Learning Environments in AR: Comparing Tablet and Head-mounted Augmented Reality Devices at Room and Table Scale

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Paul Craig

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Dr. Peter Willemsen

August 2022

Acknowledgements

I would Like to thank the following people for their help in preparing this thesis.

Thank you to my advisor Dr Peter Willemsen for all his help with preparing my experiment and writing this thesis.

Thank you to Dr Edward Downs for designing the experiment and for his help with the statistical analysis of data.

Thank you to Alex Lover for creating the data collection system for the experiment as well as the spatial test. And for his help running experiment.

Thank you to William Barber for his help running the experiment.

Thank you to Dr Tim Craig for his help with statistical analysis and editing this thesis.

Thank you to Dr Joanne Itami for her help with editing this thesis it's graphs.

Abstract

This study examined how presentation scale and device form factor affects declarative and spatial learning in augmented reality (AR) environments. The two form factors studied in this experiment were mobile-based AR using a third generation iPad Pro and head-mounted AR using a Microsoft Hololens2. These form factors were chosen to examine how the more expensive head-mounted AR displays compared to more widely available mobile-based solutions. Scale was chosen to examine how natural locomotion contributed to learning in AR. Two scales were examined: a room-scale environment that allowed participants to move freely, and a table-scale environment that participants viewed while seated. To test the impact of these factors we created a virtual cemetery based on Edgar Lee Masters' book Spoon River Anthology. We conducted a 2 (form factor) x 2 (scale) experiment on 131 participants exploring an AR learning environment across four conditions: Holoens2-Room, Hololens2-Table, iPad-Room and iPad-Table. Post experiment participants completed a 20-question multiple choice quiz to test declarative learning outcomes, a spatial reconstruction measure to test spatial learning, and a Likert scale survey to have participants self-report on their experiences with AR. Post experiment analysis found that neither form factor nor scale had a statistically significant impact on participants learning outcomes. We also noted a positive correlation between declarative learning and composited measure of participant's enjoyment on the Likert survey.

Contents

List of Tables

List of Figures

1 Introduction

Augmented reality (AR) has been shown to be a beneficial tool in educational settings. Multiple studies have shown that AR has positive effects on motivation, and engagement in classrooms as it allows for a more interactive and experiential learning (del Cerro Velázquez and Morales Méndez 2021; Garzón and Acevedo 2019; Özdemir et al. 2018). In particular AR can be beneficial to spatial understanding as it allows a user to use natural movements to explore an environment. AR allows users to change their view of a virtual object by simply walking around it (Bujak et al. 2013). Viewing 3D models of objects has been shown to be beneficial in developing student's spatial skill (Katsioloudis et al. 2014). Spatial skills are beneficial in a number of fields such as math, engineering and chemistry as it aids in students ability to view and understand 3D models in theses fields (Cai, Wang, and Chiang 2014; Onyancha et al. 2009; Tosik Gün and Atasoy 2017). For example, Lei et al. 2021 found that 3-D interactive learning environment enhanced learning by engineering students by increasing interactivity and flexibility in conducting experiments in a control system design class (Lei et al. 2021). However full natural movement requires a large space to operate in, which is not always available. In order to investigate the importance of a full scale environment for AR adapted an experiment that examined navigation in virtual reality to work in AR (Downs et al. 2022). (Downs et al. 2022) examined spatial and declarative knowledge learning outcomes related to exploring a virtual reality (VR) environment. We created two scales for our experiment: a room-scale experiment that participants could freely walk though and a table scale version viewed while seated. We hypothesized that participants in the room-scale version of our experiment would perform better in terms of spatial and declarative learning outcomes as the act of physically walking though an environment gives additional context helpfully in developing a better spatial understanding of an environment.

Our work also focuses on the affordance of a medium's modality, as defined in the MAIN model (Sundar 2008). The MAIN model defines four affordances by which technology is perceived: Modality, Agency, Interactivity and Navigability, which I will discuss in detail later. Modality refers to the form a technology takes and how it is structured. We investigated whether altering the modality to a more immersive form than has been previously used would increase learning. The majority of AR research has used less expensive mobile devices such as a smartphone or tablet AR (M. Akçayır and G. Akçayır 2017). In contrast, our experiment was initially designed to work on head-mounted display: the Microsoft Hololens2 which provides a more immersive experience. To see if our choice of modality affected learning in AR we developed a second version of the experiment than ran on a tablet: iPad Pro 3. Ultimately, our experiment sought to address the following research question.

RQ: What is the relation ship between form factor and scale on learning in Augmented Reality?

In order to answer this question, we created a 2 (Form Factor) by 2 (Scale) experiment. To the best of our knowledge we are the first to examine the differences in navigation in AR across form factors. Our dependent variables related to learning outcomes were assessed by a multiple choice narrative test about information read within the AR environment and a spatial reconstruction test developed to gauge spatial understanding. A MANCOVA analysis of results on these tests found no statistical difference between spatial or declarative learning outcomes across all conditions.

2 Background

2.1 Definition of Augmented Reality

Augmented Reality (AR) refers to technologies that combine elements of the digital world with the real world. This stands in contrast to Virtual Reality (VR) which completely immerses a user in a digital environment. We focused on AR that augments a user's vision, by overlaying digital images over a user's view. This effect can be achieved through two categories of AR as defined by (Milgram et al. 1995), monitor-based AR and see-through AR. For the purpose of the study we looked at these two categories of AR in terms of two common form factors: mobile-based AR and head-mounted AR.

Mobile-based AR is a prime example of what (Milgram et al. 1995) refers to as monitor-based AR. In monitor-based AR digital elements are overlaid onto a recorded or live video (Milgram et al. 1995). In mobile-based AR, video of the real world is captured using the camera on a mobile device such as a phone or tablet 2.1. This video is then displayed on the device's screen with digital elements drawn on top. Mobile-based AR is currently the most well represented form factor of AR in current research (M. Akçayır and G. Akçayır 2017), due to it's relatively low cost and portability as it's avalible on most current phones and tablets (Furió et al. 2013). However there are also notable draw backs to mobile-based AR including: a small display and touch interface, an inability track a user's head location, and an inability to render stereoscopic images (Zhu and Grossman 2020). As such they may not be suitable for all AR uses.

Figure 2.1: Example of augmented reality on a reality of a tabletop environment on a mobile iPad display.

An alternative to mobile-based AR are Head-Mounted AR Displays (HMD). HMD's are an example of what (Milgram et al. 1995) refers to as see through AR. In this set up a user views the world through a transparent medium that superimposes digital elements over the user's visions (Milgram et al. 1995). Head-Mounted displays (HMD) refer to devices which cover a user's eyes with AR capable lenses. Digital images are displayed on the lens overlaying the digital elements onto the wearer's vision. Many current head-mounted displays allow users to interact with digital elements by tracking their hands and gaze direction. In doing so the HMD can tell when a user's hands intercedes with a digital element, and it can also take commands through hand gestures (Rozado et al. 2014). However head-mounted displays (HMD) also have their short-comings as they have a constrained field of view (FOV) of 40°diagonal that prevents users from using their full field of vision (Ofek et al. 2015). More modern head-set such as the Microsoft HoloLens2 have a FOV of nearly 54°diagonal but this still falls short of the 122°diagonal provided by the cameras on modern tablets such as the iPad pro (*iPad Pro* 2022) and the 180°diagonal FOV normal afforded to humans (Ofek et al. 2015). A reduced FOV has been shown to negatively impact scores on spatial updating tasks (Riecke and Bülthoff 2004). Head-mounted displays have also been reported to result in more severe and longer lasting cybersickness when compared to AR on tablets (Hughes et al. 2020).

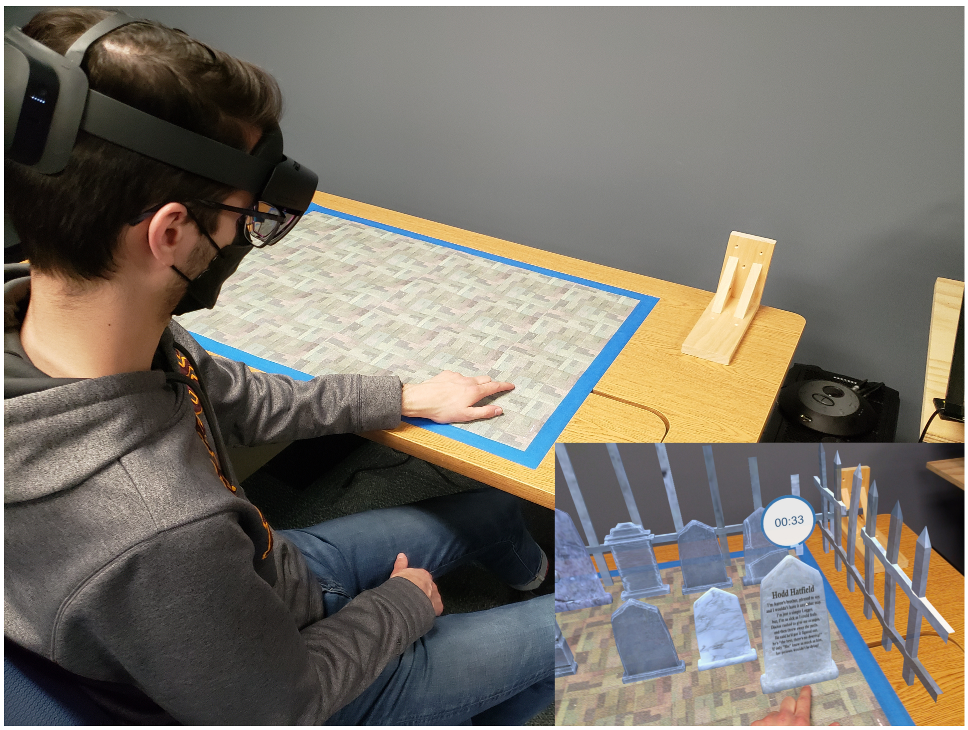

Figure 2.2: Example of augmented reality on a tabletop environment using a head-mounted display: the Microsoft HoloLens2. The wearer's view can be seen in the lower right corner.

Because of these differences in modalities we were interested in seeing if there was a measurable difference in spatial or declarative learning in AR across different scales and form factors.

Regardless of form factor augmented reality allows digital elements to be added to the digital world, which is useful when it is not practical to acquire real material (Cai, Wang, Gao, et al. 2012). AR devices are able to keep track of digital elements positions in the real world using either physical markers, such as QR codes to mark where a digital objects should be in physical space, or by using GPS coordinates to keep track of where digital object should be in physical space (Antonioli et al. 2014). As such AR allows users to change their view of a virtual object by simply walking around it (Bujak et al. 2013). Viewing 3D models of objects has been shown to be beneficial in developing students spatial skill (Katsioloudis et al. 2014). As such AR has been proposed for use in developing spatial skills.

2.2 Augmented Reality in Education

Augmented Reality has been studied as an educational tool since the early 1990s (Bell and Fogler 1995). AR has been tested in classrooms to teach a number of subjects, including biology, physics, and design courses (M. Akçayır, G. Akçayır, et al. 2016; Chien et al. 2019; Wei et al. 2015). The majority of AR has been done in a formal education setting (Garzón and Acevedo 2019). For informal environments AR has been tested as an educational tool in museums and planetariums (Jones and Lawler 2019; Jung et al. 2016a). AR was used to add additional context to museum displays (Jung et al. 2016b), and to provide an overlay to help deaf students understand a planetarium show (Jones and Lawler 2019). In these roles AR has increased engagement and motivation amongst learners, and has been generally well received by students (Balog et al. 2007; Fiorentino et al. 2021). However this may be due to the novelty of AR devices, and may not persist with long-term use (M. Akçayır and G. Akçayır 2017). A meta analysis of 15 AR studies (Özdemir et al. 2018) found that AR is an effective teaching tool across all education levels. However, these findings were disputed by (Garzón and Acevedo 2019) who found that AR is most beneficial for bachelors level students in a meta analysis of 64 AR papers. In addition, they found that AR had a large effect on teaching in engineering, manufacturing and construction. This agrees with the conclusions of (Onyancha et al. 2009) who noted that AR could be useful in the fields of engineering due to it's potential for training spatial ability.

2.2.1 Augmented Reality and Spatial Learning

As defined earlier, spatial ability is "the ability to generate, retain, retrieve, and transform well-structured visual images" (Lohman and Kyllonen 1983; Onyancha et al. 2009). Spatial ability is divided into three categories: visualization, spatial relations, and orientations (Lohman and Kyllonen 1983). Spatial ability is an important skill in a number of fields. Multiple studies have found that spatial ability is important for success in engineering (Hsi et al. 1997; Leopold et al. 2001),in chemistry for understanding complex microstructures (Cai, Wang, and Chiang 2014) and in geometry (Tosik Gün and Atasoy 2017). Viewing 3D models, either physical or digital, has been shown to result in better training in spatial abilities when compared to using 2D models. Several studies evaluated the effectiveness of 3D models in AR on training spatial abilities. (Tosik Gün and Atasoy 2017) examined the impacts AR have on training spatial abilities on 88 sixth grade math students and found that students who used AR models on a computer had a greater increase in spatial abilities than the control group that took the course without AR. Similar results were found by (Dünser et al. 2006) who performed a large-scale study on spatial learning using augmented reality. They trained 215 high school students using both a 3-D head-mounted augmented reality training program and a 2-D desktop training program. On a Objet

Perspective Test (OPT) men showed improved scores after training in AR compared to men who trained on the desktop, while women who trained in AR did worse on the OPT when compared to women who trained on the desktop. (Shelton and Hedley 2003) proposed that AR users performed better due to several factors. First, AR allows the user to view both the real world and virtual world simultaneously, enabling them to better keep track of their position in physical space. Second, AR is generally controlled by physical motion and (Shelton and Hedley 2003) argued that utilizing a user's viso-motor system would lead to better results in spatial learning. This idea was later echoed by (Bujak et al. 2013) who argued that a unique advantage of AR is how it meshes directly with the real world affording natural interaction. Traditional interfaces require users to learn how to use physical or virtual controls, such as mouse and keyboard or controllers, to rotate and zoom in on virtual objects. This corroborates other research suggesting that understanding of a space is tied to our motor functions. For example, (Rieser and Pick 2018) examined spatial ability in terms of navigation tasks which " are useful methods for illuminating how spatial layouts are represented and abstracted". They noted that when navigating an environment the use of natural locomotion causes "predictable environmental consequences" (Rieser and Pick 2018). In other words, it is easier to understand the layout of an environment if you walk through it under your own power. Creating a mental representation of a space in this way, using landmarks and distances to construct routes, is referred to as a developing a cognitive map (Rieser and Pick 2018; Ruddle et al. 2011). (Ruddle et al. 2011) showed that creating a cognitive map of an environment is more difficult in a virtual environment. However, the accuracy of participants' cognitive maps increased when navigating using a treadmill suggesting that the use of natural locomotion benefits spatial navigation in virtual environments.

The use of natural locomotion over artificial interfaces also may be beneficial in terms of spatial navigation and understanding when using VR devices. (Zanbaka et al. 2005) found that a free-walking VR setup increased participants' understanding of the layout of a virtual environment when compared to both static VR setups and more traditional digital interfaces. (Marsh et al. 2013; Zanbaka et al. 2005) found that allowing users to physically walk around a virtual environment, improved their results on spatial memory tests when compared to traditional interfaces such as keyboard and mouse. (Xie et al. 2018) found that natural locomotion was preferable to joystick controls in complex navigational tasks. , (Chance et al. 1998) found that participants who navigated a virtual maze using full natural locomotion made fewer errors than participants who were limited to turning in place or participants who controlled movement using a joystick. Finally, (Ruddle et al. 2011) found that natural locomotion led to users developing a better cognitive map of a virtual environment. However, though natural locomotion in VR has benefits, there are limitations to the technology. Within VR environments users are fully immersed and surrounded by the digital environment. They are often isolated from the physical obstacles or objects in the real environment, which may limit distractions, but which decreases their ability to interact with physical objects. In contrast, AR allows users to view the interaction of real and digital space, which increases their ability to interact with their environment. Previous work comparing AR and VR found that VR was generally more engaging and led to better spatial understanding . (Huang et al. 2019) posited that this was due to the more immersive nature of VR. It was also noted that AR seemed to have a lower cognitive load on users allowing them to better process other information such as audio. Notably, this study was limited to a single mode of AR on an iPhone. The conclusions about use of VR for spatial learning also apply to declarative learning. (Yates 1966) found that memory recall benefits from association with physical locations. In addition, the immersive positive attitude and engagement have been show to increases memory recall (Batista et al. 2020). In contrast, (Chen et al. 2019) showed that the use of an AR headset to present information resulted in lower scores on memory test when compared to a PowerPoint. However, this study was a pilot study of limited scope, and highlighted the need for further research.

The research presented here, and also that reported in (Downs et al. 2022) are based on the MAIN model presented in (Sundar 2008). Sundars' method evaluates how users perceive and evaluate the credibility of technology. The MAIN model defines four affordances by which technology is perceived: Modality, Agency, Interactivity and Navigability. Modality refers to the form a technology takes and how it is structured. Agency denotes the perceived source of information presented by the technology. Interactivity concerns to how users interact and engage with the technology. While Navigability describes how a user navigates the information provided by the technology. Modality and Navigability affordances are the most relevant for our study

When examining the differences in mobile vs head-mounted AR the most obvious affordance is modality (Sundar 2008). Given the difference in structure and presentation we hypothesized that the user's perception of an experience would differ between form factors of head-mounted and mobile AR. We also hypothesized that navigation is a key distinction between modes of AR. As previously noted, natural locomotion has been hypothesized to be beneficial to spatial learning in virtual environments (Ruddle et al. 2011; Xie et al. 2018). Thus, we felt it was important to separate out the effect of locomotion as a variable.

To fully test the effects of locomotion and form factor on learning in Augmented Reality, we developed a we created a 2 (Form Factor) by 2 (Scale) experiment, Which we will explore in the next chapter.

3 Implementation

We conducted a 2 (Form Factor) by 2 (Scale) between-subjects experiment. Scale was controlled by presenting the stimulus at full room-scale or in a miniature that fit on a table top. The room-scale experiment was designed so that a participant could freely walk around the virtual environment, and thus potentially benefit from the use of natural locomotion. Conversely the table-scale experiment required the participant to stay seated. We tested two AR form factors head-mounted AR and tablet AR, resulting in four total conditions, illustrated in Figure 3.2 and 2.1. Post experiment participants were tested on declarative learning outcomes using a multiple choice narrative test, and on spatial learning using a spatial reconstruction test.

3.0.1 Stimulus Material

The visual stimulus material consisted of a virtual cemetery containing 13 tombstones. Each stone was inscribed with a story about the life of one of the inhabitants of the fictional town Spoon River. All stories were developed in-house by Dr Edward Downs and were inspired by the Spoon River Anthology (Masters 1916). The combined reading of all tombstones informed a larger story about the inhabitants. During the experiment the stories on the tombstone where initially obscured. A glowing blue waypoint would then appear in front of a random tombstone see 3.3. Activating this waypoint would reveal the text on the tombstone. A 35-second timer would appear above the activated tombstone see 3.2. When the timer reached zero it would disappear and the text on the tombstone would be re-obscured. The waypoint would then appear in front of an unvisited tombstone. This process would repeat until every tombstone in the cemetery was viewed.

Figure 3.1: View of activated tombstone in the table scale (left) and room scale (right) versions of the experiment

Scales

For the experiment two versions of the stimulus where created: a room-scale 2.1 and a table-scale version. The room-scale version of the cemetery was scaled to fill a 21 x 33 foot lab space. Participants were able to explore this environment by freely walking among the tombstones. The table-scale environment fit on a 24 x 36 inch table space. In the room Scale condition participants viewed the stimulus from above, seated in a chair, and explored the environment by tapping on relevant tombstones, or it's associated way-point. In the table scale version the tombstone would scale up by 150% to ensure that they could be easily read see 3.2. Tombstone size would be reset after the timer expired.

3.0.2 Hardware and Software

For our head mounted form factor we used the Microsoft HoloLens 2. AR elements are superimposed over a user's field of view using a pair of 2k 3:2 aspect ratio lenses built into the headset. The HoloLens is capable of tracking both the physical environment, and the wearer's hands. In the handheld AR form factor we used the Apple iPad Pro 12.9. The tablet has a rear-mounted camera that can be used to view the environment. In addition, it is equipped with a LiDAR scanner that enables it to detect physical objects and surfaces. The 3D environment stimulus and software were programmed using the Unity game engine with executable targeted for the Hololens2 and the iPad. The cemetery was constructed in the unity game engine. The Hololens version of the experiment was created using the Mixed Reality Toolkit. Spatial anchors anchors were used to ensure that the cemetery consistently appeared in the same space across all tests.

Spatial Anchors were not used in the the iPad version of the experiment the stimulus material had to be placed manually by the experiment proctor before each session. To aid in this an interface was developing using the ARkit plane detection package. This allowed proctors to place the cemetery by tapping on the screen and using keyboard controls thoaugh a linked blue tooth keyboard. To ensure consistent placement between experiments tape markers were placed on the floor and table in the lab to help with alignment.

3.0.3 Experiment Procedure

A total of 131 participants were recruited, but due to an equipment malfunction one participant could not complete the experiment. This left us with 130 participants consisting of 54 male and 76 females. Participants were university students aged 18-

33, with a mean age of 19. Each participant was randomly assigned to one of the four conditions.

1. **Hololens2-Room:** In this condition participants viewed the experiment by wearing Microsoft Hololens 2. The cemetery was projected into the full 21 x 33 foot lab space. Participants navigated the environment by freely walking though the room, and waypoints were activated by standing in them.

2. **Hololens2-Table:** In this condition participants viewed the cemetery by wearing the Hololens 2. The cemetery was projected on to the 24 x 36 inch table space. Participants were asked to sit at the table to view the cemetery from above. Waypoints and were activated by placing a hand or finger though the waypoint or it's connected tombstone.

3. **iPad-Room:** In this condition participants viewed the cemetery on the iPad Pro. The iPad was held in the hands of participants. The cemetery was projected into the full 21 x 33-foot lab space and could be viewed on the iPad screen. Participants navigated the environment by freely walking through the room, and waypoints where activated by standing in them.

4. **iPad-Table:** In this condition participants viewed the cemetery by looking though the iPad-pro. The cemetery was projected onto the 24 x 36 inch table space. Participants were asked to sit at the table to view the cemetery from above. Waypoints where activated by tapping the waypoint or attached tombstone on the iPad screen.

To ensure consistency between participants and conditions all participants were instructed using a pre-written script. Before being admitted to the study participants were provided with an informed consent form approved by the University of

Figure 3.2: Right: an example of a user wearing the HoloLens in the lab space. Left: the users view of the lab space though the HoloLens

Minnesota IRB (Institutional Research Board). Participants filled out a demographic questionnaire. Self-reported information included: sex, GPA, past experience with AR/VR and the number of hours using various digital devices (See Appendix A.1 for full list of questions). Once they finished the questionnaire participants were instructed on how to properly don or hold the AR device and how to use it to interact with the digital elements of the environment.

Figure 3.3: The room-scale environment as viewed on the Hololens 2

Training

Immediately prior to the experiment participants where familiarized with the use of the AR device via a training scene. In this scene participants were presented with 5 virtual placards, and a blue cylinder was spawned in front of one of the placards at random. Participants were instructed on how to activate the cylinder using their AR device. Upon activation text would appear on the placard. Participants where instructed to read the text aloud. To ensure that the text was readable for each participant on the given AR device device, participants were instructed to read the text aloud. The text would remain visible for 10 seconds. Participants could view the remaining time on a timer above the placard. After the 10 seconds elapsed the timer and text would disappear. The waypoint would then spawn in front of a randomly selected unvisited placard. This process repeated until all 5 placards had been visited. Participants were also instructed to practice the full range of motion provided by their AR device. Once the participants training was completed they moved on to the experiment.

Room-scale and Table-scale Conditions

Prior to the experiment the participants were given a brief description of the experiment and informed that they would be tested after the experiments conclusion. They were then assigned to either the table-scale or room-scale experiment. In the experiment the stories on the tombstone where initially blanked out. Participants were directed to visit all tombstones. The sequence of the tombstones was randomized for each participant. A glowing waypoint marked the next tombstone in this sequence. The text on each tombstone was obscured until the waypoint was activated. After activation participants were given 35 seconds to view the text. Once the allotted time was up the text was hidden and participants were directed to the next tombstone. This process continued until every tombstone had been viewed.

For the room-scale condition, participants could freely walk about the lab space to reach each waypoint. Participants activated tombstones by standing on the waypoint directly in front of the tombstone.

Participants in the table-scale condition were seated at a 24 x 36 inch table. The cemetery was projected onto the table so that participants viewed it from it's entrance. Participants were asked to stay seated throughout the experiment. To enhance readability upon activation, tombstones would be enlarged and brought forward slightly. In the Hololens2 condition participants could adjust their view by moving their heads in relation to the projection. Tombstones were activated by placing a hand through the stone or it's waypoint marker. In the tablet condition, participants were free to maneuver the iPad to change their view of each tombstone. Tombstones were activated by tapping on the stone or it's waypoint on the iPad screen.

3.0.4 Evaluation

After exploring the stimulus three dependent measures were administered. These included: 1) a narrative test to evaluate participants declarative learning outcomes, 2) a spatial reconstruction test to measure spatial learning, and 3) a pencil-and-paper questionnaire to evaluate participants subjective experiences.

3.0.5 Narrative test

The Narrative test consisted of 20 questions about the narrative presented on the tombstones. Each question was multiple choice with 4 available answers. Examples of questions included: "Who was the nurse in spoon river?", "How many tombstones belonged to the Throckmortin family?", and "Who is responsible for most of the deaths in Spoon River?". These questions were designed to assess participants knowledge and comprehension.

3.0.6 Spatial test

For the spatial test, participants were asked to reconstruct the layout of the cemetery on a computer using a program designed by Alex Lover. They were presented with a 2D map of the Spoon River Cemetery with tombstones removed. The map included 13 orange dots each indicating the positions of a tombstone. The Celtic cross at the cemetery center was also included as a landmark. Participants were tasked with dragging images of the tombstones into the correct position on the map. Four additional images of tombstones that were not present in the experiment were added as foils. Participants were scored on the number of tombstones placed in the correct position, with a maximum possible score of 13.

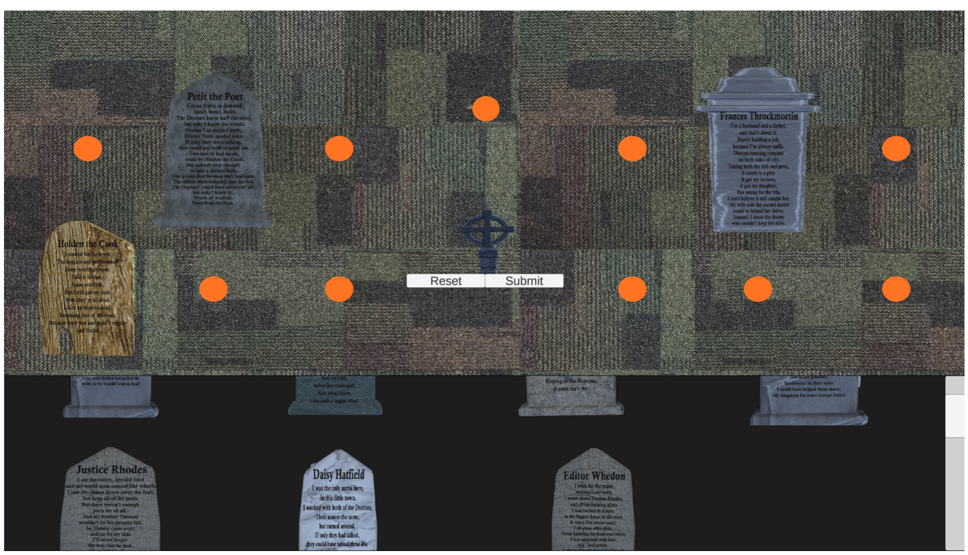

Figure 3.4: The spatial Reconstruction test. Participants dragged tombstones from the menu at the bottom of the screen onto the orange markers on the map above.

3.0.7 Likert Scale Survey Questionnaire

Before leaving participants were asked to fill out a questionnaire assessing their subjective experience in the AR environment. The questions were based on a 7-point scale with 1 indicating strongly disagree and 7 indicating strongly agree. Questions assessed the following characteristics: Novelty, Liking, Immersion and Interactivity .

Post experiment analysis showed that the Novelty and Interactivity responses were not internally consistent. The 3 questions in the Novelty group had a Cronbach's $\alpha = 0.668$, and the 8 questions in the Interactivity grouping had a Cronbach $\alpha = 0.606$. These scores indicate a low covariance among answers, and therefore these groupings were not included in our analysis. This left 2 groupings for analysis: Liking and Immersion.

Liking was covered by 6 questions such as: "I enjoyed the task I participated in today", "I had fun participating in the experiment today", and "I was disappointed in the task today". These question were found to be internally consistent with Cronbach $\alpha = 0.823$. This was improved to Cronbach $\alpha = 0.865$ by dropping one of the question items. (See Appendix A.3 for complete list of measures).

The Immersion grouping was made up of 5 questions such as: "My experience today was involving", "I felt like I was physically inside the AR environment", and "I felt immersed in the AR environment". These questions were found to be internally consistent with Cronbach $\alpha = 0.719$. This was improved to Cronbach $\alpha = 0.763$ by dropping one of the question items. (See Appendix A.3 for complete list of measures)

4 Results

Our analysis found that neither the form factor nor the scale had a significant effect on learning as evaluated by the spatial and narrative tests. Learning did not differ significantly among the four combinations of head-set and tablet forms, with the table-top and room scales. Scores on the tests were influenced by the ACT scores of students.

	Spatial Score	ACT Score	Liking	Immersion	Sex
Pearson	0.547	0.335	0.332	.009	-0.182
Sig	< 0.001	< 0.001	< 0.001	0.918	0.036

Table 4.1: Results of bi-variate correlation on narrative test scores

	Narrative Score	ACT Score	Liking	Immersion	Sex
Pearson	0.547	0.133	0.230	.023	-0.079
Sig	< 0.001	< 0.191	0.008	0.793	0.369

Table 4.2: Results of bi-variate correlation on spatial test scores

	Narrative		**Spatial**	
	Mean	SD	Mean	SD
Hololens2-Room	14.1875	3.47746	6.9375	3.49135
Hololen2-Table	14.5312	4.25012	7.1563	3.49294
iPad-Room	13.0606	4.15286	6.6970	2.88839
iPad-Table	14.0588	3.34792	6.6471	2.75111

Table 4.3: Mean score for all 4 conditons on the narrative and spatial tests. A perfect score on the Narrative test is 20 and 13 on the Spatial test.

Statistical analysis was carried out using SPSS Version 28.0.1.1(15). There were no significant differences in either the mean narrative or spatial assessment scores for each condition Table 4.3. Factorial MANOVA revealed there was neither a main effect for Form Factor, Wilks's' $\Lambda = 0.989$, F(2, 126) = 0.713, $p < 0.492$, partial $\eta^2 = 0.011$; nor for Scale, Wilks's' $\Lambda = 0.99$, F(2, 126) = 0.635, $p < 0.532$, partial $\eta^2 = 0.01$; nor was there a significant interaction effect between Form Factor and Scale: Wilks's' $\Lambda = 0.996$, F(2, 126) = 0.266, $p < 0.767$, partial $\eta^2 = 0.004$. A post hoc tukey's test found no significant differences among means.

In order to better understand the factors involved a bivariate correlation was run on all measures. Liking was found to be strongly correlated with score on both the narrative (Pearson Correlation = 0.322, $p < 0.01$) and spatial test (Pearson Correlation = 0.230, $p = 0.08$, Table 4.1, Table 4.2). A strong correlation was also found between ACT score and scores on the narrative test, (Pearson Correlation = 0.335, $p < 0.01$). Table 4.1. Additionally a correlation was found between biological sex and scores on the narrative test (Pearson Correlation = -0.182, $p = 0.036$ See apendix for full correlation table.

The MANCOVA was rerun with ACT scores and sex as covarients. However in this model sex was not significant ($p = 0.215$) and so was removed from future models. The MANCOVA was rerun with only ACT score as a covariant and showed that there was not a main effect for Form Factor, Wilks's' $\Lambda = 0.992$, F(2, 93) = 0.355, $p < 0.702$, partial $\eta^2 = 0.008$; nor for Scale, Wilks's' $\Lambda = 0.991$, F(2, 93) = 0.428, $p < 0.653$, partial $\eta^2 = 0.009$; nor was there a significant interaction effect between Form Factor and Scale: Wilks's' $\Lambda = 0.991$, F(2, 93) = 0.429, $p < 0.652$, partial $\eta^2 = 0.009$. Conversely ACT score proved to be highly significant for the narrative test F(1,11.242), $p < 0.001$.

To account for the effect of liking and immersion a polynomial regression was run

	Narrative		Spatial	
	Mean	SD	Mean	SD
Hololens2-Room	14.153	0.684	7.067	0.642
Hololen2-Table	13.668	0.641	6.656	0.602
iPad-Room	14.113	0.694	6.950	0.651
iPad-Table	14.667	0.672	6.753	0.631

Table 4.4: Results of narrative and spatial test with ACT scores as a covariate

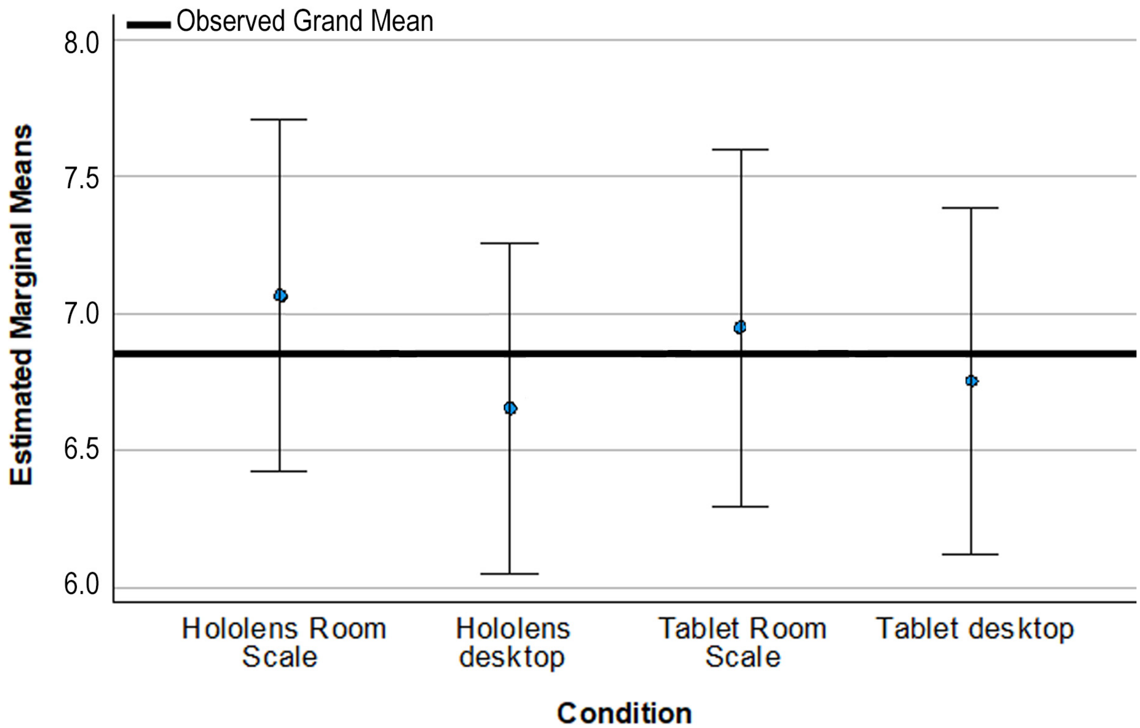

Figure 4.1: Graph of mean scores on the Spatial test organized by condition number. Covariates appearing in the model are evaluated at the following values: ACT test score = 24.0707 +/- 1 SE.

on their effects on spatial and narrative tests (see Appendix B for full table of results). Liking was a significant predictor of scores on the spatial test ($y = 6.861 + 0.230x, R^2 = 0.053, p < 0.008, d.f = 129$) see Figure 4.3. Immersion was not a significant for the spatial test under any model (see Appendix B for full table). Liking was again found to be a significant predictor of the narrative test score ($y = 13.962 + 0.332x, R^2 =$

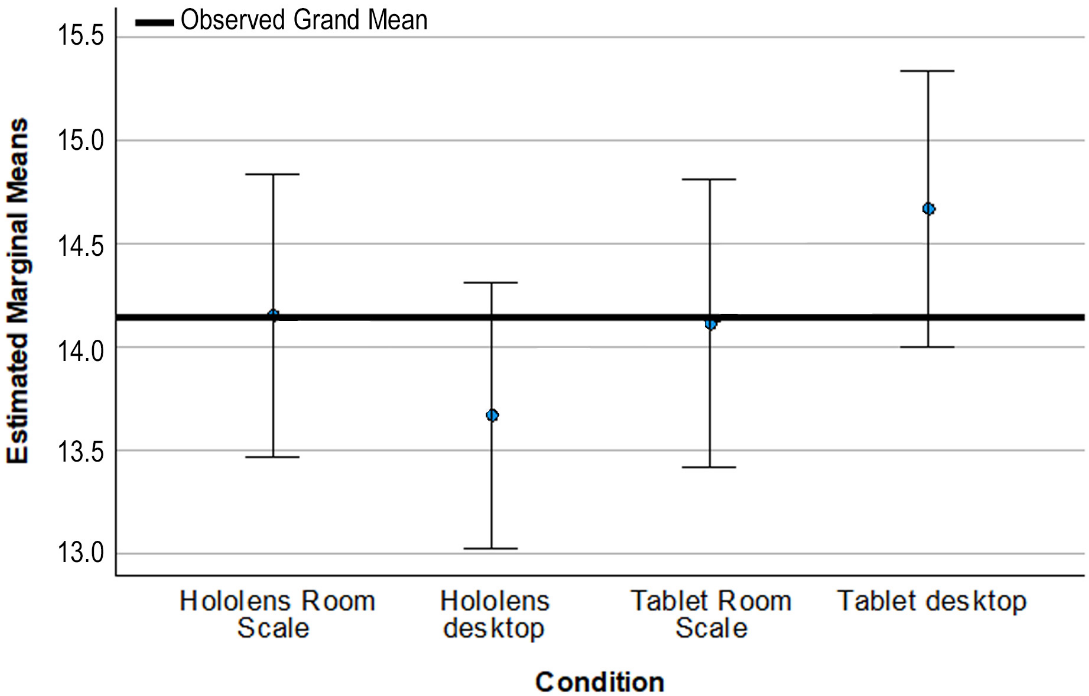

Figure 4.2: Graph of mean scores on the narrative test organized by condition. Covariates appearing in the model are evaluated at the following values: ACT test score = 24.0707 +/- 1 SE.

$0.104, p < 0.001, d.f. = 129$). See Figure 4.4. Immersion was not a significant factor under any model (see Appendix B for full table).

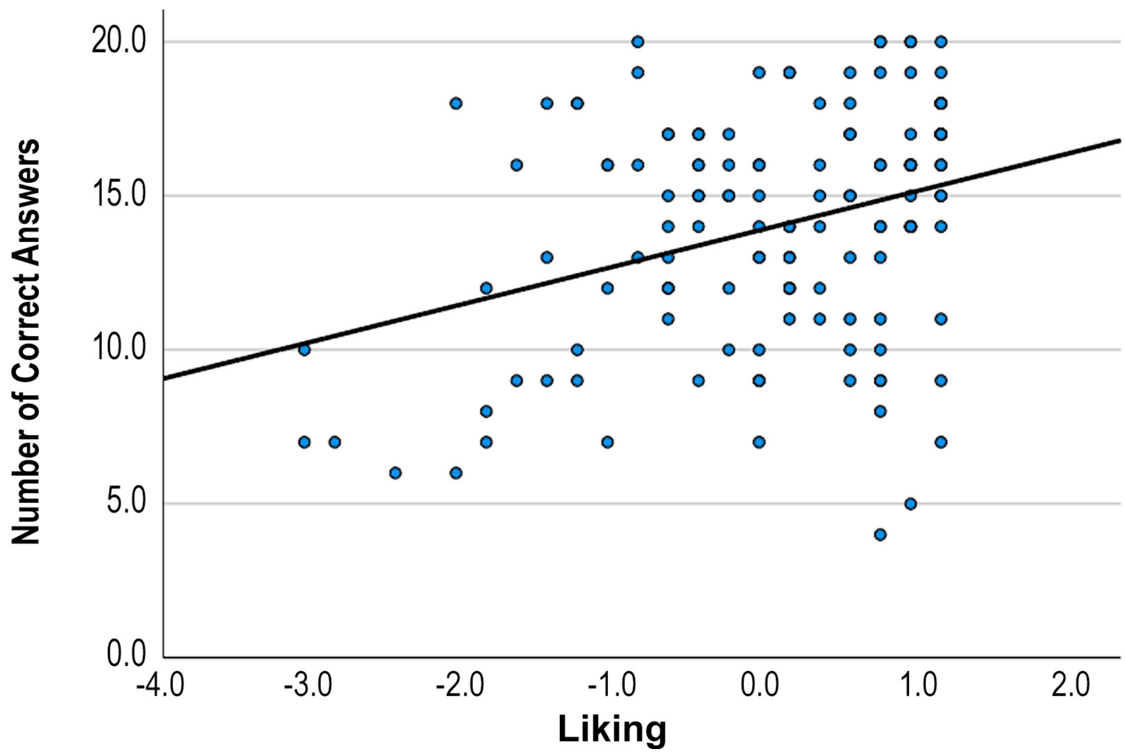

Figure 4.3: Linear regression of Likeings effect on the results of the narrative test. y=6.861+0.230x=0.53, $R^2 = 0.053$, $p = .008$, F=(1,129)

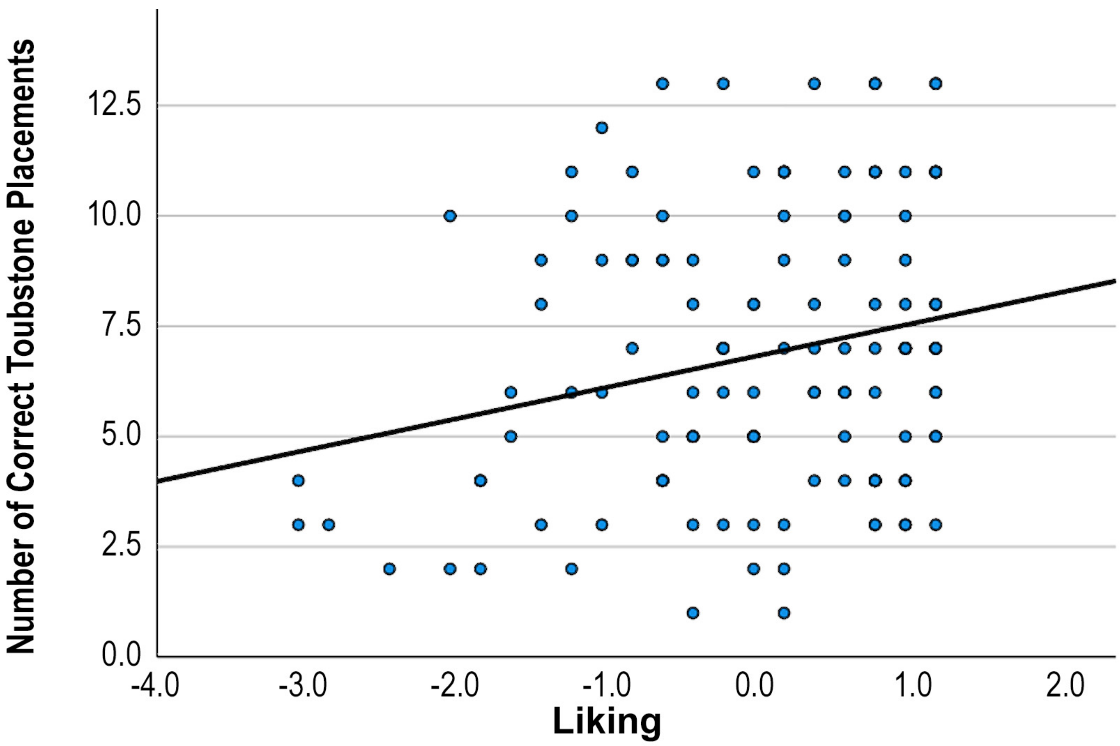

Figure 4.4: Linear regression of Likings effect on the results of the spatial test. y=13.962+0.332x=.104, $R^2 = 0.104$, $p < 0.001$, F=(1,129)

5 Conclusions

5.1 Discussion

The hypothesis that room-scale conditions or head mounted viewers improved learning outcomes related to spatial updating regardless of form factor was not supported. The ability for participants to walk around the full space of the cemetery in AR did not improve narrative or spatial scores as predicted. The prediction that a head mounted viewer would improve learning was also not supported. Declarative knowledge was found to be similar across conditions as neither form factor nor scale were differed significantly as measured by the two dependent variables in this study. Some variation was observed between the mean scores on the narrative test between the Hololens2-Table condition and the iPad room condition but this difference disappeared when variation in ACT scores were accounted for. Thus, our results indicate that scale and form factor showed no difference in their effect on spatial and declarative learning. This suggest that for AR uses in education and other settings either form factor may be used. therefore usurers can focus on choice based on off cost effectiveness and feasibility.

Our study did not agree with the results of previous research that indicated that the ability to move through the test environment improved spatial and declarative learning (Ruddle et al. 2011; Xie et al. 2018). A number of factors could explain the differences in our results and those of previous studies. First, our study and previous research differed in their experimental designs. In contrast to previous experiments, we isolated the effect of the form of user input from the scale of the experiment. This allowed us to directly compare the impact of physical versus virtual movement through an environment on learning. A second difference between our study and previous research is that we only tested differences in learning using two technological devices, and did not include a room with actual tombstones to walk through. An interaction with real objects may increase the impact of movement through the space. The potential for the interaction with physical rather than virtual objects to influence learning was indicated by the impact of user familiarity/interest in technology on learning outcomes. This may indicate that users who are not experienced or interested in using technology perform better when interacting with physical objects, and that removing this element decreased the value of moving through space. The impact of physical movement could also be dependent on the size of the space explored. Our experiment was conducted at single room size, and a larger room requiring more movement might show significant differences.

The spatial test's design may have impacted the results. The spatial test presented the participants with a top down view of the cemetery. This may have given an advantage to participants in the table condition who viewed the cemetery from above during the experiment. Conversely, the room scale participants viewed the cemetery from ground level and would have had to translate their understanding of the space to match the top down view of the spatial test. Future research may need to provide a spatial test that does not have the potential to favor either scale. In addition the scoring method may have also played a role. The spatial test was scored on number of correct tombstone placements. However it is possible that scoring based of each tombstone's distance from the correct position could have reviled more nuance between the score of each condition.

The differences in the populations tested in our study and those previous studies may have also contributed to the differences between our results and previous research. Our population consisted exclusively of undergraduate students. Previous meta analysis has found that AR is beneficial for bachelors students(Garzón and Acevedo 2019). However, previous research cited had worked with high school and middle school students (Dünser et al. 2006; Tosik Gün and Atasoy 2017). Further research on groups with different education levels and technological experience is necessary to test these hypotheses. In addition, increasing the sample size of the undergraduate population could increase the variation of students' responses and improve the ability to identifying significant factors impacting student learning.

Participants who had more negative experiences scored lower on the learning measures as indicated by the positive correlation between liking and score on both the narrative and spatial test. In contrast, those who reacted positively to the experiment had a range of scores on the final learning measure. This result is consistent with other studies that show enjoyment of a task correlated with increased learning (Goetz et al. 2006; Hou et al. 2022), although game enjoyment does not always significantly increase learning outcomes (Hou et al. 2022; Imlig-Iten and Petko 2016).

5.1.1 Limitations and Future Work

Our study looked at the interaction between different form factors and scales of AR. Ultimately these factors were not found to significantly influence learning. Notably our study did not investigate how these different modes of AR compared to other interfaces. Future work could examine how our results compare to other XR interfaces such as VR; as well as more traditional interfaces, such as XR touch screens and mouse and keyboard. The type of interface may impact the immersive experience of the AR environment. Though we found no relationship between Immersion and learning outcomes, further studies are needed in order to determine why studies in the digital interactive environment have produced conflicting results.

More information could potentially be gained in future studies by refining the questionnaire. It may be valuable to divide the enjoyment category into more than one covariant. More insight into the factors influencing the learning outcomes could be gained by including an Attitude category as suggested by (Ab Jalil et al. 2020). In addition, reworking the questions to improve the consistency scores for novelty and interactivity could provide more information about how student experiences influenced learning. (Downs et al. 2022) showed students found the VR experience to be a novel experience, but we were unable to use novelty in our analyses due to a lack of internal consistency in answers measuring this variable. In addition, as in (Downs et al. 2022), we could not use interactivity (which they termed engagement) in our analysis due to the lack of consistency between question results. By reworking the questions to improve internal consistency future work may be able to utilize these factors.

5.2 Conclusion

Our study examined how scale and form factor affect learning in augmented reality. We created a 2x2 study that examined these factors across four conditions: Hololens2-Room, Hololens2-Table,iPad-Room , and iPad table. In examining scores on narrative and spatial test, we found no statistical differences between these four conditions. Thus scale and form factor do not appear to have a significant impact on learning, and this conclusion has significant implications for the design and application of augmented reality in variety of learning environments.

A Appendix A Measures

A.1 Pre-experiment questionnaire

A. Pre-experiment questionnaire: Please answer these questions to the best of your ability. If you do not feel like answering a question, leave it blank.

1. Age in years

2. Biological Sex

3. Place an "X" by the term that best describes your relationship with the University: Student Staff Faculty

Duluth Community Member Other (specify)

4. What college are you affiliated with? College of Liberal Arts Labovitz School of Business Economics School of Fine Arts Swenson College of Science Engineering CEHSP Other (specify)

5. If you are a student, what is your major?

6. If you are a student, what academic year are you currently in? Freshman Sophomore Junior Senior

Senior + Grad student

7. ACT Score (or) SAT Score

8. Current overall college GPA (Mark here if you do not yet have a GPA)

Media Use

9. How much time do you spend watching TV or streaming (non-interactive) media on a typical:

Weekday (Monday – Friday): Hours Minutes per day

Weekend (Saturday – Sunday): Hours Minutes per day

10. How much time do you spend using a computer on a typical:

Weekday (Monday – Friday): Hours Minutes per day

Weekend (Saturday – Sunday): Hours Minutes per day

11. How much time do you play interactive digital/video games (console, computer, or phone) on a typical:

Weekday (Monday – Friday): Hours Minutes per day

Weekend (Saturday – Sunday): Hours Minutes per day

12. How much time do you spend using your smartphone on a typical:

Weekday (Monday – Friday): Hours Minutes per day

Weekend (Saturday – Sunday): Hours Minutes per day

13. How much time do you spend using virtual reality technology on a typical:

Weekday (Monday – Friday): Hours Minutes per day

Weekend (Saturday – Sunday): Hours Minutes per day

14. If you have had experience with virtual reality technology, please check all that you have used. Oculus Rift HTC VIVE PlayStation VR Samsung Gear Google

Cardboard Other

15. How much total experience do you have with virtual reality systems? Please select one option: 0-10 hours 11-20 hours 21-30 hours- 31-40 hours 41–50 hours

16. If you have had used virtual reality technology in the past, briefly describe the types of experiences (such as games, simulations, movies or other immersive activities) that you have participated in.

Medical History

17. Do you have a history of suffering from motion sickness? No Yes

If yes, circle the number that best describes the severity:

Non-existant 1 – 2 – 3 – 4 – 5 – 6 – 7 Severe

18. How are you feeling today?

Not Well / Sick 1 – 2 – 3 – 4 – 5 – 6 – 7 Great! Very healthy

19. Have you ever been diagnosed with any brain-related condition or disease: recent concussions, meningitis, spinal cord injury, Parkinson's, Alzheimer's, dementia, encephalitis, epilepsy, and/or restless leg syndrome? Yes No

20. Before we begin the experiment, are you feeling any sort of nausea, dizziness, vision impairment, or any other condition that would prohibit you from taking part in this experiment? No Yes

A.2 Narrative Questions

Part B. VR Narrative Questions: Please write the letter of the correct response in the blank next to the question number.

1. Who is responsible for most of the deaths in Spoon River? A. Holden the Cook B. Petrus Van Tassel C. Hessian Soldier D. Aaron Hatfield

2. Who was the little girl in the story that died after the cook tricked her into getting spoiled fruits and vegetables from the trash behind the store? A. Minerva Jones B. Daisy Hatfield C. Nellie Throckmortin D. Marie Rhodes

3. Which of the following residents worked as a grave-digger for Spoon River Cemetery? A. Silas Rhodes B. Butch Welty C. Blind Jack D. Peter Hatfield

4. Who owned a lumber mill? A. Aaron Hatfield B. Hodd Hatfield C. Rolf Rhodes D. Petrus Van Tassel

5. Who was the store owner in Spoon River? A. James Lindsay B. Minerva Jones C. Justice Rhodes D. Marie Rhodes

6. Which of the following people was married? A. Petit the Poet B. Minerva Jones

C. Sara Throckmortin D. Hodd Hatfield

7. Who was the nurse in Spoon River? A. Minerva Jones B. Daisy Hatfield C. Nellie Throckmortin D. Frances Throckmortin

8. Among the residents in the Spoon River Cemetery, who has a wooden grave marker? A. Blind Jack B. Hessian Soldier C. Petit the Poet D. Holden the Cook

9. Who could barely hold a job because he was always unfit? A. Editor Whedon B. Frances Throckmortin C. Blind Jack D. Petit the Poet

10. Who got caught under a wagon wheel? A. Hessian Soldier B. Blind Jack C. Butch Welty D. Aaron Hatfield

11. Who mentions that they had rich, important friends? A. Lindsey Jones B. Frances Throckmortin C. Editor Whedon D. Aaron Hatfield

12. Whose tombstone stated that "business is booming"? A. Minerva Jones B. Butch Welty C. Holden the Cook D. Editor Whedon

13. What killed the Hessian Soldier? A. Gun fire B. Cannon ball C. His horse D. Bayonet

14. Who thought they knew the whole story of what happened in Spoon River? A. Butch Welty B. Petit the Poet C. Editor Whedon D. Daisy Hatfield

15. Which of the residents died after eating an apple? A. Daisy Hatfield B. Nellie Throckmortin C. Butch Welty D. None of the above

16. Who killed the Hessian Soldier? A. Frances Throckmortin B. Holden the Cook C. Butch Welty D. None of the above

17. How many tombstones belonged to the Throckmortin family? A. One B. Two C. Three D. Four

18. How many families (of more than one person with the same last name) were in the cemetery? A. One B. Two C. Three D. Four

19. Which of the following adjectives best describes the Holden the Cook? A. Hero

B. Villain C. Hard-worker D. Father

20. What was Hodd Hatfield's profession? A. Lumberjack B. Newspaper Editor C. Doctor D. Grave digger

A.3 Likert survey

A.3.1 Novelty

1. The experience I had today using this XR technology was a new one for me.
2. The experience I had today with this XR technology was very routine for me.
3. This was the first time I have used XR technology like this before.

A.3.2 Liking

1. I enjoyed the experimental task I participated in today.
2. I thought that the task I participated in was frustrating.
3. I had fun participating in the experiment today.
4. I was disappointed participating in the task today.
5. I would like to have experiences like this again in the future.
6. I thought the task I participated in today was boring.

A.3.3 Immersion

1. My experience today was involving.
2. My experience was intense.

3. I felt like I was physically inside the XR environment.
4. I felt immersed in the XR environment.
5. I felt like I was surrounded by the XR environment.

A.3.4 Interactivity

1. The virtual world was responsive to actions that I initiated.
2. I was aware of events occurring in the lab space when I was in the virtual world.
3. It was easy to manipulate objects in the virtual world.
4. The virtual world made me feel disoriented.
5. Using the control mechanisms was intuitive.
6. I got proficient in moving around through the virtual environment.
7. The visual display interfered with my ability to perform the required activities.
8. I could concentrate on the assigned tasks in the virtual environment because the control mechanisms were easy to use.
9. I felt nauseous when I was in the virtual environment.
10. My eyes felt strained when I was in the virtual environment.

B Appendix B Tables

B.1 Bivariate Correlation Table

Correlations

		Age in years	Biological sex	Narrative Test Score
Age in years	Pearson Correlation	1	-0.161	0.065
	Sig. (2-tailed)		0.065	0.461
	N	132	132	132
Biological sex	Pearson Correlation	-0.161	1	-.182*
	Sig. (2-tailed)	0.065		0.036
	N	132	132	132
Narrative Test Score	Pearson Correlation	0.065	-.182*	1
	Sig. (2-tailed)	0.461	0.036	
	N	132	132	132
ACT test score	Pearson Correlation	0.111	-0.015	.5**
	Sig. (2-tailed)	0.275	0.880	0.001
	N	99	99	99
Current GPA	Pearson Correlation	-0.066	0.175	0.142
	Sig. (2-tailed)	0.637	0.211	0.309
	N	53	53	53
Spatial Test Score	Pearson Correlation	-0.057	-0.079	.547**
	Sig. (2-tailed)	0.518	0.369	0.000
	N	131	131	131
Likeing	Pearson Correlation	0.094	0.097	.322**
	Sig. (2-tailed)	0.286	0.270	0.000
	N	132	132	132
Immersion	Pearson Correlation	-0.072	0.135	0.009
	Sig. (2-tailed)	0.413	0.122	0.918
	N	132	132	132

Correlations

		ACT test score	Current GPA	Spatial Test Score
Age in years	Pearson Correlation	0.111	-0.066	-0.057
	Sig. (2-tailed)	0.275	0.637	0.518
	N	99	53	131
Biological sex	Pearson Correlation	-0.015	0.175	-0.079
	Sig. (2-tailed)	0.880	0.211	0.369
	N	99	53	131
Narrative Test Score	Pearson Correlation	.335**	0.142	.547**
	Sig. (2-tailed)	0.001	0.309	0.000
	N	99	53	131
ACT test score	Pearson Correlation	1	.310*	0.133
	Sig. (2-tailed)		0.045	0.191
	N	99	42	99
Current GPA	Pearson Correlation	.310*	1	0.062
	Sig. (2-tailed)	0.045		0.664
	N	42	53	52
Spatial Test Score	Pearson Correlation	0.133	0.062	1
	Sig. (2-tailed)	0.191	0.664	
	N	99	52	131
Likeing	Pearson Correlation	0.007	0.173	.230**
	Sig. (2-tailed)	0.945	0.217	0.008
	N	99	53	131
Immersion	Pearson Correlation	-0.139	0.212	0.023
	Sig. (2-tailed)	0.170	0.128	0.793
	N	99	53	131

Correlations

		Likeing	Immersion
Age in years	Pearson Correlation	0.094	-0.072
	Sig. (2-tailed)	0.286	0.413
	N	132	132
Biological sex	Pearson Correlation	0.097	0.135
	Sig. (2-tailed)	0.270	0.122
	N	132	132
Narrative Test Score	Pearson Correlation	.322**	0.009
	Sig. (2-tailed)	0.000	0.918
	N	132	132
ACT test score	Pearson Correlation	0.007	-0.139
	Sig. (2-tailed)	0.945	0.170
	N	99	99
Current GPA	Pearson Correlation	0.173	0.212
	Sig. (2-tailed)	0.217	0.128
	N	53	53
Spatial Test Score	Pearson Correlation	.230**	0.023
	Sig. (2-tailed)	0.008	0.793
	N	131	131
Likeing	Pearson Correlation	1	.480**
	Sig. (2-tailed)		0.000
	N	132	132
Immersion	Pearson Correlation	.480**	1
	Sig. (2-tailed)	0.000	
	N	132	132

*. Correlation is significant at the 0.05 level (2-tailed)

**. Correlation is significant at the 0.01 level (2-tailed)

B.2 MANCOVA

Descriptive Statistics

Condition		Mean	Std. Deviation	N
Number of correct tombstone placements in the spatial test	Hololens Room Scale	7.1667	3.61959	24
	Hololens desktop	6.7037	3.48419	27
	Tablet Room Scale	6.9130	2.85901	23
	Tablet desktop	6.6400	2.36079	25
	Total	6.8485	3.08837	99
Number of correct test answers out of 20 questions on narrative test	Hololens Room Scale	14.4583	3.16199	24
	Hololens desktop	13.8148	4.24298	27
	Tablet Room Scale	14.0000	3.28910	23
	Tablet desktop	14.3200	3.17175	25
	Total	14.1414	3.47589	99

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	0.109	5.671[b]	2.000	93.000	0.005
	Wilks' Lambda	0.891	5.671[b]	2.000	93.000	0.005
	Hotelling's Trace	0.122	5.671[b]	2.000	93.000	0.005
	Roy's Largest Root	0.122	5.671[b]	2.000	93.000	0.005
ACT	Pillai's Trace	0.123	6.516[b]	2.000	93.000	0.002
	Wilks' Lambda	0.877	6.516[b]	2.000	93.000	0.002
	Hotelling's Trace	0.140	6.516[b]	2.000	93.000	0.002
	Roy's Largest Root	0.140	6.516[b]	2.000	93.000	0.002
Condition number	Pillai's Trace	0.019	0.296	6.000	188.000	0.938
	Wilks' Lambda	0.981	.293[b]	6.000	186.000	0.939
	Hotelling's Trace	0.019	0.291	6.000	184.000	0.941
	Roy's Largest Root	0.016	.506[c]	3.000	94.000	0.679

a. Design: Intercept + ACT + Condition number

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Number of correct tombstone placements in the SR sand tray	19.046^{a}	4	4.761	0.489	0.744
	Number of correct test answers out of 20 questions	145.825^{b}	4	36.456	3.301	0.014
Intercept	Number of correct tombstone placements in the SR sand tray	51.694	1	51.694	5.307	0.023
	Number of correct test answers out of 20 questions	121.961	1	121.961	11.043	0.001
ACT	Number of correct tombstone placements in the SR sand tray	14.867	1	14.867	1.526	0.220
	Number of correct test answers out of 20 questions	139.277	1	139.277	12.610	0.001

a. R Squared = .020 (Adjusted R Squared = -.021)

b. R Squared = .123 (Adjusted R Squared = .086)

Tests of Between-Subjects Effects

Condition number	Number of correct tombstone placements in the SR sand tray	2.595	3	0.865	0.089	0.966
	Number of correct test answers out of 20 questions	12.702	3	4.234	0.383	0.765
Error	Number of correct tombstone placements in the SR sand tray	915.682	94	9.741		
	Number of correct test answers out of 20 questions	1038.196	94	11.045		
Total	Number of correct tombstone placements in the SR sand tray	5578.000	99			
	Number of correct test answers out of 20 questions	20982.000	99			
Corrected Total	Number of correct tombstone placements in the SR sand tray	934.727	98			
	Number of correct test answers out of 20 questions	1184.020	98			

a. R Squared = .020 (Adjusted R Squared = -.021)

b. R Squared = .123 (Adjusted R Squared = .086)

Estimated Marginal Means

Condition

Dependent Variable	Condition	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Spatial test score	Hololens Room Scale	7.067[a]	0.642	5.792	8.342
	Hololens desktop	6.656[a]	0.602	5.461	7.851
	Tablet Room Scale	6.950a	0.651	5.656	8.244
	Tablet desktop	6.753[a]	0.631	5.501	8.006
Narrative test score	Hololens Room Scale	14.153[a]	0.684	12.795	15.510
	Hololens desktop	13.668[a]	0.641	12.396	14.941
	Tablet Room Scale	14.113[a]	0.694	12.736	15.491
	Tablet desktop	14.667[a]	0.672	13.333	16.001

a. Covariates appearing in the model are evaluated at the following values: ACT test score = 24.0707.

B.3 Liking effects on narrative test

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.322[a]	0.104	0.097	3.61577
2	.346[b]	0.119	0.106	3.59833
3	.369[c]	0.136	0.116	3.57751

Change Statistics

Model	R Square Change	F Change	df1	df2	Sig. F Change
1	0.104	15.085	1	130	0.000
2	0.015	2.264	1	129	0.135
3	0.017	2.506	1	128	0.116

c. Predictors: (Constant), Likeing, Likeing^2, Likeing^3

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	13.962	0.315		44.366	0.000
	Likeing	1.224	0.315	0.322	3.884	0.000
2	(Constant)	14.344	0.403		35.589	0.000
	Likeing	0.858	0.397	0.226	2.161	0.033
	$Likeing^2$	-0.383	0.254	-0.157	-1.505	0.135
3	(Constant)	13.911	0.485		28.663	0.000
	Likeing	0.390	0.493	0.103	0.792	0.430
	$Likeing^2$	0.387	0.548	0.159	0.706	0.481
	$Likeing^3$	0.351	0.222	0.424	1.583	0.116

a. Dependent Variable: Number of correct test answers out of 20 questions

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	197.213	1	197.213	15.085	<.001[b]
	Residual	1699.597	130	13.074		
	Total	1896.811	131			
2	Regression	226.522	2	113.261	8.747	<.001[c]
	Residual	1670.288	129	12.948		
	Total	1896.811	131			
3	Regression	258.595	3	86.198	6.735	<.001d
	Residual	1638.216	128	12.799		
	Total	1896.811	131			

a. Dependent Variable: Number of correct test answers out of 20 questions

b. Predictors: (Constant), Likeing

c. Predictors: (Constant), Likeing, $Likeing^2$

d. Predictors: (Constant), Likeing, $Likeing^2$, $Likeing^3$

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	$Likeing^2$	-.157[b]	-1.505	0.135	-0.131	0.624
	$Likeing^3$	.256[b]	2.077	0.040	0.180	0.441
2	$Likeing^3$	.424[c]	1.583	0.116	0.139	0.094

a. Dependent Variable: Number of correct test answers out of 20 questions
b. Predictors in the Model: (Constant), Likeing
c. Predictors in the Model: (Constant), Likeing, $Likeing^2$

B.4 Likeing effect on Spatial test

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of Estimate	
1	.230a	0.053	0.045	3.06370	
2	.248[b]	0.061	0.047	3.06152	
3	.263[c]	0.069	0.047	3.06076	

Change Statistics

Model	R Square Change	F Change	df1	df2	Sig. F Change
1	0.053	7.183	1	129	0.008
2	0.009	1.184	1	128	0.279
3	0.008	1.064	1	127	0.304

a. Predictors: (Constant), Likeing

b. Predictors: (Constant), Likeing, Likeing^2

c. Predictors: (Constant), Likeing, Likeing^2, Likeing^3

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	67.420	1	67.420	7.183	.008[b]
	Residual	1210.824	129	9.386		
	Total	1278.244	130			
2	Regression	78.513	2	39.256	4.188	.017[c]
	Residual	1199.731	128	9.373		
	Total	1278.244	130			
3	Regression	88.479	3	29.493	3.148	.027d
	Residual	1189.765	127	9.368		
	Total	1278.244	130			

a. Dependent Variable: Spatial Test Score

b. Predictors: (Constant), Likeing

c. Predictors: (Constant), Likeing, $Likeing^2$

d. Predictors: (Constant), Likeing, $Likeing^2$, $Likeing^3$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6.861	0.268		25.632	0.000
	Likeing	0.719	0.268	0.230	2.680	0.008
2	(Constant)	7.095	0.343		20.689	0.000
	Likeing	0.490	0.341	0.156	1.437	0.153
	$Likeing^2$	-0.237	0.217	-0.118	-1.088	0.279
3	(Constant)	6.852	0.416		16.484	0.000
	Likeing	0.233	0.422	0.074	0.551	0.582
	$Likeing^2$	0.197	0.473	0.098	0.416	0.678
	$Likeing^3$	0.196	0.190	0.289	1.031	0.304

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	$Likeing^2$	-.118[b]	-1.088	0.279	-0.096	0.618
	$Likeing^3$	.186[b]	1.445	0.151	0.127	0.441
2	$Likeing^3$	.289[c]	1.031	0.304	0.091	0.093

a. Dependent Variable: Number of correct tombstone placements in the SR sand tray

b. Predictors in the Model: (Constant), Likeing

c. Predictors in the Model: (Constant), Likeing, $Likeing^2$

B.5 Immersions effect on Narrative test

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	
1	.009[a]	0.000	-0.008	3.81964	
2	.060b	0.004	-0.012	3.82767	
3	.184[c]	0.034	0.011	3.78369	

Change Statistics

Model	R Square Change	F Change	df1	df2	Sig. F Change
1	0.000	0.011	1	130	0.918
2	0.004	0.455	1	129	0.501
3	0.030	4.016	1	128	0.047

a. Predictors: (Constant), immersion

b. Predictors: (Constant), immersion, immersion^2

c. Predictors: (Constant), immersion, immersion^2, immersion^3

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	0.154	1	0.154	0.011	.918[b]
	Residual	1896.657	130	14.590		
	Total	1896.811	131			
2	Regression	6.820	2	3.410	0.233	.793[c]
	Residual	1889.990	129	14.651		
	Total	1896.811	131			
3	Regression	64.319	3	21.440	1.498	.218d
	Residual	1832.491	128	14.316		
	Total	1896.811	131			

a. Dependent Variable: Number of correct test answers out of 20 questions
b. Predictors: (Constant), immersion
c. Predictors: (Constant), immersion, immersion^2
d. Predictors: (Constant), immersion, immersion^2, immersion^3

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	13.962	0.332		41.997	0.000
	immersion	0.029	0.281	0.009	0.103	0.918
2	(Constant)	13.769	0.439		31.364	0.000
	immersion	0.029	0.281	0.009	0.102	0.919
	immersion^2	0.138	0.204	0.059	0.675	0.501
3	(Constant)	13.626	0.440		30.982	0.000
	immersion	-0.816	0.505	-0.255	-1.616	0.109
	immersion^2	0.239	0.208	0.103	1.150	0.252
	immersion^3	0.256	0.128	0.319	2.004	0.047

a. Dependent Variable: Number of correct test answers out of 20 questions

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	immersion^2	.059[b]	0.675	0.501	0.059	1.000
	immersion^3	.275[b]	1.776	0.078	0.154	0.316
2	immersion^3	.319[c]	2.004	0.047	0.174	0.298

a. Dependent Variable: Number of correct test answers out of 20 questions
b. Predictors in the Model: (Constant), immersion
c. Predictors in the Model: (Constant), immersion, immersion^2

B.6 Immersions effect on Spatial test

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.023[a]	0.001	-0.007	3.14699
2	.039[b]	0.002	-0.014	3.15773
3	.163[c]	0.027	0.004	3.13012

Change Statistics

Model	R Square Change	F Change	df1	df2	Sig. F Change
1	0.001	0.069	1	129	0.793
2	0.001	0.124	1	128	0.725
3	0.025	3.268	1	127	0.073

a. Predictors: (Constant), immersion

b. Predictors: (Constant), immersion, $immersion^2$

c. Predictors: (Constant), immersion, $immersion^2$, $immersion^3$

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	0.687	1	0.687	0.069	.793[b]
	Residual	1277.558	129	9.904		
	Total	1278.244	130			
2	Regression	1.923	2	0.961	0.096	.908[c]
	Residual	1276.321	128	9.971		
	Total	1278.244	130			
3	Regression	33.945	3	11.315	1.155	.330d
	Residual	1244.299	127	9.798		
	Total	1278.244	130			

a. Dependent Variable: Number of correct tombstone placements in the SR sand tray

b. Predictors: (Constant), immersion

c. Predictors: (Constant), immersion, $immersion^2$

d. Predictors: (Constant), immersion, $immersion^2$, $immersion^3$

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	$immersion^2$	.031[b]	0.352	0.725	0.031	1.000
	$immersion^3$	.260b	1.668	0.098	0.146	0.315
2	$immersion^3$	.291[c]	1.808	0.073	0.158	0.296

a. Dependent Variable: Number of correct tombstone placements in the SR sand tray

b. Predictors in the Model: (Constant), immersion

c. Predictors in the Model: (Constant), immersion, $immersion^2$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6.856	0.275		24.933	0.000
	immersion	0.061	0.232	0.023	0.263	0.793
2	(Constant)	6.772	0.363		18.642	0.000
	immersion	0.061	0.233	0.023	0.262	0.794
	$immersion^2$	0.059	0.168	0.031	0.352	0.725
3	(Constant)	6.661	0.365		18.231	0.000
	immersion	-0.574	0.420	-0.218	-1.366	0.174
	$immersion^2$	0.136	0.172	0.071	0.788	0.432
	$immersion^3$	0.191	0.106	0.291	1.808	0.073

a. Dependent Variable: Number of correct tombstone placements in the SR sand tray

References

Ab Jalil, H., Nasharuddin, N. A., Marlisah, E., Mohd Nazan, A. I. N., Ismail, I., Ma'rof, A. M., Rusdi, N., & Zaremohzzabieh, Z. (2020). Systematic review of enjoyment element in health- related game-based learning. *International Journal of Emerging Technologies in Learning (iJET)*, *15*, 40–57. https://doi.org/10.3991/ijet.v15i21.17345 (cit. on p. 29)

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, *20*, 1–11. https://doi.org/https://doi.org/10.1016/j.edurev.2016.11.002 (cit. on pp. 2, 3, 6)

Akçayır, M., Akçayır, G., Pektaş, H. M., & Ocak, M. A. (2016). Augmented reality in science laboratories: The effects of augmented reality on university students' laboratory skills and attitudes toward science laboratories. *Computers in Human Behavior*, *57*, 334–342. https://doi.org/https://doi.org/10.1016/j.chb.2015.12.054 (cit. on p. 6)

Antonioli, M., Blake, C., & Sparks, K. (2014). Augmented reality applications in education. *The Journal of technology studies*, 96–107 (cit. on p. 6).

Balog, A., Pribeanu, C., & Iordache, D. (2007). Augmented reality in schools: Preliminary evaluation results from a summer school. *International Journal of Social Sciences*, *2*(3), 163–166 (cit. on p. 6).

Batista, J. M., Muniz, K. A. S., Botelho, D., & Guimarães, B. S. (2020). I want to pay more but i don't remember what you just said: Augmented reality effects on attitude, willingness to pay and memory recall (cit. on p. 10).

Bell, J. T., & Fogler, H. S. (1995). Investigation and application of virtual reality as an educational tool (cit. on p. 6).

Bujak, K. R., Radu, I., Catrambone, R., MacIntyre, B., Zheng, R., & Golubski, G. (2013). A psychological perspective on augmented reality in the mathematics classroom. *Computers & Education*, *68*, 536–544. https://doi.org/10.1016/j.compedu.2013.02.017 (cit. on pp. 6, 8)

Cai, S., Wang, X., & Chiang, F.-K. (2014). A case study of augmented reality simulation system application in a chemistry course. *Computers in Human Behavior*, *37*, 31–40. https://doi.org/https://doi.org/10.1016/j.chb.2014.04.018 (cit. on pp. 1, 7)

Cai, S., Wang, X., Gao, M., & Yu, S. (2012). Simulation teaching in 3d augmented reality environment. *2012 IIAI International Conference on Advanced Applied Informatics*, 83–88. https://doi.org/10.1109/IIAI-AAI.2012.25 (cit. on p. 6)

Chance, S. S., Gaunet, F., Beall, A. C., & Loomis, J. M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence*, *7*, 168–178 (cit. on p. 9).

Chen, C., Zhang, L., Luczak, T., Smith, E., & Burch, R. F. (2019). Using microsoft hololens to improve memory recall in anatomy and physiology: A pilot study to examine the efficacy of using augmented reality in education. *ournal of Educational Technology Development and Exchange (JETDE)*, *12*, 105–135. https://doi.org/10.18785/jetde.1201.02 (cit. on p. 10)

Chien, Y.-C., Su, Y.-N., Wu, T.-T., & Huang, Y.-M. (2019). Enhancing students' botanical learning by using augmented reality. *Univ. Access Inf. Soc.*, *18*(2), 231–241. https://doi.org/10.1007/s10209-017-0590-4 (cit. on p. 6)

del Cerro Velázquez, F., & Morales Méndez, G. (2021). Systematic review of the development of spatial intelligence through augmented reality in stem knowledge areas. *Mathematics*, *9*(23). https://doi.org/10.3390/math9233067 (cit. on p. 1)

Downs, E. P., Willemsen, P., Leff, D., Adigun, C., Kothapalli, S. S., Miller, L., Boynton, K., & Berndt, M. M. (2022). Grave errors: Exploring the influence of motion mechanics on learning outcomes in a virtual cemetery. *Technology, Mind, and Behavior - Special Issue: Learning in Immersive Virtual Reality*, *3* (cit. on pp. 1, 10, 29).

Dünser, A., Steinbügl, K., Kaufmann, H., & Glück, J. (2006). Virtual and augmented reality as spatial ability training tools. *Proceedings of the 7th ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction: Design Centered HCI*, 125–132. https://doi.org/10.1145/1152760.1152776 (cit. on pp. 7, 28)

Fiorentino, M., Laera, F., Evangelista, A., Boccaccio, A., Manghisi, V., Gattullo, M., Uva, A., Gabbard, J., & Foglia, M. (2021). Sailing data visualization in augmented reality: Systematic review, issues, and perspectives. *Marine Technology Society Journal*, *55*, 64–80. https://doi.org/10.4031/MTSJ.55.2.2 (cit. on p. 6)

Furió, D., González-Gancedo, S., Juan, M.-C., Seguí, I., & Costa, M. (2013). The effects of the size and weight of a mobile device on an educational game. *Computers Education*, *64*, 24–41. https://doi.org/https://doi.org/10.1016/j.compedu.2012.12.015 (cit. on p. 3)

Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of augmented reality on students' learning gains. *Educational Research Review*, *27*, 244–260. https://doi.org/https://doi.org/10.1016/j.edurev.2019.04.001 (cit. on pp. 1, 6, 7, 28)

Goetz, T., Hall, N. C., Frenzel, A. C., & Pekrun, R. (2006). A hierarchical conceptualization of enjoyment in students. *Learning and Instruction*, *16*(4), 323–338. https://doi.org/https://doi.org/10.1016/j.learninstruc.2006.07.004 (cit. on p. 28)

Hou, X., Nguyen, H. A., Richey, J. E., Harpstead, E., Hammer, J., & McLaren, B. M. (2022). Learning with serious games: Is fun playing the game a predictor of learning success? *British Journal of Educational Technology*, *47*, 151–163. https://doi.org/10.1111/bjet.12226 (cit. on p. 28)

Hsi, S., Linn, M. C., & Bell, J. E. (1997). The role of spatial reasoning in engineering and the design of spatial instruction. *Journal of Engineering Education*, *86*(2), 151–158. https://doi.org/https://doi.org/10.1002/j.2168-9830.1997.tb00278.x (cit. on p. 7)

Huang, K.-T., Ball, C., Francis, J., Ratan, R., Boumis, J., & Fordham, J. (2019). Augmented versus virtual reality in education: An exploratory study examining science knowledge retention when using augmented reality/virtual reality mobile applications [PMID: 30657334]. *Cyberpsychology, Behavior, and Social Networking*, *22*(2), 105–110. https://doi.org/10.1089/cyber.2018.0150 (cit. on p. 9)

Hughes, C. L., Fidopiastis, C., Stanney, K. M., Bailey, P. S., & Ruiz, E. (2020). The psychometrics of cybersickness in augmented reality. *Frontiers in Virtual Reality*, *1*. https://doi.org/10.3389/frvir.2020.602954 (cit. on p. 5)

Imlig-Iten, N., & Petko, D. (2016). Assessing the effects of open models of learning and enjoyment in a digital learning game. *International Journal of Artificial Intelligence in Education*, *32*. https://doi.org/10.1007/s40593-021-00250-6 (cit. on p. 28)

Ipad pro. (2022). apple. https://www.apple.com/ipad-pro/specs/. (Cit. on p. 5)

Jones, M. D., & Lawler, M. J. (2019). Delivering sign language in a live planetarium show using head-mounted displays and infrared light. *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 396–401. https://doi.org/10.1145/3308561.3353809 (cit. on p. 6)

Jung, T., Tom Dieck, M. C., Lee, H., & Chung, N. (2016a). Effects of virtual reality and augmented reality on visitor experiences in museum. https://doi.org/10.1007/978-3-319-28231-2_45. (Cit. on p. 6)

Jung, T., Tom Dieck, M. C., Lee, H., & Chung, N. (2016b). Effects of virtual reality and augmented reality on visitor experiences in museum. https://doi.org/10.1007/978-3-319-28231-2_45. (Cit. on p. 6)

Katsioloudis, P., Jovanovic, V., & Jones, M. (2014). A comparative analysis of spatial visualization ability and drafting models for industrial and technology education students. *Journal of Technology Education*, *26*, 88–104. https://doi.org/10.21061/jte.v26i1.a.6 (cit. on p. 6)

Lei, Z., Zhou, H., Hu, W., Deng, Q., Zhou, D., Liu, Z.-W., & Gao, X. (2021). 3-d interactive control laboratory for classroom demonstration and online experimentation in engineering education. *IEEE Transactions on Education*, *64*(3), 276–282. https://doi.org/10.1109/TE.2020.3041070 (cit. on p. 1)

Leopold, C., Górska, R., & Sorby, S. (2001). International experiences in developing visualization abilities of engineering students. *Journal for Geometry and Graphics Volume*, *5*, 81–91 (cit. on p. 7).

Lohman, D., & Kyllonen, P. (1983). Individual differences in solution strategy on spatial tasks, 105–135 (cit. on p. 7).

Marsh, W. E., Kelly, J. W., Dark, V. J., & Oliver, J. H. (2013). Cognitive demands of semi-natural virtual locomotion. *Presence*, *22*(3), 216–234. https://doi.org/10.1162/PRES_a_00152 (cit. on p. 9)

Masters, E. L. (1916). *Spoon river anthology*. NY: Macmillan Co. (Cit. on p. 11).

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: a class of displays on the reality-virtuality continuum. In H. Das (Ed.), *Telemanipulator and telepresence technologies* (pp. 282–292). SPIE. https://doi.org/10.1117/12.197321. (Cit. on pp. 3, 4)

Ofek, E., Benko, h., Zheng, F., & Wilson, A. (2015). Fovear: Combining an optically see-through near-eye display with spatial augmented reality projections. https://doi.org/10.1145/2807442.2807493 (cit. on p. 5)

Onyancha, R., Derov, M., & Kinsey, B. (2009). Improvements in spatial ability as a result of targeted training and computer-aided design software use: Analyses of object geometries and rotation types. *Journal of Engineering Education*, *98*. https://doi.org/10.1002/j.2168-9830.2009.tb01014.x (cit. on pp. 1, 7)

Özdemir, M., Şahin, Ç., Arcagok, S., & demir, M. (2018). The effect of augmented reality applications in the learning process: A meta-analysis study. *Eurasian Journal of Educational Research (EJER)*, *74*, 165–186. https://doi.org/10.14689/ejer.2018.74.9 (cit. on pp. 1, 6)

Riecke, B. E., & Bülthoff, H. H. (2004). Spatial updating in real and virtual environments: Contribution and interaction of visual and vestibular cues. *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization*, 9–17. https://doi.org/10.1145/1012551.1012553 (cit. on p. 5)

Rieser, j., & Pick, J., H.L. (2018). Using locomotion to update spatial orientation: What changes with learning and development? *The emerging spatial mind* (pp. 77–103). Oxford University Press. https://doi.org/https://doi.org/10.1093/acprof:oso/9780195189223.003.0004. (Cit. on p. 8)

Rozado, D., Stephen, L., & Kottege, N. (2014). Interacting with objects in the environment using gaze tracking glasses and speech. *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*, 414–417. https://doi.org/10.1145/2686612.2686676 (cit. on p. 4)

Ruddle, R. A., Volkova, E., & Bülthoff, H. H. (2011). Walking improves your cognitive map in environments that are large-scale and large in extent. *ACM Trans. Comput.-Hum. Interact.*, *18*(2). https://doi.org/10.1145/1970378.1970384 (cit. on pp. 8–10, 26)

Shelton, B., & Hedley, N. (2003). Exploring a cognitive basis for learning spatial relationships with augmented reality. *ITLS Faculty Publications*, *1* (cit. on p. 8).

Sundar, S. S. (2008). The main model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73–100). The John D.; Catherine T. MacArthur Foundation Series on Digital Media; Learning. The MIT Press. (Cit. on pp. 2, 10).

Tosik Gün, E., & Atasoy, B. (2017). The effects of augmented reality on elementary school students' spatial ability and academic achievement. *TED EĞİTİM VE BİLİM*, *42*. https://doi.org/10.15390/EB.2017.7140 (cit. on pp. 1, 7, 28)

Wei, X., Weng, D., Liu, Y., & Wang, Y. (2015). Teaching based on augmented reality for a technical creative design course. *Computers & Education*, *81*, 221–234. https://doi.org/10.1016/j.compedu.2014.10.017 (cit. on p. 6)

Xie, X., Paris, R., McNamara, T., & Bodenheimer, B. (2018). The effect of locomotion modes on spatial memory and learning in large immersive virtual environments: A comparison of walking with gain to continuous motion control: 11th international conference, spatial cognition 2018, tübingen, germany, september 5-8, 2018, proceedings. https://doi.org/10.1007/978-3-319-96385-3_5. (Cit. on pp. 9, 10, 26)

Yates, F. (1966). The art of memory. *The emerging spatial mind* (pp. 77–103). Chicago: University of Chicago Press. (Cit. on p. 9).

Zanbaka, C., Lok, B., Babu, S., Ulinski, A., & Hodges, L. (2005). Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment. *IEEE Transactions on Visualization and Computer Graphics*, *11*(6), 694–705. https://doi.org/10.1109/TVCG.2005.92 (cit. on p. 9)

Zhu, F., & Grossman, T. (2020). Bishare: Exploring bidirectional interactions between smartphones and head-mounted augmented reality. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376233 (cit. on p. 3)