Iyer, S. et al. (Eds.) (2022). Proceedings of the 30th International Conference on Computers in Education.
Asia-Pacific Society for Computers in Education

Aqualab: Establishing Validity of an Adventure Game for Middle School Science

Yoon Jeon KIM[a*], Shari J. METCALF [b], Jennifer SCIANNA[a], Glenda PEREZ[b], & David GAGNON[a]
[a]*Curriculum and Instruction, University of Wisconsin-Madison, USA*
[b]*Harvard Graduate School of Education, Harvard University, USA*
*yj.kim@wisc.edu

Abstract: In this work-in-progress poster, we will present how a team including game designers, learning scientists, and assessment scientists collaborated on an online adventure game, Aqualab, with the goal of creating a comprehensive long-format game that can be used across multiple classroom sessions to support development science inquiry practices as well as assess different learning pathways within the game. In this work-in-progress poster, we discuss how the team approached design and development of the game to ensure validity of the game, and how we are planning to further investigate validity evidence of the game as a whole.

Keywords: Game-based learning, Science practices, Validity

1. Introduction

As outlined by the U.S. Next Generation Science Standards, performance of science tasks requires both understanding of core content and the ability to use science practices to investigate the world and solve problems (NGSS Lead States, 2013). To support learning of science practices, learners need to be situated in rich and authentic environments in which they can simultaneously and repeatedly engage in science practices to study phenomena (Crawford, 2012). Digital games can address this need, through immersive experiences in which participants can engage in active learning with simulated science environments and tools (Dede & Barab, 2009). Games provide a useful means of teaching complex science concepts in authentic contexts, while adding mediating scaffolds such as expanding or compressing time and perspective (Kamarainen et al., 2015).

Due to these affordances of games and simulations for better science learning, several research projects developed and implemented science games that specifically target science practices (e.g. Whyville, The Radix Endeavor). While these works report benefits of such games and simulations to support science content learning (Li & Tsai, 2013) and better assess science practices (Gobert et al., 2012), yet little is known about how learners can be immersed in rich and authentic narratives mirroring what real scientists do in the wild, and experiment with different identities in the game while honing science practices by solving comprehensive quests/tasks. In this work-in-progress paper, we present the design and development of an educational game called Aqualab, and in particular, how the team is conducting a validation study of the game.

2. Literature on Simulation and Games for Science Inquiry Practices

There have been several educational games and simulations developed to support and assess science inquiry practices of middle and high school students. For example, Radix Endeavor (Rosenheck et al., 2017), is a massively multiplayer-style online game developed by the MIT Education Arcade to engage learners with biological concepts in an alternate world. Players engaged with real-world problems such as medical diagnoses or genetic study of plant species. Epistemic games such as Nephrotex (Chesler et al., 2013) bridge the experience between more play-centric educational games and learning simulations; as a virtual internship, learners authentically participate in disciplinary practices of evaluating design tradeoffs, proposing solutions, and working within a design team. Other simulations have utilized a more open-ended, inquiry-based approach. EcoMOD (Dickes et al., 2019) and EcoMUVE (Metcalf et al., 2018) both provide students with curricular support in the form of modeling tools which assist them in making sense of the virtual environment. Students are able to collect data, conduct experiments, and aggregate their understandings through agent-based models and concept maps respectively. Observing the benefits of situating science practice within the user interactions, we moved to iterate on these prior designs and build an adventure-centric game around practice-based mechanics.

3. Context: Aqualab

In Aqualab, the student plays the role of an ocean-based research scientist on a ship, taking on "jobs" in different aquatic ecosystems in order to investigate questions relating to science phenomena. To date, 35 jobs have been created across four different ecosystems: kelp forests, bayou, coral reef, and the arctic, covering a range of life sciences phenomena, e.g., food webs, competition, photosynthesis, and adaptation. The game focuses on three science practices: experimentation, modeling, and argumentation. For example, there are jobs set in a kelp forest ecosystem in which students collect organisms, conduct experiments, and build models to determine that urchins eat kelp, and if left unchecked, can decimate a kelp forest. However, sea otters, urchins' main predators, can keep the populations balanced to form a healthy kelp forest ecosystem.

For the design of Aqualab, the team conducted a domain modeling analysis to inform game design decisions. The designers consulted with learning and assessment scientists to bring together specifications of tasks and tools related to the target science practices, a range of rich and interesting questions on aquatic phenomena and systems, and fun and engaging game mechanics and interface. The team also manipulated features of the game to vary the difficulty of in-game tasks in two ways: (1) differences in scaffolding of jobs, and (2) opportunities to engage with more advanced tools at deeper levels of complexity as they progress in game challenges.

4. Validity Study

DiCerbo et al. (2017) describe that evidence for validity in game-based learning can be established in both during the design of the game and after the game has been fully developed. There are several design frameworks (e.g. ECgD by Bob Mislevy, Game, Learning, and Assessment mechanics by Jan Plass) that can be used to achieve "validity by design" where the designers have confidence that the game design choices will produce potential evidence for target competencies and learning outcomes as well as embody theories of learning. After the game is developed, multiple approaches, ranging from traditional psychometric techniques to educational data mining practices, can be applied to provide evidence for validity. For example, Kim and Shute (2015) investigated correlation coefficients between the features generated from key performance indicators of Physics Playground and scores generated from external measure of conceptual physics understanding. Similarly, DiCerbo (2014) applied ECD to create two features (time and completion) related to persistence in an online game called Poptropica and conducted Confirmatory Factor Analysis (CFA) using scores based on the 2 features across three tasks. This work-in-progress poster addresses the question: To what extent can Aqualab be considered a valid assessment of science practices and how can we provide external validity evidence? To answer this question, the team pilot-tested Aqualab at a two middle schools in the Northeast U.S. The game was played by 336 students and five teachers for 5-7 (47-minute) class periods. The team administered an external assessment with these students, consisting of six performance-based tasks: two on each of the three science practices (experimentation, modeling, and argumentation). The tasks were adapted from existing valid assessments including the Stanford NGSS Assessment Project (SNAP) and the Trends in International Mathematics and Science Study (TIMSS) grade 8 science assessment.

Our validation study includes the following steps: (1) The team will process the game telemetry data to generate a set of features that the design team intentionally designed to align in-game actions with science practices (e.g., number of jobs completed, difficulty of jobs completed, time playing the game, incorrect choices in argument, number of experiments that the student ran). (2) The team will conduct a series of exploratory data analyses (Tueky, 1977) to examine whether the selected features seem consistent and identify opportunities for additional feature engineering. (3) The team will examine to what extent in-game features account for performance on the external assessment, including investigating correlation coefficients, factor analysis, and regressors such as Random Forest.

5. Discussion

In this work-in-progress poster, we describe the design and development of an educational game called Aqualab, particularly our effort to establish validity evidence for science practices. During the conference, we will present the game and early findings from empirical investigation of validity evidence. We are currently in the analyzing the data from the pilot study. We believe the design and development of Aqualab for the purpose of supporting and measuring science practices presents unique challenges. Thus, illuminating the design and development process, and how the team is balancing game and learning design while ensuring validity evidence, can be valuable for the field of educational games as well as the participants of the ICCE conference.

Acknowledgements

This material is based upon work supported by the National Science Foundation under grants DRL-1907384, DRL-1907398, and DRL-1907437. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Chesler, N. C., Irgens, G. A., D'Angelo, C. M., Bagley, E. A., & Shaffer, D. W. (2013). Design of a Professional Practice Simulator for Educating and Motivating First-Year Engineering Students. 30.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of educational research, 86*(1), 79-122.

Crawford, B. A. (2012). Moving the essence of inquiry into the classroom: Engaging teachers and students in authentic science. In Issues and challenges in science education research (pp. 25-42). Springer, Dordrecht.

Dede, C., & Barab, S. (2009). Emerging technologies for learning science: A time of rapid advances. *Journal of Science Education and Technology, 18*(4), 301-304.

DiCerbo, K., Shute, V., & Kim, Y. J. (2017). The future of assessment in technology rich environments: Psychometric considerations. *Learning, design, and technology: An international compendium of theory, research, practice, and policy*, 1-21.

DiCerbo, K. E. (2014). Game-based assessment of persistence. *Journal of Educational Technology & Society, 17*(1), 17-28.

Dickes, A. C., Kamarainen, A., Metcalf, S. J., Gün-Yildiz, S., Brennan, K., Grotzer, T., & Dede, C. (2019). Scaffolding ecosystems science practice by blending immersive environments and computational modeling. British Journal of Educational Technology, 50(5), 2181–2202. https://doi.org/10.1111/bjet.12806

Gobert, J. D., Sao Pedro, M. A., Baker, R. S., Toto, E., & Montalvo, O. (2012). Leveraging educational data mining for real-time performance assessment of scientific inquiry skills within microworlds. Journal of Educational Data Mining, 4(1), 104-143.

Kamarainen, A., Metcalf, S., Grotzer, T., & Dede, C.J. (2015). Exploring ecosystems from the inside: How immersion in a multi-user virtual environment supports epistemologically grounded practices in ecosystem science instruction. *Journal of Science Education and Technology. 24*(2), 148-167.

Kim, Y. J., & Shute, V. J. (2015). The interplay of game elements with psychometric qualities, learning, and enjoyment in game-based assessment. *Computers & Education, 87,* 340-356.

Li, M. C., & Tsai, C. C. (2013). Game-based learning in science education: A review of relevant research. Journal of Science Education and Technology, 22(6), 877-898.

Metcalf, S. J., Reilly, J. M., Kamarainen, A. M., King, J., Grotzer, T. A., & Dede, C. (2018). Supports for deeper learning of inquiry-based ecosystem science in virtual environments—Comparing virtual and physical concept mapping. Computers in Human Behavior, 87, 459–469. https://doi.org/10.1016/j.chb.2018.03.018

NGSS Lead States (2013)

Rosenheck, L., Clarke-Midura, J., Gordon-Messer, S., & Klopfer, E. (2017). Tipping the Scales: Classroom Feasibility of the Radix Endeavor Game. In M. Ma & A. Oikonomou (Eds.), Serious Games and Edutainment Applications (pp. 225–258). Springer International Publishing. Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2, pp. 131-160).