Title	Longitudinal effects in the effectiveness of educational virtual field trips
Author(s)	Jiayan Zhao, Jan Oliver Wallgrün, Pejman Sajjadi, Peter LaFemina, Kenneth Y. T. Lim, Jan P. Springer, and Alexander Klippel

This is the accepted author's manuscript of the following article:

Zhao, J., Wallgrün, J. O., Sajjadi, P., LaFemina, P., Lim, K. Y. T., Springer, J. P., & Klippel, A. (2022). Longitudinal effects in the effectiveness of educational virtual field trips. *Journal of Educational Computing Research,* 60(4), 1008-1034. https://doi.org/10.1177/07356331211062925

Longitudinal Effects in the Effectiveness of Educational Virtual Field Trips

Jiayan Zhao[1*], Jan Oliver Wallgrün[2], Pejman Sajjadi[2], Peter LaFemina[3], Kenneth Y. T. Lim[4], Jan P. Springer[1], and Alexander Klippel[5]

[1] Emerging Analytics Center, University of Arkansas at Little Rock, Little Rock, United States

[2] Department of Geography, The Pennsylvania State University, University Park, United States

[3] Department of Geosciences, The Pennsylvania State University, University Park, United States

[4] Centre for Research in Pedagogy and Practice, The Nanyang Technological University, Singapore

[5] Laboratory of Geo-information Science and Remote Sensing, Wageningen University & Research, Netherlands

Author Note

[*]Corresponding author: Jiayan Zhao

Email: jzhao1@ualr.edu

Telephone: +1 814-321-6982

Postal address: 1501 Rahling Road Apt. 920, Little Rock AR 72223 USA

Jiayan Zhao is a Postdoctoral Research Associate at the Emerging Analytics Center at the University of Arkansas at Little Rock. He is broadly interested in immersive technologies and their application in spatial cognition, wayfinding, geospatial data visualization, and place-based education.

Abstract

Virtual and immersive virtual reality, VR and iVR, provide flexible and engaging learning opportunities, such as virtual field trips (VFTs). Despite its growing popularity for education, understanding how iVR compared to non-immersive media influences learning is still challenged by mixed empirical results and a lack of longitudinal research. This study addresses these issues through an experiment in which undergraduate geoscience students attended two temporally separated VFT sessions through desktop VR (dVR) or iVR, with their learning experience and outcomes measured after each session. Our results show higher levels of enjoyment and satisfaction as well as a stronger sense of spatial presence in iVR students in both VFTs compared to dVR students, but no improvement in learning outcomes in iVR compared to dVR. More importantly, we found that there exists a critical interaction between VR condition and repeated participation in VFTs indicating that longitudinal exposure to VFTs improves knowledge performance more when learning in iVR than through dVR. These results suggest that repeated use of iVR may be beneficial in sustaining students' emotional engagement and compensating the initial deficiency in their objective learning outcomes compared to other less immersive technologies.

Keywords: virtual reality, novelty effect, immersion, place-based learning, STEM

Longitudinal Effects in the Effectiveness of Educational Virtual Field Trips

Introduction

Virtual field trips (VFTs) are simulated journeys transporting students to real-world places (Woerner, 1999). VFTs have been used for many years in higher education and training as a supplementary or alternative to more expensive and logistically problematic field-based activities (Hamilton et al., 2021; Tuthill & Klemm, 2002). While most VFTs use a combination of pictures, texts, videos, and hyperlinks (e.g., Cratnpton, 1999), the rising power of interactive 3D computer graphics enables educators to develop more engaging VFTs using virtual reality (VR) technologies (Han, 2021). VR has the potential to play an important role in learning and education, as realistic simulations that afford students an opportunity to visit, explore and engage in a believable environment (Klingenberg et al., 2020; Petersen et al., 2020). However, despite the promises, most evidence showing positive effects of VR is limited to single lab sessions (Klippel, Zhao, Jackson, et al., 2019; Lee & Wong, 2014; Legault et al., 2019; Zhao et al., 2020); there has been insufficient attention given to empirical research concerning its applied educational effectiveness and potential longitudinal effects (Luo et al., 2021; Makransky et al., 2020). It is pertinent to investigate how educational VR experiences can be transferred from lab environments and pilot projects into everyday teaching and learning using an evidence-based approach.

This article considers the use of VR-empowered VFTs within the context of undergraduate geoscience education. In our study, students in an introductory laboratory course participated in two temporally separated virtual trips to geological field sites as part of their regular curriculum. The rationale for this study comes from two points: (1) higher education instructors are currently facing challenges to develop their pedagogical practices of using VR on a regular basis; (2) owing to a lack of longitudinal research, very little is known about the real or long-term educational value of VFTs in improving student learning.

Since VR technology allows for experiencing VFTs through desktop screens or fully immersive head-mounted displays (HMDs), it is important to weigh benefits against constraints. Klippel et al. (2020) have proposed a trade-off between sensing capability and scalability of various VR systems: Greater sensing can create more enriching and realistic learning experiences, but at a cost of reduced scalability, that is, accessibility to fewer users or requiring substantial investments. Although lower-cost standalone HMDs such as the Oculus Quest have become readily available to many students, the use of VFTs with common desktop computers and web browsers is still the most accessible option (Luo et al., 2021). This consideration entered our experimental design as the first objective of this work, that is, to examine the outcomes of delivering VFTs through either immersive virtual reality (iVR) using an Oculus Quest HMD or desktop virtual reality (dVR) using a web browser. The results of our study can inform evidence-based recommendations for the integration of VR systems into place-based curricula.

VR can enhance students' enjoyment (Makransky & Lilleholt, 2018; Parong & Mayer, 2018; Zhao et al., 2020) and learning achievements (Klippel, Zhao, Jackson, et al., 2019; Lee & Wong, 2014). However, most studies disregard longitudinal aspects by having participants experience only a single exposure to VR learning experiences. It is possible that the previously demonstrated advantages of VR learning environments could be an interim effect due to the novelty of the technology and therefore not sustainable over an extended period of time (Pande et al., 2021). If this is the case, conclusions based on single exposures may have misjudged the benefits of VR, making longitudinal factors a critical aspect for future research. Therefore, the second objective of this study is to investigate the possible changes in students' learning experience and performance in VR over time. To achieve this objective, we had each group of students (iVR vs. dVR) participate in two different VFTs that were two weeks apart, and collected post-intervention data immediately after students' experience with each VFT. We believe that such longitudinal research examining the educational value of VFTs is especially important given the novelty of the technology and the difficulty in revealing the real effect of VR on individuals' learning over a single session.

The remainder of this article is organized as follows. First, we will summarize the theoretical framework that our work builds on. Next, we will briefly review the literature and provide testable hypotheses. The research methods will then be described. Finally, the implications of our findings will be discussed, followed by limitations and suggestions for future research.

Theoretical Framework

In our study, the preceding objectives were addressed through the theoretical framework of Biggs's (1993) model of Presage, Process, and Product (the 3-P model). Briefly, in Biggs's model, Presage factors include individual characteristics of the students such as age, gender, personality, prior knowledge, motivation, and expectations (Biggs & Moore, 1993), as well as the situational contexts which define the learning environment such as the curriculum content, teaching methods, assessment, and institutional procedures (Biggs & Moore, 1993). In terms of Process, the students' response to a learning approach and their ability to cross contexts of learning would depend, inter alia, on the learning *milieu* such as teaching practices, course structure, assessments, and curriculum content. The students' perceptions in turn derive from their individual characteristics such as their own preconceptions, motivations, and their ability to monitor, plan, and evaluate (Biggs & Moore, 1993). Finally, Product refers to the outcomes achieved. The quality of the outcomes is partially affected by the learning approaches employed by the learners. Outcomes can be categorized both cognitively and affectively. Cognitive outcomes refer to the acquisition of factual or conceptual knowledge and cognitive strategies associated with learning activities. Affective outcomes refer to how students feel about their learning (expressed satisfaction or specific perceptions of particular skill development), especially in situations of student evaluations (Lizzio et al., 2002).

Biggs's model of Presage, Process, and Product particularly informed the design of our study through its emphasis on how the passage of time and repeated exposures to the intervention may influence outcomes. For example – in terms of Process – Gordon et al. (2001) have reasonably argued that students' approaches to learning are influenced by their past success and failures. If an environment is perceived to be similar to those that students have encountered before, and if they have experienced success in that environment, it is more probable they will repeat those behaviors that they found helpful when they enter the new learning context. As a second example, in terms of Product, Biggs and Moore (1993) have argued that affective outcomes will remain with the students and are likely to have an effect on their future learning. They have elaborated that students' perceptions of their learning experiences feed back into the system (the Presage and Process aspects of the model) and thus modify learners' perceptions of future learning experiences in a continuous cycle.

In summary, we drew inspiration from Biggs's characterization of Presage, Process, and Product, when considering the impact metrics of VR-based learning, namely user experience, learning experience, and learning outcomes, respectively. In our study, user experience was evaluated using two variables: *perceived ease of use*, defined as the extent to which system use is free from additional effort (Davis, 1989); and *cybersickness*, defined as an uncomfortable side effect experienced by users of virtual interfaces (Nesbitt & Nalivaiko, 2015). The learning experience was evaluated using three variables: *spatial presence*, defined as the illusion of being physically present at a virtual field site (Lee, 2004; Slater, 2018); *field trip enjoyment*, defined as students perceiving the VFT activity enjoyable (Davis et al., 1992); and *satisfaction*, defined as students' overall pleasure and acceptance of the field-trip learning system (Liaw & Huang, 2013). Learning outcomes were measured in three different ways: *perceived learning*, defined as students' perceptions of understanding and knowledge gained (Alavi et al., 2002); *knowledge test performance*, assessed using multiple-choice questions asked during the VFT; and *lab grade*, the grade achieved in the course lab assignment.

Literature Review and Expectations

VR can be classified into two major categories: dVR and iVR. dVR involves displaying and manipulating a 3D virtual space on a regular 2D monitor with a conventional mouse and keyboard. In contrast, iVR allows users to view a 3D environment in 3D via an HMD and provides proprioceptive interaction through controllers and position tracking (Hamilton et al., 2021). An HMD has typically a larger display field of view than a desktop monitor and presents each eye with a different image, creating stereoscopic vision in environments with 3D models (Riecke et al., 2010). Compared to dVR, iVR has a higher level of technical fidelity or what is also referred to as *immersion*, defined as the capability of a system to deliver a vivid and matching illusion of reality to the senses of a user while shutting out the physical world (Slater & Wilbur, 1997). Thus, iVR affords the creation of 3D spatial representations on 3D displays (Simpson, 2020), naturalistic mapping between physical walking and virtual movements, and one-to-one scale direct manipulation with objects in the virtual world (Klingenberg et al., 2020).

The role of immersion in learning has been investigated by numerous studies designed to compare the effects of iVR with less immersive systems such as dVR. Overall, results indicate that learning in immersive virtual environments can induce a strong sense of spatial presence (Buttussi & Chittaro, 2018; Makransky et al., 2017), which in turn leads to a positive effect on the affective and motivational aspects of student engagement with the course content (Makransky et al., 2020; Parong & Mayer, 2018). However, less consistent results have been found for objective learning outcomes: Although some studies have highlighted the effectiveness of iVR in this regard (Chittaro & Buttussi, 2015; Krokos et al., 2019), other studies did not find any differences in learning performance (An et al., 2018; Leder et al., 2019; Makransky et al., 2020) or even reported opposite results, showing a better performance while using non-immersive media (Makransky et al., 2017; Makransky et al., 2020). These mixed results suggest that although iVR has a very powerful emotional impact, its actual learning effectiveness may depend on the nature of the task or knowledge domain (see Wu et al., 2020 for a review). The potential that iVR has to produce a feeling of spatial presence is a prerequisite for creating a *sense of place* (Turner & Turner, 2006). Sense of place is an authentic experience, which includes affective as well as cognitive relationships to the place (Najafi & Shariff, 2011). In place-based disciplines such as geosciences, biology, or geography, sense of place significantly contributes to the students' understanding of the studied content (Fitzsimons & Farren, 2016; Stainfield et al., 2000). Hence, we predict that for students in geosciences, iVR would result in better experiences and learning outcomes than dVR.

In human-computer interaction, novelty is considered an important variable associated with user activity in response to new technology or intervention (Card et al., 1983). VR (both dVR and iVR) is still relatively new and not commonly found in educational settings (Jerald, 2016). This “newness” may raise the enjoyment and motivation of the students, which could potentially lead to increased engagement and learning gains. Importantly, the novelty effect is inherently transitory (Huang et al., 2020): As users gain experience with the technology, their engagement and enjoyment may gradually fade (Chwo et al., 2018; Clark & Craig, 1992); learning experience and performance may also decrease as novelty wanes (Huang et al., 2020).

While there has been little research on the novelty effect of VR in the field of education, such influences are likely given the evidence from longitudinal studies that focus on training and learning via serious games (Boot et al., 2008; Hanus & Fox, 2015; Luse et al., 2013; Tsay et al., 2018). Most of these studies have relied on desktop or web-based interventions to present information. For example, Luse et al. (2013) investigated how attitudes of users of the desktop virtual world Second Life varied over a 7-week period. Results showed that users’ interest in utilizing the technology decreased over time. Tsay et al. (2018) examined the effectiveness of a gamified web learning system in a two-semester online course. They found that students’ engagement and learning performance gradually diminished in each semester. Considering the potential impact of novelty for desktop-based virtual experiences, we predict that for dVR students in our study, their affective responses and objective learning outcomes would decrease from the first to the second VFT. With respect to the novelty effect in iVR learning simulations, a prediction is much more difficult because only two longitudinal studies have been published (Huang et al., 2020; Pande et al., 2021). Pande et al. (2021) described (without statistical analysis) that students’ perceived learning and enjoyment tend to persist over multiple iVR sessions in a study about learning in environmental biology. Huang et al. (2020) engaged 50 students in different parts of an iVR tour through the solar system three times within a month.

The authors found that the motivation, spatial presence, and knowledge performance did not decline as students became familiar with the virtual environment.

Overall, the long-term impact of iVR on learning still needs to be explored and better understood, given the few studies devoted to this critical aspect. Following existing evidence, we hypothesize that iVR provides the learner with added novelty which should mitigate diminishing learning experience and outcomes compared to non-immersive media. Therefore, we further predict an interaction between level of immersion (iVR vs. dVR) and VFT (first vs. second VFT) with students' affective responses and performance decreasing over time in the dVR condition while diminishing less in the iVR condition.

Methods

Participants

In the Fall of 2020, 144 undergraduate students were recruited from an introductory geoscience laboratory course at a large American University. We attempted to ease data collection while protecting students from COIVD-19 infections. Students were given an online pre-screening survey to help identify those who were interested in iVR experiences ($N = 36$). Due to the limited number of equipment units and time constraints, we chose a subgroup of students to participate in the iVR condition ($N = 14$) with the primary goal of maintaining a balance of gender and facilitating the headset pickup process. The selected students had to confirm their agreement to use iVR equipment and pick up Oculus Quests from the library at their own convenience (as a way to maintain social distancing). Each Quest had two pre-installed VFTs (see the next section for details). The remaining students in the course were assigned to the dVR condition (*N* = 130) participating in the same VFTs using their laptop or desktop computer. All participants were compensated with course credits.

Materials

The materials consisted of a VFT entitled Sedimentary Rocks: Transition from shallow water carbonates to siliciclastics - the Ordovician Salona Formation or simply the “Salona Formation” and a VFT entitled Observing changes in sedimentary environments: The Reedsville and Bald Eagle Formations or the “Reedsville & Bald Eagle Formations”. The VFT experiences were required exercises and created based on actual field trips currently not possible due to COVID-19 constraints. The Salona lab exposed students to an outcrop of the Salona Formation consisting of shallow water carbonates to carbonate siliciclastic rock interbedded with multiple bentonite layers deposited during large volume, explosive eruptions. The bentonites help document the tectonic setting of the region and allow for numerical dating of the stratigraphic units, and the stratigraphy documents the change in source of the sediment. The Reedsville VFT centered on the Reedsville and Bald Eagle Formations that document the transition from shelf to slope sedimentary sequences (Reedsville Formation) to terrestrial sandstones (Bald Eagle Formation) from the Upper Ordovician of central Pennsylvania, roughly 458 to 450 million years ago. The content and knowledge learned in the two VFTs are not dependent upon each other.

Students in both labs investigated sedimentary rock types and depositional environments, and were able to measure stratigraphic thicknesses and construct stratigraphic columns.

The two field trips focused on different field sites and learning goals, the designs of the VFTs, though, use the same user interface, interactions, and activities. In both VFTs, students experienced the field sites through a series of interactive 360° images. Each 360° scene included short text instructions that might ask students to make an observation or perform specific activities such as navigating to a new location (Figure 1a). Additionally, some 360° scenes included audio narrations explaining specific features of the outcrop, multiple-choice questions (Figure 1b), or highlighted icons with access to additional information that was derived from textbooks or high-resolution photos (Figure 1c, d). During the VFTs, students could observe the field site from a pseudo-aerial perspective (about 8 meters above ground), control the playing of audio (e.g., pause and rewind; Figure 2a, b), and take pictures of interesting features of the outcrop (Figure 2c, d). As part of their practice of practical skills, students in each VFT were guided through a measurement activity using a 3D model of the outcrop (Figure 2e, f). The measurement results along with a screenshot of the outcrop model were sent to students after the virtual experience so that they could make a stratigraphic column as required for the lab assignment.

-Insert Figure 1 here-

-Insert Figure 2 here-

Leveraging the cross-platform capabilities of the Unity3D game engine, the Salona and the Reedsville VFTs were implemented as a dVR WebGL application that could run in a web browser, and as an iVR app experienced through the Oculus Quest VR headset. The Quest HMD has a field of view of 110°, with a refresh rate of 72Hz, and a display resolution of 2880×1600 pixels.

In the dVR condition, students sat in front of a computer screen, clicked and dragged the mouse in a 360° image to look around, and pressed the left mouse button to interact with the virtual field site. In the iVR condition, students put on the Quest HMD, stood in the center of the tracking space, and used hand-held controllers to interact with the environment. The content, features, and storyline of each VFT were the same across both VR conditions. The measurement activity was designed differently in the two conditions. In the iVR condition students operated a virtual ruler attached to their hand controller with full freedom of movement (Figure 2e). In the dVR condition, students were placed in front of the outcrop model. To measure the thickness of rock layers, dVR students hovered the mouse pointer over the outcrop to add nodes on the rock surface; each pair of nodes was connected by a straight line segment for length computation (Zhao et al., 2020; Figure 2f).

Procedure and Measures

The two VFTs, including introduction, empirical data collection, and lab assignments, were spread out over a 4-week period (October 26th-November 20th, 2020). An overview of the procedure is provided in Figure 3.

-Insert Figure 3 here-

In the first week of the study, students were provided with a URL link to the Salona lab by their teaching assistants (TAs). The link opened a website that commenced with an introduction to the lab and the consent form, in which the nature of the study was explained to the students. After reading the form and checking a box to give their consent, students proceeded to the next page to enter a pre-questionnaire, in which they provided demographic information (gender, age, and academic year) and then completed self-report, five-point Likert scale measures of spatial ability and technology enjoyment (see Appendix 1 for a list of items and sources). These basic characteristics can provide insights into the comparability of the dVR and iVR conditions.

After completing the pre-questionnaire, students went through a 20-minute VFT to the Salona Formation (VFT 1) either in a web browser (dVR condition) or using the Oculus Quest headset (iVR condition). During the VFT, students interacted with a virtual lab pad where they received six multiple-choice questions (Appendix 2) related to field observations or the factual knowledge that they just learned from the audio narration (Figure 1b). Students were allowed multiple attempts for each question with explanatory feedback until they found the correct answer. The knowledge test performance was calculated as the average number of attempts students made per question. Thus, lower scores on this test were indicative of better knowledge gained from the VFT activities (minimum = 1).

The Salona VFT was followed by a post-questionnaire with a set of self-report, five-point Likert scales to measure user experience (including *perceived ease of use* and *cybersickness*), learning experience (including *spatial presence*, *field trip enjoyment*, and *satisfaction*), and *perceived learning*. The list of items and sources is included in Appendix 1.

A complete lab session, including the Salona VFT and the pre- and post-questionnaires, took about 45 minutes in both VR conditions. Students who chose not to participate in the study were directed to the web version of the VFT without pre- or post-questionnaires. After the Salona lab experience, all students in the course were required to complete a lab assignment and return their work by the next lab period. The lab assignments were graded by the course TAs, and grades (scale 0–100) were recorded as a learning outcome variable to be evaluated in this study. The assignment asked students to report their observations of the outcrop, including a stratigraphic column, a sketch map showing the boundary between paleo-continents, and a written summary answering questions regarding depositional and tectonic settings.

In the third week of the study, students were instructed to participate in a VFT to the Reedsville & Bald Eagle Formations (VFT 2). The steps were similar to the Salona lab except that students skipped the pre-questionnaire if they had attended the previous VFT as a study participant; only those who consented to participate in both lab sessions were included for analysis. Once they had completed the Reedsville VFT, students filled out a post-questionnaire, which was identical to the one administered after the Salona VFT, and were given a week to complete the lab assignment. The assignment used in the Reedsville lab consisted of a stratigraphic column, answers to specific questions regarding the depositional environment, and a written description drawing together the observations in a discussion about the paleo-environment (Klippel, Zhao, Jackson, et al., 2019).

Results

Demographic characteristics were compared between the dVR and iVR conditions using the independent samples t-test (Mann-Whitney U test in case of non-normal distribution) or the chi-square test of independence as appropriate (see Table 1). Results highlighted that no significant differences were found on gender [X^2 (2, N = 144) = 3.06, p = .217], academic year [X^2 (3, N = 144) = 1.67, p = .643], age ($t_{(16.08)}$ = -0.86, p = .402), or spatial ability ($t_{(17.53)}$ = -1.59, p = .130). Nonetheless, students in the iVR condition reported significantly higher technology enjoyment than those in the dVR condition (U = 574.50, p = .021).

-Insert Table 1 here-

User experience (perceived ease of use and cybersickness), learning experience (spatial presence, field trip enjoyment, and satisfaction), and learning outcomes (perceived learning, knowledge test performance, and lab grade) were investigated with eight two-condition (dVR/iVR) by two-VFT (Salona/Reedsville) mixed between-within subjects ANOVAs. The results are presented in Table 2.

-Insert Table 2 here-

User Experience

For perceived ease of use, the first row of Table 2 shows that there was no significant main effect for condition ($F_{(1,142)} = 2.90$, $p = .091$, $\eta_p^2 = .02$) or VFT ($F_{(1,142)} = 2.09$, $p = .151$, $\eta_p^2 = .01$), or significant interaction ($F_{(1,142)} = 0.17$, $p = .679$, $\eta_p^2 = .001$). Similarly, for cybersickness the results in the second row of Table 2 indicate that there was no significant main effect for condition ($F_{(1,142)} = 1.38$, $p = .242$, $\eta_p^2 = .01$) or VFT ($F_{(1,142)} = 0.71$, $p = .399$, $\eta_p^2 = .005$), or significant interaction between condition and VFT ($F_{(1,142)} = 3.20$, $p = .076$, $\eta_p^2 = .02$).

Learning Experience

For spatial presence, a significant main effect for condition was observed (third row of Table 2), indicating that students reported a significantly higher level of feeling present when using iVR compared to dVR ($F_{(1,142)} = 5.86$, $p = .017$, $\eta_p^2 = .04$). The main effect for VFT ($F_{(1,142)} = 2.88$, $p = .092$, $\eta_p^2 = .02$) and the interaction ($F_{(1,142)} = 0.14$, $p = .713$, $\eta_p^2 < .001$) were not significant.

With regard to field trip enjoyment, row 4 of Table 2 shows that there were significant main effects for condition ($F_{(1,142)} = 10.86$, $p = .001$, $\eta_p^2 = .07$) and VFT ($F_{(1,142)} = 5.16$, $p = .025$, $\eta_p^2 = .04$) but not a significant interaction ($F_{(1,142)} = 0.09$, $p = .762$, $\eta_p^2 < .001$). These results indicate that students rated their enjoyment of the first (Salona) VFT higher than the second (Reedsville) VFT and enjoyed using iVR more than dVR.

For satisfaction, the results (row 5 of Table 2) indicate that students in the iVR condition were more satisfied with their lab experiences than those in the dVR condition ($F_{(1,142)} = 4.27$, $p = .041$, $\eta_p^2 = .03$), and for both conditions there was a significant drop in satisfaction from the Salona to the Reedsville VFT ($F_{(1,142)} = 9.53$, $p = .002$, $\eta_p^2 = .06$). There was no significant interaction between condition and VFT ($F_{(1,142)} = 2.37$, $p = .126$, $\eta_p^2 = .02$).

Learning Outcomes

For perceived learning (see row 6 of Table 2), the significant main effect for VFT ($F_{(1,142)} = 6.45$, $p = .012$, $\eta_p^2 = .04$) indicates that there was a decrease in perceived learning from the Salona to the Reedsville VFT for both VR conditions. The main effect for condition ($F_{(1,142)} = 1.56$, $p = .214$, $\eta_p^2 = .01$) and the interaction ($F_{(1,142)} = 0.13$, $p = .720$, $\eta_p^2 < .001$) were not significant.

For knowledge test performance, the results presented in row 7 of Table 2 (with lower scores indicating better performance) revealed significant main effects for condition ($F_{(1,142)} = 5.12$, $p = .02$, $\eta_p^2 = .04$) and VFT ($F_{(1,142)} = 15.40$, $p < .001$, $\eta_p^2 = .10$), and a significant interaction between condition and VFT ($F_{(1,142)} = 6.01$, $p = .015$, $\eta_p^2 = .04$). Bonferroni corrected pairwise comparisons for each condition individually showed that there was a significant improvement in knowledge test performance from the Salona to the Reedsville VFT for the dVR condition ($p = .023$, $g = .27$) as well as the iVR condition ($p < .001$, $g = 1.26$). Furthermore, Bonferroni corrected pairwise comparisons at each time point individually showed that dVR students significantly outperformed iVR students in the first VFT (Salona; $p = .004$, $g = -.82$) but not in the second VFT (Reedsville; $p = .807$, $g = -.07$).

We observed a different pattern for lab grades (final row of Table 2; score range 0–100, with lower scores indicating poorer performance); there was no significant main effect for condition ($F_{(1,99)} = 0.001$, $p = .974$, $\eta_p^2 < .001$) or VFT ($F_{(1,99)} = 0.15$, $p = .703$, $\eta_p^2 = .001$), or significant interaction ($F_{(1,99)} = 0.30$, $p = .585$, $\eta_p^2 < .003$). Note that some of the students did not submit their lab assignments; the statistical analysis of lab grade was thus based on the data obtained from 91 students (39 or 30.0% removed) in the dVR condition and 12 students (2 or 16.7% removed) in the iVR condition.

Discussion

Before discussing the results of our objectives-oriented evaluation, we briefly consider the dVR/iVR-comparability of user experiences. An immersive interface should be natural to use and enable learners to interact with virtual objects and tools through body movements or gestures. Such embodied interaction, however, can be unfamiliar compared to mouse and keyboard. Additionally, research indicates that cybersickness may prevent students from learning with iVR simulations (Petri et al., 2020; Rebenitsch & Owen, 2016). In response to these concerns, we kept the virtual experience short (20–25 min), avoided extended locomotion, designed straightforward interactions (pointing & selecting), used 360° images (other than videos) to provide context information, and used textual annotations along with audio instructions to guide students through the learning process. We looked into user experience factors and found that students in the two conditions expressed high ease of use and little problems with cybersickness throughout the study. The results, while limited in scope, indicate that both immersive and desktop VFTs can provide students with an efficient and easy-to-use environment for place-based learning.

The first objective investigated whether immersion could lead to better learning experiences and outcomes. Consistent with previous research (see Section ‘Literature review and expectations’), our results indicate that students who experienced immersive virtual trips (iVR) reported significantly higher ratings of spatial presence, enjoyment, and satisfaction than students who experienced the same content through a desktop computer (dVR). Regarding learning outcomes, we predicted an advantage of learning with iVR over learning with dVR. However, the multiple-choice knowledge tests show that students in the dVR condition had a better (Salona VFT) or similar (Reedsville VFT) performance than those in the iVR condition.

At first sight, our findings seem to corroborate previous studies, which indicate that students like learning in immersive virtual environments and experience a stronger sense of spatial presence, but they have similar or even inferior learning performance compared to traditional media such as desktop computers (Makransky et al., 2017; Makransky et al., 2020). This can be potentially explained by the *cognitive theory of multimedia learning* (Mayer, 2014), according to which multimedia learning is demanding and influenced by the cognitive load of learner activities (see also Petersen et al., 2020). This perspective is particularly relevant for learning in a highly immersive environment, where the perceptual realism and richness of sensory information can enhance emotional engagement and sense of presence but may create more extraneous cognitive load that distracts learners from processing essential materials (Parong & Mayer, 2018). In contrast, a desktop-based virtual environment presents 3D representations of field sites on a flat screen with very limited sensory stimuli, leaving ample cognitive resources for comprehension, observation, and recall of knowledge.

Not all of our measurements showed such a clear pattern of media effects on learning outcomes. There were no significant differences between the dVR and iVR conditions for lab-grade across both VFTs. As indicated in Section ‘Procedure and measures’, students were given one extra week to complete the lab assignment, meaning that they did have some opportunity to use other resources in addition to the virtual field experience to help them with the assignments. Nonetheless, the mixed pattern of learning outcomes is consistent with a meta-analysis performed by Luo et al. (2021), who found an overall small effect size for the comparison between HMD and non-HMD interventions, and corroborates Spector’s (2020) argument that advances in educational technology do not guarantee improved learning.

Similarly, in both conditions similar ratings of perceived learning were reported, perhaps because both the desktop and immersive versions of each VFT presented the same teaching content and factual information. It is also possible that students’ perceived learning was influenced by the perceived difficulty of knowledge-related questions. In each lab, questionnaires about perceived learning were presented after the completion of the multiple-choice knowledge test and the VFT. Of the affective variables measured in the current study, perceived learning most closely relates to objective learning outcomes (Alqurashi, 2018). Thus, students’ performance in the multiple-choice test may have affected their perceptions of knowledge gain.

The second objective of this study was to investigate if and how VR-empowered field trip experiences affect student learning over time. From the Salona to the Reedsville VFT, we observed significant declines in field trip enjoyment, perceived learning, and satisfaction regardless of the medium. This result can be explained for the dVR condition, where students attended VFTs on their own computers with limited screen space and interaction modalities. Due to their familiarity with the hardware and software setups, the novelty of the virtual experience might have quickly faded. Consequently, in the second VFT, students were more likely to feel bored and less interested in learning with the application. Surprisingly, iVR students, too, significantly lowered their ratings of learning experience from the Salona to the Reedsville VFT. Users with the HMD often need more time and energy to become familiar with the interface compared to desktop users. Huang (2020) suggests that this familiarization process helps to maintain motivation and engagement with the learning content after novelty effects wear off. However, this may not be the case in the current study as students only needed to follow through a linear sequence of field activities with simple interactions. In fact, most students perceived the technology as easy to use without extra effort, which may, in turn, result in a decreasing novelty effect for affective measures in both the iVR and dVR conditions. Another possible explanation is that it is more challenging to excite and engage students toward the end of the semester (Ertmer et al., 1996). The Reedsville VFT (VFT 2) was taken one week before the university switched to remote instruction. During that time, academic and transition pressures may have added stress on our participants, especially among residential students, thereby reducing their motivation and interest for learning from the VFT method.

Despite the impacts of novelty and end-of-semester pressure on the learning experience, we found that students in the iVR condition reported significantly higher ratings of spatial presence, enjoyment, and satisfaction compared to those in the dVR condition across both VFT sessions. This finding indicates that even after novelty effects are diminished, learning through iVR is still perceived to be more enjoyable and satisfying than learning with dVR. Furthermore, there was a significant interaction between condition and VFT for performance in the multiple-choice knowledge tests. That is, iVR students had worse test performance than dVR students in the first VFT, possibly due to the cognitive overload associated with immersive technology use or the added distractions of new technological affordances; but this difference disappeared in the second VFT, with iVR students improving their knowledge test performance more than dVR students.

Taken together and viewed through the lens of Biggs's (1993) model of Presage, Process, and Product (described in Section "Theoretical framework"), the results suggest that for learners with longitudinal exposure to VR simulations, there is an advantage of iVR over dVR with respect to fostering emotional engagement but not at the cost of learning effectiveness. iVR is often considered an overwhelming experience for students (Makransky et al., 2017; Pande et al., 2021; Parong & Mayer, 2018); our study corroborates this concern and further suggests that iVR learners may benefit specifically from longitudinal exposure to immersive virtual environments.

Some aspects are important for contextualizing these findings. First, the number of participants in the iVR condition is relatively small due to limited HMD equipment available for the class. Increasing the number of iVR participants should lead to more reliable conclusions, and put us in a better position to assess the learning effectiveness of VR at a fine-grained level (e.g., to identify the characteristics of learners that could predict the effective use of VFTs) rather than merely across the dVR/iVR distinction as a whole (e.g., Plechatá et al., 2019). Second, students in the iVR condition reported significantly higher technology enjoyment than those in the dVR condition. This may be because only students who were interested in immersive technologies would be likely to participate in the iVR condition (which is a quasi-experimental design). The fact that our iVR samples were self-selected could have introduced bias into the evaluation results (see Hauser et al., 2018 for an overview). The data reported in our previous VFT studies with voluntary (Klippel, Zhao, Oprean, et al., 2019) and mandatory (Klippel, Oprean, et al., 2019) student samples support this concern. Although employing a randomized controlled design would avoid self-selection issues, the practical value of quasi-experimental studies can be greater because they are usually conducted in more natural settings and, therefore, contribute to our knowledge of integrating iVR into formal education (Wu et al., 2020). Future quasi-experimental research that replicates our results with larger iVR samples in residential classroom settings is warranted to provide clarification on the generalizability and validity of findings from the current investigation. Third, while we have attempted to make the Salona and Reedsville VFTs as comparable as possible, some differences persist that may have contributed to the results. A repeated-measures design with the order of VFTs counterbalanced across conditions can help tease apart the effect of time from the effect of different VFTs. Fourth, our study examined VR immersion, subjective experience, and learning outcomes over two VFT sessions within one month. Although this approach is a valuable contribution, the number of sessions/checkpoints and the span of time are limited as compared to related longitudinal studies in the field of education (e.g., Tsay et al., 2018). A remaining research question is to ask how VR impacts learning over more sessions or a longer time period (i.e., in *true* longitudinal designs and timescales).

Conclusion

Evaluating the long-term efficacy of VR is challenging in educational research due to limited access to VR facilities and the laborious nature of repeated data collection. Its outcome can contribute to addressing the concerns of practitioners in using this innovative technology as a pedagogical tool in both traditional classroom and remote learning environments. Leveraging real-life educational settings, the current study investigated differences and trends in students' learning experience and outcomes for two levels of immersion, dVR and iVR, and over two temporally separated VFT sessions. Despite the limited number of sessions and the small iVR sample size, this study is one of the first to examine the longitudinal effects of VR immersion on student learning in place-based disciplines. Our results are meaningful in several ways. First, they replicate previous research (e.g., Makransky et al., 2020) indicating that students' first-time learning in iVR compared to dVR increases presence, enjoyment, and satisfaction but not learning outcomes. Second, learning in VFTs over two sessions led to improved knowledge test performance but at the cost of reduced learning experience and perceived learning in both VR conditions. Third, and most importantly, repeated use of iVR may be beneficial in maintaining student engagement and satisfaction and compensating the initial deficiency in objective learning outcomes compared to less immersive systems such as dVR. These findings highlight the necessity of such longitudinal approaches to investigate and better understand how emerging VR applications can best be designed and implemented as part of regular education programs.

References

Alavi, M., Marakas, G. M., & Yoo, Y. (2002). A Comparative Study of Distributed Learning Environments on Learning Outcomes. *Information Systems Research*, *13*(4), 404–415. https://doi.org/10.1287/isre.13.4.404.72

Alqurashi, E. (2018). Predicting student satisfaction and perceived learning within online learning environments. *Distance Education*, *40*(1), 133–148. https://doi.org/10.1080/01587919.2018.1553562

An, B., Matteo, F., Epstein, M., & Brown, D. E. (2018). Comparing the Performance of an Immersive Virtual Reality and Traditional Desktop Cultural Game. In *2nd International Conference on Computer-Human Interaction Research and Applications* (pp. 54–61). https://doi.org/10.5220/0006922800540061

Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *The British Journal of Educational Psychology*, *63 (Pt 1)*, 3–19. https://doi.org/10.1111/j.2044-8279.1993.tb01038.x

Biggs, J., & Moore, P. (1993). *The process of learning* (3rd ed.). Prentice Hall.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica*, *129*(3), 387–398. https://doi.org/10.1016/j.actpsy.2008.09.005

Buttussi, F., & Chittaro, L. (2018). Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics*, *24*(2), 1063–1076. https://doi.org/10.1109/TVCG.2017.2653117

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Erlbaum.

Chittaro, L., & Buttussi, F. (2015). Assessing Knowledge Retention of an Immersive Serious Game vs. A Traditional Education Method in Aviation Safety. *IEEE Transactions on Visualization and Computer Graphics*, *21*(4), 529–538. https://doi.org/10.1109/TVCG.2015.2391853

Chou, S.-W., & Liu, C.-H. (2005). Learning effectiveness in a Web-based virtual learning environment: A learner control perspective. *Journal of Computer Assisted Learning*, *21*(1), 65–76. https://doi.org/10.1111/j.1365-2729.2005.00114.x

Chwo, G. S. M., Marek, M. W., & Wu, W.-C. V. (2018). Meta-analysis of MALL research and design. *System*, *74*(1), 62–72. https://doi.org/10.1016/j.system.2018.02.009

Clark, R. E., & Craig, T. G. (1992). Research and Theory on Multi-Media Learning Effects. In M. Giardina (Ed.), *NATO ASI Series. Interactive Multimedia Learning Environments: Human Factors and Technical Considerations on Design Issues.* Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-77705-9_2

Cratnpton, J. W. (1999). Integrating the Web and the Geography Curriculum: The Bosnian Virtual Fieldtrip. *Journal of Geography*, *98*(4), 155–168. https://doi.org/10.1080/00221349908978875

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace1. *Journal of Applied Social Psychology*, *22*(14), 1111–1132. https://doi.org/10.1111/j.1559-1816.1992.tb00945.x

Ertmer, P. A., Newby, T. J., & MacDougall, M. (1996). Students' Responses and Approaches to Case-Based Instruction: The Role of Reflective Self-Regulation. *American Educational Research Journal*, *33*(3), 719–752. https://doi.org/10.3102/00028312033003719

Fitzsimons, S., & Farren, M. (2016). A brave new world: Considering the pedagogic potential of Virtual World Field Trips (VWFTs) in initial teacher education. *International Journal for Transformative Research*, *3*(1), 9–15. https://doi.org/10.1515/ijtr-2016-0002

Gordon, C., Simpson, T., & Debus, R. (2001). *Improving quality learning in a pre-service teacher education programme*. http://www.aare.edu.au/ 01pap/gor01441.htm

Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, *8*(1), 1–32. https://doi.org/10.1007/s40692-020-00169-2

Han, I. (2021). Immersive virtual field trips and elementary students' perceptions. *British Journal of Educational Technology*, *52*(1), 179–195. https://doi.org/10.1111/bjet.12946

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, *80*(1), 152–161. https://doi.org/10.1016/j.compedu.2014.08.019

Hauser, D., Paolacci, G., & Chandler, J. J. (2018). Common Concerns with MTurk as a Participant Pool: Evidence and Solutions. Advance online publication. https://doi.org/10.31234/osf.io/uq45c

Hegarty, M. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, *30*(5), 425–447. https://doi.org/10.1016/S0160-2896(02)00116-2

Helland, A., Lydersen, S., Lervåg, L.-E., Jenssen, G. D., Mørland, J., & Slørdal, L. (2016). Driving simulator sickness: Impact on driving performance, influence of blood alcohol concentration, and effect of repeated simulator exposures. *Accident; Analysis and Prevention*, *94*, 180–187. https://doi.org/10.1016/j.aap.2016.05.008

Huang, W. (2020). *Investigating the Novelty Effect in Virtual Reality on STEM Learning* [PhD thesis]. Arizona State University. https://repository.asu.edu/items/57391

Huang, W., Roscoe, R. D., Johnson‐Glenberg, M. C., & Craig, S. D. (2020). Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion. *Journal of Computer Assisted Learning*, *46*(2), 26. https://doi.org/10.1111/jcal.12520

Jerald, J. (2016). *The VR book: Human-centered design for Virtual Reality*. Morgan & Claypool Publishers.

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

Klippel, A., Oprean, D., Zhao, J., Wallgrün, J. O., LaFemina, P. C., Jackson, K. L., & Gowen, E. (2019). Immersive learning in the wild – a progress report. In D. Beck, A. Peña-Rios, T. Ogle, L. Morgado, C. Eckhardt, J. Pirker,. ., & J. Richter (Eds.), *5th Annual Immersive Learning Research Network Conference.* Springer International Publishing.

Klippel, A., Zhao, J., Jackson, K. L., LaFemina, P. C., Stubbs, C., Wetzel, R., Blair, J., Wallgrün, J. O., & Oprean, D. (2019). Transforming Earth Science Education Through Immersive Experiences: Delivering on a Long Held Promise. *Journal of Educational Computing Research*, *10*(2). https://doi.org/10.1177/0735633119854025

Klippel, A., Zhao, J., Oprean, D., Wallgrün, J. O., Stubbs, C., LaFemina, P. C., & Jackson, K. L. (2019). The value of being there: Toward a science of immersive virtual field trips. *Virtual Reality*, *1*(4), 24. https://doi.org/10.1007/s10055-019-00418-5

Klippel, A., Zhao, J., Sajjadi, P., Wallgrün, J. O., Bagher, M. M., & Oprean, D. (2020). Immersive Place-based Learning – An Extended Research Framework. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 449–454). IEEE. https://doi.org/10.1109/VRW50115.2020.00095

Krokos, E., Plaisant, C., & Varshney, A. (2019). Virtual memory palaces: Immersion aids recall. *Virtual Reality*, *23*(1), 1–15. https://doi.org/10.1007/s10055-018-0346-3

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*(2), 271–286. https://doi.org/10.1016/j.ssci.2018.07.021

Lee, E. A.-L., & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers & Education*, *79*, 49–58. https://doi.org/10.1016/j.compedu.2014.07.010

Lee, K. M. (2004). Presence, Explicated. *Communication Theory*, *14*(1), 27–50. https://doi.org/10.1111/j.1468-2885.2004.tb00302.x

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive Virtual Reality as an Effective Tool for Second Language Vocabulary Learning. *Languages*, *4*(1), 13. https://doi.org/10.3390/languages4010013

Liaw, S.-S., & Huang, H.-M. (2013). Perceived satisfaction, perceived usefulness and interactive learning environments as predictors to self-regulation in e-learning environments. *Computers & Education*, *60*(1), 14–24. https://doi.org/10.1016/j.compedu.2012.07.015

Lizzio, A., Wilson, K., & Simons, R. (2002). University Students' Perceptions of the Learning Environment and Academic Outcomes: Implications for theory and practice. *Studies in Higher Education*, *27*(1), 27–52. https://doi.org/10.1080/03075070120099359

Luo, H., Li, G., Feng, Q., Yang, Y., & Zuo, M. (2021). Virtual reality in K‐12 and higher education: A systematic review of the literature from 2000 to 2019. *Journal of Computer Assisted Learning*, *37*(3), 887–901. https://doi.org/10.1111/jcal.12538

Luse, A., Mennecke, B., & Triplett, J. (2013). The changing nature of user attitudes toward virtual world technology: A longitudinal study. *Computers in Human Behavior*, *29*(3), 1122–1132. https://doi.org/10.1016/j.chb.2012.10.004

Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2020). 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.* Advance online publication. 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., Terkildsen, T. S., & Mayer, R. E. (2017). 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

Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *Cambridge Handbooks in Psychology. The Cambridge handbook of multimedia learning* (pp. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, *60*(1), 48–58. https://doi.org/10.1080/02701367.1989.10607413

Najafi, M., & Shariff, M. (2011). The concept of place and sense of place in architectural studies. *International Journal of Human and Social Sciences*, *6*(3), 187–193.

Nesbitt, K., & Nalivaiko, E. (2015). Cybersickness. In N. Lee (Ed.), *Encyclopedia of Computer Graphics and Games.* Springer International Publishing. https://doi.org/10.1007/978-3-319-08234-9_252-1

Pande, P., Thit, A., Sørensen, A. E., Mojsoska, B., Moeller, M. E., & Jepsen, P. M. (2021). Long-term effectiveness of immersive VR simulations in undergraduate science learning: Lessons from a media-comparison study. *Research in Learning Technology*, *29.* https://doi.org/10.25304/rlt.v29.2482

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785–797. https://doi.org/10.1037/edu0000241

Petersen, G. B., Klingenberg, S., Mayer, R. E., & Makransky, G. (2020). The virtual field trip: Investigating how to optimize immersive virtual learning in climate change education. *British Journal of Educational Technology*, *51*(6), 2099–2115. https://doi.org/10.1111/bjet.12991

Petri, K., Feuerstein, K., Folster, S., Bariszlovich, F., & Witte, K. (2020). Effects of Age, Gender, Familiarity with the Content, and Exposure Time on Cybersickness in Immersive Head-mounted Display Based Virtual Reality. *American Journal of Biomedical Sciences*, 107–121. https://doi.org/10.5099/aj200200107

Plechatá, A., Sahula, V., Fayette, D., & Fajnerová, I. (2019). Age-Related Differences With Immersive and Non-immersive Virtual Reality in Memory Assessment. *Frontiers in Psychology*, *10*, 1330. https://doi.org/10.3389/fpsyg.2019.01330

Rebenitsch, L., & Owen, C. (2016). Review on cybersickness in applications and visual displays. *Virtual Reality*, *20*(2), 101–125. https://doi.org/10.1007/s10055-016-0285-9

Riecke, B. E., Bodenheimer, B., McNamara, T. P., Williams, B., Peng, P., & Feuereissen, D. (2010). Do We Need to Walk for Effective Virtual Reality Navigation? Physical Rotations Alone May Suffice. In C. Hölscher, Shipley, Thomas F., Belardinelli, Marta Ilivettti, & Bateman, John A., Newcombe, Nora S. (Eds.), *Lecture Notes in Computer Science: Vol. 6222, Spatial Cognition VII: International Conference, spatial cognition 2010, Mt. Hood/Portland, OR, USA, August 15-19, 2010. Proceedings* (pp. 234–247). Springer. https://doi.org/10.1007/978-3-642-14749-4_21

Simpson, M. (2020). *Scale and Space: Representations in Immersive Virtual Reality* [Doctoral dissertation]. The Pennsylvania State University. https://etda.libraries.psu.edu/catalog/17663mbs278

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., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Spector, J. M. (2020). Remarks on progress in educational technology. *Educational Technology Research and Development*, *68*(3), 833–836. https://doi.org/10.1007/s11423-020-09736-x

Stainfield, J., Fisher, P., Ford, B., & Solem, M. (2000). International Virtual Field Trips: A new direction? *Journal of Geography in Higher Education*, *24*(2), 255–262. https://doi.org/10.1080/713677387

Tsay, C. H.-H., Kofinas, A., & Luo, J. (2018). Enhancing student learning experience with technology-mediated gamification: An empirical study. *Computers & Education*, *121*(3), 1–17. https://doi.org/10.1016/j.compedu.2018.01.009

Turner, P., & Turner, S. (2006). Place, Sense of Place, and Presence. *Presence: Teleoperators and Virtual Environments*, *15*(2), 204–217. https://doi.org/10.1162/pres.2006.15.2.204

Tuthill, G., & Klemm, E. B. (2002). Virtual field trips: Alternatives to actual field trips. *International Journal of Instructional Media*, *29*(4), 453–468.

References

Chou, S.-W., & Liu, C.-H. (2005). Learning effectiveness in a Web-based virtual learning environment: A learner control perspective. *Journal of Computer Assisted Learning*, *21*(1), 65–76. https://doi.org/10.1111/j.1365-2729.2005.00114.x

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

Hegarty, M. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, *30*(5), 425–447. https://doi.org/10.1016/S0160-2896(02)00116-2

Helland, A., Lydersen, S., Lervåg, L.-E., Jenssen, G. D., Mørland, J., & Slørdal, L. (2016). Driving simulator sickness: Impact on driving performance, influence of blood alcohol concentration, and effect of repeated simulator exposures. *Accident; Analysis and Prevention*, *94*, 180–187. https://doi.org/10.1016/j.aap.2016.05.008

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, *60*(1), 48–58. https://doi.org/10.1080/02701367.1989.10607413

Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, F., . . . Jäncke, P. (2004). MEC Spatial Presence Questionnaire (MECSPQ): Short Documentation and Instructions for Application. Report to the European Community, Project Presence: MEC (IST-2001-37661). Retrieved from https://academic.csuohio.edu/kneuendorf/frames/MECFull.pdf

Woerner, J. J. (1999). Virtual field trips in the earth science classroom. In P. A. Rubba, J. A. Rye, & P. F. Keig (Chairs), *Proceedings of the Annual Conference of the Association for the Education of Teachers in Science,* Austin, TX.

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

Zhao, J., LaFemina, P., Carr, J., Sajjadi, P., Wallgrun, J. O., & Klippel, A. (2020). Learning in the Field: Comparison of Desktop, Immersive Virtual Reality, and Actual Field Trips for Place-Based STEM Education. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 893–902). IEEE. https://doi.org/10.1109/VR46266.2020.1581091793502

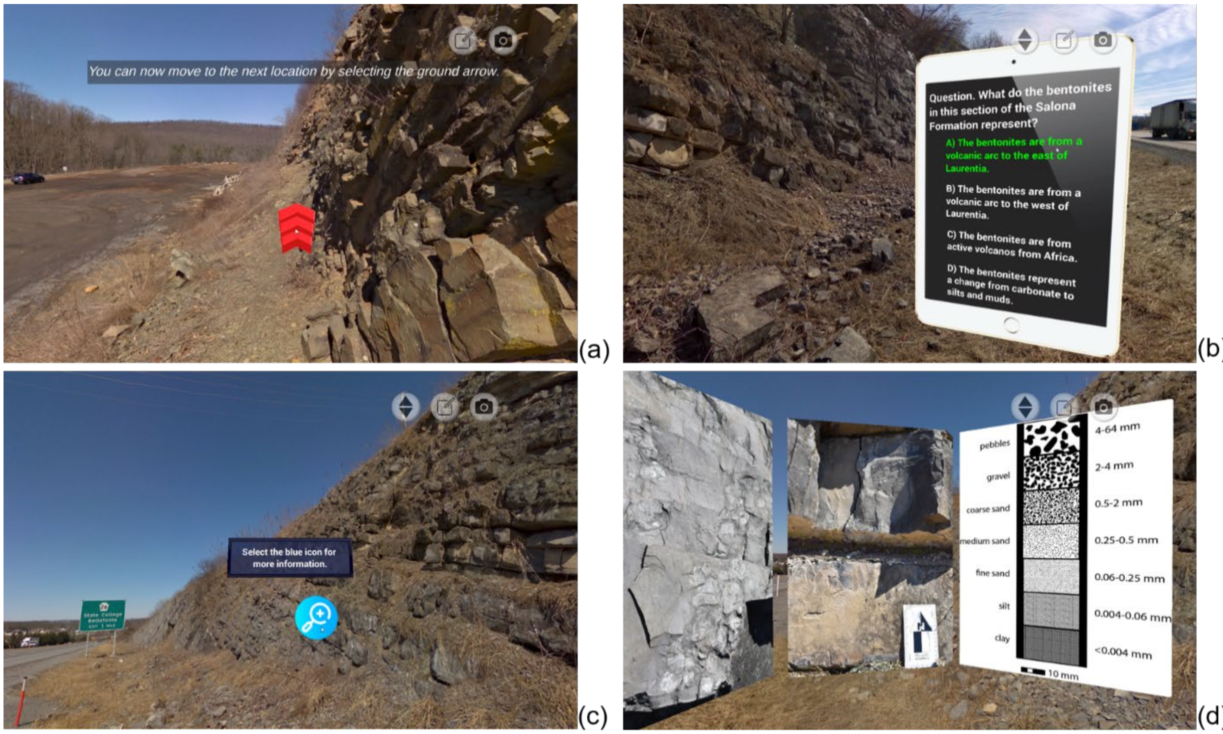

Figure 1. Screenshots of students' activities in the VFTs: (a) The students selected arrows on the ground to navigate through the 360° images; (b) multiple-choice questions were embedded in the VFT to assess content knowledge; (c) the students clicked on blue icons to open (d) additional information that illustrated key concepts or showed details of geological structures.

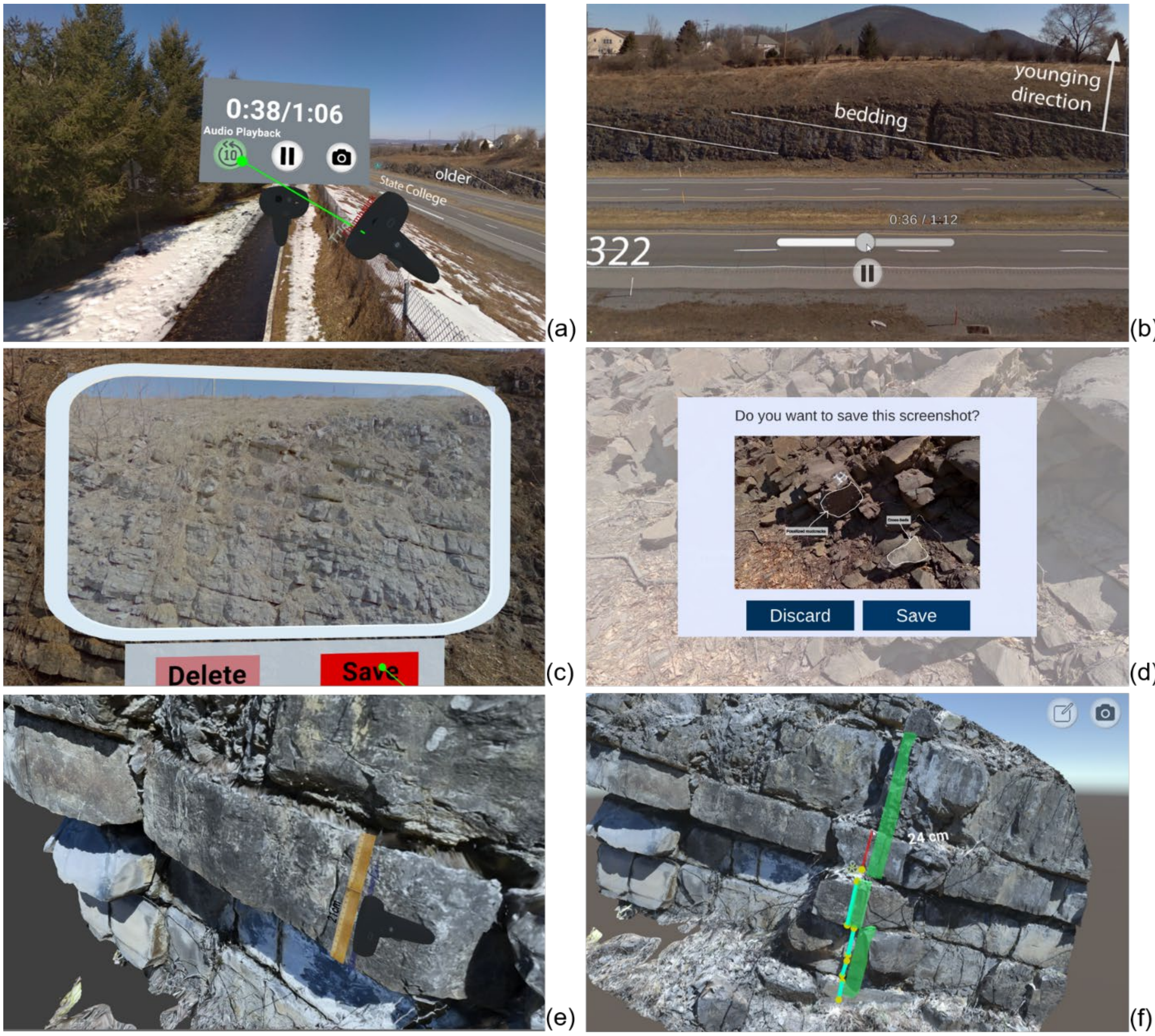

Figure 2. Screenshots from iVR (left) and dVR (right): (a+b) audio control; (c+d) picture taken; and (e+f) stratigraphic measurement using a 3D outcrop model.

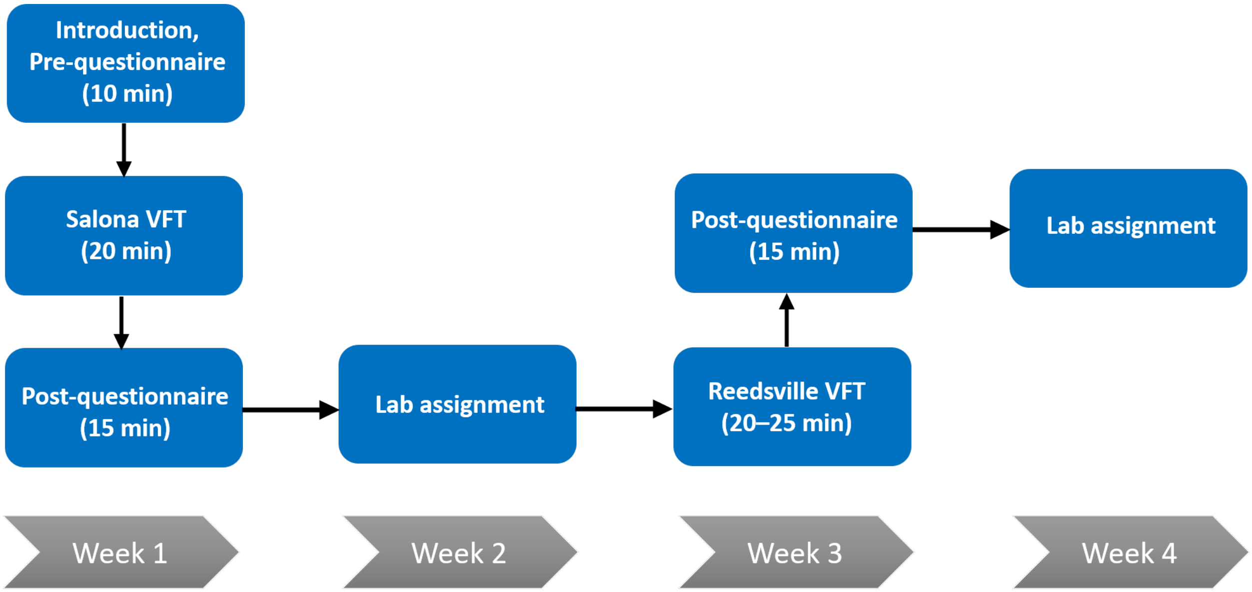

Figure 3. The procedure of the current study.

Table 1. Summary of demographic differences between the desktop virtual reality (dVR) and immersive virtual reality (iVR) conditions

Variables		*dVR condition*	*iVR condition*	*Significance*
Gender	Female (*N*)	36	7	p = .217
	Male (*N*)	93	7	
	Other (*N*)	1	0	
Academic year	Freshman (*N*)	15	1	p = .643
	Junior (*N*)	32	5	
	Senior (*N*)	4	1	
	Sophomore (*N*)	79	7	
Age	Mean	19.57	19.79	p = .402
	SD	0.91	0.89	
Spatial ability	Mean	3.36	3.62	p = .130
	SD	0.71	0.58	
Technology enjoyment	Mean	3.77	4.32	**p = .021**
	SD	0.94	0.80	

Note. Significance values from the t-test analysis of independent samples or the chi-square test of independence are presented in the final column. Bold values are significant at $\alpha = .05$.

Table 2. Means, standard deviations (SD), and ANOVAs for the dependent variables measured in the study

	Salona (VFT 1)		*Reedsville (VFT 2)*		*Significance*		
	dVR	*iVR*	*dVR*	*iVR*	*Condition*	*VFT*	*Interaction*
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)			
Perceived ease of use	3.88 (0.97)	4.32 (0.67)	3.74 (0.93)	4.07 (0.87)	$p = .091$	$p = .151$	$p = .679$
Cybersickness	1.83 (1.08)	1.79 (0.75)	1.74 (1.00)	2.21 (1.10)	$p = .399$	$p = .242$	$p = .076$
Spatial presence	2.74 (0.94)	3.36 (1.26)	2.55 (0.97)	3.07 (0.98)	**$p = .017$**	$p = .092$	$p = .713$
Field trip enjoyment	3.02 (1.10)	3.93 (1.07)	2.70 (1.14)	3.68 (1.05)	**$p = .001$**	**$p = .025$**	$p = .762$
Satisfaction	3.28 (1.01)	4.04 (0.93)	3.07 (1.06)	3.39 (1.04)	**$p = .041$**	**$p = .002$**	$p = .126$
Perceived learning	3.39 (0.83)	3.63 (1.03)	3.09 (0.91)	3.41 (0.97)	$p = .214$	**$p = .012$**	$p = .720$
Knowledge test performance	1.29 (0.33)	1.56 (0.32)	1.21 (0.24)	1.23 (0.18)	**$p = .020$**	**$p < .001$**	**$p = .015$**
Lab grade	88.4 (14.6)	86.8 (15.3)	87.9 (12.2)	89.3 (10.8)	$p = .974$	$p = .703$	$p = .585$

Note. Significance values from mixed between-within subjects ANOVAs (condition, VFT, and interaction) are presented in the final three columns. Bold values are significant at $\alpha = .05$.