

# Leveling Up: Measuring and Leveraging Implicit STEM learning in Games

**Abstract:** Games provide an important vehicle for educators to promote and study learning. This symposium will examine research on measuring implicit game-based learning and teachers leveraging its relationship for explicit (e.g. school-based) STEM learning. The authors have developed a series of learning games that simulate authentic scientific phenomena, providing a learning mechanic for players to dwell in that phenomena and build their implicit understandings. The data logs generated through digital gameplay were mined to understand the patterns of play that may be related to implicit learning—the development of knowledge that is not yet explicitly formalized. Teachers used examples from games to help bridge implicit game-based learning to explicit STEM concepts taught in class.

## Presentation 1: Framing of Implicit Learning in Games

Jodi Asbell-Clarke, Educational Gaming Environments group at TERC  
Elizabeth Rowe, Educational Gaming Environments group at TERC

The theoretical framing that guides the research is based in a model of implicit learning, explored more commonly in psychology, philosophy, and sociology (e.g., Collins, 2010; Polanyi, 1966; Reber, 1993). Implicit knowledge is, by definition, largely unexpressed by the learner. Explicit knowledge is what educators typically attempt to measure in learning assessments. Implicit learning is considered foundational to all knowledge (Polanyi, 1966), but has not made headway in educational research because until now, it has been particularly difficult to measure. This framework differentiates between explicit knowledge, what we can *express*, from implicit knowing, what we are able to *do*. Cook and Brown (1999) argue that implicit and explicit knowledge can aid one another, claiming that a dynamic affordance of the interaction between acquisition and usage of knowledge such that knowledge and knowing (the *doing* that is associated with knowledge) are linked. A classic example used in implicit knowledge literature is learning to ride a bicycle. One does not need to formalize the physics to ride a bike, but familiarity with the sensations of riding often help students learn the physics.

Games present a rich opportunity to support and measure implicit learning (Thomas & Brown, 2011). Players are often immersed in problem-solving situations where they experiment with the mechanics to understand the rule system, using trial and error with helpful feedback and rewards for motivation and sustained engagement (NRC, 2011). Reber (1993) suggests that experimental procedures to measure implicit learning should be (a) novel to the learner, (b) complex enough to not be “cracked” easily, (c) emotionally neutral to the learner, and (d) synthetic and arbitrary. Many games fulfill these criteria quite nicely.

We argue, however, that games must be designed with attention to learning and measurement. Plass and his colleagues (2011) suggest designers must carefully identify and align the game mechanics, learning mechanics, and assessment mechanics. *Game mechanics* are what the player does in the game, *learning mechanics* are the activities through which the player learns a construct, and *assessment mechanics* are the diagnostics that provide evidence of that learning.

Our work considers game, learning, and assessment mechanics as part of the overall game design. Our games use simple game mechanics found in many popular games (e.g., get a ball to a goal without crashing into other balls; or solving puzzles to point lasers to hit a target) within a scientifically accurate simulation. By creating increasingly complex situations in which a player must grapple with the consequences of scientific laws and phenomena, we are creating an environment in which the game and learning mechanics are well aligned. Presentations 2 and 3 describe our methods for developing the in-game assessment mechanics, relying on observed play patterns rather than pre-defined metrics.

## Impulse

In Impulse, players are immersed in what is known to physicists as an n-body simulator, where all the balls have mass and obey Newton’s laws of motion (Figure 1). Players must use an impulse (a click or touch on the screen) to move their ball into the goal without crashing into ambient balls. As the levels of the game increase, more ambient balls are introduced, with varying mass.

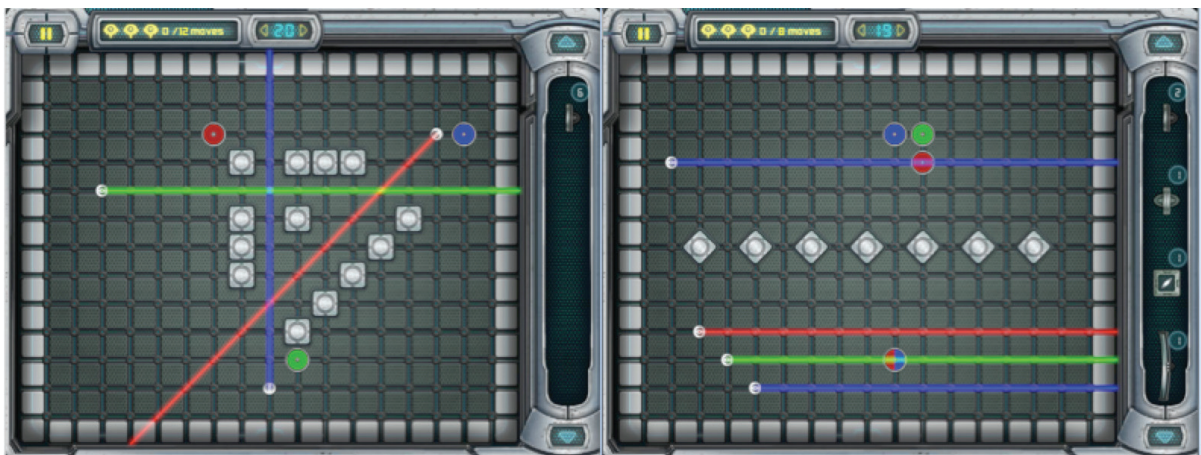


**Figure 1: A screenshot from Impulse. The player is the green particle and is going towards the cyan goal in the bottom-left corner. Red, blue, and white particles have different masses.**

As players reach higher levels with greater numbers and variety of masses of balls, they need to “study” the balls’ behavior to predict the motion of balls so that they can guide their ball to the goal, not run out of energy, and avoid collision with other balls.

### Quantum Spectre

Quantum Spectre is a puzzle-style game designed to immerse players in a simulated optics bench and improve their implicit understanding of the concepts of focal length, angle of incidence equals angle of reflection, and slope. Each level requires the player to direct one or more laser beams to targets while (potentially) avoiding obstacles (Figure 2). For each level, an inventory provides the player with access to resources, such as flat and curved (concave, convex, and double-sided) mirrors, (concave and convex) lenses, beam-splitters, and more, that can be placed and oriented within the puzzle and that interact with and direct the laser beams in a scientifically accurate manner. When the appropriate color laser beam(s) have reached all the targets, a level is complete.



**Figure 2: Two Quantum Spectre puzzles**

The player earns three “stars” if the puzzle has been solved in the fewest possible moves, two “stars” for a low number of extra moves, and one “star” for any solution. A player can go onto to the next level as soon as a puzzle is complete, regardless of the number of moves used, but the stars system provides an incentive for level replay and an understanding of the puzzle’s solution.

## Presentation 2: Strategic Moves as Measures of Implicit Learning

Elizabeth Rowe, EdGE at TERC  
Jodi Asbell-Clarke, EdGE at TERC  
Ryan Baker, Teachers College, Columbia University

There are three steps to our emergent method of building in-game measures of implicit science learning: (a) Video coding as ground truth; (b) Building automated detectors; and (c) Validating detectors with pre-post assessment data. We chose this emergent approach due to the open-ended nature of the game space (millions of paths through each level of Impulse) and not wanting to a priori select specific player behaviors as evidence of an implicit understanding without a detailed observation of how they played.

### Video Coding as Ground Truth

Strategic moves are the actions (clicks) players take within a game that have an intended outcome consistent with the goal of the game. In Impulse, the goal of the game is moving the player ball to the goal without colliding with other balls. Using three-minute segments of videos with screen capture of 69 high school students playing Impulse, we identified and reliably coded six strategic moves (Table 1). This video coding later grounds the detectors with meaningful human labels.

Strategic Move	Definition	Kappa
Float	The player ball was not acted upon for more than 1 second	0.759
Move toward goal	The learner intended to move the player ball toward the goal	0.809
Stop/slow down	The learner intended to stop or slow the motion of the player ball	0.720
Keep player path clear	The learner intended to move non-player balls to keep the path of the player ball clear	0.819
Keep goal clear	The learner intended to move non-player balls to keep goal clear	0.832
Buffer	The learner intended to create a buffer between the player and other balls to avoid collision	0.772

Source: Rowe, Baker & Asbell-Clarke (2014)

**Table 1: Strategic moves, definitions, and Cohen's Kappas**

Two of these strategic moves, Float and Stop/Slow Down, are consistent with an implicit understanding of Newton's First Law—an object will keep moving unless acted upon by a force. Float is the passive version of Newton's First Law, requiring no action on the part of the player. Stop/Slow Down requires players to actively oppose the motion of the player ball. The remaining four strategies, while useful game strategies, were not hypothesized to support implicit science learning.

Newton's Second Law—that different mass particles react differently to the same force—required examining sequences of fast moves. Besides the player ball, there were four other types of balls each with a color signifying a different mass (in order from least to most massive): blue, red, white, dark grey. The blue, red, and white balls also increased in size (consistent with the same density of ball) but the dark grey ball was most massive and smallest in size. This was to ensure that mass was being differentiated in players' behaviors rather than size.

To analyze whether students were behaving as if they understood Newton's Second Law, we coded information about the target of a click and whether or not the target of the current click was the same as the previous click (Table 2). Kappas for these codes exceeded 0.80 (Rowe et al., 2014b). These codes were combined to determine if players consistently used more force (clicks) to move the heavier balls than the lighter ones. From these codes, the number of consecutive clicks (e.g., sequence length) for each color target was calculated. We found that players impart more force for the heavier balls, even the grey balls that are much smaller in diameter, indicating that it is indeed mass, rather than size, that motivates their increase of force (Rowe, Asbell-Clarke & Baker, in press).

### Building Automated Detectors

Impulse logs every game event as well as the location of every object in the game space. Recorded game events include level starts/ends, pausing and resuming the game, as well as moves and states salient to the individual game. From this raw game log, we distill features such as the speed of the player ball and the time since the last click. These distilled features are added to the original clickstream data. Using the synchronized timestamps, these

features are then aggregated at the click level to map to the labels provided by the video coder.

Classifiers (strategic move 0=absent; 1=present) were created using J48 decision trees within RapidMiner 5.3 that mapped the player behaviors in the features distilled from the clickstream data to the human labels, cross-validating at the student level. The goal of these classifiers was to develop an automated, algorithmic way of analyzing the logs of student interaction that would come to the same judgments as a human being. All detectors for the strategic moves discussed here had cross-validated Kappas between 0.51 and 0.86 and A' between 0.78 and 0.97 (Rowe et al., 2014b).

### **Validate Detectors with Pre-Post Assessment Data**

We applied automated detectors of strategic moves to a new and larger sample of gameplay data. These data were collected as part of national implementation study of Impulse. This study compared 213 students in 21 classrooms that only played the game and 180 students in 18 classrooms where the players' teacher used game examples to bridge the implicit science learning in the game with explicit science content covered in class. Path analyses suggest the mediating role of strategic moves on students' implicit science learning is different between the two conditions (Rowe, Baker & Asbell-Clarke, 2015).

### **Presentation 3. Interaction networks to measure implicit science learning**

Michael Eagle, North Carolina State University  
Elizabeth Rowe, Educational Gaming Environments Group at TERC  
Rebecca Brown, North Carolina State University  
Andrew Hicks, North Carolina State University  
Tiffany Barnes, North Carolina State University  
Jodi Asbell-Clarke, Educational Gaming Environments Group at TERC  
Teon Edwards, Educational Gaming Environments Group at TERC

Understanding user behavior in complex problem solving tasks is important for both assessing learning and for the design of content. Problem solving is an important skill across all STEM (science, technology, engineering, and math) fields. One strong benefit of digital learning environments is the large amounts of student log-data that we can collect. This data provides measurements of problem-solving behavior at a detail that were not possible before. However, the data is not easy for developers or instructors to use in ways that inform teaching and learning.

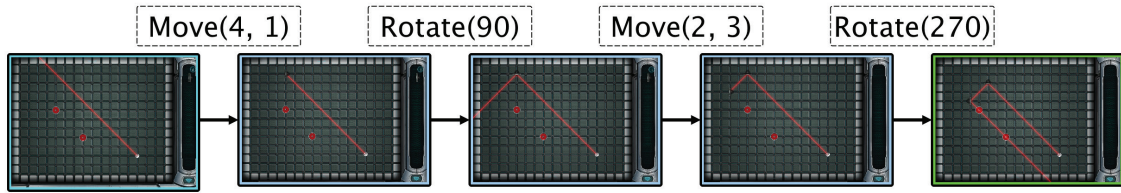
To provide insight into student problem-solving behavior in interactive systems, we have developed a complex Interaction Network (IN) representation of student-game interactions (Eagle, Brown, Rowe, Asbell-Clarke, Barnes, & Edwards, 2015). An *Interaction Network* is a complex network representation of all observed student-game interactions for a given problem. We define an *interaction* as a  $\{Start\_State, Action, End\_State\}$  tuple. A state represents the moves a student has made up until a given point. As a simplified example, consider a checkers game. At the beginning of the game, the *Start\_State* is the set of standard locations of all checker pieces. The *Actions* are all of the possible moves a player could make (e.g., one or two spaces forward). The *End\_States* are the set of checker locations after one checker piece had been moved. These *End-States* become the *Start\_States* for the next set of moves the player makes. The interaction network would be all possible combinations of *Start\_States*, *Actions*, *End\_States* from the beginning to end of a game.

These *Interaction Networks* have been used to provide instructors and game developers with visualizations of their user's problem-solving behaviors (Johnson, Eagle, & Barnes, 2013, Eagle, Johnson, Barnes, & Boyce, 2013). Clustering similar states within Interaction Networks together, we can observe differences student's high-level approaches towards solving the problem (Eagle & Barnes, 2014; Hicks, Peddycord & Barnes, 2014). The resulting clustered representation is called an *Approach Map*. Within a propositional logic tutor, this Approach Map technique was able to help demonstrate significant between-group differences in problem solving approaches for an experimental and control group (Eagle & Barnes, 2014; Hicks, Peddycord, & Barnes, 2014).



## Interaction Networks in Quantum Spectre

In this work, we apply Interaction Networks to data from Quantum Spectre to derive understanding and insight about (a) what common correct strategies students use (b) what common science (and puzzle errors students make; and (c) where in the game-play sequence students are most likely to quit. This paper reports Interaction Networks built from gameplay data of 195 high school students playing Level 18 of Quantum Spectre. Of these students, 101 (52%) were in the Bridge group and 94 (48%) were in the Games group. In Level 18, the science concepts of interest are the angle of incidence equals the angle of reflection and slope.



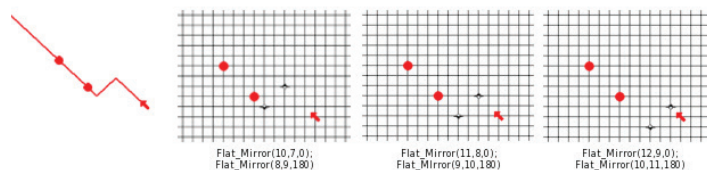
**Figure 3: Quantum Spectre, Level 18 puzzle state-to-state transitions for one solution path**

Figure 3 shows an example of a single student attempt of level 18 in Quantum Spectre, mapped into a sequence of states (screenshots) with edges (arrows) for each action (move or rotate) the student took. Each edge is labeled with the action the student player took to change the game state; for example Move (2,3) represents moving a mirror to position (2,3). To construct an Interaction Network for a problem, we take the union of all of student puzzle solution attempts and merge actions (edges) and states (vertices) that are the same according to a *matching function*. The usefulness of an Interaction Network for visual analytics is determined by the quality of the state representation, the granularity of the actions, and the matching function.

The first step in modeling gameplay data as an Interaction Network is to determine a suitable state representation and matching function. In games, our first attempt is to serialize the game state – in other words, record everything the game needs to recreate the current status of the game. In Quantum Spectre, this is a list of the game objects that players can interact with and their current position and rotation. Level 18 has 2 flat mirrors. We ignore the distinction between objects of the same type, so the order of placement does not matter.

For level 18, using this representation for 6145 student-transactions produced an Interaction Network with 916 states and 1614 edges. We applied basic filtering, removing states that occurred for only one student, to simplify the network further to 322 states and 874 edges. We first applied the Approach Map technique (Eagle & Barnes, 2014) to cluster these interactions into 18 region states (clusters of highly-connected states) and 30 edges. Since each region represents multiple states with varying types of errors (i.e., the clusters were not sorting on the types of errors making the regions difficult to interpret), we then developed an Approach Map with laser shape representation for regions. A “laser shape” representation consists of a list of the targets that are hit by a particular colored laser and a list of the angles that the laser beam takes on its path. Laser shapes are what player is trying to alter to solve the puzzle and it is the shape of the path that provides feedback about the accuracy of their placement and rotation of the mirrors, so it is not surprising laser shapes provided a more parsimonious, interpretable visualization.

We found that some game states in the puzzle were equivalent in terms of their correctness, and showed the same player proficiencies or errors. To group these equivalent states and reduce the number of states, we calculated the shape of the laser as it passed through any objects on the board. Figure 4 shows one laser shape and the three states it represents, where red circles are targets the laser beam should pass through, the red arrow is the laser source, and the black curves are mirrors. This approach preserves the relevant properties of a board state while ignoring distance traveled, which does not matter for correctness.



**Figure 4: The far left is the *LaserShape* and represents all three of the other states**

We combined the Interaction Network extraction, reduction, graph mining, and visualization to build the Approach Map shown in Figure 5. The regions have been outlined according to the correctness of states they represent, with correct states contributing green to the edge and region outline colors, and incorrect states contributing orange. Regions with full orange outlines represent incorrect solutions, and those with green outlines are correct. Blue-outlined regions have a combination of correct and incorrect, or not yet complete, actions. We have grouped what we call a “confusion region” with a dashed line, to illustrate the various incorrect attempts students make.

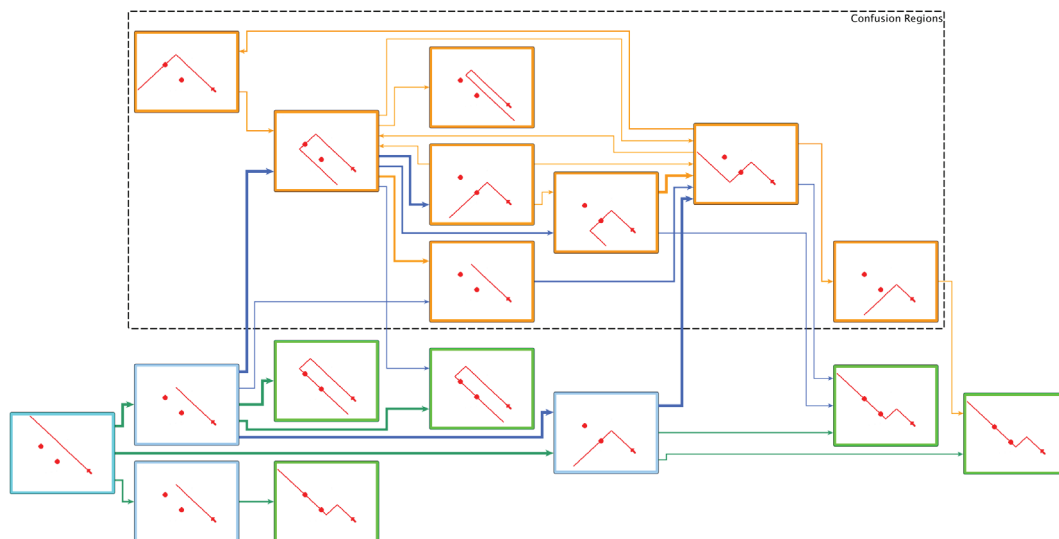


Figure 5: The Approach Map for Level 18.

These representations enable game developers and learning scientists to better understand the broad patterns in the behaviors of players solving the puzzle. Players who start and stay in the confusion region seem to be able to solve the puzzle for one target but not for two. Instructors and developers who want to explore the individual regions can “zoom in” on the internal states of each cluster region. Learning scientists will label these internal states for their evidence of a lack of science understanding. Labeling clusters in using these visualizations saves a large amount of coding time and makes reliable coding easier to achieve. In our future work, we plan to apply these data-driven visualizations and graph mining techniques across several levels of Quantum Spectre, and look for ways to provide summaries across problems and look for differences in learning between groups.

#### Presentation 4. Teachers Bridging Implicit to Explicit Learning

Elizabeth Rowe, Educational Gaming Environments Group at TERC  
 Erin Bardar, Educational Gaming Environments Group at TERC  
 Jodi Asbell-Clarke, Educational Gaming Environments Group at TERC

The connections that people make socially and cognitively in other web spaces (game affinity sites) and in person (e.g., in a classroom) around a game are a large part of game-based learning. Jim Gee (2008) refers to this as the Big “G” Game. Teaching methods that leverage game-based implicit learning must provide tools for teachers to “see” the learning and respond to it. The teachers may not be as likely as their students to be playing the newest game in the app store, but teachers are embracing game-based education and would like more opportunities to use games as vehicles for learning (Cooney, 2012).

We conducted a national implementation study with three groups of high school learners:

- a) the **Games** group whose teachers encourage students to play Impulse outside of class;
- b) the **Bridge** group whose teachers encourage students to play Impulse outside of class and use examples from the game (bridge activities) when they teach related content in class;
- c) the **Control** group that does not play the game or use bridge activities.

As reported previously at GLS, this study has shown significant STEM learning gains in Bridge classes, with the largest effect among students in non-honor/AP classes (Rowe et al., 2014a). When a teacher uses game examples to help bridge the game-based implicit learning to STEM content discussed explicitly in the classroom, students have higher gains on pre/post tests about related science content than students in the Control classes. To

unpack these findings further, we have been analyzing logs of teacher activity modeled after the SCOOP Notebook developed by CRESST (Borko, Stecher & Kuffner, 2007). The coding system developed for this study focuses on the *type of classroom activities* (e.g., direct instruction, hands-on activities, etc.); the *science content taught* (e.g., Newton's Laws, forces, etc.), and *game-based pedagogies* used (e.g., using game examples during instruction, modeling game play, discussing the game, etc.). To confirm the reliability of the coding system, 10 of the 50 teacher logs are were double-coded by our independent evaluators with an average Cohen's Kappa of 0.71.

Two types of analyses using these logs are being conducted. The first set of analyses compares student demographics, teacher background, and science pedagogy and content covered in Honors/AP and non-Honors/AP across the Control, Game, and Bridge groups. The second set of analyses describes the game-based pedagogies used in the 18 Bridge classrooms. The game-based pedagogies (e.g., frequency with which examples were used, the amount of teacher or student modeling of game play) in 6 Honors/AP classrooms will be compared to the pedagogies used in the 12 non-Honors/AP classrooms.

Since submission, we found few group differences in student demographics, teacher background, and science pedagogy/content coverage that might explain these findings (Rowe, Bardar, Asbell-Clarke, Shane-Simpson, Roberts, in press). There were significant differences, however, between Honors/AP and non-Honors/AP classes in their use of specific game-based pedagogies.

## References

- Borko, H., Stecher, B., & Kuffner, K. (2007). Using Artifacts to Characterize Reform-Oriented Instruction: The Scoop Notebook and Rating Guide. CSE Technical Report 707. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.
- Collins, H. (2010). *Tacit and explicit knowledge*: University of Chicago Press.
- Cook, S. D., & Brown, J. S. (1999). Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing. *Organization science*, 10(4), 381-400.
- Cooney Center (2012). *Teacher Attitudes about Digital Games in the Classroom*. New York: Joan Ganz Cooney Center.
- diSessa, Andrea A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, 10(2/3), 105-225. doi: 10.2307/3233725
- Eagle, M., Brown, R., Rowe, E., Asbell-Clarke, J., Barnes, T., & Edwards, T. (2015, June). Exploring Problem Solving Behaviors in an Optics Game. Poster to be presented at the Eighth international conference on Educational Data Mining Society in Madrid, June 25-30.
- Eagle, M., Hicks, A., Peddycord, B., & Barnes, T. (2015, April). Exploring networks of problem-solving interactions. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge (LAK '15)*, Poughkeepsie, NY, USA, pp. 21-30.
- Eagle, M., & Barnes, T. (2014). [Exploring differences in problem solving with data-driven approach maps](#). In proceedings of *Educational Data Mining (EDM2014)*, London, UK, pp. 76-83.
- Eagle, M., Johnson, M., Barnes, T., & Boyce, A. (2013, April). [Exploring Player Behavior with Visual Analytics](#). In proceedings of *Foundations of Digital Games 2013*: pp. 380-383.
- Gee, J. (2008). Learning and Games. In K. Salen (Ed). *The Ecology of Games: Connecting Youth, Games, and Learning*. Cambridge, MA: The MIT Press. 21-40.
- Hicks, A., Peddycord, B., & Barnes, T. (2014). Building games to learn from their players: Generating hints in a serious game. In proceedings of *Intelligent Tutoring Systems (ITS2014)*: 312-317.
- Johnson, M., Eagle, M., & Barnes, T. (2013). [InVis: An Interactive Visualization Tool for Exploring Interaction Networks](#) In proceedings of *Educational Data Mining (EDM 2013)*: pp. 82-89.
- McCloskey, M. (1983). Intuitive Physics. *Scientific American*, 248(4), 122-130.
- 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.). Washing-

ton, DC: National Academies Press.

- Plass, J., Homer, B.D., Kinzer, C.K., Chang, Y.K., Frye, J., Kaczetow, W., Isbister, K., Perlin, K. (2013). Metrics in Simulations and Games for Learning. In M. Seif El-Nasr, Drachen, A., & Canossa, A. (Eds.), *Game Analytics: Maximizing the Value of Player Data* (pp. 694-730). London: Springer-Verlag.
- Polanyi, M. (1966). *The Tacit Dimension*. London: Routledge. (University of Chicago Press. ISBN 978-0-226-67298-4. 2009 reprint).
- Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford, UK: Oxford University Press.
- Rowe, E., Asbell-Clarke, J. & Baker, R. (in press). Serious game analytics to measure implicit science learning. In C.S. Loh, Y. Sheng, & D. Ifenthaler *Serious Game Analytics: Methodologies for Performance Measurement, Assessment, and Improvement*. To be published by Springer Science+Business Media in Summer 2015.
- Rowe, E., Bardar, E., Asbell-Clarke, J., Shane-Simpson, C., & Roberts, S. (under review). Building Bridges: Teachers Leveraging Game-Based Implicit Science Learning in Physics Classrooms. To appear in D. Russell & J. Laffey (Eds). *Handbook of Research on Gaming Trends in P-12 Education*, to be published by IGI-Global in Fall 2015.
- Rowe, E., Baker, R., & Asbell-Clarke, J. (2015, June). Strategic game moves mediate implicit science learning. Paper to be presented at the Eighth international conference on Educational Data Mining Society in Madrid, June 25-30.
- Rowe, E., Asbell-Clarke, J., Bardar, E., Kasman, E., & MacEachern, B. (2014). Crossing the bridge: Connecting game-based implicit science learning to the classroom. Paper to be presented at the 10<sup>th</sup> annual Games+Learning+Society conference in Madison, WI, June 11-13.
- Rowe, E., Baker, R., & Asbell-Clarke, J. (2014, July). Building automated detectors of gameplay strategies to measure implicit science learning. Poster presented at the Seventh international conference on Educational Data Mining Society in London, July 4-7.
- Sao Pedro, M.A., Baker, R.S.J.d., Gobert, J., Montalvo, O. Nakama, A. (2013). Leveraging Machine-Learned Detectors of Systematic Inquiry Behavior to Estimate and Predict Transfer of Inquiry Skill. *User Modeling and User-Adapted Interaction*, 23 (1), 1-39
- Thomas, D. and Brown, J. S. (2011). *A New Culture of Learning: Cultivating the Imagination for a World of Constant Change*. Lexington, KY: CreateSpace.

## Acknowledgments

This research was funded as part of a DRK12 grant from NSF (#1119144) to develop and study the *Leveling Up* games. We gratefully acknowledge the rest of the EdGE team: Jamie Larsen, Barbara MacEachern, and Katie Stokinger. Our evaluators, the New Knowledge Organization, helped establish the reliability of video and teacher log coding.