Assessment Design for Emergent Game-Based Learning

Jodi Asbell-Clarke
EdGE at TERC
2067 Massachusetts Ave
Cambridge, MA 02140
jodi_asbell-clarke@terc.edu

Elizabeth Rowe
EdGE at TERC
2067 Massachusetts Ave
Cambridge, MA 02140
elizabeth_rowe@terc.edu

Elisabeth Sylvan
EdGE at TERC
2067 Massachusetts Ave
Cambridge, MA 02140
sylvan@terc.edu

CHI 2013 Extended Abstracts, April 27–May 2, 2013, Paris, France.
ACM 978-1-4503-1952-2/13/04.

Abstract
Educational games may lend themselves to innovative forms of learning assessment. This paper reports on game-based science learning assessments that explore how to measure the emergent learning that takes place in games by revealing tacit knowledge development. This research combines video analysis and educational data mining to identify cognitive strategies that emerge through gameplay. By studying the video and click data from high school learners playtesting the game, *Impulse*, we identify systematic ways of predicting the observed strategies and making possible connections to formal science learning.

Author Keywords
Game-based learning; tacit knowledge; indwelling; assessment; video; educational data mining.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction
Nearly all youth and most adults participate in Internet-based games [1]. Games have been shown to foster scientific inquiry, problem-solving, and have enabled the public to make breakthrough scientific discoveries [2, 3]. As a result, many educators and researchers see digital games as potential learning and assessment environments for the 21st century [4].

Our research group is designing digital web and mobile games that focus on high-school science concepts drawn from U.S. standards for science education. These games use simplified game mechanics that emphasize the laws of nature and principles of science, thereby allowing players to dwell in scientific phenomena in order to build and solidify their tacit knowledge over time.

It is not our intent that these games *teach* science content, but rather they engage the learner with scientific phenomena in the effort to build their intuitive understandings about these phenomena. We then attempt to a) measure and b) leverage those intuitive (tacit) understanding through formal instruction.

Newton's First Two Laws

(1) An object that is in motion will stay in motion unless acted upon by an opposing force

(2) the motion exhibited by a particle under a force is dependent on the object's mass.

To measure learning in games, we explore the extent to which we can relate the development of cognitive strategies we see players building in the games to classroom learning of similar content. Thus, we address the question: *Does players' advancement in the game correspond to increased understanding of the science content driving the game mechanics?* Success in this design will result in a new way to think about game-based assessments, starting not from prescribed learning outcomes but from watching what types of learning actually take place. The first step of this research, reported in this paper, is to accurately predict the strategic moves that emerge in a physics-based game from the click data generated by that game.

Background

Game-based Learning Assessment

For the past decade, researchers have been exploring how to assess learning occurring in games [6, 7, 8]. The feedback mechanisms used by many game developers to move players through complex game tasks are similar to formative assessments educators call for [9].

A common research design is pre-post tests before and after gameplay. There is, however, a disconnect between traditional validated assessments and the styles of learning observed in games [6]. Stealth assessments, those integrated in a non-disruptive manner as part of the game, on the other hand, "strive to support learning, maintain flow, and remove (or seriously reduce) test anxiety, while not sacrificing validity and reliability" [6, p. 10] by making the assessment an integral part of the gameplay. Stealth assessments are typically developed using the evidence-centered design (ECD) framework that seeks to establish a logically coherent argument between the domain being assessed (competency model), assessment task design (task model), and interpretation (evidence model) [10].

This paper presents a new evidentiary argument for assessment based on the emergent strategies observed as learners play the game, *Impulse.* In our model, we observe game strategies that emerge as players dwell in simulated physical phenomena and then relate those strategies to the learning of the physics content itself. This evolving model of building evidence of learning is counter to most applications of the ECD model where explicit learning outcomes are defined in advance and assessment tasks stem from those outcomes [10].

The behavior and motion of the particles underlying the game mechanic simulate Newtonian motion under gravitational forces. We use the game to seek evidence of player's tacit understanding of general behaviors associated with Newton's Laws of motion. We believe gameplay strategies may reveal the development of tacit understanding of scientific concepts and, as such, represent the task model for assessment.

We use video-recorded playtesting data with dozens of players to code strategic gameplay moves. With coded clicks as "ground truth," we then use educational data mining techniques to predict those strategic moves and describe how strategies evolve as players advance in the game [11].

Indwelling and Tacit Knowledge
Indwelling involves the process of revealing complex ideas, practices, and processes that learners understand intuitively, but not yet explicitly [5, 12]. As players become immersed in games, not only may their tacit knowledge grow, but they may also demonstrate their unexpressed knowledge by virtue of their activities within the medium. Digital games also allow the recording and analysis of these activities, providing great potential for assessment of intuitive knowledge.

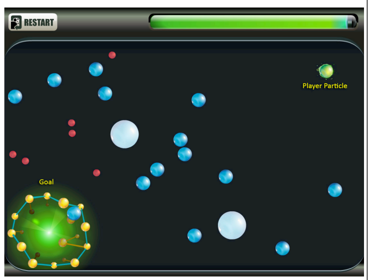

Figure 1: Impulse game

Game mechanics often rely on intuitive knowledge of physics and other STEM content. For example, in the game *Angry Birds* [13], players launch a bird with a slingshot to knock down a structure. That basic mechanic never changes, but as players progress, they use it in increasingly sophisticated ways. By higher levels, they cohere their various game experiences into a larger understanding, an aspect of indwelling where learners take intellectual control over these experiences [14], and begin to understand how their subsidiary experiences build a systemic picture of the environment. Players use these intuitive understandings to predict and navigate their way through the environment, which may provide measurable evidence of their implicit knowledge. Our research strives to describe how this measurement of emergent learning takes place in games can lead to the development of innovative learning assessment for a new generation of learners.

Investigation
Our team designed the game *Impulse* to scaffold and measure players' tacit knowledge of forces and motions (Figure 1). In *Impulse,* particles have different masses and thus behave differently under the corresponding gravitational forces. Players use an impulse (upon a click) to apply a force to particles, with the goal of moving their one particle to the goal while avoiding the ambient particles. If the player's particle collides with any ambient particle, she loses that round. In terms of the science, the player is immersed inside an N-body simulation with accurate gravitational interactions and elastic collisions among up to 30 ambient particles.

As players reach higher levels, they require cognitive strategies to predict the motion of the particles so that they can get to the goal, not run out of energy, and avoid collision with other particles. While navigating a sea of (elastically) colliding particles that are attracted to each other (through gravity), players need to "study" the particles' behavior in order to win. They need to predict ambient particles' motion and interactions with each other to avoid them as their own particle is guided to the goal.

Since there is no known best way for learners to build intuitive understanding of these physics phenomena in games, our research captures the myriad of strategies players develop during gameplay that may reveal tacit knowledge. As a first step of this work, we have identified an initial set of strategic moves that we observe players making in the game *Impulse.* In this paper, we describe these moves and why we think they have promise as an evidence model, linking the gameplay to tacit understandings of the underlying physics.

Data Sources

Data were collected over six workshops conducted in November and December 2012 with 15 high school students (3 female) from urban and suburban schools in the Northeastern United States. Players were recorded with Silverback software [15] that captures players' onscreen game activities, and audio and video of their faces. They typically worked in pairs, inducing conversation about gameplay and they were asked to 'think aloud" to explain their activities.

Two coders, a designer of Impulse with a physics background and a researcher without, independently watched the videos and coded two randomly selected three-minute segments from each player. The coding system was developed through repeated coding of hundreds of clicks with different play styles. It has evolved from focusing on the direction and subject of the impulse (player particle or ambient particle) to identifying moves as proactive and reactive, and then describing various dimensions separately (e.g., direction, subject, actual impact). *Proactive strategies* are typically clicks players use to move their particle to the goal and sometimes conserving energy to earn more points in the game. *Reactive strategies* are clicks they use to protect their particle when under threat of collisions. This distinction is important because players may act more thoughtfully on their tacit understandings when they are not threatened. We also anticipate that as players become better predictors of the phenomena, they will use more proactive strategies and need the reactive strategies less frequently.

The coding system also captures: 1) the target of the impulse—player, ambient particle, or both, 2) the direction of the impulse in relation to the path of the target particle, 3) whether or not the impulse happens at the beginning of a new level, 4) whether or not there is a time lag between impulses, and 5) the intended impact of the impulse on the target particle (e.g., toward goal, stopping, toward wall, avoid collision, buffer player particle, clearing ambient particles from path or goal). Player intentions are judged based on audio commentary and mouse over behaviors. Often players hold their mouse over spots, ready to click if needed, providing visible clues of their intended path. While not in the click data, these behaviors are observable in video and aid interpretation.

These refinements have led to significant increases in inter-rater agreement from less than 50 percent to over 70 percent across six dimensions with coders agreeing on the two of the dimensions almost 100 percent of the time. More refinement of the remaining four dimensions is planned to reach a high level of agreement after accounting for chance (Cohen's Kappa) across all 30 segments (2 segments x 15 players).

Strategic Moves as an Evidence Model

Because of the open-ended nature of *Impulse*, we need to take great care in identifying the assumptions we make and evidence we observe leading to our claims about the cognitive strategies that emerge from its gameplay. For each strategic move we identify for

Impulse, we first present the competency model describing intended learning, the task model describing the indwelling environment, and the evidence model linking the two.

Examples of several pro-active moves were evident in the gameplay of many different players in the videos. Three of the most prevalent are Push-to-Goal, Float-then-Redirect, and Oppose. Each of these strategic moves occurred repeatedly in video across multiple players. We are starting to separate the strategic moves that indicate mastery of a concept of the game mechanic, but not tacit understanding of the physics phenomena that underlies the game, from those strategic moves that we can start to build into cognitive strategies that have usefulness to science learning.

For example, a Push-to-Goal move, where the player uses the impulse to push the particle directly towards the goal, is used by nearly all players. With experience, players often begin to use the Float-then-Redirect strategic move in which they use the impulse and then wait, letting their particle float in the new direction. A similar move is using the impulse to Oppose the motion of the particle by clicking on the side opposite of the direction of its motion to redirect it toward the goal or avoid a collision with another particle.

Examples of reactive moves include Buffering where players preserve the space between the player and immediate ambient particles, and Blocking where players head off incoming particles from afar. We also see players use the impulse to clear ambient particles out of their path. We hypothesize that player's use of this strategy will evolve as the ambient sea of particles becomes more complex in higher levels.

To distinguish between strategies that indicate only advanced understanding of game mechanics and those that are indicative of intuitive physics, we rely on players' commentary on video and clinical interviews about their reasons for using the strategies.

Players comment that they notice differences in how particles of different masses behave, however in the coding of early levels we have not yet observed strategic moves in their gameplay that show preferential treatment of one mass over another. We predict players may shift strategies in later levels, demonstrating a tacit understanding of Newton's Second Law.

Identifying Potential Predictors of Strategy Use

The next step to studying the strategy-based learning in *Impulse* is detecting and distinguishing specific strategic moves from click data. Using our observations of the video data, we have identified *50+ gameplay features*—potential predictors like distance between particle and goal and the time since the last impulse. Related features are calculated for every player click.

Using Decision Trees to Predict Strategic Moves

We are creating decision trees using the *Rapid Miner* data-mining software. Decision trees are hierarchical classification rules that best predict an outcome of interest [11]. Sample classification rules for a strategic move might be (a) time since last impulse <=X seconds or (b) distance between particle and goal < Y pixels. The outcomes of interest are strategic moves coded from Silverback recordings. We create decision trees with the gameplay features using a cross-validation process in which the decision tree models are trained with half of the data. Then, the best fitting decision tree is used to predict specific strategic moves in the second half. The accuracy of those predictions is compared to what would be expected by chance.

Conclusions

During the playtest of *Impulse,* we saw strategies that the designers anticipated and emergent ones that they did not. We are attempting to identify patterns in the click data that predict players' strategic moves. These are early steps in developing an evidence model of tacit physics knowledge demonstrated via gameplay.

Acknowledgements

We are grateful for NSF grant #1119144 and our research group, EdGE at TERC, which includes Erin Bardar, Teon Edwards, Jamie Larsen, and Barbara MacEachern. Jasmine Williams assisted with coding and Ryan Baker consulted on the analyses.

References

[1] Lenhart, A., et al. (2010) *Social Media & Mobile Internet Use Among Teens and Young Adults*. Washington, DC: Pew Internet & American Life Project.

[2] Steinkuehler, C. and Duncan, S. (2008). Scientific Habits of Mind in Virtual Worlds. *Journal of Science Education and Technology* 17(6): 530-543.

[3] Cooper, S., et al. (2010). Predicting protein structures with a multiplayer online game. *Nature, 466*(7307), 756-760.

[4] National Research Council. (2011). *Learning Science Through Computer Games and Simulations. Committee on Science Learning: Computer Games, Simulations, and Education*. M. Honey & M. Hilton (Eds.). Washington, DC: National Academies Press.

[5] Thomas, D. and Brown, J. S. (2011). *A New Culture of Learning: Cultivating the Imagination for a World of Constant Change*: Printed by CreateSpace.

[6] Shute, V. and Becker, B. J. (2010). *Innovative Assessment for the 21st Century*. New York: Springer-Verlag.

[7] Fisch, S.M., Lesh, R., Motoki, E., Crespo, S., and Melfi, V. (2011). Children's mathematical reasoning in online games: Can data mining reveal strategic thinking? *Child Development Perspectives. 5*(2), 88-92.

[8] Halverson, R., Wills, N. and Owen, E. (2012). CyberSTEM: Game-Based Learning Telemetry Model for Assessment. Presentation at 8th GLS, Madison, WI, USA

[9] Isbister, K., Flanagan, M., and Hash, C. (2010). Designing games for learning: Insights from conversations with designers, *Proc. of CHI 2010*.

[10] Mislevy, R. and Haertel, G. (2006). Implications of Evidence-Centered Design for Educational Testing. *Educational Measurement: Issues and Practice, 25*(4), 6-20. Malden, MA: Blackwell Publishing.

[11] Romero, C., Ventura, S., Pechenizkiy, M., Baker, R.S.J.d. (Eds.) (2010). *Handbook of Educational Data Mining*. Boca Raton, FL: CRC Press.

[12] Polanyi, M. (1966). *The Tacit Dimension*. London: Routledge.

[13] Mauro, C. (2011, February 6). Why Angry Birds is so successful and popular [Blog]. Retrieved from http://www.mauronewmedia.com/blog/2011/02/why-angry-birds-is-so-successful-a-cognitive-teardown-of-the-user-experience/.

[14] Yu, Z. (2001-2002). Two Cultures Revisited: Michael Polanyi on the Continuity Between the Natural Sciences and the Study of Man. *Tradition and Discovery, XXVIII, No. 3*(28:3), 6-19.

[15] Clearleft Ltd. (2013) Silverback (Version 2.0) [Software]. Available from http://silverbackapp.com.