Docens Series in Education
ISSN: 2583-1054

Annetta & Newton, 2024

Volume 7, pp. 25-43

Received: 27th August 2023

Revised: 9th November 2023, 28th November 2023, 1st December 2023

Accepted: 14th September 2023

Date of Publication: 15th June 2024

This paper can be cited as: Annetta, L. & Newton, M. (2024). Perceived Cognitive Load of Extended Reality Serious Educational Games about Climate Change. Docens Series in Education, 7, 25-43

PERCEIVED COGNITIVE LOAD OF EXTENDED REALITY SERIOUS EDUCATIONAL GAMES ABOUT CLIMATE CHANGE

Leonard Annetta, Ph.D.
Taft Distinguished Professor of Science Education, East Carolina University, Flanagan Bldg., Greenville, NC, 27858 USA
Annettal16@ecu.edu

Mark Newton, Ph.D.
Assistant Professor of Science Education, East Carolina University, Flanagan Bldg., Greenville, NC, 27858 USA
newtonm19@ecu.edu

Abstract

Integrating extended reality (XR) into undergraduate classrooms is not a new concept. However, comparing identical content in subdomains of XR is unique. This study compared two undergraduate courses with objectives about climate change on the Outer Banks of North Carolina coast at a large university in the Mid-Atlantic region of the United States. The purpose of these courses was to examine human and environmental impacts of global climate change in a local context. Investigating the challenges facing North Carolina barrier islands, the class took a 5-day field trip to the Outer Banks of North Carolina and visited four sites where they used augmented reality (MR) to learn about the impact on climate change at those respected locations. The *comparison class immersed in virtual reality (iVR) of the four sites using the same information provided in the MR. 24 (6 MR and 18 iVR) participants completed the National Aeronautics and Space Agency Task Load Index (TLX) immediately after completion of either the respective MR or iVR based game. Independent samples Mann-Whitney U testing rejected the null hypotheses for temporal, effort, and performance only. An explanation for possible reasons for these results are discussed.*

Keywords

Extended Reality, Climate Change, Undergraduate, Science

1. Introduction

Climate change is a global phenomenon grounded in many social and scientific topics. This is a topic that often conjures contentious arguments around its origin and human influence to stop or at least slow its impact on the environment. Climate change's complexity creates challenges for providing authentic experiences for learners while making it a difficult concept to teach. As a result, many do not get to see the impact of climate change beyond overly sensationalized videos online leading to misconceptions about the causes and validity of climate change. International science educators have identified the need to explicitly teach about climate change in K-12 science, which was identified by the United Nations Educational, Scientific, and Cultural Organization (UNESCO). UNESCO seeks to create innovative teaching approaches through interdisciplinary practices (Mermer, 2010).

Controversial topics, like climate change, meld into a research paradigm known as Socioscientific Issues (SSI). Implementing an SSI pedagogical approach addresses the goals set forth by UNESCO. SSI is a pedagogical framework for which educators can introduce such seemingly antagonistic topics (e.g., climate change, evolution, cloning, etc.) while affording the learner a safe experience to access or collect data to make conclusions on their own and create origins for an argument favoring a perspective within the issue. This framework incorporates personally relevant and complex issues that are interdisciplinary and develop functional scientific literacy where individuals consider various lines of evidence, weigh a range of arguments, and consider the moral and ethical implications of potential resolutions to an issue (Zeidler & Newton, 2017). A further challenge is integrating SSI into a personalized, place-based learning experience. It has been argued that embedding SSI instruction in a place-based course allows students to become immersed in the community impacted by the contentious issue, which leads to greater personal relevance, perspective-taking, and development of empathy (Newton & Zeidler, 2020; Sadler, 2009). Ubiquitous technologies have the potential to supplement SSI; especially in place-based SSI.

Place-based SSI is not only challenging, but these experiences are expensive and often logistically demanding. These experiences require an abundance of planning, resources, and can be impacted by the weather (Dolphin et al., 2019; Zhao et al., 2020). To overcome these difficulties, Extended Reality (XR), a term that overarches such technologies as Virtual Reality (VR), Augmented Reality (AR), or Mixed Reality (MR), have potential to complement place-based learning. It is also important to operationally define what we call, immersive Virtual Reality (iVR). Many inexpensive VR applications provide the user a 360-degree image only. Although interesting and fun, it is not completely immersive or often engaging. In iVR, there are 6-degrees of viewpoint freedom in the virtual surroundings with audio and interactive hotspots creating a interesting, fun, and engaging immersive user experience. In this study, we use the term iVR.

Serious Educational Games (SEG) (Annetta, 2008) are digital game-based experienced designed for K-20 educational learners. The term is a derivative of Serious Games, which are defined as games designed for purposes other than entertainment. The "serious" adjective refers to video games used by defense, education, medicine, first responders, etc. industry sectors (Abt, 1970).

1.1. Research Questions

With respect to the immersive nature of a technology-enable and place-based SSI, researchers must begin to question cognitive demands on the learners as they attempt to assimilate new knowledge while learning how to navigate a technology. To this end, the research question of this study became:

What is the perceived cognitive load of participants while engaged in place-based Serous Educational Game delivered through either Mixed Reality or immersive Virtual Reality?

2. Rationale

2.1. Extended Reality (XR)

The technologies that fall under XR are quickly becoming a more relevant topic than ever because of the ubiquitous nature of the hardware used to deploy the experiences. In recent years, smart device enabled AR and MR educational activities have been developed. Additionally, low-cost VR experiences have also been adopted on smart phones or inexpensive VR headsets. The use of XR has the potential to either immerse learners in a technology-rich environment (AR/MR) at a physical location (place-based) and/or virtually take students to a physical location (VR). Environmental issues are topics that easily assimilate into XR applications because these issues often occur on large scales of both time and physical space, while also being potential harmful to humans. The personal relevance and embodied cognitive of this combination of SSI and XR is an untapped research paradigm for science teaching and learning.

In a study on VR immersion and interactivity, Petersen et. al., (2022) reported effects of interactivity and/or immersion on cognitive load, situational interest, and a sense of physical presence along with effects between immersion and interactivity on agency and embodied learning. Other studies have suggested that XR has impacted student learning and indicators of learning, such as self-efficacy while not imposing unnecessary cognitive demands during the learning process (X. Huang et al., 2023).

As in this study, comparing cognitive impact of VR to AR or MR is becoming an essential research agenda. To date, studies have shown VR to provide a greater sense of presence when compared and is also more immersive and engaging when it comes to a sense of user presence.to MR (Allcoat et al., 2021). AR, however, has shown to be a more effective medium for delivering auditory information through spatial presence. A possible explanation could be due to the increased cognitive demands associated with immersive experiences in VR (Huang et al., 2019). In a 2019 study, AR produced more individual excitement and activation than VR even though users felt a greater sense of presence in the VR condition (Giglioli et al., 2019). In an exploratory study comparing VR to AR, (Newton, Annetta & Bressler, 2023) found the VR group made better connections between the physical impacts of climate change and the social, political, economic ramifications of climate change. This study began to shed light on the importance of both instructional design and the design of the technology conditions and how cognitive load might affect those conditions.

2.2. Cognitive Load

Cognitive load research is well-established. Cognitive load is said to account for the key reason for the ineffectiveness of problem solving because cognitive processes insufficiently overlap and causes a great deal of cognitive processing during problem which is unavailable for schema acquisition (Sweller, 1988). As technology and multimedia increasingly become popular tools for teaching required school age content, cognitive load needs to be accounted for in the instructional design.

Investigating cognitive load on XR users is necessary to inform how the user experiences should be designed to access learners to acquire schema most effectively and efficiently. Wenk (2023) found cognitive load did not differ when comparing VR to AR. The instructional design in this study did, however, suggest VR was more motivating and usable than AR. A potential resolution to this conundrum is stratifying learners based on their learning style. Students who identify as visual learners tend to have better learning effectiveness than the verbal type students when learning through XR conditions. Sequential and global learners showed no learning differences by XR condition (Chen & Huang, 2023). AR also did not show learning gains in conceptual knowledge; however, it nonetheless suggested a significant lower extraneous cognitive load than traditional teaching and learning (IbİlĬ & BİllĬnghurst, 2019; Thees et al., 2020).

Gender is another consideration as it pertains to instructional design using XR. Males tend to perceive a relationship between ease of use and extraneous cognitive load while females show a perceived ease of use with intrinsic cognitive load. Females also show a perceived natural interaction strongly correlated with perceived usefulness. Both genders perceived ease of use and germane cognitive load similarly, however (IbİlĬ & BİllĬnghurst, 2019). Removing gender as a variable, VR has a marked effect on extraneous and germane cognitive load, but no influence on intrinsic cognitive load (Y. C. Chen et al., 2022). Overall, motor learning has a direct effect on cognitive load in VR that often result in learning and long-term motor memory (Juliano et al., 2022).

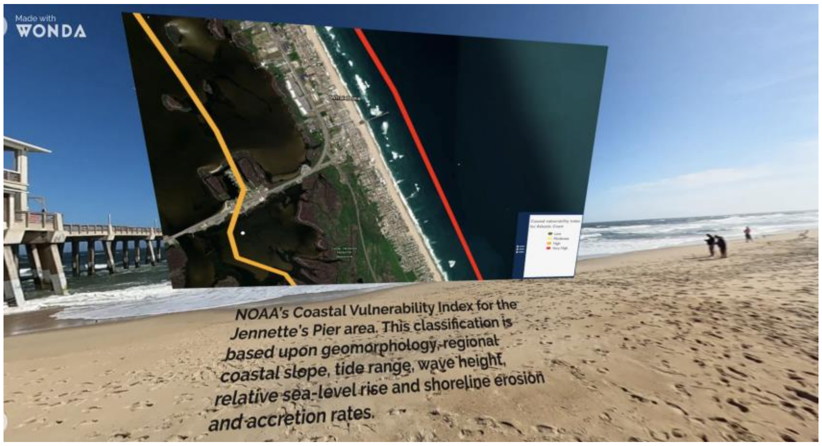
3. Methodology

3.1. Setting

This study took place during the spring semester of 2023 at a large southeastern university in the United States. Two groups of students engaged with four locations of the Outer Banks (OBX) barrier islands in North Carolina (Figure 1) through either an iVR condition or an MR condition. Each condition was a game-based scenario where students received in world points or artifact collection based on interaction with science content designed to teach about coastal resiliency including unique geological and historical sites from the chronological impacts of climate change in each location. The OBX physical locations were Jockey's Ridge State Park, Jennette's Pier, Oregon Inlet, and Cape Hatteras. These locations highlighted the fragility of these islands from human and storm impacts. A reoccurring theme for the experience was also discussing the role resiliency strategies like bridge construction near impacted areas and on-going road maintenance caused by sea level rise and increased storm frequency as a result of climate change. As students experienced the delicate ecosystems, they made cross-curricular connections and informed decisions about the complex issues facing North Carolina coastal communities.

Figure 1: *The North Carolina Outer Banks Barrier Island Chain*
(Source: https://www.visitob.com/plan-your-trip/outer-banks-nc-map/)

While on the islands, students first used a mobile MR SEG created in *Adobe Aero* for each of the four locations. Each MR SEG had embedded events (e.g., images, videos, etc.) to teach about specific aspects of climate change impact on the given location (figure 2). These embedded events provided historical data as well as future projections for the barrier islands based on current climate change data. Each student engaged with the MR SEG through their personal mobile phone that had enabled cellular data. A second group of students were not able to make the 2-hour trip to the place-based locations. A second group of students were not able to make the 2-hour trip to the place-based locations. This group interacted with the exact same content but in an iVR SEG condition created in WondaVR. The iVR SEGs were created using 360-degree images with the same events embedded at the same four locations in the OBX as the MR SEGs and deployed through a Meta Quest 2 headset. Ambient sound of wind and sea birds were included in the iVR game design to build a sense of presence and this auditory addition would complement the nature sounds the MR group experienced. Ambient sound of wind and sea birds were included in the iVR game design to build a sense of presence and this auditory addition would complement the nature sounds the MR group experienced.



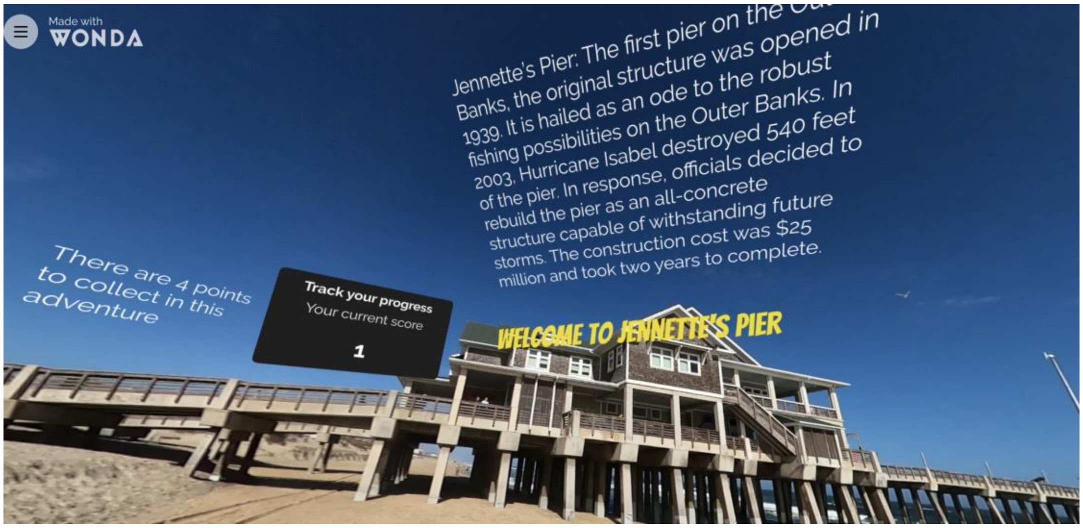
NOAA's Coastal Vulnerability Index for the Jennette's Pier area. This classification is based upon geomorphology, regional coastal slope, tide range, wave height, relative sea-level rise and shoreline erosion and accretion rates.

Figure 2: *Screen Capture from the iVR SEG Depicting Embedded Science Content (Source: Self from iVR Game)*

The MR game challenged players to walk the four locations in the OBX and unlock science content with the goal of collecting gold coins left by pirates. Each location had a minimum of three content artifacts in the form of a still image, video, or combination of the two. Players knew they had unlocked all content when all the coins were collected in the given location (figure 3).

Figure 3: *Gold coin collection at Cape Hatteras light house.*
(Source: Self from MR Game)

Because the WondaVR software had an interface that greatly differed from Adobe Aero, coin collecting was not an option to embed in the game logic. Instead, WondaVR allows the designer to create a scoreboard for the player and thus, gold coin collection in MR became a simple point total in the iVR game (figure 4).



Jennette's Pier: The first pier on the
Banks, the original structure was opened in 1939. It is hailed as an ode to the robust fishing possibilities on the Outer Banks. In 2003, Hurricane Isabel destroyed 540 feet of the pier. In response, officials decided to rebuild the pier as an all-concrete structure capable of withstanding future storms. The construction cost was $25 million and took two years to complete.

There are 4 points to collect in this adventure

Figure 4: *iVR SEG scoreboard at the Jennette's Pier Location*
(Source: Self from iVR Game)

3.2. Participants

Six MR condition student participants (4 female and 2 male) enrolled in an undergraduate honors seminar course traveled to experience the onsite locations and interacted in the MR SEG at each location. The focus of the course was on coastal resiliency to climate change on the Outer Banks of North Carolina. This topic was intentionally selected by the researchers to align the SSI's requirement that issues should be personally relevant to students, who in this case, live in proximity to the OBX. The course met weekly on campus for the first seven weeks of the semester for 75 minutes per meeting and then spent 5 days in the Outer Banks. The second group of 18 iVR condition students (12 female and 6 male) from an undergraduate earth science for teachers' class was not able to travel to the OBX and thus, met for two 75-minute classes per week in a traditional classroom environment. These students engaged in the iVR SEG that mirrored the content and interactions in MR SEGs. It is worth noting that there was no mention of either technology condition in the course description, meaning students' decision to enroll in either course was not influenced by the presence of technology.

3.3. Data Collection

Cognitive load data was collected from each study participant upon completing each XR condition using the National Aeronautics and Space Administration (NASA) Task Load Index (TLX) (Hart & Staveland, 1988). The NASA TLX is a multi-dimensional rating procedure providing an overall workload score based on weighted average of ratings on 6 subscales (Table 1). Participants used a sliding scale within a Qualtrics online survey instrument to illustrate their perceived workload magnitude from 0-100/low-high for each subscale.

Ratings of factors participants deem most important in creating workload for them during their experience of a task are given more weight in computing overall workload. This two-part evaluation consists of both self-reported weights and magnitude ratings for each subscale. These weights account for two potential sources of inter-rater variability in the sources of the workload between tasks. To calculate the 15 possible pair-wise comparisons of the 6 subscales, each pair must designate the subscale of each pair provided as part of the index that contributed more to their workload of that task.

Table 1*: NASA TLX Subscale Definitions*

Subscale	Definition
Mental Demand	How much mental/perceptual (thinking, deciding, calculating, remembering, looking, searching) activity was required? Was the task easy or demanding; simple or complex; exacting or forgiving?
Physical Demand	How much physical activity (pushing, pulling, turning, controlling, activating) was required? Was the task easy or demanding; slow or brisk; slack or strenuous; restful or laborious?
Temporal Demand	How much time pressure did you feel due to the rate or pace at which the tasks occurred? Was the pace slow and leisurely; rapid and frantic?
Self-Performance	How successful do you think you were in accomplishing the goals of the task set by the experience? How satisfied were you with your performance in accomplishing these goals?
Self-Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration Level	How insecure, discouraged, irritated, stressed and/or annoyed are you versus feeling secure, gratified, content, relaxed, and complacent during the task?

(*Source: Hart & Staveland 1988)*

3.4. Data Analysis

The overall workload score is calculated by multiplying each rating by the weight given to the factor by that participant. The sum of the weighted ratings for each task is divided by the 15 possible pair-wise comparisons (the sum of the weights). The NASA TLX requires a tally sheet to be constructed to collect each time a participant indicates information on the subscales for each task to compute the weighting. Because we asked participants in this study to complete the NASA TLX immediately upon completing the entire XR condition, each subscale was given an equal weight in the overall workload calculation.

The mean average of the load ratings was taken from the XR condition group on each subscale. The weighted mean averages were compared per the NASA TLX manual and due to the low N, non-parametric statistical analyses were performed. Independent samples Mann-Whitney U testing was conducted to test the null hypothesis for subscale differences between each XR condition.

4. Results

Result of the independent samples Mann-Whitney U testing rejected the null hypotheses for temporal, effort, and performance only (table 1) as compared across each XR condition whereas temporal, effort, and performance were statistically significant ($p<.05$). A comparison for each statistically significant subscale follows describing individual Mann-Whitney U testing on each XR condition for each subscale.

Table 2: *Independent Samples Mann-Whitney U Hypothesis Test Summary*

S. No.	Null Hypothesis	Sig.^{a,b}	Decision
1	The distribution of Mental Demand is the same across XR condition.	.310^{c}	Retain the null hypothesis.
2	The distribution of Physical Demand is the same across XR condition.	.119^{c}	Retain the null hypothesis.
3	The distribution of Temporal Demand is the same across XR condition.	<.001^{c}	Reject the null hypothesis.
4	The distribution of Performance is the same across XR condition.	.015^{c}	Reject the null hypothesis.
5	The distribution of Effort is the same across XR condition.	<.001^{c}	Reject the null hypothesis.
6	The distribution of Frustration is the same XR condition.	.056^{c}	Retain the null hypothesis.

a. The significance level is .050.
b. Asymptotic significance is displayed.
c. Exact significance is displayed for this test.

(Source: Self from SPSS Output)

Drilling down into each subscale, we can see the comparison between the XR conditions by subscale. As Table 2 shows, statistical significance of the Mann-Whitney U testing was seen in the temporal, performance, and effort subscale. At an alpha level of .05, the p-value in the Mann-Whitney U test displays both the 2-tailed Asymptotic significance and the exact significance level. Table 3 indicates the summary for the Mann-Whitney U and Wilcoxon W results for each condition on temporal demand. Comparing the N=24 across both conditions, the Mann-Whitney U was 104.500 with an exact p-value of <.001 with a Wilcoxon W value of 125.50.

Table 3: *Independent-Samples Mann-Whitney U Test Summary on Temporal Demand*

Total N	24
Mann-Whitney U	104.500
Wilcoxon W	125.500
Asymptotic Sig. (2-tailed test)	<.001
Exact Sig. (2-tailed test)	.000

(Source: Self from SPSS Output)

The mean ranks for each temporal demand condition is 9.69 (iVR) and 20.92 (MR). This suggests a higher cognitive load for the MR group on temporal demand.

Table 4 shows the summary for the Mann-Whitney U and Wilcoxon W results on performance demand. Comparing the N=24 across both conditions, the Mann-Whitney U was 90.00 with an exact p-value of .015 with a Wilcoxon W value of 111.00.

Table 4: *Independent-Samples Mann-Whitney U Test Summary for Performance Demand*

Total N	24
Mann-Whitney U	90.000
Wilcoxon W	111.000
Asymptotic Sig. (2-tailed test)	.016
Exact Sig. (2-tailed test)	.015

(Source: Self from SPSS Output)

The mean ranks for each performance demand condition is 10.50 (iVR) and 18.50 (MR). This also suggest a higher cognitive load for the MR group on performance.

Table 5 shows the summary for the Mann-Whitney U and Wilcoxon W results on performance demand. Comparing the N=24 across both conditions, the Mann-Whitney U was 102.50 with an exact p-value of <.001 with a Wilcoxon W value of 123.50.

Table 5: *Independent-Samples Mann-Whitney U Test Summary*

Total N	24
Mann-Whitney U	102.500
Wilcoxon W	123.500
Asymptotic Sig. (2-tailed test)	.001
Exact Sig. (2-tailed test)	.000

(Source: Self from SPSS Output)

The mean ranks for each effort demand condition is 9.81 (iVR) and 20.58 (MR). Again, this result suggests a higher cognitive load for the MR group on effort.

5. Discussion

In this study with undergraduate students, many of whom aspire to be teachers, results are beginning to show the cognitive load impact when using integrated XR SEGs as a part of the science teaching and learning process. The researchers recognize the sample size is small and not evenly distributed but the findings are significant as a baseline for future studies. We also consider that the weight given by the users were evenly distributed across the six subscales, which has a potential impact on the NASA TLX calculations. When hypothesizing that there will be no difference in cognitive load between Mixed Reality SEGs and Virtual Reality SEGs, this study suggests that a self-reported measure, such as the NASA TLX, insinuates that Mixed Reality SEGs create a larger cognitive demand than Virtual Reality SEGs with respect to temporal, performance, and effort. This study also indicates there is no statistically significant differences in cognitive load on mental demand, physical demand, or frustration level.

These findings were a bit surprising to the researchers. We anticipated that the physical demand of walking a physical location while assimilating the nature surrounds and looking at a MR SEG on one's mobile device was surely going to show a higher cognitive workload on mental and physical demand for the user as opposed to user of the iVR SEG where the user either sat or stood in a location and simply looked around. The results of this study suggested otherwise. It was, however, exciting to see the frustration demand was not statistically significant across the XR conditions. In previous studies, we did see high frustration in users of an Augmented Reality-based SEG but results of that study suggest the design of the SEGs in each XR condition was insufficient and this producing a high frustration level.

Referring to Table 1 and the definitions the developers of the NASA TLX provided for each cognitive demand subscale, we can begin to make inferences on possibly why there was statistical significance for the MR SEG group. The Temporal demand definition was*: How much time pressure did you feel due to the rate or pace at which the tasks occurred? Was the pace slow and leisurely; rapid and frantic?* The design of each course may have played a role in the temporal demand. Students on location in the Outer Banks that activated the MR SEGs were on a time schedule. The first SEG, at Jockey's Ridge State Park was the last stop of day 1 of the 5-day experience. The class had ample time to play the MR SEG being the end of a long day could have caused some internal pressure to complete the game so they could move on to dinner and checking into their hotel. Day 2 began with stops at the subsequent three locations and playing the MR SEG at each. There could have been a sense of, "What's next?", and student felt pressure to complete the game in a timely manner so they could board the bus and move to the next stop. In contrast, the students using the iVR did not have to worry about the logistics of a field experience and had the allocated class time to complete the tasks set forth in the iVR SEG.

With respect to Performance, the NASA TLX definition stated: *How successful do you think you were in accomplishing the goals of the task set by the experience? How satisfied were you with your performance in accomplishing these goals?* Again, results showed statistical significance for the MR condition over the iVR condition. It can be inferred that students who completed the SEG and collected all the gold coins in a timely manner felt a sense of accomplishment and thus, performed well in the game/on the task. It can also be suspected that students playing the MR SEG used a familiar device to activate the games. Most everyone now has a smart phone and knows how to use it effectively and efficiently. Conversely, students playing the iVR SEGs likely engaged with the Meta Quest 2 headset for the first time. Although they were given instructions and guided through a tutorial before playing the first game, it was likely still a foreign experience for most and thus, they likely did not feel a sense of high performance in this task.

Finally, the NASA TLX defined Effort as: *How hard did you have to work (mentally and physically) to accomplish your level of performance?* The result on this subscale speaks to our original hypothesis that the MR condition group would have a higher physical and mental workload. The NASA TLX instrument does have separate subscales for mental and physical demand and does define each differently than under the Effort subscale, combining mental and physical demand into one subscale likely encouraged students to report this combination as higher for the MR condition group since they did walk and focus on their mobile device during game play. The iVR group did not have such a physical demand during these games, which is why the iVR condition group scored Effort lower than those in the MR condition group.

6. Conclusion

Infusing technology into any educational setting has varying effects on both the instructor and the learner. Technology can engage, allow students to explore environments in a different way or locations they would never get to explore in person. Technology applications can even serve as a formative assessment, but technology in and of itself is not a panacea. Technology cannot, and should not, replace good teaching.

This study indicated the need for thoughtfully designed Serious Educational Games. Juxtaposing instructional design with game design is a challenge unto itself, but considering games in XR conditions adds another layer to the complexity. Understanding the cognitive demand each SEG has on the learner is important considerations in both the instructional design and SEG design. Clearly MR tasks the user to explore their physical surroundings while focusing their attention on a smart device while iVR users do not have such a physical demand and are arguably more immersed in the game world. This supports work done by Nur'amalia, Supriatna, & Ilfiandra (2023) that higher task loads increase cognitive and mental fatigue. In doing so, learners lose motivation to learn and increased boredom, stress, and anxiety.

Another important implication for practice is how we consider XR technology embedded within place-based SSI versus remote SSI. First, based on the findings, we postulate that familiarity with the technology is an important consideration when implementing SSI instruction with XR technology. Participants in this study who used their smart phones felt more successful and satisfied with their performance than did those who used a new piece of equipment (i.e., Meta Quest). Teachers should provide opportunities for students to become familiar with the technology in a low stake setting prior to engaging in content. Second, educators embedding MR into place-based experiences must be judicious in the use of MR. The findings from this study indicate that participants using the MR felt negatively impacted by time, which may have been the result of designing too many MR experiences coupled with the perceived increased self-effort needed to move from one MR experience to another. For example, in this study, reaching some of the MR content artifacts required participants to walk some distance and up inclines. While this was done to maximize the pedagogical impact of the location, it may have been a zero-sum game in that any potential benefits of the location and/or technology were canceled out by the increased effort required to reach the artifact. These findings are in potential conflict with those reported by Wenk (2023) and Cheng and Huang (2023) where there were no cognitive load differences reported between AR or VR. There is much more research to be done into the synthesis of science instruction, SSI, XR, and SEGs and this study should provide a baseline for future work.

The research limitations of this study are few but important. First, the cognitive demand instrument used is dated and possibly did capture the true cognitive loads since they were self-reported. Because cognitive load is difficult to measure, especially in low-cost technologies, a self-report measure such as the NASA TLX used here is appropriate. However, self-reported data is often exaggerated or embellished (i.e., the Hawthorne effect). Self-reported data should be paired with other sources of data to ensure reliability.

This study will set the stage for a scope of future research. Most notably, the results of this study suggested that users of these games must first be familiar with the technology before engaging in the educational practices embedded in the games. Future studies will also include bioinformatics as data sources. Such data points as eye and hand tracking, galvanic skin response, heart rate, and fNIR brain blood flow information will support self-reported cognitive load information. We would also see a future direction working with learners with special needs (Mekacher, 2019).

REFERENCES

Abt, C. (1970). Serious Games. New York: The Viking Press.

Allcoat, D., Hatchard, T., Azmat, F., Stansfield, K., Watson, D., & Von Mü Hlenen, A. (2021). Education in the Digital Age: Learning Experience in Virtual and Mixed Realities. *Journal of Educational Computing Research, 59*(5), 795–816. https://doi.org/10.1177/0735633120985120

Annetta, L. A., (2008). Serious Educational Games: From Theory to Practice. Amsterdam, The Netherlands: Sense Publishers. pp. 83. https://doi.org/10.1163/9789087903817

Chen, C. C., & Huang, P. H. (2023). The effects of STEAM-based mobile learning on learning achievement and cognitive load. *Interactive Learning Environments*, *31*(1). https://doi.org/10.1080/10494820.2020.1761838

Chen, Y. C., Chang, Y. S., & Chuang, M. J. (2022). Virtual reality application influences cognitive load-mediated creativity components and creative performance in engineering design. *Journal of Computer Assisted Learning*, *38*(1). https://doi.org/10.1111/jcal.12588

Climate Change Education. In Shepardson, D.P., Roychoudhury, A., & Hirsch, A.S. (Eds.). (2017). Teaching and Learning about Climate Change: A Framework for Educators (1st ed.). Routledge. https://doi.org/10.4324/9781315629841

Dolphin, G., Dutchak, A., Karchewski, B., & Cooper, J. (2019). Virtual field experiences in introductory geology: Addressing a capacity problem, but finding a pedagogical one. *Journal of Geoscience Education*, *67*(2), 114–130. https://doi.org/10.1080/10899995.2018.1547034

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

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. *Cyberpsychology, Behavior, and Social Networking*, *22*(2). https://doi.org/10.1089/cyber.2018.0150

Huang, X., Huss, J., North, L., Williams, K., & Boyd-Devine, A. (2023). Cognitive and motivational benefits of a theory-based immersive virtual reality design in science learning. *Computers and Education Open*, *4*, 100124. https://doi.org/10.1016/J.CAEO.2023.100124

Ibili, E., & Billinghurst, M. (2019). Assessing the Relationship between Cognitive Load and the Usability of a Mobile Augmented Reality Tutorial System: A Study of Gender Effects. *International Journal of Assessment Tools in Education*, *6*(3). https://doi.org/10.21449/ijate.594749

Juliano, J. M., Schweighofer, N., & Liew, S. L. (2022). Increased cognitive load in immersive virtual reality during visuomotor adaptation is associated with decreased long-term retention and context transfer. *Journal of NeuroEngineering and Rehabilitation*, *19*(1). https://doi.org/10.1186/s12984-022-01084-6

Mekacher, L. (2019). Augmented Reality (AR) and Virtual Reality (VR): The Future of Interactive Vocational Education And Training For People With Handicap. *PUPIL:*

International Journal of Teaching, Education and Learning, *3*(1), 118–129. https://doi.org/10.20319/pijtel.2019.31.118129

Mermer, T. (2010). *The UNESCO climate change initiative.* UNESCO.

Newton, M. H., & Zeidler, D. L. (2020). Developing socioscientific perspective taking. *International Journal of Science Education*, *42*(8), 1302–1319. https://doi.org/10.1080/09500693.2020.1756515

Newton, M.H., Annetta, L.A., & Bressler, D. (2023). Using extended reality technology in traditional and place-based environments to study climate change. Journal of Science Education and Technology. https://doi.org/10.1007/s10956-023-10057-w

Nur'amalia, Y., Supriatna, M., & Ilfiandra, (2023). ONLINE LEARNING DIFFICULTIES AS IMPACT OF COVID-19 IN INDONESIA: Received: 09th August 2021; Revised: 29th November 2022, 18th January 2023; Accepted: 20th January 2023. *PUPIL: International Journal of Teaching, Education and Learning*, *6*(3), 48–56. https://doi.org/10.20319/pijtel.2023.63.4856

Petersen, G. B., Petkakis, G., & Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers and Education*, *179*. https://doi.org/10.1016/j.compedu.2021.104429

Sadler, T. D. (2009). Situated learning in science education: Socio-scientific issues as contexts for practice. In *Studies in Science Education* (Vol. 45). https://doi.org/10.1080/03057260802681839

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2). https://doi.org/10.1016/0364-0213(88)90023-7

Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior*, *108*. https://doi.org/10.1016/j.chb.2020.106316

Wenk, N., Penalver-Andres, J., Buetler, K. A., Nef, T., Müri, R. M., & Marchal-Crespo, L. (2023). Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment. *Virtual Reality*, *27*(1). https://doi.org/10.1007/s10055-021-00565-8

Zeidler, D. L., & Newton, M. H. (2017). Using a socioscientific issues framework for climate change education: An ecojustice approach. In Teaching and learning about climate change (pp. 56-65). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315629841-5/using-socioscientific-issues-framework-climate-change-education-dana-zeidler-mark-newton

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*. 893–902. https://doi.org/10.1109/vr46266.2020.00012