Supplementary material

Table S1: Primary Education

REF	Field	Platform / Game	Method	Instruments	Evaluation
Barab et al. (2005)	Science, Mathematics, Social Studies	Quest Atlantis	Qualitative	Design-based research, Design Ethnography	Engage users with lessons from educational research on learning and motivation through experiential learning, inquiry-based learning, and portfolio assessment
Battal & Tokel (2020)	Computer Science	OpenSim	Mixed	Multiple case study research, semi-structured Interviews, two Questionnaires [1]	Factors affecting satisfaction across different educational programs
Fokides & Chachlaki (2020)	Environmental Education	OpenSim	Quantitative	Quasi-experimental design, Pre-test, two post-tests, delayed post-test, revised Toad Attitude Questionnaire [2], New Ecological Paradigm scale [3], Learning Experience Questionnaire [4]	Examine whether MUVEs have a more positive impact on students' knowledge, attitudes toward seals and eco problems, and views regarding the use of MUVEs for learning
Jakoš & Verber (2016)	Computer Science	OpenSim	Quantitative	Pre-post-tests (isomorphic) validated with Elliott Taw's methodology [5]. Debriefing sessions with Questions from reflection stems [6]	Impact of student outcomes (achievement of minimal standards of knowledge) Impact of predetermined factors, such as gender, assessment marks in science subjects and the amount of entertainment games play time on participants' outcomes
Kamarainen et al. (2015)	Environmental Education	EcoMUVE	Qualitative	Video and audio Recordings, Student work	Student engagement in modeling practices that reflect authentic practices in the domain of ecosystems science
Kim & Ke (2017)	Mathematics	OpenSim	Quantitative	Pre-post-tests	Mathematical performance
Lim (2006)	Science	Quest Atlantis	Qualitative	Observation, Interview, Student work, Reflections based on the taxonomy of [7]	Engagement

Merchant (2010)	Language (Literacy)	Active Worlds	Qualitative	Observation, Interview, Field notes, Meeting minutes and documents, Questionnaire	Implications for curriculum and pedagogy
Mystakidis & Berki (2018)	Language (Literacy)	Second Life	Mixed	Questionnaire, Interviews	Examination of UW, impact on cognitive, affective and psychomotor aspects of learning
Tüzün (2007)	Culture; Health, ICT	Quest Atlantis	Qualitative	Design-based research, Observation, Interview, Field notes, Student work, Reflections	Examine core issues and challenges when video games are used in the classroom
Wang et al. (2018)	Communication	iSocial	Quantitative	Observations coded according to the ESP framework [8]	Interaction patterns, verbal and nonverbal, for learning social skills that facilitate the development of sense of embodied social presence (ESP)
Yeh & Lan (2018)	Second Language learning (English)	OpenSim	Mixed	Quasi-experimental design, Questionnaire based on Adediwura [9] and Ding and Stapleton [10], Observation, Interview, Student work	Autonomy

Table S2: Secondary Education

REF	Field	Platform / Game	Method	Instruments	Evaluation
Barab et al. (2012)	Language	Plague: Modern Prometheus	Mixed	Pre-post-tests, Interview and Observation	Learning gain and engagement
Dede et al. (2017)	Humanities	EcoXPT	Qualitative	Interview	Engagement
Grotzer et al. (2013)	Science	EcoMUVE	Mixed	Pre-post-tests, and Interview with teachers	Learning with event-based explanations as they did explanations focused on patterns
Jacobson et al. (2015)	Science	EcoMUVE	Mixed	Pre-post-tests, and Interview using adapted from validated and reliable items from Burns, Okey, and Wise [11]	Learning computational scientific inquiry (CSI) as an innovative model for learning important scientific knowledge and new practices for "doing" science. To assess learning outcomes using multiple-choice questions about science inquiry (items Q1–Q12) adapted from validated and reliable items from Burns, Okey, and Wise (1985)

Ketelhut (2006)	Science	River City	Quantitative	Self-efficacy using the SETS instrument [12] and Control Predictors [13]	How differences in self-efficacy can affect students' participation in scientific inquiry
Loula et al. (2014)	Science	Caatinga biome	Qualitative	Interview and observations	Students' understanding of thermal regulation in ectothermic animals
Metcalf et al. (2018)	Science	EcoMUVE	Quantitative	Post-test of students learning outcomes (concept mapping)	The use of EcoMUVE to support the construction of concept maps representing causal relationships in a dynamic system
Moon et al. (2020)	Science	OpenSim	Quantitative	Classifiers Attention-switching Multimodal representation Pattern development Pattern contextualization	To maintain learners' cognitive and affective engagement in a highly interactive digital learning environment
Nelson et al. (2007)	Science	River City	Quantitative	Test score gains of control based on students' learning outcomes	Investigation on (a) how students who viewed greater numbers of guidance messages performed on content tests compared to those who viewed less, and (b) whether boys and girls accessed the guidance to the same degree, and whether there was an interaction between gender and guidance use on content scores
Pellas (2014)	Computer Science	OpenSim+Scratch4SL	Quantitative	Students' engagement [14]	Positive or negative impact on student engagement (cognitive, social emotional)
Pellas & Persoutseas (2016)	Computer Science	Second Life+Scratch4SL	Quantitative	Students' engagement Fredricks et al. [15] and Appleton et al. [16] and user experience	The impact of Second Life+Scratch4SL on students' engagement and experience
Pellas (2017)	Computer Science	OpenSim+Scratch4SL	Quantitative	Correlations of Community of Inquiry (CoI) presence indicators (social, cognitive, teaching) Swan et al. [17]	The contribution of CoI model presence indicators as theoretical principles for analyzing students' interactions in a 3D game-like environment becomes to this notion as crucial parameters for and instructional process
Pellas & Vosinakis (2018)	Computer Science	OpenSim+Scratch4SL	Mixed	A closed-ended pre-questionnaire was adopted by Lahtinen et al. [18]	The effect of simulation games developed in Scratch and OpenSim on students' learning performance by assessing computational problem-solving strategies

				A closed-ended (pre-and-post-questionnaire) Computational Thinking Scales (CTS) questionnaire based on Korkmaz et al. [19] A think-aloud protocol Interview	
Rico et al. (2011)	Computer Science	OpenSim	Quantitative	Correlation between knowledge degree of VWs and programming skills for a typical group of high school students	Students' educational experience and programming skills
Şimşek (2017)	Maths	Second Life	Quantitative	Students' pre-experiment and post-experiment scores, taken from Mathematics Attitude Scale (Ministry of turkey Education)	Second Life environment has an important effect on students' attitude towards mathematics
Twining	Science	Second Life	Qualitative	Observations and interviews	The Schome Park Programme (SPP), which was established with the specific aim of extending our thinking about schome inside Second Life to provide opportunities offered by 3D virtual worlds
Young et al. (2012)	STEM	Second Life	Quantitative	Pre-post-tests	To assess the educational gains made by the students when they can play Second Life-supported games for STEAM subjects
Zheng et al. (2009)	English Language learning	Quest Atlantis	Qualitative	Negotiation for Action Codes using CMDA and ethnographies [20,21]	To explain how interaction within avatar-embodied collaboration contexts between native English speakers and non-native English speakers can be provided as resources for English language acquisition.

References

1. Chou, S.; Liu, C. Learning Effectiveness in Web-Based Technology-Mediated Virtual Learning Environment. Proceedings of the 38th Annual Hawaii International Conference on System Sciences. 6 January 2005. Big Island, HI, US. https://doi.org/10.1109/HICSS.2005.385
2. Tomažič, I. Reported experiences enhance favourable attitudes toward toads. *Eurasia J. of Math. Sci. & Tech. Ed*. **2011**, *7(4)*, 253–262.

3. Dunlap, R.; Van Liere, D.; Mertig, G.; Jones, R. New trends in measuring environmental attitudes: measuring endorsement of the new ecological paradigm: a revised NEP scale. *J. Soc. Iss*. **2000**, *56 (3)*, 425–442.
4. Fokides, E.; Kaimara, P.; Deliyiannis, I.; Atsikpasi, P. Development of a scale for measuring the learning experience in serious games. Preliminary results. In Proc. of the digital culture & audiovisual challenges international conference. Corfu, Greece: Ionian University. 10−11 May 2019. pp. 1-2
5. Elliott Tew, A. Assessing fundamental introductory computing concept knowledge in a language independent manner. Ph.D. Thesis. School of Interactive Computing Georgia Institute of Technology, Georgia, US, December 2010
6. Discover Sensors. Discover sensors. Retrieved from http://www.discoversensors.ie/junior_certificate/%20reflection_stems1, 2016
7. Bangert-Drowns, R.; Pyke. C. A Taxonomy of Student Engagement with Educational Software: An Exploration of Literate Thinking with Electronic Text. *J. Educ. Comput. Res*., **2016**, *24 (3)*, 213–234.
8. Mennecke, B.; Triplett, J.; Hassall, M.; Conde, J.; Heer, R. An examination of a theory of embodied social presence in virtual worlds. *Dec. Scienc*. **2016**, *42*, 413–450.
9. Adediwura, A. Effect of peer and self-assessment on male and female students' self-efficacy and self-autonomy in the learning of mathematics. *Gen. Behav*. **2012**, *10(1)*, 4492–4508.
10. Ding, F.; Stapleton, P. Walking like a toddler: Students' autonomy development in English during cross-border transitions. *Syst*. **2016**, *59*, 12–28.
11. Burns, J.; Okey, J.; Wise, K. Development of an integrated process skill test: TIPS II. Journal of Research in Science Teaching, Wiley, New Jersey, US. **1985**, *22*, 169–177.
12. Ketelhut, D. Assessing science self-efficacy in a virtual environment: A measurement pilot. Harvard University, Cambridge, MA, December **2004**.
13. Lawless, A.; Kulikowich, M. Understanding hypertext navigation through cluster analysis. *J. Educ. Comput. Res*. **1996**, *14(4)*, 385–399.
14. Kong, Q.; Wong, N.; Lam, C. Student engagement in mathematics: Development of instrument and validation of a construct. *Math. Educ. Res. J*. **2003**, *54*, 4–21.
15. Fredricks, J. A.; Blumenfeld, P. C.; Paris, A. H. School engagement: Potential of the concept, state of the evidence. Review of Educational Research. **2004**, *74(1)*, 59–109.
16. Appleton, J.; Christenson, S.; Kim, D.; Reschly, A. L. Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *J. Sch. Psychol*., Elsevier, Amsterdam, Netherlands: **2006**, *44(5)*, 427–445.
17. Swan, K.; Richardson, J.; Ice, P.; Garrison, R.; Cleveland-Innes, M.; Arbaugh, B. Validating a measurement tool of presence in online communities of inquiry. E-Ment. **2008**, *2 (24)*, 1–12.
18. Lahtinen, E.; Ala-Mutka, K.; Järvinen, H. A study of the difficulties of novice programmers. In: Proceedings of the 10th Annual SIGCSE Conference on innovation and Technology in Computer Science Education, Caparica, Portugal, 27-29 June 2005.
19. Korkmaz, Ö. Çakir, R.; Özden, M. A validity and reliability study of the Computational Thinking Scales (CTS). *Comput. Hum. Behav*. **2017**, *72*, 558-569. DOI: http://dx.doi.org/10.1016/j.chb.2017.01.005.
20. Chapelle, C. *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology*. John Benjamins: Philadelphia, US. 2003; Volume 7, 1-211
21. Herring, C. Computer-mediated discourse analysis: An approach to researching online behavior. In J. H. Gray (Ed.), *Designing for virtual communities in the service oaf learning*. Cambridge University Press, Cambridge, UK. **2004**, pp. 338–376