EMBRY-RIDDLE Aeronautical University
SCHOLARLY COMMONS

International Journal of Aviation, Aeronautics, and Aerospace

Volume 11 | Issue 1 Article 1

2024

Evaluating the Effectiveness of Game-Based Virtual Reality in Satellite Ground Control Operations Education and Training

Lana Laskey
Embry-Riddle Aeronautical University, laskeyl@erau.edu
Joseph R. Keebler
Embry Riddle Aeronautical University, keeblerj@erau.edu
Paul M. Cairns
Embry-Riddle Aeronautical University, cairnsp@erau.edu
Geovanny Lopez
Embry-Riddle Aeronautical University, lopezg25@my.erau.edu

Scholarly Commons Citation

Laskey, L., Keebler, J. R., Cairns, P. M., & Lopez, G. (2024). Evaluating the Effectiveness of Game-Based Virtual Reality in Satellite Ground Control Operations Education and Training. *International Journal of Aviation, Aeronautics, and Aerospace, 11*(1). DOI: https://doi.org/10.58940/2374-6793.1841

Globally, there is a growing demand for satellite amenities such as navigation, communications, weather reporting, disaster management, agricultural operations, and humanitarian assistance (United Nations Office for Outer Space Affairs, 2019). As a result, satellite launches and deployments have increased exponentially over the last decade. Likewise, human space flight and the space tourism market are also expanding (Chang, 2020). With this increased space traffic, there is a need for continued advancements in education and training of satellite ground control operations. Novel and innovative training methods may benefit this demanding field.

Background

A satellite operator's role involves managing all satellite and ground systems, including signal communications, antenna alignment, telemetry processing, command data handling, onboard computers, attitude control, onboard sensors, electrical power, and thermal control (Wertz et al., 2011). Since satellite technology is an essential element of modern global infrastructures, any failures of these systems must be quickly resolved. Therefore, the most critical responsibility of a satellite operator is early anomaly detection and rapid real-time resolution of any problems that may arise. Satellite anomalies may result from causes such as manufacturing or design imperfections, initial orbit insertion or activation mishaps, hardware or software defects, electromagnetic radiation during increased space weather, long-term degradation due to the high-plasma space environment, exposure to extreme thermal conditions, atmospheric drag, operator error, space debris, or intentional anti-satellite interference by bad actors (Galvan, 2014). Anomaly scenarios involving an operator with a keen sense for a mission-degrading event and strong system recovery skills may ultimately have little to no impact on the satellite customer. However, the inability to visualize the remote spacecraft and promptly execute anomaly resolution could lead to satellite outages or even permanent mission failure. Robust training programs are paramount for this complex discipline but may challenge instructional designers.

Problem Statement

Conventional training for satellite operators includes classroom instruction and traditional two-dimensional (2D) computer simulation replicating the real-time operational console (Sellmaier et al., 2022). While satellite operators are responsible for maintaining the health and safety of remotely orbiting spacecraft, most operators are never exposed to physical satellites for comprehension of the geometric scale and capability of the respective systems. Instead, satellite ground operators typically train, observe, and respond to anomalies based on obscure streaming telemetry data viewed on a 2D computer display. Therefore, the operation of a distant orbiting satellite can be abstract and difficult to visualize or interpret. Due to this complexity, operators may experience imbalanced cognitive workloads, causing traditional console simulation training methods to be timeconsuming and require a steep learning curve to gain proficiency (Laskey, 2022). Consequently, satellite operations present a challenge in instructional design and the development of robust training scenarios.

Purpose Statement

The current study employed an immersive form of education and training that combines two instructional methods: gamification and virtual reality (VR) simulation. Gamification, or game-based instruction, enhances user motivation and facilitates cognitive engagement (Plass et al., 2015). Similarly, VR simulation provides an immersive three-dimensional (3D) environment promoting user presence and prolonged cognitive engagement (Wang et al., 2018). Integrating these two instructional approaches produces an advanced strategy of game-based virtual reality (GBVR), where serious instructional content is merged with gameplay (Shi et al., 2022). The current study employed GBVR in a satellite ground control training scenario by adding game challenges to a 3D virtual satellite environment. The effects of GBVR were then evaluated when integrated as a supplemental activity during the training scenario. The GBVR training phase complements the 2D training scenario, heightening the comprehension and visualization of an otherwise inaccessible satellite. Accordingly, the study answers the *research question* of whether GBVR is feasible and effective regarding proper cognitive loading during a satellite training scenario.

Historically, *controller-to-satellite interaction* is unavailable to ground operators but may improve visualization and understanding of the satellite's size, scope, and physical operation. Therefore, the purpose of the study was to observe the cognitive effects of operating a satellite at the ground console in the traditional manner, followed by a more immersive method involving physical interaction with a full-scale virtual satellite. The GBVR method is intended to supplement traditional 2D training, offering practice in the GBVR environment, which may assist the operator during eventual real-time operations outside of the GBVR environment. Instructional difficulties of satellite operations may benefit from this kind of *controller-to-satellite interaction*, offering a viable alternative for satellite ground training.

Prior research examined the *feasibility* of GBVR when deployed in a satellite training scenario within a university classroom and laboratory setting. The prior study measured system usability, workload suitability, and user experience with satisfactory results (Laskey & Keebler, 2023). The current study repeats the prior feasibility measurements of system usability and user satisfaction, along with a new examination of the *effectiveness* of GBVR on student learning by incorporating a cognitive load study. Student learning was operationalized with a previously validated survey instrument measuring cognitive load to determine any significant differences in student learning with and without GBVR.

The current study also measures and accounts for any simulator sickness experienced by the participants. Simulator sickness is a well-known phenomenon where users experience psychological or physiological disturbances during simulated motion (Frank et al., 1983). Motion sickness due to simulated motion is much less severe than sickness related to actual movement. However, it may still impact user experience during a simulated scenario with symptoms such as eyestrain, headache, disorientation, or nausea (Kennedy et al., 1993). Therefore, simulator sickness was observed during this research study.

Significance

Since satellite and spacecraft mission control operators are in high demand, robust training programs for this complex discipline must continue to emerge and advance. The current research analyzes GBVR, an enhanced form of instruction and training, when applied to satellite ground operations. The components of GBVR, gamification and immersive VR, are each well supported in the current research literature. The alternative form of instruction presented here, combining gamification and VR in satellite operator training, may further the body of knowledge and contribute to current literature. The results of this study may benefit stakeholders, including aerospace industry personnel, educators, and researchers.

Limitations

The study comprised two main limitations. The first limitation involved a smaller sample size, which may limit generalizability to a larger population. The smaller sample size is attributed to the specialized nature of participants with unique skill sets in satellite operations. Follow-on studies may support larger sample sizes to enhance external validity. The second limitation involves the use of Likert scale data with parametric statistics. Although Likert scales generally provide ordinal data, the interval-like usage of the Likert scales is a commonly accepted practice in the extant literature (Braly et al., 2019; Kim et al., 2018; Shelsted et al., 2019). Therefore, with this limitation acknowledged, the study proceeded under the assumption of equal intervals within the Likert scales.

Literature Review

The focus of this study investigates the advanced instructional design model of GBVR, combining gamification and immersive VR. Both instructional methods are well supported in the existing literature. Therefore, unifying the two techniques may reveal an even more comprehensive approach for complex training scenarios. First, one of the main theories supporting gamification is the *goal-setting theory*, where motivation is influenced by game mechanics requiring goal completion (Huang & Hew, 2018). Second, one of the most prominent theories supporting immersive VR is the *cognitive theory of multimedia learning*, where including multiple forms of media can improve cognitive processing (Mayer, 2017). Lastly, according to the *cognitive load theory*, proper balancing of instructional components can expand the learner's ability to process information and achieve learning objectives (Leppink et al., 2013). Therefore, the *cognitive load scale*, developed from this theory, was used to measure participant learning during GBVR training scenarios.

Gamification & Goal-Setting Theory

Historically, games consisted of goals and challenges designed for entertainment purposes. However, in recent years, games have been applied to learning environments meant for a more serious purpose, also called *serious games* (Abt, 1970). Gamification includes game mechanics in a non-game setting for educational purposes (Detering et al., 2011). Game elements should comprise achievable goals and challenges, instant feedback, and attractive visual and audio aesthetics for multimedia applications (Cheng et al., 2015). When appropriately developed, these game mechanics will likely contribute to continued player motivation and engagement (Krath et al., 2021).

Goal-setting theory (GST) states that the mere presence of goals will likely affect effort, and higher goals typically lead to even higher effort and performance (Locke & Latham, 2002). This heightened effort can be especially beneficial if practiced within complex disciplines since completing challenging tasks requires increased activity and engagement. Goal mechanisms can also direct attention toward goal-relevant activities, encourage persistence and prolonged effort, and inspire strategic problem-solving to meet goal challenges (Huang & Hu, 2018). Like game mechanics, effective goal mechanics must be designed appropriately to the individual's skill level with attainable tasks and immediate feedback. Goals that are either extremely easy or excessively hard can counter the intent of GST by losing the individual's attention. Furthermore, immediate feedback allows individuals to gauge their performance, competence, or necessary improvement (Locke & Latham, 2002). Game mechanics developed with the attributes surrounding GST will likely benefit game-based instructional models.

Virtual Reality & Cognitive Theory of Multimedia Learning

VR provides an immersive environment resembling real-world settings, allowing visual observation, problem-solving, or hands-on experimentation not otherwise available (Krath et al., 2021). For example, in a study by Arents et al. (2021), medical students were first able to observe childbirth through Caesarean Sections before their first experience inside an operating room. This alternative instructional method offered a safe and repeatable process for surgery observation, where no patient was impacted. Virtual immersion facilitates prolonged engagement through experiential and dynamic learning compared to conventional instructional approaches, such as teacher and whiteboard or 2D computer applications (Christopoulos et al., 2020). Active learning is instrumental in complex training scenarios where intricate details could be less evident in a passive instructional environment.

The cognitive theory of multimedia learning (CTML) conveys the advantage of multiple forms of media in the learning environment. For example, a classroom presentation with text and graphics may improve student learning rather than text alone (Moreno & Mayer, 1999). Delivery of the classroom presentation could be further developed with the addition of audio or video. According to CTML, cognitive learning is based on human memory. With the appropriate stimulus, short-term working memory can be converted into long-term stored memory, where learning occurs (Mayer, 2017). Like CTML, VR offers an environment with multiple media stimuli, including on-screen text, audible strings, visual cues, action-reaction feedback, and several degrees of movement freedom (Holly et al., 2021). The appropriate combination of these media formats within the VR environment could facilitate improved cognitive outcomes.

Cognitive Load Theory

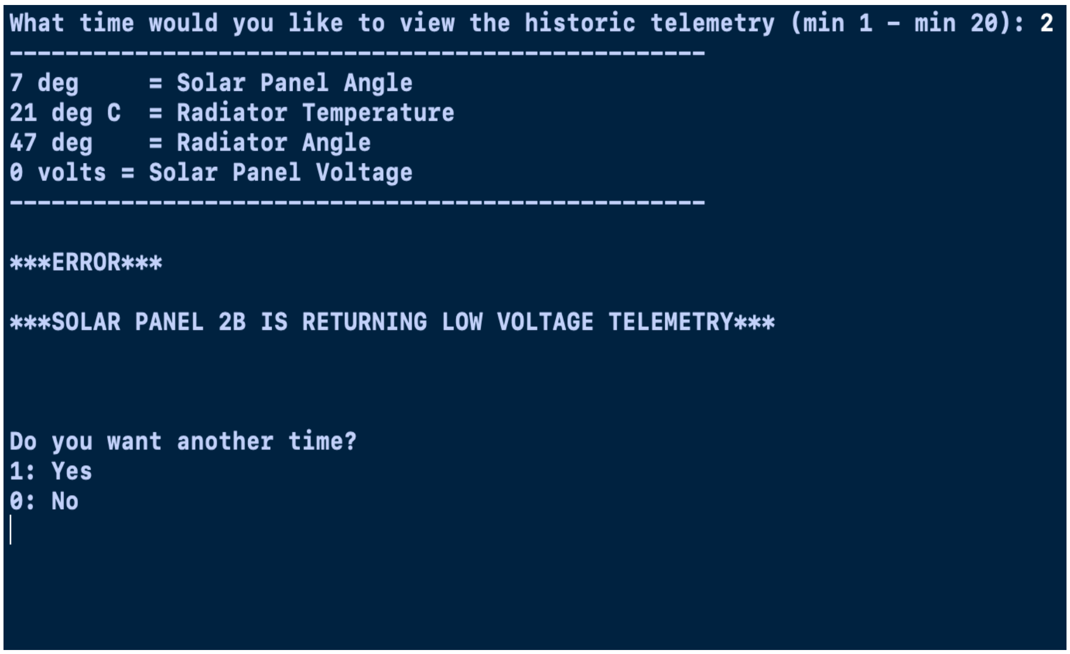
According to cognitive load theory (CLT), learners can experience three types of cognitive loading, including intrinsic (IL), extraneous (EL), and germane loading (GL) (Leppink et al., 2013). First, intrinsic loading represents the complexity of instructions provided relative to the learner's skill level. Instructional material that is either too easy or too complex may relinquish the mental engagement of the learner. Second, extraneous loading represents any instructional components that deter from the ultimate learning objectives. These non-essential instructional elements should be minimized to the extent possible as they may distract the learner from the beneficial cognitive process. Lastly, germane loading represents the elements most valuable to learning and should be the focus of an instructional designer. According to Sweller (2010), instructional design can be significantly enhanced by properly balancing these three types of cognitive loading. As a result, the *Cognitive Load Scale* (CLS) was developed by Leppink et al. (2013).

Method

A quantitative experimental design was used to evaluate the effectiveness of GBVR when incorporated into a satellite training scenario. The study took place within a university laboratory setting. Participants comprised N = 28 college-level students enrolled in the senior capstone course for space operations. The average age of participants is 23.4 years (SD = 3.2). The participants were divided into two groups of n = 14 students. Both groups were exposed to the initial 15-minute traditional 2D simulation of a satellite ground control training scenario. The initial scenario involved resolving an electrical anomaly, shown as operated from the traditional 2D ground control displays (see Figures 1 and 3). Then, only one group was exposed to a second treatment involving a 15-minute GBVR simulation of the same training scenario, as seen from the site of the anomaly aboard the virtual satellite (see Figures 2 and 3). Instead of a replacement for traditional training, this method demonstrates GBVR as a supplement to traditional training scenarios.

Figure 1

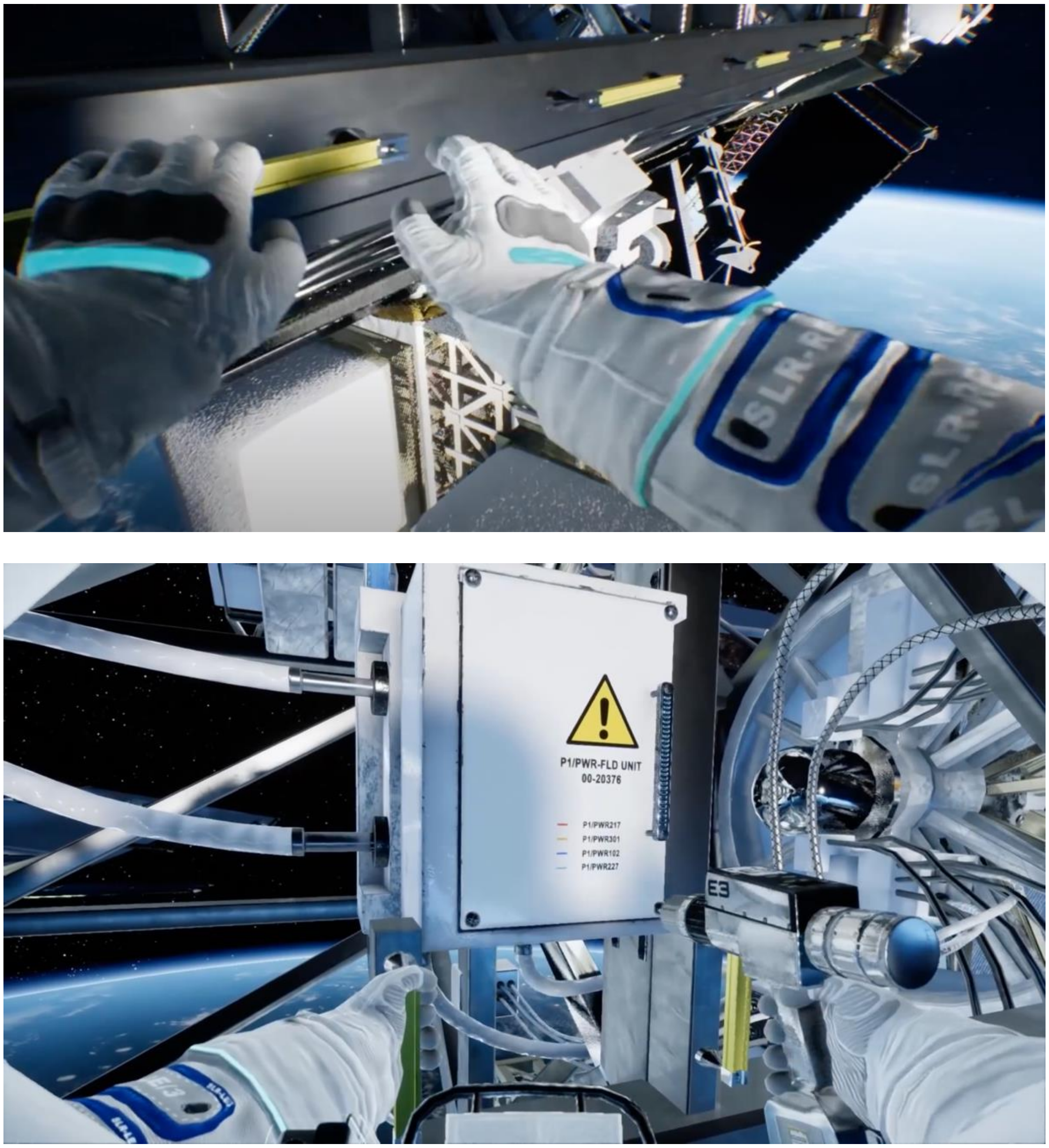
Conventional Two-Dimensional Ground Control Display Simulation – Electrical Anomaly



Note. Screenshot of anomalous telemetry showing a low voltage status.

Figure 2

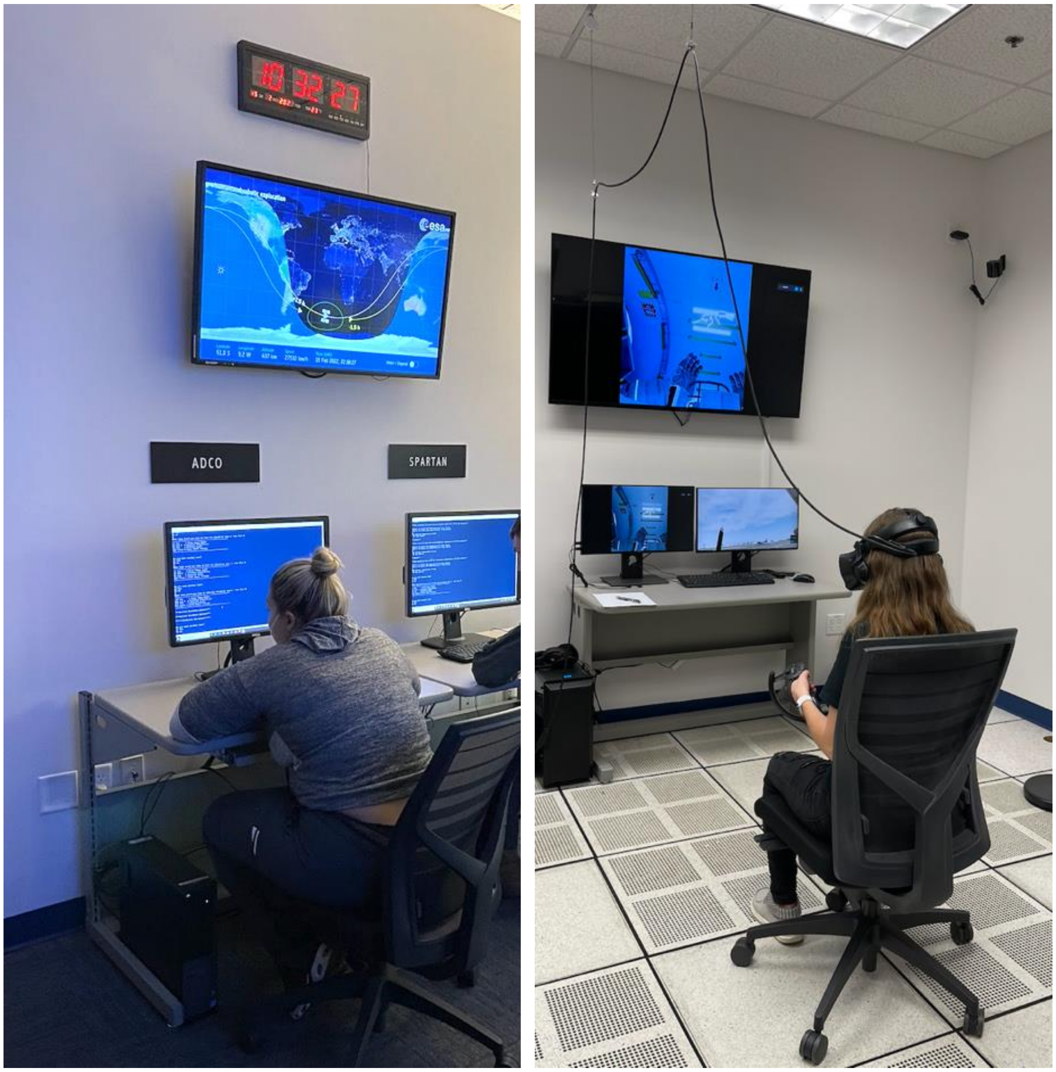
GBVR Simulation of Translation to Work Site and Equipment Repair – Electrical Anomaly



Note. Photos depict (top) virtual translation (navigation) to the work site via handrails and (bottom) interactive repair of electrical anomaly. From "Earthlight: Spacewalk - Release Trailer" by Opaque Space, 2017 (https://www.youtube.com/watch?v=juiY8rcAtEQ). Copyright 2017 by Opaque Space LLC.

Figure 3

Traditional Ground Control Simulation Setting (left) vs. GBVR Simulation Setting (right)



Note. Photos taken by the author during the experiment.

The study involved the use of commercially available *EarthLight* software simulating virtual interaction with the International Space Station (ISS) (Opaque Media Group, 2016). The ISS spacecraft simulated an orbiting satellite with an exterior electrical anomaly. GBVR participants were equipped with *Valve Index VR Kits* comprising a head-mounted display and two hand controllers (Valve Corporation, n.d.). Using this equipment, participants of the GBVR group translated along the exterior structure to the physical work site and performed equipment repair based on audio instructions heard in the headset. Lastly, all training scenarios were accomplished from a seated position, which is known to help avoid symptoms of simulator sickness (Hu et al., 2021). After the simulation scenarios, each group completed surveys regarding their respective experiences. The surveys consisted of multiple validated scales addressing system usability, user satisfaction, cognitive loading, and simulator sickness. An explanation of each scale is provided in the following sections.

System Usability

The *System Usability Scale* (SUS) was used in this study to measure the usability of the multimedia applications. The SUS measures user perception of system complexity, ease of use, functionality, and user confidence regarding hardware devices and software applications (Usability.gov, 2013). The SUS survey is comprised of ten statements rated on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree), with final combined scores rated on a scale of 0 (negative) to 100 (positive) (Bangor et al., 2009). According to research, a value of $M = 68$ is the accepted average score (Usability.gov, 2013). The SUS survey validates the proper setup of the laboratory equipment, operational inputs, and user interface. The required knowledge must match the assigned users' skill level for successful operations.

User Experience

The *Game User Experience Satisfaction Scale* (GUESS) was used to measure the participant experience in this study. The GUESS is a 55-question survey designed to measure user enjoyment and satisfaction during gameplay (Phan et al., 2016). The GUESS-18 used in this study is an abbreviated 18-question version of the larger 55-question scale (Keebler et al., 2020). User responses are ranked on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree), measuring nine constructs: usability, narratives, play engrossment, enjoyment, creative freedom, audio aesthetics, personal gratification, social connectivity, and visual aesthetics. Final composite scores are ranked from 0 (worst) to 100 (best). Shelsted et al. (2019) measured six popular video games, resulting in an average score of $M = 78.7$. While this score is used as a standard in this study, it must be noted that video games are created for fun and entertainment purposes, which sets a high benchmark for comparison of educational games designed for a more serious objective (Krath et al., 2021). Likewise, wording within the survey questions was changed from “play/playing” to “operate/operating” and “game” to “sim.”

Cognitive Load

The CLS was used to measure the cognitive loading of participants in this study. The CLS measures the relationship between the three types of cognitive loading experienced by the learner during instruction and training (Leppink et al., 2013). The 10-item survey measuring IL, EL, and GL ranks user responses on a scale from 0 (not the case at all) to 10 (completely the case). Survey items 1, 2, and 3 measure IL with questions regarding the learner's perception of instructional complexity. In keeping with CLT, medium to low IL scores would provide the best outcome, indicating an appropriate loading of complex information for the learner. Second, survey items 4, 5, and 6 measure EL with questions concerning the ineffectiveness of the instructions and explanations. Low scores for EL would indicate that students perceived material delivery as clear and compelling. Lastly, survey items 7, 8, 9, and 10 question whether learners perceive the material as knowledge-enhancing. Achieving the highest scores possible in the GL category should be the goal of the instructional designer. Final composite scores for each category are ranked on a scale from 0 (low) to 10 (high), and proper balancing of all three cognitive load types should provide the best opportunity for enhanced student learning (Leppink et al., 2014).

Simulator Sickness

Finally, the *Simulator Sickness Questionnaire* (SSQ) was also used in the study to examine whether overall results were affected by symptoms of simulator sickness. The SSQ is a 16-item survey initially developed in 1993 to measure symptoms caused by simulated motion. Symptoms were organized into three categories: nausea, oculomotor disturbance (eye strain and headache), and disorientation (Kennedy et al., 1993). With the modern addition of VR to simulation, a new Virtual Reality Sickness Questionnaire (VRSQ) was adapted from the original SSQ in 2018 to address VR applications. During research trials for the VRSQ, investigators found that nausea symptoms were rarely reported (Kim et al., 2018). Therefore, VRSQ researchers adopted nine items from the original SSQ for eye strain and disorientation but excluded the seven items covering nausea symptoms. The 16-item SSQ and the 9-item VRSQ were utilized during this research experiment. The GBVR participants will complete the VRSQ, and the non-GBVR participants will complete the SSQ. Both instruments rank each user symptom on a 4-point Likert scale (0 = none, 1 = slight, 2 = moderate, 3 = severe). Final scores for each symptom are rated on a scale from 0 (no symptoms) to 100 (severest symptoms).

Furthermore, according to Jaeger and Mourant (2001), post-test symptom severity is affected by the duration of exposure to a simulated scenario. Based on an approximate average of 15-minutes of exposure for each trial in the current study, the accepted maximum symptom severity score is $M = 18.5$ (Jaeger & Mourant, 2001). Although the non-GBVR participants did not experience simulated motion and the SSQ scores were expected to be very low, the data was still collected for consistency across groups.

Results

For comparison of participant results to the standard benchmarks, one-sample *t*-tests were performed for the SUS and GUESS-18 responses (see Table 1, Figures 4 and 5). The SUS score for GBVR participants ($M = 83.4$) was significantly higher than the average standard score ($M = 68.0$), where $t(13) = 4.693$, $p < .001$. A large effect size of $d = 1.25$ was revealed, demonstrating that the participants found the GBVR simulation generally easy to use. Likewise, the SUS score for non-GBVR participants ($M = 93.2$) was significantly higher than the average standard score ($M = 68.0$), where $t(13) = 16.397$, $p < .001$. A large effect size of $d = 4.38$ was indicated, illustrating that the participants also found the non-GBVR simulation easy to use.

The GUESS-18 score for GBVR participants ($M = 86.3$) was significantly higher than the average popular video game score ($M = 78.7$), where $t(13) = 3.277$, $p = .003$. A large effect size of $d = .88$ resulted, signifying a high level of user satisfaction and enjoyment for the simulation with GBVR applications. Conversely, the GUESS-18 score for non-GBVR participants ($M = 79.9$) did not significantly differ from the average popular video game score ($M = 78.7$).

The mean comparisons of the CLS, VRSQ, and SSQ results to the benchmark standards can be found in Table 1 and Figures 6 and 7. The mean CLS scores for the GBVR group ($M_{IL} = 2.7$, $M_{EL} = 1.7$, and $M_{GL} = 8.8$) placed within the approximate accepted ranges ($2 < M_{IL} < 5$, $0 < M_{EL} < 2$, and $5 < M_{GL} < 10$). Likewise, the mean CLS scores for the non-GBVR group ($M_{IL} = 2.8$, $M_{EL} = 1.5$, and $M_{GL} = 8.5$) also placed within the approximate accepted ranges ($2 < M_{IL} < 5$, $0 < M_{EL} < 2$, and $5 < M_{GL} < 10$). The VRSQ oculomotor and disorientation mean scores for the GBVR group (M_{Ocu} = 12.8, M_{Dis} = 9.5) placed within the accepted range (0 < M < 18.5). Lastly, the SSQ oculomotor, disorientation, and nausea scores for the non-GBVR were very low (M_{Ocu} = 0.7, M_{Dis} = 1.0, M_{Nau} = 2.0), also placing within the accepted range (0 < M < 18.5).

Table 1

Results vs. Standard Benchmarks

Variable	*n*	Min	Max	*SD*	*M*	Standard (*M*)
SUS $_{GBVR}$	14	55.0	100.0	12.3	83.4	68.0
SUS $_{no\ GBVR}$	14	80.0	100.0	5.8	93.2	68.0
GUESS-18 $_{GBVR}$	14	70.6	98.4	8.7	86.3	78.7
GUESS-18 $_{no\ GBVR}$	14	55.6	94.4	12.7	79.9	78.7
CLS – IL $_{GBVR}$	14	1.0	6.0	1.5	2.7	Approx. 2-5
CLS – IL $_{no\ GBVR}$	14	1.0	6.0	1.8	2.8	Approx. 2-5
CLS – EL $_{GBVR}$	14	1.0	4.7	1.0	1.7	Approx. 0-2
CLS – EL $_{no\ GBVR}$	14	1.0	4.0	0.8	1.5	Approx. 0-2
CLS – GL $_{GBVR}$	14	6.5	10.0	1.1	8.8	Approx. 5-10
CLS – GL $_{no\ GBVR}$	14	1.0	10.0	2.4	8.5	Approx. 5-10
VRSQ – Oculomotor $_{GBVR}$	14	0	66.7	16.4	12.8	0-18.5
SSQ – Oculomotor $_{no\ GBVR}$	14	0	8.0	2.0	0.7	0-18.5
VRSQ – Disorientation $_{GBVR}$	14	0	40	12.2	9.5	0-18.5
SSQ – Disorientation $_{no\ GBVR}$	14	0	13	4.0	1.0	0-18.5
SSQ Only – Nausea $_{no\ GBVR}$	14	0	29	7.7	2.0	0-18.5

Note. The participant total was N = 28, where the first group (n = 14) was exposed only to the non-GBVR training scenario, and the second group (n = 14) was exposed to both types of training, including the non-GBVR scenario followed by the supplemental GBVR scenario.

Figure 4

System Usability Scale (SUS)

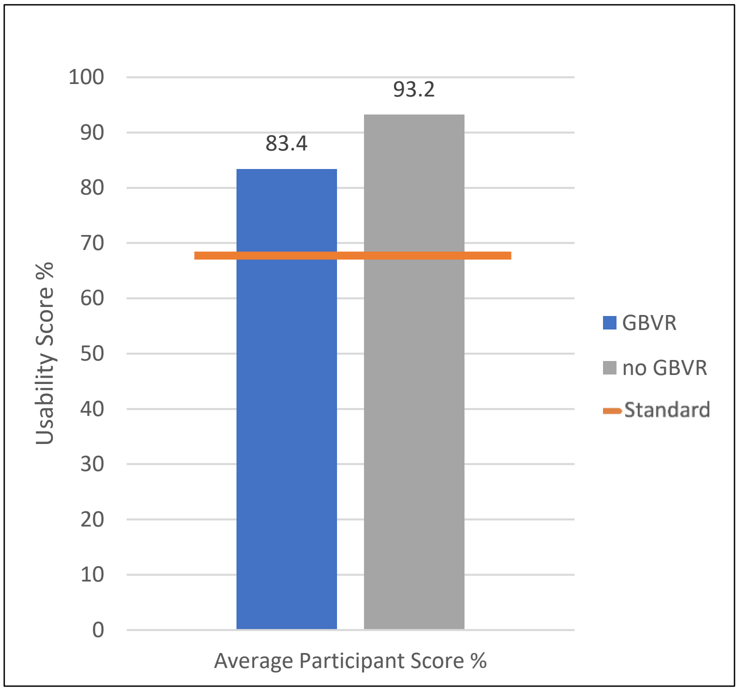
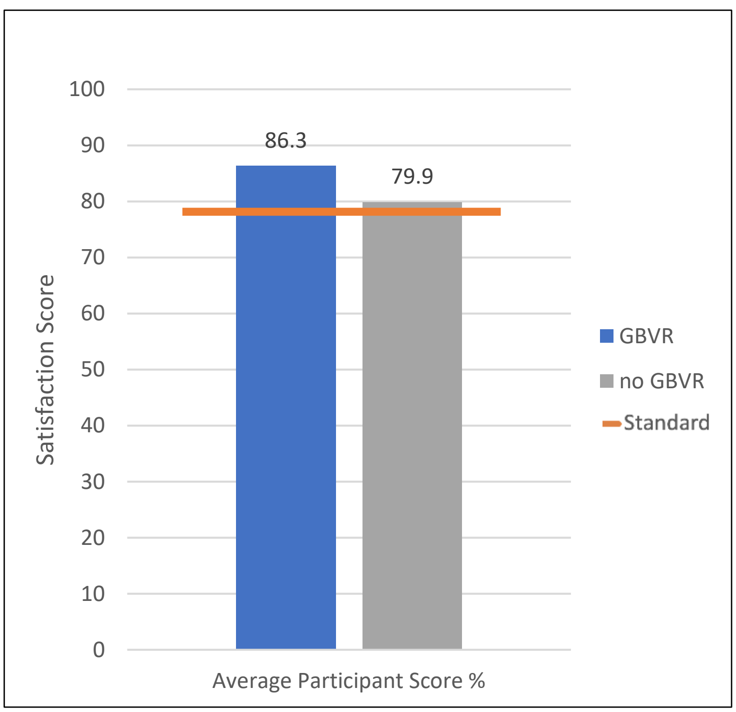


Figure 5

Game User Experience Satisfaction Scale (GUESS-18)



DOI: 10.58940/2374-6793.1841

Figure 6

Cognitive Load Scale (CLS)

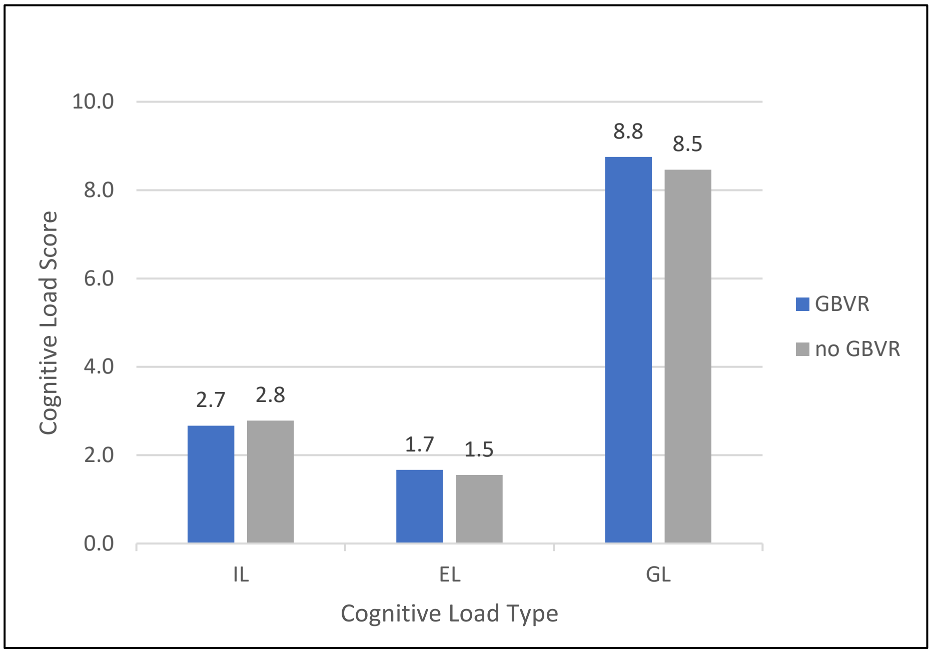
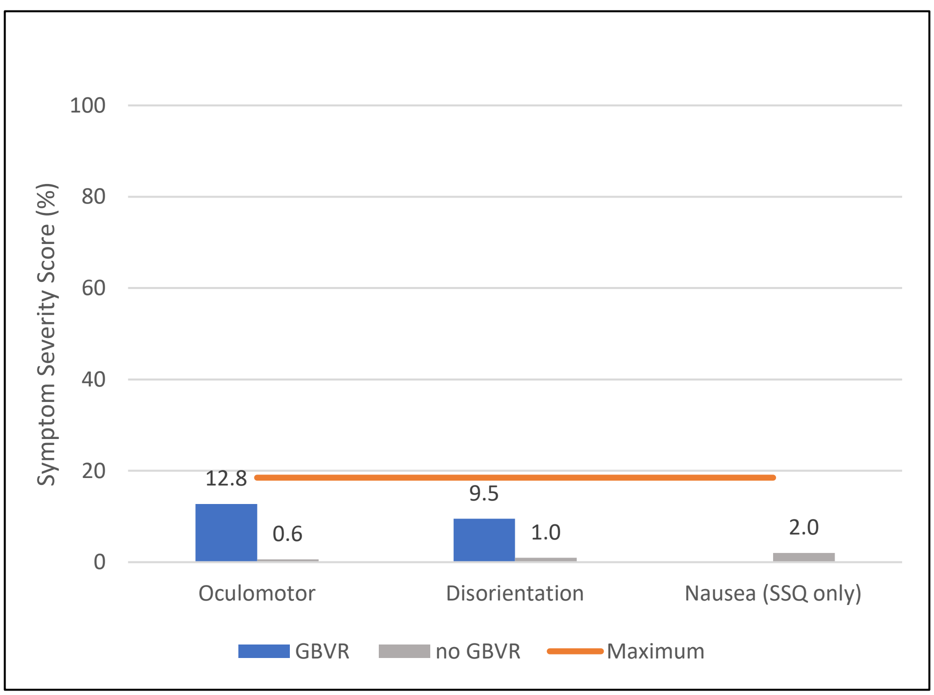


Figure 7

Virtual Reality Sickness Questionnaire (VRSQ) & Simulator Sickness Questionnaire (SSQ)



Discussion

According to the data analysis, the training scenario for satellite ground operations involving GBVR demonstrated satisfactory system usability, user satisfaction, cognitive loading, and simulator sickness compared to the benchmark standards. However, each element requires further explanation compared to the non-GBVR training scenario. The GBVR SUS scores were significantly higher than the benchmark, denoting adequate levels of complexity, ease of use, user confidence, and functionality (see Figure 4). The non-GBVR scores ranked even higher than the GBVR usability scores. The lower score for the GBVR scenario, compared to non-GBVR, is likely due to the inherent intricacy of VR equipment and the necessary learning curve for navigating the VR environment.

Despite users ranking the GBVR training scenario as more complicated than the non-GBVR session, this can be viewed as a positive outcome, especially since GBVR was significantly higher than the benchmark. As stated in the goal-setting theory, a reasonable level of complexity is necessary to challenge and inspire the learner for significant learning to occur. In other words, achieving a more challenging goal can be more rewarding than a less demanding goal, leading to intrinsic motivation and prolonged engagement.

In terms of user experience, according to the results of the GUESS-18 instrument, users found the GBVR training scenario enjoyable and satisfying (see Figure 5). Unlike the non-GBVR scenario, the GBVR session ranked significantly higher than the popular game score, created to compare video game entertainment value. The higher score for the GBVR scenario is likely attributed to the appropriate GBVR laboratory setup and proper inclusion of game mechanics and goal setting. Additionally, even though the non-GBVR score was not significantly higher than the benchmark, the results indicate that users still found the non-GBVR training scenario to be a generally positive experience.

Employing the CLS to measure the cognitive loading of GBVR participants, the scores revealed that both types of training lead to proper cognitive loading (see Figure 6). First, intrinsic loading, a measure of complexity, showed appropriate results for GBVR and non-GBVR, implying that the skill levels of the learners were adequately matched. Secondly, extrinsic loading, considered counter-productive to learning, introduced only a small amount of irrelevant information in either case. Lastly, both scenarios produced very high germane loading necessary for meaningful learning. Therefore, both the GBVR and non-GBVR training scenarios exhibited similarly effective student learning.

Furthermore, mean comparisons of simulator sickness were examined for both simulation scenarios. The GBVR participants were evaluated based on the VRSQ for virtual simulations, and the non-GBVR participants were evaluated based on the SSQ used for non-VR applications. Mean comparisons for simulator sickness ranked below the maximum accepted threshold for both simulation scenarios. Furthermore, as expected, the non-GBVR participants ranked very low since the traditional console training scenario did not involve simulated motion. Therefore, symptoms due to simulator sickness likely had no impact on the overall outcome of the study.

Conclusion

The purpose of this study was to evaluate the *effectiveness* of GBVR on student learning when integrated into the complex scenario of satellite ground control operations education and training. First, *feasibility* tests showed satisfactory outcomes in all categories, including system usability, user experience, and simulator sickness. Participants ranked the system as generally easy to use with high user enjoyment and satisfaction scores, and symptoms of simulator sickness were minimal. Therefore, incorporating GBVR into the satellite training scenario is likely feasible and practical. Next, the effectiveness of GBVR on student learning was evaluated. During the GBVR training scenario, participants revealed appropriate cognitive loading necessary for meaningful learning, demonstrating the effectiveness of GBVR within this complex discipline.

Furthermore, when investigating the efficacy of the training scenario without GBVR, the results also showed proper cognitive loading of participants. While results indicated a satisfactory application of GBVR, the current study did not demonstrate GBVR as more effective than without GBVR regarding cognitive loading and student learning. However, the additional feasibility study illustrated a more positive user experience for the GBVR participants than those without GBVR, compared to the benchmark standard. A positive and satisfying learning experience can lead to enhanced motivation and increased cognitive engagement, foundations for meaningful learning. Therefore, the results of the trials establish GBVR as a viable and effective training option for satellite ground control operations and may benefit other complex educational or training environments.

A limitation of this study is the small sample size ($N = 28$). Therefore, a recommendation would be to repeat the study with a larger sample size to increase generalizability over the target population. While the sample was small, most results showed statistical significance and large effect sizes within the participant groups. Although the target population represents a small community of spacecraft operators, increasing the sample size may improve external validity and enhance findings. Future research recommendations might also include the evaluation of GBVR as a stand-alone training method rather than a supplement to traditional training. Lastly, a further recommendation would include additional goals of progressing difficulty within the GBVR scenario. The accomplishment of increasingly more challenging goals has shown benefits to increased effort, prolonged engagement, and enhanced learning. However, while this may demonstrate a learning benefit, the instructional designer must maintain an appropriate level of complexity and cognitive loading suitable for the learner.

References

Abt, C. (1970). *Serious games*. The Viking Press.

Arents, V., de Groot, P. C. M., Struben, V. M. D., & van Stralen, K. J. (2021). Use of 360° virtual reality video in medical obstetrical education: A quasi-experimental design. *BMC Medical Education, 21*(1), 202–202. https://doi.org/10.1186/s12909-021-02628-5

Bangor, A., Kortum, P., T., & Miller., J., T. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies, 4*(3), 114-123.

Braly, A. M., Nuernberger, B., & Kim, S. Y. (2019). Augmented reality improves procedural work on an international space station science instrument. *Human Factors, 61*(6), 866–878. https://doi.org/10.1177/0018720818824464

Chang, E. Y.-W. (2020). From aviation tourism to suborbital space tourism: A study on passenger screening and business opportunities. *Acta Astronautica, 177*, 410–420. https://doi.org/10.1016/j.actaastro.2020.07.020

Cheng, M.-T., Chen, J.-H., Chu, S.-J., & Chen, S.-Y. (2015). The use of serious games in science education: a review of selected empirical research from 2002 to 2013. *Journal of Computers in Education (the Official Journal of the Global Chinese Society for Computers in Education), 2*(3), 353–375. https://doi.org/10.1007/s40692-015-0039-9

Christopoulos, A., Pellas, N., & Laakso, M.-J. (2020). A learning analytics theoretical framework for STEM education virtual reality applications. *Education Sciences*, *10*(317), 317–. https://doi.org/10.3390/educsci10110317

Detering, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September 28). From game design elements to gamefulness: Defining "gamification.” *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040

Frank, L. H., Kennedy, R. S., Kellogg, R. S., & McCauley, M. E. (1983). Simulator sickness: A reaction to a transformed perceptual world. 1. Scope of the Problem. *Naval Training Equipment Center*, Orlando, FL.

Galvan, D. A., Hemenway, B., Welser, W., & Baiocchi, D. (2014). *Satellite anomalies : Benefits of a centralized anomaly database and methods for securely sharing information among satellite operators*. RR-560-DARPA. RAND Corporation.

Holly, M., Pirker, J., Resch, S., Brettschuh, S., & Gutl, C. (2021). Designing VR experiences – Expectations for teaching and learning in VR. *Educational Technology & Society*, *24*(2), 107–119.

Huang, B., & Hew, K. F. (2018). Implementing a theory-driven gamification model in higher education flipped courses: Effects on out-of-class activity completion and quality of artifacts. *Computers and Education*, *125*, 254–272. https://doi.org/10.1016/j.compedu.2018.06.018

Jaeger, B. K., & Mourant, R. R. (2001). Comparison of simulator sickness using static and dynamic walking simulators. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *45*(27), 1896–1900. https://doi.org/10.1177/154193120104502709

Keebler, J. R., Shelstad, W. J., Smith, D. C., Chaparro, B. S., & Phan, M. H. (2020). Validation of the GUESS-18: A short version of the game user experience satisfaction scale (GUESS). *Journal of Usability Studies*, *16*(1), 49 – 62.

Kennedy, R. S., Lane, N. E., Berbaum K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International*

Journal of Aviation Psychology, 3(3), 203-220, https://doi.org/10.1207/s15327108ijap0303_3
Kim, K., H., Park, J., Choi, Y., & Choe, M. (2018). Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics*, *69*, 66–73. https://doi.org/10.1016/j.apergo.2017.12.016
Krath, J., Schürmann, L., & von Korflesch, H. F. O. (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior, 125*. https://doi.org/10.1016/j.chb.2021.106963
Hu, Y., Elwardy, M., & Zepernick, H.-J. (2021). On the effect of standing and seated viewing of 360° videos on subjective quality assessment: A pilot study. *Computers (Basel)*, *10*(6), 80. https://doi.org/10.3390/computers10060080
Laskey, L. (2022, April). Space games: Evaluating game-based virtual reality in higher education. In *4th Symposium on Space Educational Activities*. Universitat Politècnica de Catalunya. 10.5821/conference-9788419184405.057
Laskey, L., & Keebler, J. R. (2023). Evaluating system usability, workload suitability, and user experience of game-based virtual reality in spaceflight education and training. *Journal of Aviation/Aerospace Education & Research, 32*(1). https://doi.org/10.15394/jaaer.2023.1948
Leppink, J., Paas, F., Vleuten, C. P. M. v. d., Gog, T. A. J. M. v., & Merrienboer, J. J. V. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058-1072. https://doi.org/10.3758/s13428-013-0334-1
Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *The American Psychologist*, *57*(9), 705-717. https://doi.org/10.1037//0003-066X.57.9.705
Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning*, *33*(5), 403-423. https://doi.org/10.1111/jcal.12197
Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*(2), 358–368. https://doi.org/10.1037/0022-0663.91.2.358
Opaque Media Group, (2016). *Earthlight*. https://www.opaque.media
Opaque Space. (2017, November 22). *Earthlight: Spacewalk - release trailer*. https://www.youtube.com/watch?v=juiY8rcAtEQ
Phan, M. H., Keebler, J. R., & Chaparro, B. S. (2016). The development and validation of the game user experience satisfaction scale (GUESS). *Human Factors*, *58*(8), 1217-1247. https://doi.org/10.1177/0018720816669646
Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, *50*(4), 258-283. https://doi.org/10.1080/00461520.2015.1122533
Sellmaier, F., Uhlig, T., & Schmidhuber, M. (Eds). (2022). *Spacecraft operations* (2nd ed.). Springer.
Shelstad, W. J., Chaparro, B. S., & Keebler, J. R. (2019). Assessing the user experience of video games: Relationships between three scales. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 1488-1492. https://doi.org/10.1177/1071181319631300

Shi, A., Wang, Y., & Ding, N. (2022). The effect of game-based immersive virtual reality learning environment on learning outcomes: Designing an intrinsic integrated educational game for pre-class learning. *Interactive Learning Environments*, *30*(4), 721–734. https://doi.org/10.1080/10494820.2019.1681467

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*(2), 123-138. https://doi.org/10.1007/s10648-010-9128-5

United Nations Office for Outer Space Affairs. (2019, October). *Benefits of space for humankind*. UNOOSA. https://www.unoosa.org/oosa/en/benefits-of-space/benefits.html

Usability.gov. (2013). *System usability scale* (SUS). https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

Valve Corporation. (n.d.). *Valve index® - Upgrade your experience - Valve Corporation*. https://www.valvesoftware.com/en/index

Wang, P., Wu, P., Wang, J., Chi, H., & Wang, X. (2018). A critical review of the use of virtual reality in construction engineering education and training. *International Journal of Environmental Research and Public Health*, *15*(6). https://doi.org/10.3390/ijerph15061204

Wertz, J. R., Everett, D. F., & Puschell, J. J. (2011). *Space mission engineering : The new SMAD*. Microcosm Press.