Where the Rubber Meets the (Cross)Road: Insights into Game-Based Learning & Assessment Design

Abstract: Human agency—the power to shape a course of action (Bandura, 2006)—has long been held central to empowered, self-regulated learning (Pintrich, 1999). Implicit in this choice-centered learning is the ability to take different pathways towards a solution, a key affordance of games as a medium to deliver engaging, effective educational experiences (c.f. Salen & Zimmerman, 2004). The work in this symposium reflects a holistic design goal in game-based learning—integrated learning and assessment which can support multiple pathways to learning. This integrated design is discussed from four perspectives: teacher-facing UI design, learning game design (integrated instruction & assessment), psychometric assessment analysis, and exploratory data mining of emergent learner patterns.

One Theme, Multiple Pathways: An Overview

V. Elizabeth Owen, GlassLab Games

Human agency—the "power to shape…circumstances" and a "course" of action (Bandura, 2006, p. 164)—has long been held central to empowered, self-regulated learning (Pintrich, 1999; Glaser, 1996). Implicit in this choice-centered learning is the ability to take different pathways towards a solution (Pintrich, 2000). Indeed, multiple pathways of strategy in problem solving is a trademark of authentic assessment in rich, interactive learning contexts (c.f. Schank, 2011; Mayrath et al., 2012). One powerful vehicle for self-regulated learning can be the scaffolded, engaging microworlds of games (Rieber, 1996). Good games provide well-ordered problems in which students are the drivers of their own experience, able to customize strategy with low cost of failure (Gee, 2005). As data-rich environments which offer roles, goals, and agency (e.g. Norton, 2008; Steinkuehler et al., 2012), digital games can integrate instruction and assessment for optimal, adaptive support of individual learner choices.

The work in this symposium reflects our core design goal—integrated learning and assessment which can support multiple pathways to learning, a key affordance of games as a medium to deliver engaging, effective educational experiences (Salen & Zimmerman, 2004). The four proposed symposium presentations explore challenges and insights into core components of this process: teacher-facing UI design, learning game design (integrated instruction & assessment), psychometric assessment analysis, and exploratory data mining of unexpected learner patterns.

Each themed around multiple pathways, these studies are based in current learning game design, development, and analysis efforts at GlassLab (in partnership with Educational Testing Service and Pearson). The first discusses design efforts around GlassLab's online dashboard, optimizing visibility of game-based learning to educators (critical constituents in UI design). "Teacher Portal Design" discusses considerations and lessons learned around designing multiple representations of data that are simple yet richly informative. Next, "Designing on Mars" centers on a design process that fosters and assesses students' learning, and yields, rich multi-path learning data for the dashboards. In alignment with designed game tasks, "Game-based Assessment of Argumentation Skills" next shares psychometric analysis of event-stream data. It represents varied player choices in the form of evidence fragments, and evaluates the relationship of specific action combinations to learning outcomes. Lastly, "Mining Multiple Learner Pathways" uses an exploratory approach. Leveraging event-stream data beyond the initially designed evidence fragments, this analysis uses Educational Data Mining (Baker & Yacef, 2009) to mine organic (and unexpected) learner trajectories.

References

Baker, R., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3–17.

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, 1(2), 164–180. doi:10.1111/j.1745-6916.2006.00011.x

Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-Learning and Digital Media*, 2(1), 5–16.

Glaser, R. (1996). Changing the agency for learning: Acquiring expert performance. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 303–311). Mahwah, NJ: Lawrence Erlbaum Associates.

Mayrath, M. C., Clarke-Midura, J., & Robinson, D. H. (Eds.). (2012). *Technology-based assessments for 21st Century Skills: Theoretical and practical implications from modern research*. Charlotte, NC: Information Age Publishing.

Norton, D. (2008). A practical model for separating games and simulations. Presented at *5th Annual Games+Learning+Society Conference*. Retrieved from http://www.previous.glsconference.org/2008/session.html?id=139

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, 31(6), 459–470.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92(3), 544.

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development*, 44(2), 43–58.

Salen, K., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. MIT Press.

Schank, R. (2011). *Teaching minds: How cognitive science can save our schools*. New York, NY: Teachers College Press.

Schwartz, D. L., & Arena, D. (2009). *Choice-based assessments for the digital age*. MacArthur 21st Century Learning and Assessment Project.

Steinkuehler, C., Barab, S., & Squire, K. (Eds.). (2012). *Games, learning, and society: Learning and meaning in the digital age*. New York: Cambridge University Press.

Teacher Portal Design: Making Game-Based Learning Visible to Educators

Kristen E. DiCerbo, Pearson
Elizabeth Kline, GlassLab Games
Michelle M. Riconscente, GlassLab Games

There is great interest in the use of data by teachers to inform instruction. Research has repeatedly shown that student achievement improves dramatically when teachers are provided with information about their students' learning (Hattie, 2009). The use of learning games in classrooms exponentially increases the amount of student learning data available. However, these same systems also remove the teacher from direct interaction with students' work products. If the promise of game data to reshape education is to be met, we must find ways to communicate the information from digital activities to teachers in ways that help them make instructional decisions in the classroom (Fishman et al., 2015).

Indeed, as the conduit for opening game-based learning to students across the country, teachers are critical constituents in user-friendly UI design. A main goal of our work here was informed design of an educator portal for visibility of learner progress, thus optimizing classroom facilitation tools. We describe the teacher interview process, key feedback takeaways, and resulting iterative UI designs used to make learning visible to educators who are facilitating game-based education in the classroom. Multiple representations of data are considered, both in terms of showing variations in student play patterns, as well as optimizing final visuals for an educator audience with different forms of data transparency.

To begin the process, we conducted a series of iterative think-aloud studies with teachers in an effort to understand how they interpret data displays of their students' game activity. We began by identifying particularly instructional decisions to target, including: what should I teach next, how should I group students, and what support do particular students need? We designed displays targeting these questions and conducted 3 cycles of think-aloud and design iteration. Results of these iterations led to many designs, including those shown below (Figure 1). These utilize simple color schemes, improved intuitive shape representations, and class-wide as well as individual student views.

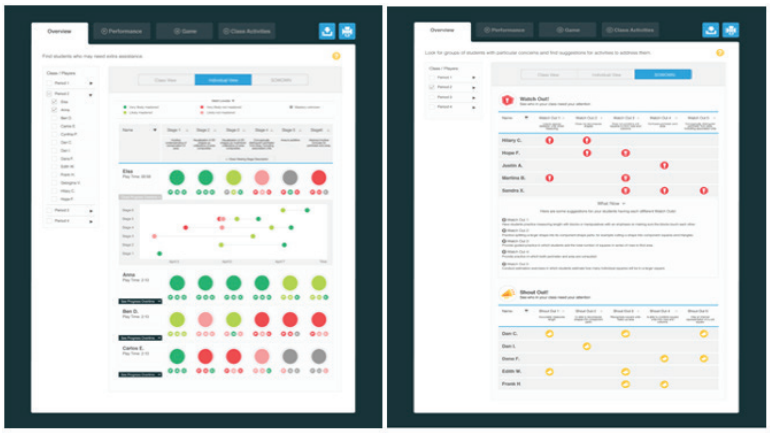

Figure 1: Iterative UI designs of game-based learning portal for educators.

Some particular lessons learned fall under four main themes. First, considering the *nature of the data*, one challenge is educator difficulty in understanding visualizations that represent the probability of mastery. *Information overload* emerged as another theme, as hover interaction or "tooltip" prove difficult to teachers to make use of in whole class views. Next, in *showing learner progress*, teachers often confound level of gameplay, formative learning progress, and summative content mastery. Lastly, in distinguishing these kinds of learner progressions, *actionable data* become important in two ways: first, teachers need tools that quickly help them spot trends in the class group, so that they can quickly target groups of students for re-teaching or extension; second, in order to apply the information about student progress, teachers also benefit from suggestions regarding instructional strategies that will be most effective in addressing the learning issues surfaced in the reports.

Overall, this teacher interview process, ensuing feedback, and ongoing refinement loop supported optimized, iterative UI design. Throughout the process, these insights helped greatly to help make learning visible to educators facilitating game-based education in the classroom.

References

Fishman, B., Riconscente, M., Snider, R., Tsai, T., & Plass, J. (2014). *Empowering educators: Supporting student progress in the classroom with digital games*. Ann Arbor: University of Michigan. gamesandlearning.umich.edu/agames

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Taylor & Francis.

Designing on Mars: Learning, Assessment, & Game Designed Together

Erin Hoffman, GlassLab Games

This section of the symposium focuses on the GlassLab learning game design process and how it brings together designers from the learning, assessment, and game design spaces to collaborate on a game product. In *Mars Generation One: Argubot Academy*, GlassLab learning, assessment, and game designers worked together on a hybrid process based on evidence-centered design (Mislevy, Corrigan et al, 2014; Mislevy, Steinberg, & Almond, 2003; Riconscente, Mislevy & Corrigan, in press) to create a new game design methodology that first considers a learning performance, then matches it to a game mechanic, then tunes that game mechanic to reveal a student's thinking. Within the learning game, the content being assessed should guide the selection of relevant tasks, as well as the rational development of content-based scoring criteria and rubrics (c.f. Messick, 1994).

A product of this integration is an assessment-aligned design process for GlassLab learning games. Generally, our game design begins with learning and assessment designers mapping a competency and teaching it to the game team. Once the competency is identified, fundamental design work involves mapping it to a core loop of learning tasks in-game. Specifically: 1) a game is always built around a core mechanic; 2) the mechanic is a set of verbs that create a core loop; and 3) if the core doesn't match the learning, the learning never happens. Assessment is inherent here, where event-stream data from the core loop provide task-aligned learning evidence. Narrative variations on the loop provide occasions for multiple play pathways. An example of the core loop in MGO is shown in Figure 2. In iterative development of this loop and game narrative, the approach to playtesting is discussed, as well

as strengths and weaknesses of the process as we perceive them at this stage of development. Common pitfalls will be discussed as well as ways to mitigate them, including strategies now being used to develop GlassLab's latest learning game: an HTML5 product to teach proportional reasoning to 7th graders.

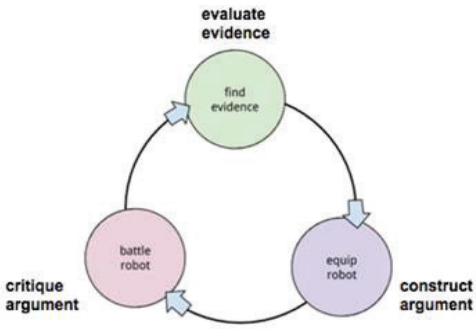

Figure 2: Learning Game Core Loop: Mars Generation One

At the end of the presentation we will discuss our revised thinking—how the methodology has changed between products—and discuss further modifications being made to the process with the current math product. We'll also provide a brief comparison between developing for two similar-in-difficulty but widely variant in content competencies (argumentation and proportional reasoning).

References

GlassLab, Inc. (2014). *Mars Generation One: Argubot Academy*.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Education Researcher*, 32(2), 13-23.

Mislevy, R.J., Corrigan, S., Oranje, A., Dicerbo, K., John, M., Bauer, M.I., Hoffman, E., von Davier, A.A., Hao, J. (2014). *Psychometric considerations in game-based assessment*. Redwood City, CA: GlassLab.

Mislevy, R.J., Steinberg, L.S., & Almond, R.G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1, 3-66.

Riconscente, M.M., Mislevy, R.J., & Corrigan, S. (in press). Evidence- centered assessment design. To appear in S. Lane, M. Raymond, & T. Haladyna (Eds.), *Handbook of Test Development*. Mahwah, NJ: Lawrence Erlbaum Associates.

Game-based Assessment of Argumentation Skills

G. Tanner Jackson, Educational Testing Service
Maria Bertling, Educational Testing Service

Game-based assessment (GBA) is progressing beyond just promises (Gee, 2007; Klopfer, Osterweil, & Salen, 2009), emerging as an empirically supported and reasoned approach to learning and assessment (Mislevy et al., 2014). GBA takes particular advantage of the interactivity and engagement games are built around to provide rich data streams and multiple sources of evidence for a given construct. An important driver to this work is that significant investment in the development and validation of cognitive models (e.g., learning progressions) provides a robust basis for underlying game design and allows for making detailed and actionable claims about student performance. Particularly, leveraging the development guidelines of *Evidence Centered game Design* (ECgD; Hoffman, John, Makany, 2014) provides educators and researchers with the necessary arguments and principled design practices to produce quality assessment games.

Mars Generation One: Argubot Academy (MGO; developed using ECgD by GlassLab and Educational Testing Service) teaches and assesses students' argumentation skills through an RPG-adventure-based educational game for the iPad. Targeted argumentation skills were selected from an existing learning progression (Song, Deane, Graf, & van Rijn, 2013) and implemented within MGO as different game mechanics (i.e., explore, equip, battle) involving aspects of argument identification, organization, and evaluation.

Approximately 590 middle school students interacted with MGO and completed a pre- and post-test on argumen-

tation knowledge. Results showed mean differences in test-scores with students' performing significantly higher on the post-test than the pre-test, t(588)=10.779, p<.001, Cohen's d=0.44. Additionally, significant correlations among in-game evidence and the external argumentation measures (displayed in Figure 1) indicate that previously hypothesized telemetry evidence fragments (identified during the ECgD process) are relevant to the construct of argumentation and can be used to draw connections between student performance in the game and their knowledge of argumentation. Additionally, subgroups of students (high or low levels of prior argumentation knowledge) were compared and different patterns of relations were discovered between learning gains (post – pre) and in-game evidence fragments. Thus, students' prior knowledge contributes significantly to their learning gains from different components within the game.

	1	2	3	4	5	6	7	8	9	10	11	12
1 Pre Arg Measure	-	.489**	.359**	.105*	.261**	.273**	.329**	.129**	0.078	0.056	0.074	.142**
2 Post Arg Measure		-	.435**	0.061	.278**	.240**	.377**	.228**	.084*	.083*	.120**	.085*
3 Strong Cores			-	0.008	.247**	.237**	.339**	.211**	0.071	.090*	.087*	.129**
4 Relevant Data				-	-0.007	0.016	.121**	.099*	0.031	-0.014	.086*	0.036
5 Core Attack Not Related					-	.149**	.247**	.119**	-0.056	0.044	-0.031	0.055
6 Core Attack Not Supporting						-	.179**	-.087*	0.053	.092*	0.046	0.054
7 Critical Question Attack							-	.467**	.142**	0.04	.138**	.099*
8 Player's Bots in Battle								-	0.056	-0.017	0.066	0.078
9 ConServatrons' Defeated									-	.217**	.243**	.313**
10 Authoritrons' Defeated										-	.183**	.213**
11 Comparidroids' Defeated											-	.364**
12 Correct Tests												-

Figure 1: Correlations among in-game evidence fragments and argumentation anchor measures. *p < .05. **p < .01.

These findings extend previous research by demonstrating not only the feasibility of tackling hard-to-measure constructs and 21st century skills within game-based assessments, but also the importance of providing a multitude of interaction mechanisms with rich, diverse sources of evidence. We showed examples of evidence fragments for a latent construct (argumentation) that were derived from in-game process data and that can serve as significant indicators of learning.

References

Gee, J. P. (2007). *Good video games and good learning*. New York, NY: Peter Lang. (2007)

Hoffman, E., John, M. & Makany, T. (2014). How do game design frameworks align with learning and assessment design frameworks? Paper presented at *National Council for Measurement in Education*, Philadelphia, PA.

Klopfer, E., Osterweil, S., & Salen, K. (2009). Moving learning games forward: Obstacles, opportunities & openness. Retrieved July 14, 2012 from http://education.mit.edu/papers/MovingLearningGamesForward_EdArcade.pdf.

GlassLab, Inc. (2014). *Mars Generation One: Argubot Academy*.

Mislevy, R.J., Corrigan, S., Oranje, A., Dicerbo, K., John, M., Bauer, M.I., Hoffman, E., von Davier, A.A., Hao, J. (2014). *Psychometric considerations in game-based assessment*. Redwood City, CA: GlassLab.

Song, Y., Deane, P., Graf, E. A., & van Rijn, P. (2013). Using argumentation learning progressions to support teaching and assessments of English language arts. *R&D Connections*, No. 22.

Adventure on the Red Planet: Mining Multiple Learner Pathways in Mars Generation One

V. Elizabeth Owen, GlassLab Games
Kristen E. DiCerbo, Pearson

Upon landing on the mysterious Red Planet, players can begin to make a series of individual choices that affect civic decisions in the game *Mars Generation One: Argubot Academy* (MGO; GlassLab, Inc. , 2014). In the game, players learn critical argumentation skills and decide the fate of their city by sending Argubots into battles of persuasion. Since there is more than one pathway through the game, and each player brings their own approach

to play (e.g. Bartle, 1996; Yee, 2006), ways of capturing these paths must center on moment-to-moment student decisions (not one pre-defined "right" track). In other words, the analysis methods used to understand this player experience need to be as sensitive to organic player choice as the game itself. One aligned field is Educational Data Mining (EDM), which has developed methods to mine organic patterns from digital educational settings (Baker & Yacef, 2009).

This study explores multiple, organic play pathways in relationship to enjoyment and learning outcomes in *Mars Generation One*, as built on fine-grained event-stream data of each student using EDM. Since these player paths can consist of many different kinds of actions (not just a few predetermined learning moments), the identification of salient game events for the analysis (also called feature selection or feature engineering, e.g., Guyon & Elisseeff, 2003) presents a significant challenge (c.f. Halverson & Owen, 2014). The exploration of *MGO* has two parts: 1) The feature engineering process—distilling hundreds of click-stream event types into a set optimal for analysis; and 2) Building a predictive learner model—mining organic play patterns and multiple pathways based on these foundational data features. These two layers fuse to focus on the question: what organic in-game trajectories are most characteristic of learning and enjoyment in Mars Generation One?

Data Collection

In fall 2014, over 500 middle-school students across the country played *MGO* as part of their regular school day, spanning 5 hours over the course of a week. Outcome measures included the CBAL, a validated external measure of argumentation created by ETS, as well as a shorter argumentation measure that was piloted with the game. The latter was a close argumentation measure was designed to be an assessment of learning more tightly aligned to specific game skills than the more distally-related CBAL measure. An additional survey gathered information on student engagement (including enjoyment, effort, and self-reported learning). During play, hundreds of unique player actions and game events were logged for a recorded total of over 2 million click-stream events.

Feature Engineering

To help focus on event-stream actions salient to analysis, we looked at play through three lenses: base game progress, in-game success and failure as aligned with core design, and exploration and boundary testing. Features measuring base progress through the digital world (e.g. time elapsed, missions complete, etc.) have been features successfully used in game-based EDM modeling (e.g. DiCerbo & Kidwai, 2013; Baker & Clarke-Midura, 2013), while understanding success and failure relationships has yielded insights into failing productively (e.g. Kapur, 2006; Juul, 2013). Similarly, game research suggests that certain kinds of players have paths of exploration (Salen & Zimmerman, 2004) and failure-filled boundary testing (Owen, 2014) positively related to learning.

Outcome Variables

The outcome variables of learning and enjoyment were also refined for analysis. The CBAL scores and the close argumentation measure post-scores were selected as learning outcomes. The self-report engagement survey, however, measured multiple construct: enjoyment, competition, effort, difficulty, and self-reported learning. An exploratory factor analysis revealed three constructs: enjoyment, effort, and self-reported learning. Each of these three subscores then became a distinct outcome variable to be used in the predictive model.

Mining Play Patterns: Building a Predictive Learning Model

Next, we established predictive modeling to explore students' play patterns in relationship to these final outcome variables: argumentation skill, enjoyment, effort, and self-reported learning. Core EDM predictive algorithms (Baker & Siemens, 2014) were used from the family of classification and regression trees (CART)—specifically M5Prime, JRip, Naïve Bayes, J48, and PART. Preliminary results suggest strong multiple play pathways in relationship to learning and enjoyment as predicted by event-stream player actions (themed along progress, success, failure, and exploration). Results will be discussed in terms of prominent click-stream events in the model, emergent play trajectories, and outcome variables of best fit.

Mining the game data—themed along progress, success, failure and exploration—to predictively model learning and enjoyment can impact learning, game, UI, and curriculum design. For example, where organic (and perhaps unexpected) play pathways arise in positive relationship to learning gives the game designer opportunity to scaffold in just-in-time support. Emergent learning-related data features augments Evidence Centered Design (Mislevy & Haertel, 1996) assessment models, supplementing initial lists of hypothesized evidence fragments. Overall, both the methods and results here provide insight into the relationship between micro-level game actions and large scale learning goals.

References

Baker, R., & Clarke-Midura, J. (2013). Predicting Successful Inquiry Learning in a Virtual Performance Assessment for Science. In *Proceedings of the 21st International Conference on User Modeling, Adaptation, and Personalization* (pp. 203–214).

Baker, R., Siemens, G. (2014) Educational data mining and learning analytics. In Sawyer, K. (Ed.) *Cambridge Handbook of the Learning Sciences: 2nd Edition*, pp. 253-274.

Baker, R., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3–17.

Bartle, R. (1996). Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD Research*, 1(1), 19.

DiCerbo, K. E., & Kidwai, K. (2013). Detecting Player Goals from Game Log Files. Presented at the 6th *International Conference on Educational Data Mining*. Retrieved from http://www.educationaldatamining.org/EDM2013/papers/rn_paper_58.pdf

GlassLab, Inc. (2014). *Mars Generation One: Argubot Academy*.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, 3, 1157–1182.

Halverson, R., & Owen, V. E. (2014). Game Based Assessment: An Integrated Model for Capturing Evidence of Learning in Play. *International Journal of Learning Technology [Special Issue: Game-Based Learning]*, 9(2), 111-138.

Juul, J. (2013). *The art of failure: An essay on the pain of playing video games*. Cambridge: MIT Press.

Kapur, M. (2006). Productive failure. In S. Barab, K. Hay, & D. Hickey (Eds.), *Proceedings of the International Conference on the Learning Sciences* (Vol. 0, pp. 307–313).

Mislevy, R. J., & Haertel, G. D. (2006). Implications of Evidence Centered Design for Educational Testing. *Educational Measurement: Issues and Practice*, 25(4), 6–20.

Owen, V. E. (2014). *Capturing In-Game Learner Trajectories with ADAGE (Assessment Data Aggregator for Game Environments): A Cross-Method Analysis* (Unpublished doctoral dissertation). Madison, WI: University of Wisconsin-Madison.

Salen, K., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. Cambridge: MIT Press.

Yee, N. (2006). Motivations for play in online games. 9(6), 772-77. *CyberPsychology & Behavior*, 9(6), 772–77.