

Situating *Deep Multimodal* Data on Game-Based STEM Learning

Craig G. Anderson, John V Binzak, Jennifer Dalsen, Jenny Saucerman, Anna Jordan-Douglass, Vishesh Kumar, Aybuke Turker, Matthew Berland, Kurt Squire, and Constance Steinkuehler
cganderson4@wisc.edu, binzak@wisc.edu, jdalsen@wisc.edu, jenny.saucerman@gmail.com,
annamjordan@gmail.com, vishesh.kumar@wisc.edu, aybketrker@gmail.com, mberland@wisc.edu,
kurt.squire@gmail.com, constances@gmail.com
University of Wisconsin – Madison

Abstract: As STEM embedded games become more prevalent in classrooms, the need for teachers and researchers to understand the ways in which students learn in these complex environments increases. This paper describes a multimodal datastream approach to understanding student learning in an informal game-embedded curriculum. Through a multi-stream approach, we have more information on what students are using and how they are improving, which in preliminary analyses, proves to be more complex than one might think. Importance of multiple data streams in analyzing complex learning environments and future directions for more complex analyses are discussed.

Introduction

Video games are gaining attention for their potential to engage learners' interests in STEM and achieve learning outcomes (Mayo, 2009), yet they also present an exciting opportunity to extract formative information about how learning occurs through these digital devices. Games have the capability to amass a large body of click-stream data representing the actions, choices, successes, and failures of the player without disrupting the engagement that games evoke. Games often incorporate effective learning strategies such as scaffolding, adapting challenges to learning curves, engaging curiosity, and situating content in interactive environments with social surroundings. In addition, gameplay can mimic the scientific process of discovery, as players develop hypotheses and experiment with different actions. These similarities to other educational processes have not gone unnoticed by educators, who have produced educational video games to capitalize on the inquiry-based nature of gaming (Barab, Gresalfi, & Ingram-Goble, 2010).

Although players of commercial games have been known to engage in STEM practices outside of more traditional learning contexts (e.g., Steinkuehler & Duncan, 2008; Martinez-Garza, 2015), many STEM games are just one component of a larger learning curriculum. Incorporating STEM games into a larger learning curriculum allows players to work in groups, reflect upon their gameplay with their peers, and draw connections between their gameplay and the broader scientific content. This approach also allows researchers of scientific games to collect more targeted data regarding players' learning.

The data that researchers in this area collect is as varied as the STEM games themselves. At the very least, researchers may collect pre to post data on domain content knowledge in order to determine whether their scientific games lead to learning gains (e.g., Papastergiou, 2009; Collier & Scott, 2009). Participants' attitudes regarding the game are also frequently captured. Less frequently captured, however, is the connection between participants' in-game behaviors and their learning gains. Even more rarely captured is the relationship between learning gains, in-game behaviors, and participant discourse surrounding the game. In this paper, we explore these multiple data streams in order to form a more cohesive understanding of our participants' experience with our STEM game.

Methods

Intervention: Game-a-Palooza

To examine how participants engage in science learning through games, we developed a spring break camp for middle school students dubbed *Game-a-Palooza* (GaP). During the 5-day event, participants cycled through three educational video game sessions including our intervention of interest, titled *Virulent*. *Virulent* is a strategy game for ages nine and up designed for players to learn about viral replication and how the body's immune system reacts to infection. In the game, players control the 'Raven Virus', a fictional virus with biology-content-relevant characteristics. Players move through levels by infecting host cells, stealing precious resources, and fighting the immune system with viral proteins. To further engage the participants, *Virulent* sessions were framed in a role-playing activity.

The activity started with a video from actors roleplaying as a mock Center for Disease Control (CDC) team asking for help to stop the Raven Virus epidemic. Participants were grouped into twelve research teams of 3-4 players each. Participants were also given the opportunity to switch teams after assignment to allow friends or siblings to participate together, as the camp was designed to be both fun and informal. Across our groupings, we collected large and diverse amounts of data including clickstream gameplay data, discourse on and around the game, and pre- and post-assessments. On the first day, teams investigated preliminary recommendations on how to stop the Raven virus from spreading by observing the virus and the cells' defenses in the first few levels. The second day focused more on gameplay and constructing a visual model of virus and cell interactions. Teams updated their models and recorded presentations of their models to be sent to the mock-CDC for review on the third day. The fourth day consisted of cohort presentations to other groups and ended with an "emergency" call from the CDC reporting that they too have become infected by the Raven virus. On the last day, teams were presented three hypothetical solutions for stopping the Raven virus. These were based on articles from journals, media, and textbooks that were adapted to the content. Teams presented their findings to the cohort, debated, and voted for which solution to use.

Virulent acted as a source of click-stream data for our research by logging all game actions, their context, the use of an in-game almanac. The almanac provided a virtual directory where players could look up definitions and corresponding pictures of every cell part and virus component found in the game, allowing players to connect biological knowledge with game knowledge.

Data collection

Pre and post assessments

Before and after the *Virulent* curriculum, all participants completed identical short assessments measuring their knowledge about cellular biology and virology. Items measuring content knowledge included 8 multiple choice questions, 2 open response prompts, and a model with cell parts to label. Open response items were scored using a rubric from 0 to 1 based on accuracy and completeness of response. An aggregate score of *Biological knowledge* was calculated from these items for each participant. Participants were allowed to leave responses blank, as the assessment was not treated as a formal test. As such, biology knowledge scores were calculated as a proportion of correct responses out of the number of questions attempted.

Talk data

We collected a complete stream of discourse (verbal) data via a lavalier audio recorder worn by every student on their name badge. In addition, we video-recorded the classroom presentations of models of the virus and debate of proposed solutions for the epidemic. Lastly, we conducted daily interviews with the participant teacher.

Click-stream data

For each participant across all five 90-minute sessions, we recorded a complete stream of in-game telemetry data via our backend data framework: the Assessment Data Aggregator for Game Environments ADAGE (Owen & Halverson, 2014). ADAGE is a tool for logging and tagging clickstream data for every event within the game, allowing for analysis of player-actions to triangulate those in-game play patterns against other external measures.

Hypotheses

Intuitively, we expected our primary data streams to align. Increased use of the almanac to look up cell parts and virus components should lead to increased use of biology terminology in discourse and increased biology content knowledge post assessment scores. We would also expect that increases in group engagement (measured via amount of general talk) would align with increases in post assessments scores, as we anticipate player engagement to track with learning outcomes.

Results

Pre and post assessment

Forty-three players completed a pre and post assessment. Overall, players showed limited knowledge of cell biology and virology facts on the pretest ($M = .233$, $SD = .207$). After the intervention, players attempted a similar proportion of posttest items ($M = .753$, $SD = .325$), but their accuracy on those attempts was much higher ($M = .491$, $SD = .213$), indicating a significant increase in biology and virology knowledge over the 5 day intervention ($t(42) = 5.47$, $p < .05$).

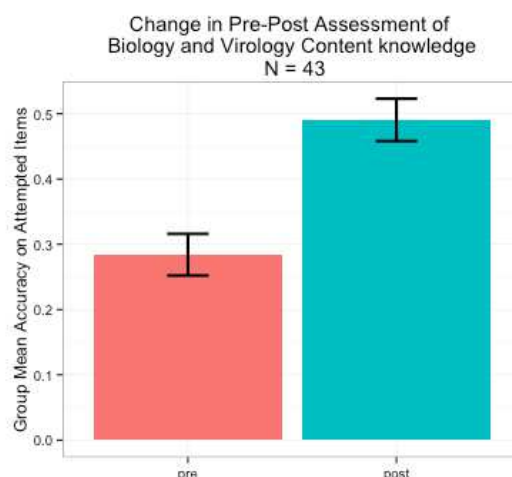


Figure 1. Aggregate change in biology knowledge[—]

Discourse data

To assess how discourse changed between the participants in the *Virulent* sessions, we began by looking at the length of words used as a rough proxy for sophistication of talk (after Flesch, 1948), and examined how this discourse changed across the five days within each research team. As seen in *Figure 3(a)*, research teams conversed with longer words on average during the fourth and fifth days of the intervention. Group means across all five days show that word length increased for all participants. Looking across the five day curriculum in *Figure 2(a)*, we see differences in how the curriculum may have prompted more discussion, though the data are skewed by one outlier group. During the penultimate day, participants used more biology vocabulary words as a percentage of their total spoken words.

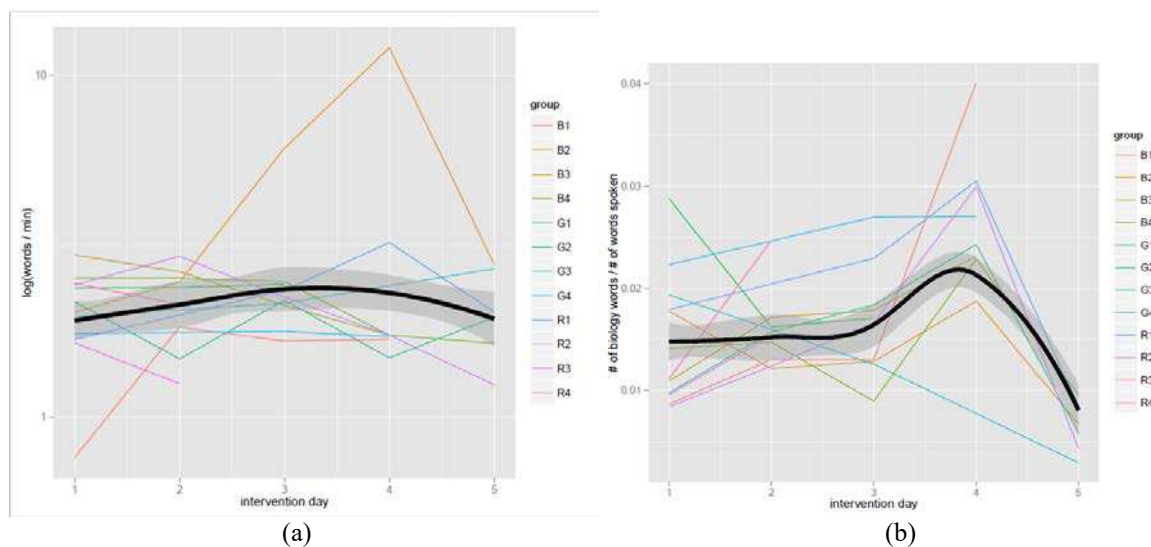


Figure 2. Overall use of biology terminology (a) and proportional use of biology words (b).

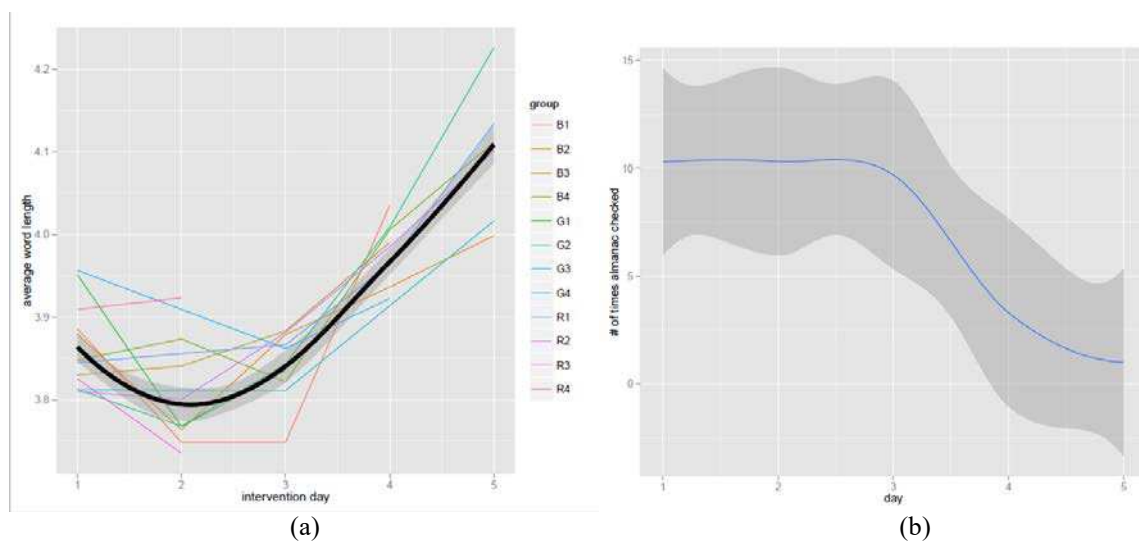


Figure 3. Average word length (a) and almanac usage (b).

Discussion

These preliminary analyses of multiple adjacent data streams in a scaffolded informal learning environment show the importance of each data source for understanding how students construct knowledge in a complex space. Being able to contrast game play behaviors with talk audio as well as pre/post assessment provides a granularity in the data that shows a larger picture of student development and understanding. However, from these data, we find patterns that are much more complex than expected. While we do not see the expected correlations between data streams, we still find that virology-specific knowledge increased 18% across participants, and participants used more complex language towards the end of the camp. This suggests that students did learn the embedded biology content but this was not captured in ways we would expect. Perhaps the increase in biology terminology in discourse and decrease in almanac usage suggest a transition away from reliance on the almanac as players become more familiar with the content. The complexity of the data and the importance of the individual differences only became clear as a result of triangulating seemingly clear findings; this is a challenge to the field.

Future research

Several lines of inquiry are currently being pursued to further complexify our data, including: participant interest, biology content gains, argumentation, model making, play patterns viewed through telemetry, differences in formal and informal play, and motivational factors. The multiple data streams from this study allow us to merge audio, video, artifacts and clickstream data in ways that will provide a more complete, albeit more complicated narrative.

References

- Barab, S. A., Gresalfi, M., & Ingram-Goble, A. (2010). Transformational play using games to position person, content, and context. *Educational Researcher*, 39(7), 525-536.
- Coller, B. D., & Scott, M. J. (2009). Effectiveness of using a video game to teach a course in mechanical engineering. *Computers & Education*, 53(3), 900-912.
- Martinez-Garza, M. (2015). "For!! SCIENCE!!": Examining Epistemic Practices of the Community of Players of Dwarf Fortress. *International Journal of Gaming and Computer-Mediated Simulations*, 7, 46-64.
- Mayo, M. J. (2009). Video games: a route to large-scale STEM education?. *Science*, 323(5910), 79-82.
- Owen, V. E., Ramirez, D., Salmon, A., & Halverson, R. (2014). Capturing Learner Trajectories in Educational Games through ADAGE (Assessment Data Aggregator for Game Environments): A Click-Stream Data Framework for Assessment of Learning in Play. In *Annual Meeting of the American Educational Research Association, Philadelphia, April* (pp. 3-7).
- Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education*, 52(1), 1-12.
- Steinkuehler, C., & Duncan, S. (2008). Scientific habits of mind in virtual worlds. *Journal of Science Education and Technology*, 17(6), 530-543.