathy for others within the VR environment (e.g. Frentzel-Beyme & Krämer, 2023; Martingano et al., 2021; Nakamura, 2020). To trigger generative learning activities that engage learners with the learning material on an emotional level, we have designed the prompts in such a way that learners should always establish a connection between what is represented in VR and the reality of their own life experiences. Hence, we have expanded upon Fiorella and Mayer (2016) generative learning activities by adding the aspect of *ego-involvement* (i.e. making connections/comparisons between the self and the virtual agent) in the task prompts.

2.4. Hypotheses of the present study

Our main aim was to empirically examine an immersive VR application in the field of commemorative history education, with the goal of investigating its effects on affective as well as cognitive learning outcomes. We furthermore checked whether learners would benefit from being prompted to engage in generative learning activities after a VR experience, compared to experiencing the same without a generative learning activity. Three study conditions are employed: no generative learning activity (control group), learners producing two drawings (*drawing* group), and learners creating a written statement for themselves (*self-explaining* group). We recorded learning on a cognitive (i.e. knowledge acquisition) and an affective level (i.e. perspective taking).

We assume that in each of the three experimental conditions posttest scores are higher than pretest scores (hypothesis 1). We expect this effect of the immersive technology for both cognitive and affective indicators. Based on generative learning theory, we furthermore predict that learners engaging in generative learning activities will display significantly higher posttest scores than those in the control group (hypothesis 2). Next, we formulated the prompts for the generative learning activities in such a way that the learners had to make a connection to the reality of their own lives. This was particularly aimed at addressing the learners' affect. We therefore also assume that learners in the generative learning conditions will show a greater improvement between the pretest and posttest in perspective taking than in knowledge (hypothesis 3).

3. Methods

3.1. Participants and design

To verify our hypotheses, we employed an experimental design in a laboratory setting. Participants were randomly assigned to one of the three study conditions: (1) no generative learning activity, (2) *drawing*, (3) *self-explaining*.

A total of 75 bachelor students at the University of Duisburg-Essen in Germany were recruited as participants, one of whom was excluded from analysis since they did not fill in both questionnaires. Out of the remaining 74, 17 identified as male, 55 as female and 2 as non-binary. Gender and age by test condition are listed in Table 1. Participants largely had a background in Education, as 43 indicated their major subject as Education Sciences, with a further 17 indicating Teaching. Five more participants were studying Psychology, the remaining 9 indicated other major subjects (i.e. Software Engineering, Pedagogy, Law). Most participants ($n = 45$) had no prior experience with VR, while 22 specified that they had only tried VR shortly. The remaining 5 indicated that they used VR rarely, the other options labeled *occasionally* and *regularly* were not ticked by any participant.

3.2. VR application

We utilized the free application *Anne Frank VR House* which was developed cooperatively by the *Anne Frank Foundation Amsterdam* and the Game development studio *Force Field VR*. The VR application allows learners to explore the hiding place of Anne Frank, her family, and four other people, all of them of Jewish origin. Anne Frank is representative of millions of Jewish people who had been discriminated against, persecuted and systematically murdered.

In 1929, Anne was born as a daughter to Jewish parents in Frankfurt, Germany. At that time, living conditions in Germany were marked by economic instability, accompanied by a rise in antisemitism, which was both exacerbated and deliberately incited. For these reasons, Anne Frank and her family fled to Amsterdam. When German forces occupied the Netherlands in 1930, ten-year-old Anne Frank witnessed the increasing discrimination and persecution of Jews under the Nazi regime. Therefore, Anne Frank and her family went into hiding from 1942 to 1944 in a concealed back house in Amsterdam. In the hiding place, eight individuals lived under restrictive conditions, in complete silence and darkness, to avoid discovery. Anne Frank documented these harsh living conditions, along with her thoughts, worries, and hopes, in her diary. After the hiding was exposed, Anne Frank and her family were deported. Anne Frank died in the Bergen-Belsen concentration camp in 1945 at the age of 15. The only survivor, Anne Frank's father Otto Frank later published her diary, which has been translated into over 75 languages. The hiding place is now part of a museum.

The *Anne Frank VR House* is a faithful replica of the hideout in VR and allows learners to explore the hiding place of Anne Frank and immerse themselves in their living conditions at that time. Moreover, each of the eight rooms of the hiding place has been recreated in detail. The application can be explored using HMDs and controllers from various providers. By using the controllers, learners can interact with different elements and obtain auditory information.

In our study, we focus solely on one of these rooms, namely that of Anne Frank, which she had to share with a non-related middle-aged Jewish friend of the family named Fritz Pfeffer. In total, the room was 6.5 square meters in size. In that room, there are four interactive elements: Anne Frank's diary, a postcard, a book resting on the bed, and a pair of binoculars (see

Table 1. Age and gender by test condition.

		Age			Gender		
	N	*Min*	*M* (*SD*)	*Max*	Male	Female	Non-binary
Full sample	74	18	24.82 (8.41)	66	17 (23.0 %)	55 (74.3 %)	2 (2.7 %)
Drawing	24	18	23.50 (7.44)	55	5 (20.8 %)	18 (75.0 %)	1 (4.2 %)
Self-explaining	24	19	27.79 (11.61)	66	6 (25.0 %)	17 (70.8 %)	1 (4.2 %)
Control	26	19	23.31 (4.38)	40	6 (23.1 %)	20 (76.9 %)	0 (0.0 %)

Figure 1). Whenever one of the elements is manually picked up by the learner, the voice of a young woman is heard, suggesting that Anne Frank herself is speaking. For example, when the diary is picked up, she may say:

> Writing liberates me from everything. My sorrows vanish, my courage reignites. Yet, the main question lingers: Will I ever produce something truly great? Can I become a journalist and a writer? I fervently hope so, as writing allows me to document everything: my thoughts, my ideals, and my fantasies (Vertigo Games & Knucklehead Studios, 2019).

The entire experiment lasted approximately 25 min per person. The VR experience was presented to the participants using HMDs, more precisely *Meta Quest 2*, and two associated controllers which were used for movement and interaction with objects.

The selection of this particular VR application for the present study was based on its usage in a previous study that investigated the effects of different VR devices, instructional methods and learning processes such as flow and presence (Mulders, 2023, 2024). It is worthy to note that participants in our study were German students who may have a different historical frame of reference than international students would have had. Other researchers also utilize VR in history education because this technology provides the opportunity to travel to past locations and engage in conversations with deceased individuals (e.g. Nachtigall et al., 2022; Parong & Mayer, 2021).

Figure 1. The interior of the *Anne Frank VR House* application. *Note*. Picture a shows Anne Frank's room. Picture b shows Anne Frank's diary. Picture c shows Anne Frank's binoculars. Picture d shows the wall which Anne Frank decorated with pictures.

3.3. Generative learning activities

We have chosen two out of the total eight generative learning activities posited by Fiorella and Mayer (2016). We chose these two because previous studies have predominantly investigated other activities, like teaching (e.g. Klingenberg et al., 2020), summarizing (e.g. Parong & Mayer, 2018), or collaborating (e.g. Petersen et al., 2023). The two selected activities are also suitable for individual work where no other people are needed, which suited our study design. We developed prompts for these two activities. The worksheets containing the prompts for the generative learning activities are available online[1]. For the *drawing* activity, we asked the students to draw Anne Frank's room from a bird's eye view. The students were given a checkered sheet of paper for this purpose. They were instructed to assume that four squares correspond to one square meter. Objects or furniture should be drawn on the sheet. The second part of the task involved drawing their own room. An exemplary solution of one student is given in Figure 2.

For the *self-explaining* activity, we instructed the students to imagine themselves living for a week under the conditions in hiding similar to those Anne Frank endured. They were asked to describe in their own words how they would have felt in that situation, what problems they would have faced the most, and what would have scared them the most. Below are two exemplary quotes from the students:

- Participant SO18: *I would feel very cramped, especially the loss of privacy would bother me. Permanently being with so many people (especially family) in such a small space. Above all, I would suffer not being allowed to go out anymore, being limited in all my freedoms […].*

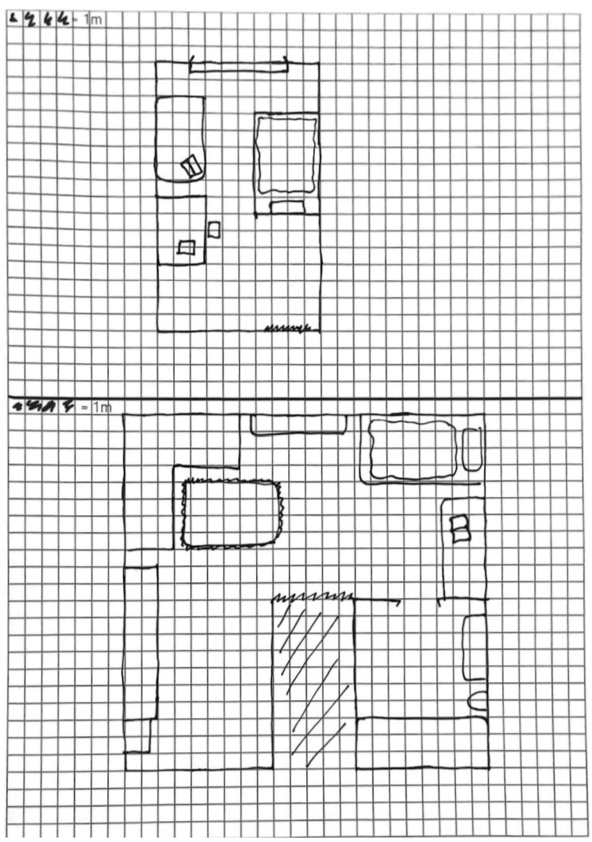

Figure 2. Generative learning activity *drawing* of participant HE15. *Note*. Top section: Anne Frank's room. Bottom section: own room.

- Participant IM03: *What is happening outside my room? Will I be discovered? Will anybody hear me when I use the toilet? My situation is marked by hopelessness: the confinement within the small room, the lack of contact with the outside world, and the inability to communicate. Being constantly surrounded by past experiences, photos, and memories. The prospect of a future existence hinges solely on hope for assistance, luck, and trust. Yet, I would be inundated by fears that I couldn't shake, suppress, or conquer. They persist relentlessly.*

While previous prompts for generative learning activities were mostly formulated to achieve cognitive learning outcomes (e.g. Parong & Mayer, 2018; Petersen et al., 2023), we endeavored to ensure that the emotional experiences of the learners were also addressed when creating the prompts. We attempted to encourage learners to establish a connection to their own life experiences or make comparisons between Anne Frank's life and their own within both the *drawing* and *self-explaining* activities. Hence, during the generative learning activity, the learner has to involve their own person in order to even be able to complete the task. It is worth noting that this paper focuses on whether doing the activity itself can add value to learning. The quality of the execution when drawing and explaining will be reported in another paper (Mulders et al., 2024).

3.4. Questionnaires

The self-reported data were collected using a pre- and a post-questionnaire. Both were completed by the students using a conventional laptop provided by the project team. We used *Sosci Survey* as a standard tool for academic online surveys. All questionnaire materials are accessible online[2].

To measure cognitive learning outcomes, we assessed students' knowledge. Immediately before the VR session, prior knowledge was measured. Therefore, students rated their knowledge level on a scale from 1 (*not available*) to 10 (*extensive knowledge*). Furthermore, students were asked to answer four knowledge questions (e.g. *What does Anne's diary look like? Try to name characteristics.*) after the VR experience. A maximum of 2 points each for the first three questions and 3 points for the fourth question was reachable, resulting in a maximum point value of 9. Two raters awarded points independently of one another, reaching an interrater reliability of .91 according to weighted Cohen's κ (Cohen, 1968). For all further calculations, the ratings of the more experienced rater were used.

As an affective learning objective, perspective taking in Anne Frank was determined. Perspective taking is interpreted as the capacity to empathize with another person's feelings. It is a prerequisite for the emergence of empathy or compassion (Roberts et al., 2014; Wolgast et al., 2020). To measure the extent of perspective taking, an adapted version of the historical perspective taking questionnaire (Hartmann, 2008) was implemented. The questionnaire used in the current study consists of nine items (e.g. *As a daughter Anne couldn't object to her parents' strict rules in the hiding place and had to comply with their wishes.*) to be answered on a scale from 1 (*does not fit her situation at all*) to 4 (*fits her situation very well*). The questionnaire was filled out immediately after the VR experience. For a previous study (Mulders et al., 2022), the questionnaire originally developed by Hartmann (2008) was already adapted to the topic of Anne Frank. Because the scale formerly exhibited low internal consistency in Mulders et al. (2022) and even in the original study by Hartmann (2008) itself, we further refined the items and pretested them with students in the corresponding age group. For additional validation of perspective taking, students had to assess how well they could empathize with Anne Frank's living situation in the hiding place. A ten-point scale from 1 (*not at all*) to 10 (*completely*) was utilized. This question was asked twice, before and after the VR experience.

Age, gender, bachelor subject, and prior technological experience were considered as control variables. To capture these variables, closed-ended questions were asked at the first measurement time point.

3.5. Procedure

Before the project began, we submitted an ethics application. On 2nd March 2023, our project was approved by the Ethics Committee of the University of Blinded for Review. We have pre-registered our study *via Open Science Framework*[1].

The data collection started in April 2023 and was completed in January 2024. The students were invited to participate *via* posters around the university or calls on social media networks. Participation was voluntary and unpaid. The study was carried out at a laboratory at the University of Blinded for Review. An individual appointment was scheduled with each student. The experimental procedure is depicted in Figure 3. Firstly, the students were greeted by a research assistant. The students were informed that they could terminate the experiment at any time (e.g. due to feeling unwell). Prior to the experiment, students read and signed a consent form that contained further information regarding the experiment. Following randomization, students responded to the pre-questionnaire. Next, the research assistant read aloud a brief informational text about Anne Frank and her hiding place in Amsterdam during the Nazi occupation. This introductory information text and the consent form are available online[1]. Thereafter, the research assistant provided each student with an HMD and a set of controllers. To accommodate students' varying levels of experience with VR, all students received instruction in how to use the controllers and tried on the HMD before starting the application. Once equipped with the HMD, the VR was launched automatically.

After the VR experience, those students in the two experimental conditions worked on the worksheets provided to them. These worksheets contained prompts for the generative learning activities, as described earlier in this section. The students in the control group immediately completed the post-questionnaire, whereas those in the experimental conditions filled it out only after finishing the generative learning activity. In general, students completed both questionnaires, the pre- and the post-questionnaire, at their own pace.

4. Results

First, descriptive statistics will be reported, followed by the prerequisite check for a multivariate analysis of covariances (MANCOVA) and results of the analyses of hypotheses. Finally, some

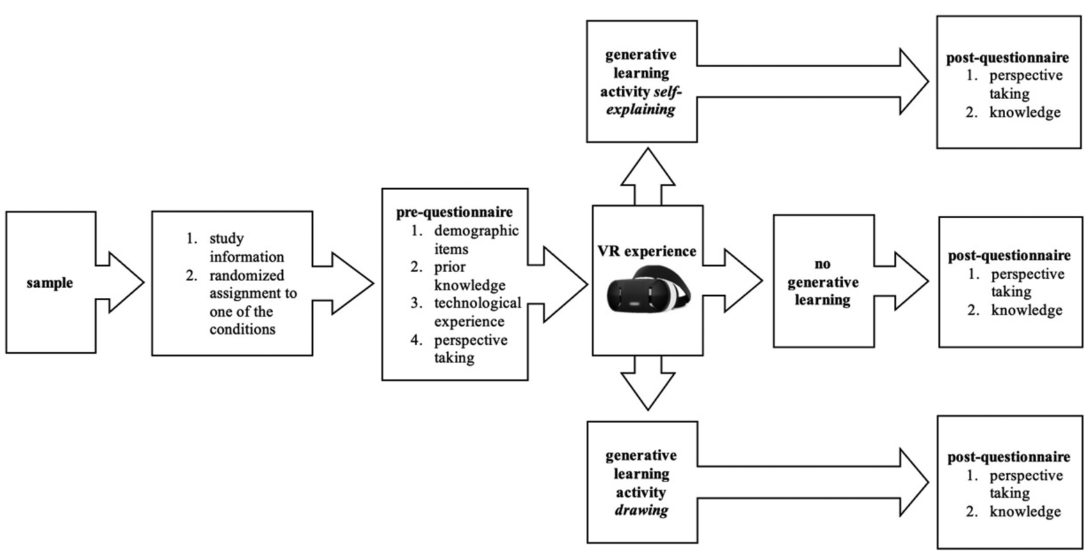

Figure 3. Overview of the experimental procedure and the three conditions.

exploratory facets of the present data will be presented. All analyses were performed in *R* version 4.2.2 (R Core Team, 2022).

4.1. Descriptive statistics

Descriptive statistics for the five affective and cognitive variables (1) self-assessed pretest knowledge, (2) posttest knowledge score, (3) self-assessed perspective taking pretest, (4) self-assessed perspective taking posttest, and (5) Hartmann scale perspective taking (posttest) by test condition are shown in Table 2. The Hartmann scale and the knowledge score were the only measures composed of multiple items. Reliability for the Hartmann scale was mediocre, with Cronbach's $\alpha = .69$. Weak reliability values are a known flaw of the Hartmann scale (Hartmann, 2008; Mulders et al., 2022). Using the Kruskal-Wallis test (Kruskal & Wallis, 1952), no pretest group differences were found for self-assessed knowledge ($\chi^2 = 2.72$, $p = .257$) or self-assessed perspective taking ($\chi^2 = 5.24$, $p = .073$).

4.2. Prerequisites

To be able to perform a MANCOVA, certain prerequisites need to be met. The dependent variables knowledge score and perspective taking (Hartmann scale) did not show multicollinearity, $r = .024$. In fact, the dependent variables did not seem to correlate at all, a peculiarity that will be discussed later. Normal distribution was found for the posttest knowledge score ($W = 0.98$, $p = .158$), but not for either of the perspective taking measures (Hartmann: $W = 0.95$, $p = .009$). Homogenous variances were found across test conditions for both knowledge score ($F(2, 71) = 1.69$, $p = .192$) and Hartmann scale perspective taking ($F(2, 71) = 0.30$, $p = .743$). For the following analyses of hypotheses, an alpha-level of 5% was assumed. The Holm-Bonferroni method (Holm, 1979) was used to adjust the p-values of t-tests to account for alpha error accumulation.

4.3. Hypotheses testing

Hypothesis 1 assumed that posttest values would be higher than pretest values for knowledge and perspective taking, across all conditions. This proved correct in the entire sample for perspective taking ($t(73) = 9.16$, $p < .001$, $d = 1.06$), but not for knowledge ($t(73) = 2.27$, $p = .065$, $d = 0.26$). Looking at the data by experimental condition, self-assessed perspective taking increased in every group (see Figure 4, part a), while the difference between the self-assessed pretest knowledge and the posttest knowledge score did not show statistical significance in either condition (see Figure 4, part b). Exact statistics are shown in Table 3. In total, hypothesis 1 can only be partially corroborated.

Table 2. Descriptive statistics for the cognitive and affective variables.

		Min	*Md*	*Max*	*M*	*SD*
Self-assessed knowledge (pretest)	Drawing	0.0	3.0	6.0	2.83	1.52
	Self-explaining	0.0	4.0	7.0	3.71	1.85
	None	0.0	3.0	8.0	3.65	2.12
Knowledge score (posttest)	Drawing	0.5	3.5	7.0	3.81	2.00
	Self-explaining	1.0	4.0	9.0	4.04	1.76
	None	2.5	3.75	6.5	4.23	1.31
Self-assessed perspective taking (pretest)	Drawing	0.0	4.0	7.0	3.79	1.86
	Self-explaining	1.0	4.0	9.0	3.88	1.87
	None	0.0	2.5	7.0	2.73	2.24
Self-assessed perspective taking (posttest)	Drawing	2.0	6.0	7.0	5.17	1.76
	Self-explaining	2.0	6.0	9.0	5.83	1.69
	None	2.0	6.0	9.0	5.85	2.13
Hartmann scale perspective taking (posttest)	Drawing	2.33	2.89	3.56	2.88	0.34
	Self-explaining	2.44	2.89	3.67	2.91	0.33
	None	2.33	2.89	3.67	2.87	0.30

Hypothesis 2 assumed that the perspective taking scores as well as knowledge scores would be higher in the two experimental conditions than in the control condition. A MANCOVA was implemented, with knowledge score and the Hartmann scale as dependent variables, while test condition was the independent variable and gender, prior VR experience, and bachelor subject served as covariates. Neither of the covariates influenced the outcome, as shown in Table 4. The three groups also showed no significant differences, $F(4, 136) = 0.23$, $p = .921$, Pillai's trace $V = 0.01$. Hence, hypothesis 2 must be rejected.

Hypothesis 3 referred only to the two experimental conditions and stated that the difference between pre- and posttest would be higher for the self-assessed perspective taking item than for knowledge. Differences were calculated by subtracting the pretest value from the posttest value. One-sided, pairwise t-tests were utilized. Across both experimental conditions, the pre-post difference for perspective taking was $M = 1.67$ ($SD = 1.69$), while the difference for knowledge was $M = 0.66$ ($SD = 2.23$). This difference between variables is statistically significant, $t(47) = 2.86$, $p = .019$, $d = 0.41$. Looking at the test conditions in isolation, the difference remained significant in the *self-explaining*, but not the *drawing* condition. Interestingly, the difference between perspective taking and knowledge also showed significance in the control group. Exact statistics are shown in Table 5. Hypothesis 3 can at least partially be corroborated.

4.4. Exploratory analyses

An analysis of the control variables gender, bachelor subject, and prior technological experience did not show a difference between groups. Male and female students did not differ in knowledge score or Hartmann scale perspective taking. Neither did Education Sciences students from participants studying other subjects, nor students with prior VR experience from students without experience. Participants' age was found to correlate with perspective taking on the Hartmann scale ($r = .355$, $p = .002$) as well as the self-assessed item in the pretest ($r = .269$, $p = .021$), meaning older participants showed higher perspective taking.

As mentioned above, knowledge score and Hartmann perspective taking did not seem to correlate at all in the full sample. This remained the same on a group level. Further exploratory analyses found a correlation between self-assessed perspective taking in the pre- and posttest for the full sample ($r = .461$, $p < .001$), a pattern which persisted when looking at the groups isolated. The pretest perspective taking also correlated with Hartman scale perspective taking ($r = .339$, $p = .003$), which only remains intact in the *self-explaining* group ($r = .509$, $p = .011$). The two posttest perspective taking measures did not seem to correlate, further calling the accuracy of our instruments into question.

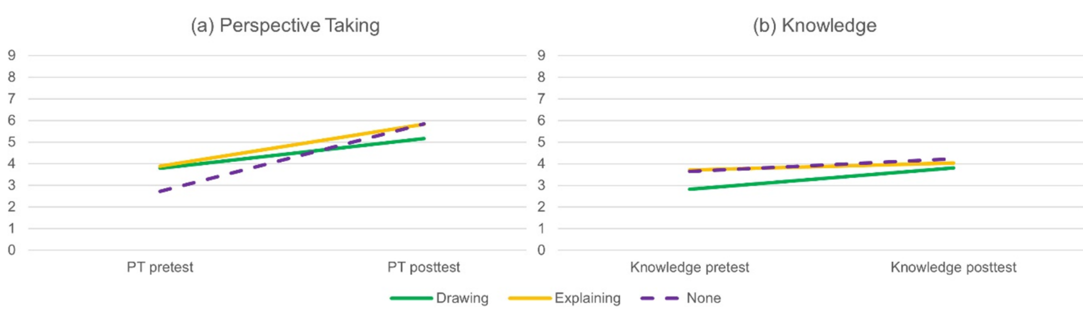

Figure 4. Difference between pretest and posttest averages for (a) perspective taking (PT) and (b) knowledge.

Table 3. Test statistics for hypothesis 1 by group.

		t	*df*	*p*	*d*
Perspective taking	Drawing	3.82	23	**.004**	0.78
	Explaining	5.99	23	**< .001**	1.22
	None	6.84	25	**< .001**	1.34
Knowledge score	Drawing	2.15	23	.085	0.44
	Explaining	0.74	23	.449	0.15
	None	1.09	25	.427	0.21

Note. The *p*-values have been corrected using Holm-Bonferroni method. Statistical significance is marked in boldface.

Table 4. MANCOVA results: multivariate tests.

	F	*df*	*p*	*V*
Test condition	0.23	4, 136	.921	0.01
Gender	0.82	2, 67	.446	0.02
VR experience	0.10	2, 67	.907	0.00
Bachelor subject	0.40	2, 67	.672	0.01

Note. V: Pillai's trace.

Table 5. Test statistics for hypothesis 3 by group.

	Δ*M (SD)*					
	PT	K	*t*	*df*	*p*	*d*
Drawing	1.38 (1.76)	0.98 (2.23)	0.77	23	.449	0.16
Explaining	1.96 (1.60)	0.33 (2.22)	3.52	23	**.006**	0.72
None	3.12 (2.32)	0.57 (2.69)	3.80	25	**.004**	0.75

Note. Δ*M*: Mean difference, PT: perspective taking, K: knowledge. The *p*-values have been corrected using Holm-Bonferroni method. Statistical significance is marked in boldface.

5. Discussion

5.1. Empirical contributions

The present study demonstrated that the selected VR environment improved the ability to take on the perspective of the VR protagonist (role taking) but did not lead to an increase in declarative knowledge. This finding is at least somewhat consistent with other research results, which also suggest that VR is more suitable for promoting affective learning than cognitive learning (e.g. Finken & Wölfel, 2023; Jensen & Konradsen, 2018).

Regarding hypothesis 1, we found that the students' ability to empathize with Anne Frank improved from before to after the VR experience. This increase in perspective taking is significant in all three groups and meaningful with an average effect size of $d > 1$. In the original study, Hartmann (2008), who examined the extent of historical perspective taking in an unmarried woman in the Middle Ages who is forced to enter a convent, reported a total mean score of 2.15. In the current study as well as in the former study by Mulders et al. (2022), the overall scale value ($M = 2.89$) averaged across all conditions is more than one standard deviation unit above this value. From this, it seems that the immersive VR technology itself already engages students on an emotional level and prompts them to attempt to empathize with the VR protagonist, in this case Anne Frank. In other words, the technology in itself appears to be adequate to trigger an emotional state. As a result, it seems to be easier to relate to the role of Anne Frank. A significant increase in knowledge from before to after the VR experience, however, is not observed in any of the three groups even if the results were numerically indicative of an increase in self-assessed knowledge. This may be due to a biased assessment of knowledge by the students themselves. Self-assessment may not always be accurate, especially for individuals who lack knowledge (Dunning, 2011). This means that students might have overestimated their knowledge in the pretest, whereas the knowledge score in the posttest which included four knowledge questions should be somewhat accurate. The high standard deviations in their prior knowledge also suggest that students' self-assessment is not precise. If we assume overestimation in the pretest, the measured difference to the posttest is smaller than the actual difference.

Regarding hypothesis 2, we failed to find the expected positive effects on knowledge and perspective taking through the execution of generative learning activities. Neither for *drawing* nor for *self-explaining* an additional value was observed. Possible explanations for the non-significant results could be that our study is a cross-sectional study. For effects to occur on an affective level, it may take more time than a few minutes to process and reflect on the emotionally moving content. This might especially be true considering that most of our participants were using VR technology for the first time during this experiment, giving them an additional experience to process. Furthermore, we cannot ensure that the prompts for the activities were formulated appropriately enough. It is possible that the input of a historian or history teacher is missing here. In addition, methodological concerns can be uttered because of small sample sizes and questionable measurement reliability and validity (see section 5.3) and therefore effects may not have been detected.

Regarding hypothesis 3, we expected that the effects for the affective learning objective of perspective taking would be greater than those for the cognitive learning objective of knowledge. This expectation was partially confirmed by our results. In both the *self-explaining* and the control group, the increase in perspective taking was significantly higher than the increase in knowledge. However, this effect was not found for *drawing*. In total, the results for hypothesis 3 align with those of hypothesis 1: Perspective taking appears to be a learning objective that can be achieved through immersive VR technology itself. However, as the results for all three hypotheses indicate, an additional generative learning activity is not necessarily required for this.

Regarding exploratory analyses, we found that age correlated positively with perspective taking, indicating that older participants showed higher perspective taking. Considering the age range of the present sample, this is in line with results that show that empathy is highest in middle-aged adults, and lower in younger and older adults (O'Brien et al., 2013). We further found that the Hartmann scale correlated with single-item perspective taking in the pretest, but not the posttest. This implies a somewhat inaccurate measurement of perspective taking, which, as pointed out before, is a known problem (Hartmann, 2008; Mulders et al., 2022).

Overall, our results indicate that affective role taking in a VR actor is possible through an authentic VR environment. At least in our study, adding generative learning activities to a VR session is not as effective as we originally anticipated. Yet, other studies consistently found positive effects using additional learning activities (e.g. Klingenberg et al., 2020; Parong & Mayer, 2018; Petersen et al., 2023). In these studies, however, other activities (e.g. teaching) and more cognitive learning objectives (e.g. retention) were examined. We also need to acknowledge the possibility that the quality of our prompts for the generative learning activities may have been insufficient. It is possible that previous prompts (e.g. Petersen et al., 2023) are more effective than those we designed. This raises the general question of what constitutes an appropriate prompt. Nevertheless, further research is needed to confirm or refute the additional value of generative learning activities when learning with immersive VR. We acknowledge that limitations of our study (e.g. weak measurement instruments due to low internal consistency) may have prevented us from uncovering positive effects of the generative learning activities. We will discuss these limitations in more detail later in this section.

5.2. Theoretical and practical contributions

This study contributes to the expanding body of literature examining the efficacy of immersive VR technology to engender compassion and empathy with the characters portrayed in the virtual world (e.g. Bertrand et al., 2018; Martingano et al., 2021). Contrary to the present paper, an increasing number of studies have shown the potential benefits of incorporating generative learning activities, such as summarization (Parong & Mayer, 2018), teaching (Klingenberg et al., 2020), and collaborating (Petersen et al., 2023), to facilitate learners' comprehension of the learning material presented in VR. However, further research is warranted to delineate the specific conditions under which the integration of generative learning activities enhances learning. For instance, Klingenberg et al. (2023) recently highlighted how the effects of summarization in immersive VR may vary depending on the timing of its implementation. While we implemented the generative learning activities after the VR experience, students in the Parong and Mayer (2018) study took off their HMDs, wrote their summaries, and then proceeded to use the HMDs for the subsequent sections of the VR. Petersen et al. (2023) even implemented generative learning activities within the VR environment. From a research perspective, it would be interesting to replicate our study but implement the activities of *drawing* and *self-explaining* within VR. For instance, the prompt for *self-explaining* could be presented by the female voice, suggesting being Anne Frank, and participants could verbally perform the task instead of in written form.

One could also entertain the possibility that some generative learning tasks are more effective than others. The present comparison only allows us to draw conclusions on two of the eight proposed generative tasks (Fiorella & Mayer, 2016). A larger scale study that encompasses all eight might be able to uncover a hierarchy of effectiveness. Meta-analyses should also be able to show which generative learning tasks end up being most effective. However, the literature basis for such a review might be too thin at the current stage of research.

Furthermore, it should also be noted that the present paper investigated if one generative learning activity works better than the other, but not how. Such qualitative aspects could encompass surface-level variables like time or word count, as well as deeper criteria derived from subject-specific pedagogies.

In the context of the present study, we expanded upon the existing generative learning activities derived from the article by Fiorella and Mayer (2016) to include the aspect of *ego-involvement*. Since immersive VR, particularly in history education, engages its users on an emotional level, this extension seemed meaningful to us. When it comes to affective learning outcomes, future studies should develop further generative learning activities that prompt learners to make connections to their personal experiences and empirically investigate them.

We were unable to find a significant effect of generative learning tasks on knowledge acquisition. This does not align with the research laid out in the literature review that suggests there could be a connection between the two (Parong & Mayer, 2018; Petersen et al., 2023). It should be noted that we did not measure cognitive load in the present study. High cognitive load in virtual environments may lead to cognitive overload, which may overwhelm learners and interfere with the learning process (Makransky et al., 2019; Mayer et al., 2023). Future endeavors should include a measurement of cognitive load to be able to unravel how it interacts with affective and cognitive learning in the context of generative tasks and VR.

Our study points toward immersive VR as a practical way of learning about things that cannot be explored in traditional learning settings. Hence, VR should be reserved for experiences that are impossible, dangerous, or expensive in the real world (Bailenson, 2018), such as looking at the human blood stream (Parong & Mayer, 2018), talking to witnesses (Simine & Ch'ng, 2023), or visiting a museum in Amsterdam. In the case of our study, it would also have been possible to explore Anne Frank's hiding place using a desktop VR system (Anne Frank Stichting, 2018) but the level of immersion, presence, and sense of body ownership provided by a desktop VR system would not be the same (Slater et al., 2022).

5.3. Limitations

There are four primary limitations of the present study. First, although we were aware that the scale by Hartmann (2008) for assessing perspective taking exhibits only low internal consistency, due to a lack of suitable alternatives, we once again opted for this scale, assisted by the singular self-assessment items. While we revised the items of the Hartmann scale, the internal consistency in this study remains on the borderline of acceptability, thereby limiting the interpretation of the results on perspective taking. In addition, the partially low correlations between the instruments for perspective taking indicate that perspective taking could not be adequately assessed in our study. This is also due to the fact that the second self-assessment for perspective taking is a single item measure, contributing to the limited validity and reliability. In the future, new instruments for empathy and perspective taking should be developed, validated, and tested in empirical studies. Since there seems to be no valid instrument to assess perspective taking (that we know of) apart from the one by Hartmann (2008), which aims more toward historical perspective taking in a text-based manner, perhaps different methods of measurement should be explored. For example, an evaluation of the generative learning activities themselves or behavioral observations also appear to be more valid methods for recording perspective taking. The utilization of proxy variables should also be considered.

A second limitation is that the present study solely depended on the immediate assessment of learning outcomes and did not include delayed assessment. In a recent review, Parong (2021) noted that immersive VR has a more pronounced positive effect on delayed assessments compared to immediate assessments. In another study conducted by the authors (Mulders et al., 2023, 2025), where we attempted to induce attitude changes among students toward biodiversity in the Amazon rainforest following a VR training, we found that students require time and, importantly, opportunities for reflection to process the learning material. While immediate post-questionnaires did not capture any attitude changes, several weeks later, focus group interviews with the participants revealed initial shifts in attitudes and even behaviors. While this did not go along with mentions of higher general interest in learning, participants expressed their wish to have more VR-based interventions included in their regular lessons (Mulders et al., 2023). In general, future research should include more delayed assessments in order to be able to measure learning effects beyond the single VR experience. We also encourage future researchers to investigate more motivational aspects. The novelty effect (Miguel-Alonso et al., 2024), which in our study may have ensured that VR inexperienced learners in particular were motivated by the use of the technology to engage with the learning content, may be considered in follow-up studies, as well as the longer-term motivational effects on attitudes and behavior in relation to the learning content.

Third, as previously mentioned, students' self-assessment of their knowledge prior to the VR experience might be biased and potentially overestimated. Consequently, self-assessment appears to be only partially suitable for measuring prior knowledge and knowledge in general. While self-assessment is a less time-consuming way of measurement, future studies should employ more valid methods to assess cognitive learning outcomes (e.g. identical pre- and posttest measures, validated knowledge screenings).

Fourth, the generalizability of this study is limited by the somewhat homogenous sample consisting mainly of students in the field of Education Sciences and with little prior experience in VR. Future endeavors might want to expand the sample to include participants from different courses with a higher affinity for VR.

6. Conclusion

Our study looked at affective learning in a highly immersive environment. It delved into the instructional relevance of *drawing* and *self-explaining* as generative learning activities within an emotionally engaging VR environment, in comparison to a session without a generative activity. The findings reveal that the immersive VR itself does have a strong impact on affective learning and, thus, the targeted objectives of commemorative history education: The environment is able to foster perspective taking (seen as a measure of empathy and affective learning), but it did not contribute to knowledge acquisition.

Contrary to our expectations and previous research findings, the groups with additional generative learning activities after the VR experience did not yield better results in perspective taking and knowledge acquisition. Additional generative learning activities were not able to further enhance learning. Still, the question remains to what degree established principles of learning and instructional design can be applied to an environment that works under different preconditions. The results of our study partly contradict previous studies that provide evidence for the additional value of generative learning activities. This underscores the importance to further investigate the interdependencies of the variety of generative learning activities when learning with immersive VR and regarding a highly emotionally engaging topic. Follow-up studies should examine the conditions under which the activities are implemented (e.g. timing, inside or outside the VR) in more detail. They should contribute to a reexamination of the instructional feasibility and value of established principles of multimedia learning in immersive environments.

In total, our study focuses on affective learning in the context of history education and, thus, differs from previous studies that mostly have been conducted in less emotionally engaging environments and on less stressful topics primarily addressing cognitive learning objectives, like science (e.g. Parong & Mayer, 2018; Petersen et al., 2023). Immersive technologies, however, have the potential to evoke strong emotions, provide high levels of engagement, and include an attractive environment to address affective learning objectives which are often difficult to implement purposefully in everyday educational settings.

Notes

1. https://osf.io/ks28j/?view_only=77fcc7d0359b407dacce7e0e18e84dcd.
2. https://osf.io/ks28j/?view_only=77fcc7d0359b407dacce7e0e18e84dcd.

Disclosure statement

Q3

No potential conflict of interest was reported by the author(s).

Notes on contributors

Miriam Mulders is a postdoctoral researcher and teacher at the Learning Lab of the University of Duisburg-Essen. Her main research area is the implementation of virtual reality in education settings.

Kristian H. Träg is an academic staff member at the Learning Lab of the University of Duisburg-Essen. His main area of research is the effect of artificial intelligence on affective learning outcomes.

Lilly Kaninski is a student assistant at the Learning Lab. She is studying social sciences and computer science.

Lara Kirner is a research assistant at the Learning Lab and supports virtual reality research projects.

Michael Kerres holds the chair of Educational Technology and Instructional Design (Learning Lab) at the University of Duisburg-Essen. His expertise covers design-based research, media pedagogy, vocational, higher, and adult education, among others.

ORCID

Miriam Mulders http://orcid.org/0000-0003-0683-2310
Kristian H. Träg http://orcid.org/0009-0004-4557-2026
Lilly Kaninski http://orcid.org/0009-0009-3174-2318
Lara Kirner http://orcid.org/0009-0000-1298-8640
Michael Kerres http://orcid.org/0000-0002-7419-3023

References

Abadia, R., Calvert, J., & Dasika, R. (2019). *Effectiveness of using an immersive and interactive virtual reality learning environment to empower students in strengthening empathy and mastery learning* [Paper presentation]. ICCE 2019 - 27th International Conference on Computers in Education, Proceedings, Kenting, Taiwan, 495–504. https://doi.org/10.58459/icce.2019.495

Albus, P., Vogt, A., & Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Computers & Education*, *166*, 104154. https://doi.org/10.1016/j.compedu.2021.104154

Anne Frank Stichting (2018, September 25). The Secret Annex. *Anne Frank House*. https://www.annefrank.org/en/anne-frank/secret-annex/

Baceviciute, S., Terkildsen, T., & Makransky, G. (2021). Remediating learning from non-immersive to immersive media: Using EEG to investigate the effects of environmental embeddedness on reading in virtual reality. *Computers & Education*, *164*, 104122. https://doi.org/10.1016/j.compedu.2020.104122

Bachen, C. M., Hernández-Ramos, P. F., & Raphael, C. (2012). Simulating REAL LIVES: Promoting global empathy and interest in learning through simulation games. *Simulation & Gaming*, *43*(4), 437–460. https://doi.org/10.1177/1046878111432108

Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. W. W. Norton & Company.

Barrett, A., Pack, A., Monteiro, D., & Liang, H.-N. (2024). Exploring the influence of audience familiarity on speaker anxiety and performance in virtual reality and real-life presentation contexts. *Behaviour & Information Technology*, *43*(4), 787–799. https://doi.org/10.1080/0144929X.2023.2186145

Bertrand, P., Guegan, J., Robieux, L., McCall, C. A., & Zenasni, F. (2018). Learning empathy through virtual reality: Multiple strategies for training empathy-related abilities using body ownership illusions in embodied virtual reality. *Frontiers in Robotics and AI*, *5*, 26. https://doi.org/10.3389/frobt.2018.00026

Bowman, D. A., & McMahan, R. P. (2007). Virtual Reality: How much immersion is enough? *Computer Magazine*, *40*(7), 36–43. https://doi.org/10.1109/MC.2007.257

Bunnenberg, C. (2020). Mittendrin im historischen Geschehen? *geschichte für heute*, *13*(4), 45–59. https://doi.org/10.46499/1596.1602

Buttussi, F., & Chittaro, L. (2023). Acquisition and retention of spatial knowledge through virtual reality experiences: Effects of VR setup and locomotion technique. *International Journal of Human-Computer Studies*, *177*(C), 103067. https://doi.org/10.1016/j.ijhcs.2023.103067

Calvert, J., & Abadia, R. (2020). Impact of immersing university and high school students in educational linear narratives using virtual reality technology. *Computers & Education*, *159*, 104005. https://doi.org/10.1016/j.compedu.2020.104005

Chandler, P. (2009). Dynamic visualisations and hypermedia: Beyond the "Wow" factor. *Computers in Human Behavior*, *25*(2), 389–392. https://doi.org/10.1016/j.chb.2008.12.018

Cheng, K.-H., & Tsai, C.-C. (2020). Students' motivational beliefs and strategies, perceived immersion and attitudes towards science learning with immersive virtual reality: A partial least squares analysis. *British Journal of Educational Technology*, *51*(6), 2140–2159. https://doi.org/10.1111/bjet.12956

Chittaro, L., Corbett, C. L., McLean, G. A., & Zangrando, N. (2018). Safety knowledge transfer through mobile virtual reality: A study of aviation life preserver donning. *Safety Science*, *102*, 159–168. https://doi.org/10.1016/j.ssci.2017.10.012

Christopoulos, A., & Mystakidis, S. (2023). Gamification in education. *Encyclopedia*, *3*(4), 1223–1243. https://doi.org/10.3390/encyclopedia3040089

Cipresso, P., Giglioli, I. A. C., Raya, M. A., & Riva, G. (2018). The past, present, and future of virtual and augmented reality research: A network and cluster analysis of the literature. *Frontiers in Psychology*, *9*, 2086. https://doi.org/10.3389/fpsyg.2018.02086

Coban, M., Bolat, Y. I., & Goksu, I. (2022). The potential of immersive virtual reality to enhance learning: A meta-analysis. *Educational Research Review*, *36*, 100452. https://doi.org/10.1016/j.edurev.2022.100452

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220. https://doi.org/10.1037/h0026256

Dengel, A., Buchner, J., Mulders, M., & Pirker, J. (2023). Levels of immersive teaching and learning: Influences of challenges in the everyday classroom. In *Immersive education* (pp. 107–122). Springer International Publishing AG. https://doi.org/10.1007/978-3-031-18138-2_7

Dunning, D. (2011). Chapter five: The Dunning–Kruger effect: On being ignorant of one's own ignorance. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. *44*, pp. 247–296). Academic Press. https://doi.org/10.1016/B978-0-12-385522-0.00005-6

Filter, E., Eckes, A., Fiebelkorn, F., & Büssing, A. G. (2020). Virtual reality nature experiences involving wolves on YouTube: Presence, emotions, and attitudes in immersive and nonimmersive settings. *Sustainability*, *12*(9), 3823. https://doi.org/10.3390/su12093823

Finken, J., & Wölfel, M. (2023). Influence of immersive virtual reality on cognitive and affective learning goals. In D. Guralnick, M. E. Auer, & A. Poce (Eds.), *Innovative approaches to technology-enhanced learning for the workplace and higher education* (pp. 510–521). Springer International Publishing. https://doi.org/10.1007/978-3-031-21569-8_48

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

Fiorella, L., & Mayer, R. E. (2021). The generative activity principle in multimedia learning. In L. Fiorella & R. E. Mayer (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 339–350). Cambridge University Press. https://doi.org/10.1017/9781108894333.036

Freina, L., & Ott, M. (2015). A literature review on immersive virtual reality in education: State of the art and perspectives. *Proceedings of eLearning and Software for Education (eLSE)*, *1*, 133–141. https://doi.org/10.12753/2066-026X-21-020

Frentzel-Beyme, L., & Krämer, N. C. (2023, Spring). Historical time machines: Experimentally investigating potentials and impacts of immersion in historical VR on history education and morality. *Technology, Mind, and Behavior*, *4*(1). https://doi.org/10.1037/tmb0000099

Han, I., Shin, H. S., Ko, Y., & Shin, W. S. (2022). Immersive virtual reality for increasing presence and empathy. *Journal of Computer Assisted Learning*, *38*(4), 1115–1126. https://doi.org/10.1111/jcal.12669

Hartmann, U. (2008). [*Perspektivenübernahme als eine Kompetenz historischen Verstehens*]. [Doctoral thesis]. University of Göttingen. https://doi.org/10.53846/goediss-290

Hassenfeldt, C., Jacques, J., & Baggili, I. (2020). Exploring the learning efficacy of digital forensics concepts and bagging & tagging of digital devices in immersive virtual reality. *Forensic Science International Digital Investigation*, *33*, 301011. https://doi.org/10.1016/j.fsidi.2020.301011

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Horton, D., McIlraith, S. A., Wang, N., Majedi, M., McClure, E., & Wald, B. (2022). Embedding ethics in computer science courses: Does it work? Proceedings of *The 53rd ACM Technical Symposium on Computer Science Education*, *1*, 481–487. https://doi.org/10.1145/3478431.3499407

Janssen, D., Tummel, C., Richert, A., & Isenhardt, I. (2016). Virtual environments in higher education – Immersion as a key construct for learning 4.0. *International Journal of Advanced Corporate Learning (iJAC)*, *9*(2), 20. https://doi.org/10.3991/ijac.v9i2.6000

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515–1529. https://doi.org/10.1007/s10639-017-9676-0

Klingenberg, S., Fischer, R., Zettler, I., & Makransky, G. (2023). Facilitating learning in immersive virtual reality: Segmentation, summarizing, both or none? *Journal of Computer Assisted Learning*, *39*(1), 218–230. https://doi.org/10.1111/jcal.12741

Klingenberg, S., Jørgensen, M. L. M., Dandanell, G., Skriver, K., Mottelson, A., & Makransky, G. (2020). Investigating the effect of teaching as a generative learning strategy when learning through desktop and immersive VR: A media and methods experiment. *British Journal of Educational Technology*, *51*(6), 2115–2138. https://doi.org/10.1111/bjet.13029

Knoch, H. (2021). Das KZ als virtuelle Wirklichkeit. *Geschichte und Gesellschaft*, *47*(1), 90–121. https://doi.org/10.13109/gege.2021.47.1.90

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Leder, J., Horlitz, T., Puschmann, P., Wittstock, V., & Schütz, A. (2019). Comparing immersive virtual reality and powerpoint as methods for delivering safety training: Impacts on risk perception, learning, and decision making. *Safety Science*, *111*, 271–286. https://doi.org/10.1016/j.ssci.2018.07.021

Lewers, E. (2022). Durch Raum und Zeit?: Medienkritische Auseinandersetzung mit Virtual Reality im Geschichtsunterricht. *Medienimpulse*, *60*(2), 2. https://doi.org/10.21243/mi-02-22-20

Madary, M., & Metzinger, T. K. (2016). Real virtuality: A code of ethical conduct. recommendations for good scientific practice and the consumers of VR-technology. *Frontiers in Robotics and AI*, *3,* Article 3. https://doi.org/10.3389/frobt.2016.00003

Maister, L., Slater, M., Sanchez-Vives, M. V., & Tsakiris, M. (2015). Changing bodies changes minds: Owning another body affects social cognition. *Trends in Cognitive Sciences*, *19*(1), 6–12. https://doi.org/10.1016/j.tics.2014.11.001

Makransky, G. (2021). The immersion principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 296–302). Cambridge University Press.

Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology*, *113*(4), 719–735. https://doi.org/10.1037/edu0000473

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, *66*(5), 1141–1164. https://doi.org/10.1007/s11423-018-9581-2

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, *33*(3), 937–958. https://doi.org/10.1007/s10648-020-09586-2

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Martingano, A. J., Hererra, F., & Konrath, S. (2021). Virtual reality improves emotional but not cognitive empathy: A meta-analysis. *Technology, Mind, and Behavior*, *2*(1). https://doi.org/10.1037/tmb0000034

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

Mayer, R. E., Fiorella, L., & Stull, A. (2020). Five ways to increase the effectiveness of instructional video. *Educational Technology Research and Development*, *68*(3), 837–852. https://doi.org/10.1007/s11423-020-09749-6

Mayer, R. E., Makransky, G., & Parong, J. (2023). The promise and pitfalls of learning in immersive virtual reality. *International Journal of Human–Computer Interaction*, *39*(11), 2229–2238. https://doi.org/10.1080/10447318.2022.2108563

Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, *140*, 103603. https://doi.org/10.1016/j.compedu.2019.103603

Miguel-Alonso, I., Checa, D., Guillen-Sanz, H., & Bustillo, A. (2024). Evaluation of the novelty effect in immersive virtual reality learning experiences. *Virtual Reality*, *28*(1), 27. https://doi.org/10.1007/s10055-023-00926-5

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education*, *56*(3), 769–780. https://doi.org/10.1016/j.compedu.2010.10.020

Milk, C. (2015, March). Chris Milk: How virtual reality can create the ultimate empathy machine | TED Talk. *TED2015*. https://www.ted.com/talks/chris_milk_how_virtual_reality_can_create_the_ultimate_empathy_machine
Mulders, M. (2022). Jenseits von Medienvergleichen: Komplexe Zusammenhänge des Lernens in Virtual Reality am Beispiel des Anne Frank VR House [Dissertation thesis, University of Duisburg-Essen]. https://doi.org/10.17185/duepublico/76929
Mulders, M. (2023). Learning about victims of holocaust in virtual reality: The main, mediating and moderating effects of technology, instructional method, flow, presence, and prior knowledge. *Multimodal Technologies and Interaction*, *7*(3), 28. https://doi.org/10.3390/mti7030028
Mulders, M. (2024). Beyond comparing learning technologies: experiencing flow in virtual reality. In D. G. Sampson, D. Ifenthaler, & P. Isaías (Eds.), *Smart learning environments in the post pandemic era: Selected Papers from the CELDA 2022 conference* (pp. 39–56). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-54207-7_3
Mulders, M., Buchner, J., & Kerres, M. (2020). A framework for the use of immersive virtual reality in learning environments. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(24), 208–224. https://doi.org/10.3991/ijet.v15i24.16615
Mulders, M., Buchner, J., & Kerres, M. (2024). Virtual reality in vocational training: a study demonstrating the potential of a VR-based vehicle painting simulator for skills acquisition in apprenticeship training. *Technology, Knowledge and Learning*, *29*(2), 697–712. https://doi.org/10.1007/s10758-022-09630-w
Mulders, M., Träg, K. H., Kaninski, L., & Kirner, L. (submitted). *Past lives, present learners: Future directions for history education in virtual reality* [Manuscript submitted for publication].
Mulders, M., Träg, K., & Kirner, L. (2023). Artenvielfalt im Unterricht - Ergebnisse qualitativer Befragungen von Schüler:innen zum Einsatz von XR. Workshops der 21. Fachtagung Bildungstechnologien (DELFI), Aachen, Germany, 89–98. https://doi.org/10.18420/wsdelfi2023-24
Mulders, M., Träg, K., & Kirner, L. (2025). Go green: Evaluating an XR application on biodiversity in German secondary school classrooms. *Instructional Science*. https://doi.org/10.1007/s11251-024-09697-1
Mulders, M., Träg, K., Kirner, L., & Kaninski, L. (2024). Gestaltung immersiver Lernszenarien: Generative Aktivitäten beim Lernen mit dem Anne Frank VR House. Proceedings of DELFI Workshops 2024, Fulda, Germany. https://doi.org/10.18420/delfi2024-ws-02
Nachtigall, V., Yek, S., Lewers, E., Brunnenberg, C., & Rummel, N. (2022). Fostering cognitive strategies for learning with 360° videos in history education contexts. *Unterrichtswissenschaft*, *50*(4), 615–638. https://doi.org/10.1007/s42010-022-00154-x
Nakamura, L. (2020). Feeling good about feeling bad: Virtuous virtual reality and the automation of racial empathy. *Journal of Visual Culture*, *19*(1), 47–64. https://doi.org/10.1177/1470412920906259
Nash, K. (2018). Virtual reality witness: Exploring the ethics of mediated presence. *Studies in Documentary Film*, *12*(2), 119–131. https://doi.org/10.1080/17503280.2017.1340796
O'Brien, E., Konrath, S. H., Gruhn, D., & Hagen, A. L. (2013). Empathic concern and perspective taking: Linear and quadratic effects of age across the adult life span. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *68*(2), 168–175. https://doi.org/10.1093/geronb/gbs055
Parong, J. (2021). Multimedia learning in virtual and mixed reality. In L. Fiorella & R. E. Mayer (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 498–509) Cambridge University Press. https://doi.org/10.1017/9781108894333.051
Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785–797. https://doi.org/10.1037/edu0000241
Parong, J., & Mayer, R. E. (2021). Learning about history in immersive virtual reality: Does immersion facilitate learning? *Educational Technology Research and Development*, *69*(3), 1433–1451. https://doi.org/10.1007/s11423-021-09999-y
Passig, D., Tzuriel, D., & Eshel-Kedmi, G. (2016). Improving children's cognitive modifiability by dynamic assessment in 3D immersive virtual reality environments. *Computers & Education*, *95*, 296–308. https://doi.org/10.1016/j.compedu.2016.01.009
Patterson, T., Han, I., & Esposito, L. (2022). Virtual reality for the promotion of historical empathy: A mixed-methods analysis. *Theory & Research in Social Education*, *50*(4), 553–580. https://doi.org/10.1080/00933104.2022.2118091
Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787. https://doi.org/10.1016/j.concog.2013.04.016
Pellas, N., Kazanidis, I., & Palaigeorgiou, G. (2020). A systematic literature review of mixed reality environments in K-12 education. *Education and Information Technologies*, *25*(4), 2481–2520. https://doi.org/10.1007/s10639-019-10076-4
Petersen, G. B., Stenberdt, V., Mayer, R. E., & Makransky, G. (2023). Collaborative generative learning activities in immersive virtual reality increase learning. *Computers & Education*, *207*, 104931. https://doi.org/10.1016/j.compedu.2023.104931
R Core Team (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/
Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778. https://doi.org/10.1016/j.compedu.2019.103778

Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022). What is XR? Towards a framework for augmented and virtual reality. *Computers in Human Behavior*, *133*, 107289. https://doi.org/10.1016/j.chb.2022.107289
Roberts, J., Lyons, L., Cafaro, F., & Eydt, R. (2014). *Interpreting data from within: Supporting humandata interaction in museum exhibits through perspective taking* [Paper presentation]. Proceedings of the 2014 Conference on Interaction Design and Children, New York, USA, 7–16. https://doi.org/10.1145/2593968.2593974
Schutte, N. S., & Stilinović, E. J. (2017). Facilitating empathy through virtual reality. *Motivation and Emotion*, *41*(6), 708–712. https://doi.org/10.1007/s11031-017-9641-7
Segovia, K. Y., & Bailenson, J. N. (2009). Virtually true: Children's acquisition of false memories in virtual reality. *Media Psychology*, *12*(4), 371–393. https://doi.org/10.1080/15213260903287267
Sherman, W. R., & Craig, A. B. (2018). *Understanding virtual reality: Interface, application, and design*. Morgan Kaufmann.
Simine, S. A., & Ch'ng, E. (2023). Distributed remembering: Virtual reality testimonies and immersive witnessing. In S. Jones & R. Woods (Eds.), *The Palgrave handbook of testimony and culture* (pp. 301–323). Springer International Publishing. https://doi.org/10.1007/978-3-031-13794-5_14
Sirkkunen, E., Väätäjä, H., Uskali, T., & Rezaei, P. P. (2016). *Journalism in virtual reality: Opportunities and future research challenges* [Paper presentation]. Proceedings of the 20th International Academic Mindtrek Conference, New York, USA, 297–303. https://doi.org/10.1145/2994310.2994353
Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology (London, England: 1953)*, *109*(3), 431–433. https://doi.org/10.1111/bjop.12305
Slater, M., Banakou, D., Beacco, A., Gallego, J., Macia-Varela, F., & Oliva, R. (2022). A separate reality: An update on place illusion and plausibility in virtual reality. *Frontiers in Virtual Reality*, *3*. https://doi.org/10.3389/frvir.2022.914392
Slater, M., Lotto, B., Arnold, M. M., & Sanchez-Vives, M. V. (2009). How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicología*, *40*(2), 193–210.
Snelson, C., & Hsu, Y.-C. (2020). Educational 360-degree videos in virtual reality: A scoping review of the emerging research. *TechTrends*, *64*(3), 404–412. https://doi.org/10.1007/s11528-019-00474-3
Sousa Santos, B., Dias, P., Pimentel, A., Baggerman, J.-W., Ferreira, C., Silva, S., & Madeira, J. (2009). Head-mounted display versus desktop for 3D navigation in virtual reality: A user study. *Multimedia Tools and Applications*, *41*(1), 161–181. https://doi.org/10.1007/s11042-008-0223-2
Sweller, J. (2011). Cognitive load theory. In *The psychology of learning and motivation: Cognition in education* (Vol. *55*, pp. 37–76). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8
Thiele, M. (2023). Erinnern mit Instagram? @ichbinsophiescholl zwischen Identifikation, parasozialer Interaktion und Anmaßung. In G. Schweiger & M. Zichy (Eds.), *Zwischenmenschliche Beziehungen im Zeitalter des Digitalen: Ethische und interdisziplinäre Perspektiven* (pp. 79–93). Springer. https://doi.org/10.1007/978-3-662-66931-0_7
Thiele, M., & Thomas, T. (2023). Really?! Sophie Scholl on Instagram. An analysis of the journalistic discourse. *Journalism Research*, *6*(1), 6–31. https://doi.org/10.1453/2569-152X-23021-11525-en
Träg, K., & Mulders, M. (in press). *What the Amazon can't deliver: Lessons learned from Virtual Reality-based sustainability education*. Immersive Learning Research Network (iLRN). 2024
Vertigo Games & Knucklehead Studios (2019). *Anne Frank House VR* (Version 2.0) [Meta Quest]. https://www.meta.com/experiences/quest/1958100334295482/
Wittrock, M. C. (2010). Learning as a generative process. *Educational Psychologist*, *45*(1), 40–45. https://doi.org/10.1080/00461520903433554
Wohlgenannt, I., Simons, A., & Stieglitz, S. (2020). Virtual reality. *Business & Information Systems Engineering*, *62*(5), 455–461. https://doi.org/10.1007/s12599-020-00658-9
Wolf, M., Wehking, F., Montag, M., & Söbke, H. (2021). 360°-based virtual field trips to waterworks in higher education. *Computers*, *10*(9), 118. https://doi.org/10.3390/computers10090118
Wolfartsberger, J. (2019). Analyzing the potential of virtual reality for engineering design review. *Automation in Construction*, *104*, 27–37. https://doi.org/10.1016/j.autcon.2019.03.018
Wolgast, A., Tandler, N., Harrison, L., & Umlauft, S. (2020). Adults' dispositional and situational perspective-taking: A systematic review. *Educational Psychology Review*, *32*(2), 353–389. https://doi.org/10.1007/s10648-019-09507-y
Wong, E. Y. C., & Lee, P. T. Y. (2024). Virtual reality in transportation and logistics: A clustering analysis of studies from 2010 to 2023 and future directions. *Computers in Human Behavior*, *153*, 108082. https://doi.org/10.1016/j.chb.2023.108082
Wu, B., Yu, X., & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology*, *51*(6), 1991–2005. https://doi.org/10.1111/bjet.13023
Yang, W., & Wang, X. (2021). Why do generative learning strategy improve memory in VR?—Based on ICALM. *International Journal of Information and Education Technology*, *11*(12), 646–650. https://doi.org/10.18178/ijiet.2021.11.12.1576
Zhang, G. (2019). *Virtual simulation for history education* [Paper presentation]. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan. https://doi.org/10.1109/VR.2019.8797734