Journal of Aviation/Aerospace Education & Research

Volume 32 | Number 1 Article 2

2023

Evaluating System Usability, Workload Suitability, and User Experience of Game-Based Virtual Reality in Spaceflight Education and Training

Lana Laskey , M.S.
Embry-Riddle Aeronautical University, laskeyl@erau.edu

Joseph R. Keebler , Ph.D.
Embry Riddle Aeronautical University, keeblerj@erau.edu

Introduction

Satellites provide many essential and value-added amenities for end-users on Earth. Examples include satellite-based navigation systems, internet service, and severe weather monitoring. As a result, global dependence on space-based amenities has grown exponentially in the last decade, thus significantly increasing the number of satellites in Earth's orbit and the need for more skilled operators (Ryan-Mosley et al., 2019). This upward trend is expected to continue, emphasizing advanced proficiencies and situational awareness of satellite controllers operating in this crowded environment. Therefore, satellite operator training must also evolve to meet this increasing demand for space-based technology.

Spaceflight education and training is a complex discipline where it is customary to operate a remote and inaccessible object. The space operations environment is one where small mistakes can lead to costly and catastrophic mission failures. In the case of satellite ground control operators, training can be time-consuming and likely requires a steep learning curve to grasp complicated concepts. This learning curve may be further limited due to the inaccessibility of the satellite for physical inspection by ground operators. More advanced instructional methods might bridge this disconnect between the satellite operator and the physical spacecraft equipment.

The enhanced instructional model of game-based virtual reality (GBVR) was investigated in this study when applied to a complex discipline, such as spaceflight education and training. Game-based instruction, or *gamification*, is defined as the application of "game design elements in non-game contexts" (Deterding et al., 2011, p. 9). This form of instruction has been known to improve learners' cognitive engagement and situational awareness (Plass et al., 2015). Likewise, immersive virtual reality (VR) simulations have been found to improve skill development within a safe, repeatable, and low-cost environment where, for example, no actual spacecraft equipment would be necessary (Papanikolaou et al., 2019). In this environment, satellite operators could access full-scale satellite models for more advanced interactive training.

Based on the rise in demand for skilled operators, this research study examined the feasibility of the innovative GBVR training technique. The known benefits of GBVR within complex disciplines were analyzed when applied to a complicated satellite ground control training scenario. The feasibility of GBVR was assessed through measurements of system usability, workload suitability, and user experience. Understanding the effects of GBVR within this challenging discipline may serve to evolve and refine training techniques for the space operations industry.

Theoretical Foundation

Gamification

A *game* is defined as "structured play with rules, goals, and challenges for the purpose of entertainment" (Krath et al., 2021, p. 2). However, in educational design, games are employed for a more serious or constructive purpose, and therefore, the term *serious games* is now accepted. This term is synonymous with gamification, or the application of game mechanics in game-free applications and has continued to gain relevance since 2008 (Detering et al., 2011).

The most widely accepted theories supporting research in gamification include (1) self-determination theory and (2) flow theory (Krath et al., 2021). Self-determination theory states that gameplay can satisfy three basic psychological needs of the player, including competence, independence, and socialization (Ab Jalil et al., 2020). Achievement of these basic needs promotes emotional enjoyment and satisfaction when experienced during gameplay. According to Greipl et al. (2021), whether positive or negative, human emotions are directly related to human cognition. Positive emotion typically inspires pleasure and motivation for continued cognitive engagement. Although negative emotion might cause frustration, this is a valuable feature in game design, as frustration can encourage overcoming challenges. In either scenario, heightened emotional experiences facilitate a higher degree of memory and information retention than less emotional events (Greipl et al., 2021).

Like self-determination theory, flow theory is also related to the level of enjoyment and satisfaction experienced by the user (Csikszentmihalyi & Asakawa, 2016). More specifically, *flow theory* is defined as the immersion of the learner in meaningful learning objectives, leading to peak performance through a methodical flow of content (Zainuddin et al., 2020). The flow of activities and material should be characterized by full engagement of the learner with well-defined objectives, instant feedback, progressively advancing challenges, and a robust reward system (Huang & Hew, 2018). Flow theory is a well-supported model specifically in educational game design due to the focus on learners' motivation, effort, and achievement of learning outcomes (Ab Jalil et al., 2020).

Game-Based Virtual Reality

According to Wang et al. (2018), VR can be described as an immersive application with various magnitudes of *virtual* and *real* components that, when combined, produce a three-dimensional *visualization* experience with multiple degrees of motion and freedom. VR is considered an enhanced approach to activity visualization compared to a traditional two-dimensional method with static pictures. In addition to suspension in virtual space, the game-based element is the necessary completion of challenges and learning objectives (Shi et al., 2022). This environment stimulates cognitive performance, solidifies information retention, and prolongs player engagement and satisfaction.

Current Study

GBVR research has developed within several educational and training environments. Examples of GBVR studies include solar energy education, medical training, language instruction, electrical theory application, textile instruction, classroom mathematics, and emergency evacuation scenarios (AlQallaf et al., 2022; Butt et al., 2018; Chen & Hsu, 2020; Frieß et al., 2021; Lau et al., 2017; Shi et al., 2022; Snopková et al., 2022). Each study demonstrated positive educational advantages due to the incorporation of GBVR. However, there remains a gap in the literature regarding GBVR simulation and training for space mission ground control operations. Therefore, the current study examined the use of GBVR within a satellite ground control training scenario. The known benefits of game-based learning and virtual immersion may aid in advancing training techniques for ground control operators.

College-level student participants were the focus of the current study, as they represent potential entry-level trainees for space mission ground control operations. The study took place in a university classroom and laboratory setting. The study was quantitative and based on the perceptions of the student participants after experiencing the GBVR scenarios. Participants were surveyed using multiple validated scales. Results were analyzed for consistency across scales, and three main attributes were evaluated, including perceived system usability, workload suitability, and user experience.

Method

Participants

Participants included 10 university students. The age of participants ranged from 21 to 30 years (M = 23.4, SD = 2.7). At the time of the study, all participants were within six months of graduating with Bachelor of Science degrees in Spaceflight Operations. All students were enrolled in the senior capstone course covering space mission control operations at the university where the study was held. The university institutional review board (IRB) was contacted for research approval with human subjects. However, the study was considered exempt by the IRB office since the research trials were a part of regularly scheduled classroom activities. Therefore, no IRB approval or informed consent was required.

Materials

Materials used in this study included computers, VR equipment, and survey instruments. The study was conducted in a classroom laboratory in two parts. The first part included simulation with a computer, keyboard, mouse, and two-dimensional monitor display, while the second part involved simulation using VR head-mounted displays and hand control equipment. After completing all activities, participants were surveyed using multiple validated scales, including the System Usability Scale (SUS), the National Aeronautics and Space Administration (NASA) Task Load Index (TLX), and the Game User Experience Satisfaction Scale (GUESS-18).

Procedure

Participants underwent three 75-minute lectures over the two weeks leading up to the day of simulation activity. The lectures covered the nuances of satellite systems necessary for mission operations. Participants were briefed on the day of the simulation activity and then completed the two-part simulation sequence. Part A simulates the typical experience of a ground control operator observing anomalous telemetry readouts. Conversely, Part B provides physical interaction with virtual spacecraft hardware not typically available to ground operators. Part A was conducted in the mission control lab while seated at a computer monitor observing digital readouts for 10-15 minutes (see Figure 1). Part B was conducted in the VR lab in a seated position, using hand controls, wearing a head-mounted device, and operating the GBVR simulation for 10-15 minutes (see Figure 2). During the GBVR experience, participants were provided a 2-minute tutorial on the use of the controls. Participants were then instructed to search for the damaged hardware and physically inspect the spacecraft component that produced anomalous telemetry in the previous scenario (observed in Part A). After completing the simulation activities, participants were provided surveys concerning the GBVR portion of their experience.

It should be noted that all lecture and simulation activities were a normal part of the capstone course and were required to complete the course objectives. Conversely, the surveys conducted after completing the simulation activities were not required, and the voluntary nature of the surveys was explained to all participants. All participants provided consent by completing the online surveys, as described in the form directions. Weblinks for the surveys were provided to the participants and administered via Google forms. All surveys were conducted anonymously, and no personally identifiable data was collected for any participant.

Figure 1

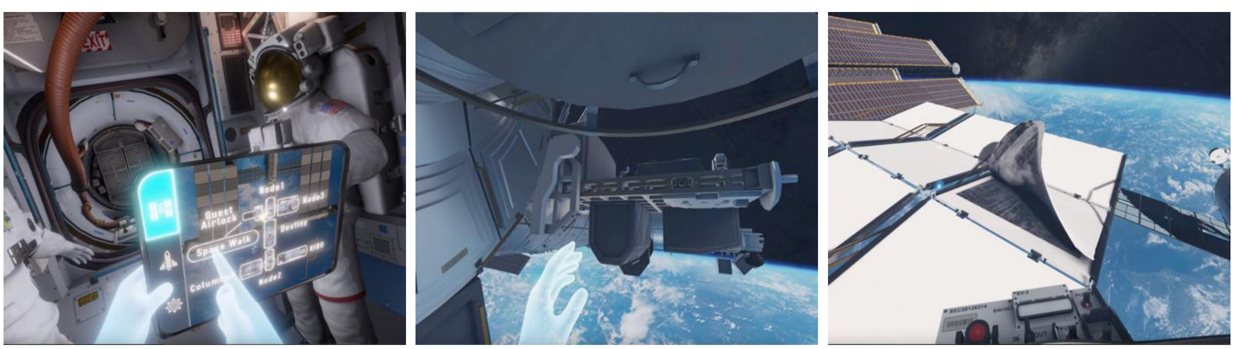
Sim Part A: Ground Control Computer Console Simulation – Telemetry Inspection of Anomaly

```
MAIN MENU:
Press 1 to Gather Telemetry
Press 0 to EXIT the Program
1
What time would you like to view the telemetry (min 1 - min 20): 4
--------------------------------------------------
9 deg     = Solar Panel Angle
97 deg C  = Radiator Temperature
48 deg    = Radiator Angle
154 volts = Solar Panel Voltage
--------------------------------------------------

***ERROR***

***The Radiators are above the acceptable temperature of 90 deg celsius***

Press 1 to reduce the temperature by 5 deg celsius
Press 2 to reduce the temperature by 10 deg celsius
```

Note. Anomalous telemetry shows that radiator temperatures are above acceptable ranges.

Figure 2

Sim Part B: Game-Based Virtual Reality (GBVR) Simulation – Physical Inspection of Anomaly



(a) *(b)* *(c)*

Note. Photos depict (a) decision menu selection, (b) handrail (virtual) navigation to the site, and (c) physical (virtual) inspection of anomaly (damaged radiator panel). From "*Mission ISS*" by Magnopus, 2017 (https://www.magnopus.com/mission-iss). Copyright 2017 by Oculus VR LLC.

Measures

Survey instruments were constructed based on three validated scales exploring participant perception of the GBVR system usability, workload suitability, and user experience.

Perceived Usability

Developed in 1986, the System Usability Scale (SUS) is a subjective evaluation tool for assessing the usability of hardware, software, or devices (Usability.gov, 2013). A total of 10 questions formulate the SUS survey. Each question is ranked on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree) regarding system complexity, ease of use, functionality, and user confidence. Final combined usability scores were rated on a scale from 0 (lowest) to 100 (highest). A SUS score of 68.0 is the published average used as the standard for this study (Usability.gov, 2013).

Perceived Workload

The National Aeronautics and Space Administration (NASA) Task Load Index (TLX) was developed at NASA Ames Research Center in the 1980s and provides a subjective measurement of operator workload regarding human-machine interface systems. The TLX questions are ranked on a 20-point scale ranging from 0 (very low) to 20 (very high) and comprised of six subscales, including mental demand, physical demand, temporal demand, performance, effort, and frustration (NASA, 2020). Final composite scores rank from 0 (lowest) to 100 (highest). The *work* that is *loaded* onto a user must be appropriate and reasonable, and therefore, a workload perceived as excessive by an operator might lead to reduced performance (Hart & Staveland, 1988). Based on Hertzum's (2021) analysis of 556 TLX tests performed over 30 years (1990-2019), TLX workload scores were evenly distributed around a mean of 42.0. Therefore, this value (M = 42.0) was the accepted workload score used during this study.

https://commons.erau.edu/jaaer/vol32/iss1/2
DOI: https://doi.org/10.15394/jaaer.2023.1948

User Experience

The Game User Experience Satisfaction Scale (GUESS) was created in 2016 by Phan et al. as a 55-question survey and later re-validated as an 18-question survey (Keebler et al., 2020). The tool was designed using nine constructs to assess user experience and satisfaction during gameplay, including usability/playability, narratives, play engrossment (engagement), enjoyment, creative freedom, audio aesthetics, personal gratification, social connectivity, and visual aesthetics (Keebler et al., 2020). Each question in the GUESS-18 survey is rated on a 7-point Likert scale (1 = Strongly Disagree, 7 = Strongly Agree). Scores were obtained by summing the average ratings within each of the nine subscales and dividing by the maximum score of 63. Final composite scores were tabulated as percentages on a scale of 0 to 100% (0 = lowest, 100 = highest), allowing simple comparisons between all scales employed in this study. In a study by Shelstad et al. (2019), six popular video games were examined with the GUESS-24 scale, resulting in an average score of 49.6 (raw score). This converts to 78.7% when divided by the maximum raw score of 63. This score (M = 78.7) was used as a comparable standard for game user experience scores for this study.

For this educational study concerning the *operation of a simulator*, the verbiage within the GUESS-18 survey consisting of “play/playing” and “game” was changed to “operate/operating” and “sim,” respectively.

Results

One-sample t-test results were used to compare participant scores to the accepted standards of the three measurement scales shown in Table 1. Regarding the SUS data, the results denoted significantly higher scores for the simulation group (M = 81.8) than the benchmark score (M = 68.0), $t(9) = 5.76$, $p < .001$. A large effect size of $d = 1.82$ resulted, signifying that participants perceived the simulation as easy to use. The GUESS-18 survey score (M = 82.1) placed 3.4% higher than the average GUESS-24 score of six popular games (M = 78.7) (Shelstad et al., 2019). Although this suggests a high level of enjoyment, this mean was not significantly different from the benchmark. Finally, the mean NASA TLX survey score (M = 40.5) placed within 1.5% of the accepted mean score (M = 42.0). In this case, scoring as close as possible to the accepted mean value is ideal, revealing no significant difference, indicative of an acceptable workload level.

Table 1

Validated Scale Results and Accepted Benchmarks

Scale	*N*	Min	Max	Mean	*SD*	Accepted Benchmarks
SUS	10	67.5	90.0	81.8	7.6	68.0 = average score
GUESS-18	10	73.0	98.4	82.1	8.3	78.7 = popular game score [a]
NASA TLX	10	22.0	60.0	40.5	12.3	42.0 = mean score

Note. [a] An average GUESS-24 score of M = 78.7 resulted when examining six popular video games (Shelstad et al., 2019).

Evidence of convergent validity emerged when comparing the system usability scores of the SUS and GUESS-18 scales. In particular, the SUS average usability score (M = 81.8) and the GUESS-18 usability subscale score (M = 79.3) only differed by 2.5%, demonstrating an overlap between scales. This parallel provides additional legitimacy that both scales captured accurate perceptions of system usability.

DOI: https://doi.org/10.15394/jaaer.2023.1948

Discussion

When integrated into spaceflight education and training, GBVR simulation exhibited satisfactory system usability, workload suitability, and user experience. SUS scores were statistically significant and ranked well above average within all subscales, including complexity, ease of use, user confidence, and functionality. Successful equipment selection and effective laboratory setup may have contributed to the above-average usability scores.

Although the GUESS-18 user experience scores were not statistically significant, the benchmark comparison was based on the scoring of popular video games designed for entertainment purposes. Additionally, the video games selected included an element of in-game social connectivity with other players. Using the same benchmark to evaluate a serious game meant for educational purposes and without social connectivity posed an exceptionally high evaluation standard. Despite this high standard, the GBVR training scenario still attained an overall average score remarkably similar to the popular video game score. The resulting high score in user experience translates to the users' interest in their overall performance and motivation to succeed during the challenge, both necessary elements for effective learning.

While the SUS and GUESS-18 scores are viewed positively when scoring a high value, the NASA TLX workload scores are not considered successful if ranked too high or too low. Therefore, the best workload scores should be distributed as close as possible to the accepted value for ideal user performance. The study revealed an average workload score similar to the benchmark, demonstrating that the GBVR training scenario provided a workload level suitable for optimal user performance.

The small sample size (n = 10) may have posed a limitation for the study by restricting generalizability over a larger population. However, the study still met the purpose of answering the question of GBVR feasibility when applied to a satellite ground control training scenario. Based on the results collected with a small sample, the study may benefit from being repeated with a larger sample size for more generalizability. Lastly, consideration of a fourth scale on simulator sickness, known as the Simulator Sickness Questionnaire (SSQ) by Kennedy et al. (1993), should be included in any future trial as one participant reported feelings of sickness.

Conclusion

The findings indicate that GBVR is a feasible tool for applications involving complex disciplines, such as spaceflight education and training. First, when employed in a satellite ground control training scenario, the use of GBVR ranked well above average in *system usability*, indicating ease of system use. Secondly, the GBVR training experience ranked similarly to the *user experience* score of popular video games designed for entertainment, suggesting user satisfaction, enjoyment, and motivation to succeed. Lastly, the system *workload* scored within close range of the preferred mean TLX value, demonstrating an appropriate workload for prolonged engagement. Both self-determination theory and flow theory support the idea that when user enjoyment and satisfaction are heightened, motivation, prolonged cognitive engagement, and skill retention will likely follow.

https://commons.erau.edu/jaaer/vol32/iss1/2
DOI: https://doi.org/10.15394/jaaer.2023.1948

References

Ab Jalil, H., Nasharuddin, N., Marlisah, E., Nazan, A., Ismail, I., Ma'rof, A., 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*(21), pp. 40-57. http://dx.doi.org/10.3991/ijet.v15i21.17345

AlQallaf, N., Chen, X., Ge, Y., Khan, A., Zoha, A., Hussain, S., & Ghannam, R. (2022, March). Teaching solar energy systems design using game-based virtual reality. In *2022 IEEE Global Engineering Education Conference (EDUCON)*, 956-960. IEEE. https://doi.org/10.1109/EDUCON52537.2022.9766460

Butt, A. L., Kardong-Edgren, S., & Ellertson, A. (2018). Using game-based virtual reality with haptics for skill acquisition. *Clinical Simulation in Nursing*, *16*, 25–32. https://doi.org/10.1016/j.ecns.2017.09.010

Chen, Y.-L., & Hsu, C.-C. (2020). Self-regulated mobile game-based English learning in a virtual reality environment. *Computers and Education*, *154*. https://doi.org/10.1016/j.compedu.2020.103910

Csikszentmihalyi, M., & Asakawa, K. (2016). Universal and cultural dimensions of optimal experiences: Flow, culture, and human evolution. *Japanese Psychological Research*, *58*(1), 4-13. https://doi.org/10.1111/jpr.12104

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September 28). From game design elements to gamefulness: Defining "gamification". *MindTrek '11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040

Frieß, R., Voigt, T., Gnadlinger, F., Holtmann, C., & Steinicke, M. (2021). *Design-based research on a cooperative educational VR game about Ohm's law*. Academic Conferences International Limited.

Greipl, S., Klein, E., Lindstedt, A., Kiili, K., Moeller, K., Karnath, H. O., Bahnmueller, J., Bloechle, J., & Ninaus, M. (2021). When the brain comes into play: Neurofunctional correlates of emotions and reward in game-based learning. *Computers in Human Behavior*, *125*. https://doi.org/10.1016/j.chb.2021.106946

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 139-183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hertzum, M. (2021). Reference values and subscale patterns for the task load index (TLX): A meta-analytic review. *Ergonomics*, *64*(7), 869-878. https://doi.org/10.1080/00140139.2021.1876927

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 & Education*, *125*, 254-272. https://doi.org/10.1016/j.compedu.2018.06.018

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: A new method for quantifying simulator sickness. *International Journal of Aviation Psychology*, *3*, 203-220. https://doi.org/10.1207/s15327108ijap0303_3

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

Lau, K. W., Kan, C. W., & Lee, P. Y. (2017). Doing textiles experiments in game-based virtual reality: A design of the stereoscopic chemical laboratory (SCL) for textiles education. *The International Journal of Information and Learning Technology*, *34*(3), 242–258. https://doi.org/10.1108/IJILT-05-2016-0016

Magnopus. (2017). *Mission ISS*. https://www.magnopus.com/mission-iss

National Aeronautics and Space Administration. (2020, December 15). *NASA TLX: Task load index*. Human Systems Integration Division @ NASA Ames. https://humansystems.arc.nasa.gov/groups/tlx/

Papanikolaou, I. G., Haidopoulos, D., Paschopoulos, M., Chatzipapas, I., Loutradis, D., & Vlahos, N. F. (2019). Changing the way we train surgeons in the 21th [sic] century: A narrative comparative review focused on box trainers and virtual reality simulators. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *235*, 13-18. https://doi.org/10.1016/j.ejogrb.2019.01.016

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

Ryan-Mosley, T., Winick, E., & Kakaes, K. (2019). The number of satellites orbiting Earth could quintuple in the next decade: The coming explosion of constellations. *MIT Technology Review*. https://www.technologyreview.com/2019/06/26/755/satellite-constellations-orbiting-earth-quintuple/

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

Snopková, D., Ugwitz, P., Stachoň, Z., Hladík, J., Juřík, V., Kvarda, O., & Kubíček, P. (2022). Retracing evacuation strategy: A virtual reality game-based investigation into the influence of building's spatial configuration in an emergency. *Spatial Cognition and Computation: An Interdisciplinary Journal*, *22*(1-2), 30–50. https://doi.org/10.1080/13875868.2021.1913497

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

Wang, P., Wu, P., Wang, J., Chi, H.-L., & 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

Zainuddin, Z., Chu, S. K. W., Shujahat, M., & Perera, C. J. (2020). The impact of gamification on learning and instruction: A systematic review of empirical evidence. *Educational Research Review, 30*. https://doi.org/10.1016/j.edurev.2020.100326