Learning Analytics for Games

V. Elizabeth Owen[1] & Ryan S. Baker[2]

Abstract

Computer-based educational games can provide engaging designed experiences for learning (Squire, 2006; Gee, 2003), affording rich performance data situated in a meaningful learning context (Mislevy, 2011; Clark et al., 2012). This kind of big data in education (cf. U.S. DoE, 2012) has fostered emergent fields like educational data mining (Baker & Yacef, 2009) and learning analytics (Siemens & Long, 2011). In the design of these game environments, there is increasing evidence that players rarely interact in exactly the way designers envision, highlighting the need for early, iterative user-testing (Schell, 2008; Salen & Zimmerman, 2004). Adding the element of content-specific learning goals, or concrete growth over time in a domain-specific skill, attending and adjusting to organic play patterns becomes even more vital (cf. Shute, 2011; Institute of Play, 2013). Thus, educational game design needs to leverage learning-specific assessment mechanisms and sophisticated techniques to understand nuanced learner patterns in play—informing development from the earliest design stages. Recent work in learning analytics and educational data mining (Baker & Siemens, 2014) identify a host of methods for analysis of log file data from these systems, thus enabling insights into event-stream

---

[1] V. E. Owen
University of Wisconsin-Madison
v.elizabeth.owen@gmail.com

[2] R. S. Baker
University of Pennsylvania, Graduate School of Education
rybaker@upenn.edu

playful learning and data-driven design. This chapter presents an overview of learning analytics and educational data mining (LA/EDM) for the analysis of nuanced game data for learning insights and potential data-driven design. Specifically, it aligns event-stream assessment and applied LA/EDM with three development stages—Alpha (inception), Beta (mid-development), and late Beta/final release—and reviews recent applications of these methods in game-based empirical research.

Introduction: Learning Analytics and Educational Game Applications

Learning Analytics (LA) and Educational Data Mining (EDM) represent a host of education-specific methods for exploring and mining big data (U.S. DoE, 2012), which can be used to enhance learning design and learning outcomes. In recent literature, EDM and LA have been discussed together as a converging set of methods for interpreting large streams of data from educational contexts (Baker & Siemens, 2014); while there are differences between the research questions these two communities ask, for the purposes of this chapter they can be treated as interchangeable. (For brevity in subsequent sections of this chapter, therefore, we will refer to the collective set of methods as Learning Analytics or LA.)

EDM and LA have drawn from methods originally developed in a range of communities, from data mining and analytics in general, and from psychometrics and educational measurement (Baker & Siemens, 2014), as well as increasingly producing methods unique to these research communities. The methods used in these communities can be divided into five major categories: prediction, structure discovery, relationship mining, discovery with models, and visualization. Prediction modeling infers an outcome or measure of interest (i.e. a predicted variable) when given input data (i.e. predictor variables) via a range of potential algorithms. In contrast, structure discovery “attempts to find structure in the data without an a priori idea of what should be found," using methods like clustering, factor analysis, and network analysis (p. 258). Relationship mining is used to discover relationships between variables in a large data set, leveraging approaches like correlation mining, association rules, and sequential pattern mining. Discovery with models involves layering methods, often utilizing the results of one data mining analysis within another data mining analysis to optimize insights. Finally, visualization is designed to express data visually to elucidate patterns (e.g. color-coded heat maps and graphics of trajectories over time with learning curves). Explored more deeply in the next section, these five categories provide important insight into game-based learning, and map to specific development phases for optimal data-driven insights throughout the design process.

In the context of serious games, a substantial base of recent empirical research has utilized many of these learning analytics methods—particularly visualization, structure discovery, relationship mining, and prediction. As learning designers increasingly attend to event-stream data to inform iterative design (e.g. Kerr, 2015), these methods can be mapped to various stages of game development to support data-driven design for learning and engagement. In early-development Alpha stages, when game design may be in nascent stages, implementing basic data collection and using visualization can help surface basic player interactions for improved core mechanics, UI/UX and learning design. Structure discovery and relationship mining can uncover deeper player patterns as mechanics are solidified in Beta phases; for example, results can help isolate points of attrition or bottlenecks in the game, identifying larger patterns of player navigation or strategy across game levels. Finally, late Beta/final release analyses utilizing prediction can identify key predictors of target behaviors (e.g. game success, strategy, or engagement) to support final polish, as well as provide insight that enables player-adaptive personalized paths through the learning space.

These learning analytics methods are discussed in greater depth below, setting a foundation for a review of current research in game-based learning analytics with implications for data-driven design. The following pages discuss overall methods of LA/EDM, potential alignment with learning game development stages, and review applications of these analysis methods in recent game-based learning research.

Overview of Learning Analytics / Educational Data Mining

Baker and Siemens (2014) divide learning analytics into a set of five main categories, building off of an earlier review by Baker and Yacef (2009).

The first of these five categories is prediction. In prediction modeling, the researcher's goal is to create a model which can make inferences about a single variable, the predicted variable, from some combination of other variables, the predictor variables. The predicted variable may be a variable that can be easily collected for a small sample of data, but cannot be collected at larger scale. Alternatively, it may be some future outcome that is desirable to predict before it comes to pass, for example to drive early intervention. Either way, a model is created based on this sample of data, validated to give confidence that it will function correctly on new data, and is then applied to new data. Three types of prediction are common in LA/EDM: classification, where a binary variable or multi-category variable is predicted; regression, where a number is predicted; and latent knowledge estimation, where student knowledge is assessed (typically as a probability between 0 and 1, typically based on correctness data that is itself binary).

The second of the categories in Baker and Siemens (2014) is structure discovery, algorithms that attempt to discover structure in data with no specific variable as a focus. Within LA/EDM for game-based learning, the categories of cluster analysis, network analysis, and domain structure discovery are particularly prominent. In cluster analysis, the researcher attempts to use automated processes to discover which data points group naturally together, dividing the data set into groups of data points referred to as clusters. Cluster analysis is of particular value when the categories of interest among a data set are not known, a priori. In domain structure discovery, the structure of content is discovered automatically. For example, in a set of items, problems, or tasks, it may be possible to determine which problems involve some of the same content (perhaps skills, concepts, or strategies), such that performing well on one problem implies performing well on the other problem. It is possible, in such a framework, to search for partial overlap of content—situations where problems A and B share skill Alpha, but problem B also shares skill Beta with problem C. It is also possible to find prerequisite relationships, where successful performance on problem A implies successful performance on problem B, but not vice-versa. In network analysis, more complex networks of relationships between data points are investigated. For example, the paths a player might take through a specific puzzle might be turned into a graph and then subjected to network analysis to predict the best positive move a player might take next. A fourth type of structure discovery, common in other areas of LA/EDM but less common in game-based learning, is factor analysis, where the relationship between variables is analyzed in order to determine which variables can be combined into a smaller number of latent factors. (Factor analysis is sometimes used to analyze test data for domain structure discovery, but is not frequently used in more complex game-based contexts where students are learning as well as demonstrating their skill).

The third of the categories in Baker and Siemens (2014, p. 260) is relationship mining. Referred to in that review as the “most common category of EDM research,” relationship mining is generally common in LA/EDM research on game-based learning as well. There are four broad categories of relationship mining, each of which has been conducted in the context of educational games. In the first, association rule mining, the software automatically finds if-then relationships where if a specific variable/value pair (or set of pairs) is seen, another specific variable/value pair usually accompanies it. In the second, sequential pattern mining, association rules are found, with the additional criteria that the "then" part of the rule must occur after the "if" part of the rule. In correlation mining, a large number of variables are checked for correlation relationships between them, with post-hoc statistical controls used to reduce the probability of finding spurious findings. Finally, in causal data mining (a method whose conclusiveness remains under debate), patterns of covariance are used to determine if one event in a sequence of events is statistically likely to be the "cause" of a second, later event.

The fourth of the categories in Baker and Siemens (2014) is discovery with models. Within discovery with models, a variable or set of variables are created through LA/EDM—using prediction modeling or clustering, for instance—and then used in a second analysis. For example, building a model of student disengagement for a game or simulation, and then studying how that variable correlates to eventual student success in the game, would be an example of discovery with models.

The fifth of the categories in Baker and Siemens (2014, p. 260) is visualization, referred to in that paper as "distillation of data for human judgment". Visualizations of data can elucidate patterns in a way that is easily visually processed, at best representing a high-dimension data in a simple, digestible presentation (Tufte, 2001). In the context of learning analytics, these can take the form of descriptive statistical charts, simple learning curves, heat maps, and radial visualizations. These have been used in LA/EDM for games, often in conjunction with other method categories discussed above; several examples are given in the section to follow, which discusses specific application of these methods to serious games for learning insights and data-driven design. Methods categories commonly used in recent research are highlighted below, including visualization, structure discovery, relationship mining, and prediction.

Learning Analytics for Serious Games: Recent Application of Methods

Learning analytics methods—especially visualization, structure discovery, relationship mining, and prediction—can support deep insight into playful learning patterns, as well as enhance design iteration for optimal learning and engagement when applied during various stages of game development. These investigations can be mapped to phases of design and game production in sync with pre-existing game refinement cycles to fuel data-driven, iterative design for engaged learning. In early stages of development (i.e. the Alpha phase), data framework definition and visualization analytics can be valuable in supporting formative design; structure discovery and relationship mining can uncover deeper player patterns as mechanics are solidified in Beta phases; and predictive modeling can support final game production in providing prediction of learning and behavior detection for in-game adaptivity in support of learning pathways. It's worth noting that, like best practices of Agile game development[3], this alignment is flexible, with potential for application of analysis methods extended into multiple stages of development to support design as needed. In doing so, these event-stream analyses can complement ongoing qualitative research (e.g. observations, think-alouds, and interviews) in informing iterative improvement. The following section reviews recent research using these types of learning analytics methods in serious games to investigate student play patterns, with discussion of the insights provided into game-based learning and potential implications for data-driven design.

[3] https://en.wikipedia.org/wiki/Agile_software_development

Early Development: Learning Data Collection & Visualization

In early phases of game development, in which core mechanics and basic design may still be in nascent stages, data framework design and visualization of basic user interactions can help game designers understand how players initially are approaching the game and support formative thinking about learning, game, and assessment mechanics. This can be particularly beneficial in conjunction with qualitative user testing (e.g. observations, think-alouds, and interviews) to support a well-rounded understanding of initial playful learning experiences, informing effective iterative design. This section discusses the benefits of a strong learning data foundation in early game design, as well as applications of learning analytics visualization in serious games.

Learning game data frameworks. Event-stream data collection in serious games is an important undertaking, and foundational to analyses that provide actionable insight. Any process of making meaning out of data, whether involving thorough feature engineering or more bottom-up processes, is dependent on the integrity, quality, and scope of the original data. Recent efforts in structuring learning game data delineate the need for comprehensive, clearly organized, and design-aligned data collection (cf. Chung, 2015; Danielak, 2014; Serrano-Laguna et al., 2017; Hao et al., 2016). ADAGE (Assessment Data Aggregator for Game Environments) provides one approach tailored to serious games, an event-stream data framework designed specifically to support embedded assessment and educational data mining (Halverson & Owen, 2014). ADAGE collects comprehensive game events and player interactions enriched with contextual data, while providing salient performance data aligned with key learning mechanics. This kind of comprehensive data allows for multiple methods of analysis in game-based learning investigations. Clear, design-aligned data output provides clear reference to the game's design of learning mechanics; when data is interpretable in this fashion, outcomes of analysis can be more easily translated to direct feedback into design. A consistently labeled series of event-stream interactions supports aggregation of data for analysis and feature engineering—a critical element of robust modeling in many approaches to learning analytics (e.g. Guyon & Elisseeff, 2003; Sao Pedro et al., 2012). This applies for analytics in a single game, and is also vital for scalable analysis and adaptivity across a system of multiple games that interplay to jointly support student learning.

Finally, early implementation of a strong data collection framework can support good learning design practices in clearly aligning data-producing game mechanics with targeted learning objectives. A well-designed game will have game events that can be interpreted directly in terms of the types of competencies and learning that the designer wants to measure (e.g. Shute & Kim, 2014). Consideration of this alignment during early design stages can support good learning design and more robust event-stream data for analysis.

Data visualization. These comprehensive, design-aligned data structures in early design enable analysis for data-driven design in the alpha game development phases. In particular, visualizations and descriptive statistics can support early game development in surfacing basic player interaction with the game (e.g. identification of bugs, bottlenecks, and core mechanic interaction) for improved UI/UX and learning design. Visualization methods can consistently support subsequent stages of game development as well.

Learning analytics for serious games, as a growing field, has set a foundation in data visualization for game analysis, including capturing movement within the game space (UI and game map), interaction with core learning mechanics at different stages of the game, and even aiding capture of biometrics and metacognitive student behavior. Towards this end, Wallner and Kriglstein (2015) detail a taxonomy of visualization types for comparative analysis of serious game data based on juxtaposition (e.g. comparing two player groups side by side), superposition (stacking visualization of each group), explicit encoding (visually encoding differences between data sets); particularly using star plots, network diagrams / graph analysis, heat maps and color overlays.

Indeed, such visualizations have supported analysis of player movement within the gamespace in related research. For example, Kim and colleagues (2008) developed a game analysis method which combined player survey pop-ups and heat mapping, which allowed identification of game areas of frustration and high failure. This tool was used to fix areas in a real-time strategy game with abnormally high rates of player death; the authors found that the modifications increased both player performance and player engagement. Similarly, Games+Learning+Society (GLS)[4] research has used visualization to improve the early design of serious games, creating heat maps of main game level usage in order to intuit areas of high traffic for optimal placement of critical player resources and to iterate on map design (e.g. Ramirez, 2016). Data visualization of player navigation through game levels can also be utilized for early game development, especially network diagrams or "state space diagrams" that show paths through a network of game states. In a study of interaction with a level selection menu, network diagram visualizations isolated game maps with low traffic, and subsequently informed improvements of UI design in the early development stages to support higher usage (e.g. Beall et al., 2013). Similarly, network diagrams were used as part of a suite of visualizations to understand play in a fractions game (Butler & Banerjee, 2014) utilizing a node-edge visualization along with heat maps of game tool use to compare progress between players in the same level. In the physics puzzle game *Quantum Spectre*[5], descriptive statistics of player

[4] A learning game development and research group at UW-Madison; http://www.gameslearningsociety.org/
[5] https://terctalks.wordpress.com/tag/quantum-spectre-game/

interactions (including game error types and number of moves in a level) were employed, along with a state space diagram, to better understand player dropout and improve design (Hicks et al., 2016). *SimCityEDU*[6] was also studied using state space diagrams to show archetypal student paths through the simulation space (Institute of Play, 2013). Additional visualization of progress in non-linear learning games builds on this idea to visualize different possible states of play. Aghababyan, Symanzic, and Martin (2013) move beyond basic nodes and edges to customize a tree visualization incorporating timeline, progress along visually fixed markers specific to each game level, and the student win state. These visualizations, which tend to focus on user interaction one game level at a time, have utility in early stages of development, and in subsequent stages, to inform iterative design about where and how players struggle, and how they can be scaffolded in reaching successful performance.

In related analysis, other methods of visualization have been utilized to show student progress (often related to performance) across multiple stages within a game. Using a radial sunburst style visualization, for example, Cooper and team (Cooper et al., 2010) showed different player strategies across multiple levels of the science game *Foldit*[7], designed to enable player production of accurate protein structure models. Dimensions of the radial visualization included time elapse, summative puzzle performance, and tool usage during different slices of play—information valuable to iterative design targeted towards supporting multiple play pathways to success. GLS researchers have used similar descriptive statistics (paired with discourse analysis) to investigate multi-modal data streams for game-based learning during multi-day play workshops for a middle-school biology game (Anderson et al., 2016). Results suggested that students initially looked up more key words in the in-game almanac and tapered

[6] https://www.glasslabgames.org/games/SC
[7] https://fold.it/

off this behavior towards the end of play, transitioning from seeing these words in a glossary to adopting these biology terms in social discourse over time. Other visualizations summarizing progress across learning game levels have been used as student-facing communication to encourage future success. As part of an intervention to support growth mindset in players of a fractions game[8], a summary screen of progress for students (given at key points in play), paired with reward points (O'Rourke et al., 2016), resulted in greater student retention and persistence. Other player-facing progress visualizations across game levels includes work in commercial games like *Civilization* (a game used for learning in classroom contexts in recent research—e.g. Squire, 2011). *Civilization V*[9], for example, has persistent player progress visualizations in the form of network diagrams (for technology researched) and simple totals of vital game resources (e.g. gold, science points, and cultural strength). Visualizations across game levels, including those which are user-facing, have strong potential to inform game progression design and support desirable player behavior.

Some game data visualizations sweep further, aiming to provide data visualizations across games. In analyzing differences between populations across two games used by different populations, a state space visualization and descriptive stats were used to elucidate game level interaction for juxtaposition of groups (O'Rourke et al., 2013), ultimately showing that younger users were interacting with the game in a less focused way (thus limiting success) in comparison with older players. Generalizable game visualization tools have also been created—including *Playtracer*, which was built to visually analyze play traces, creating a generalized heatmap that applies to any game with discrete state spaces (Andersen et al., 2010). Although not applicable to all genres, and difficult to scale with highly complex games with many possible actions, it can

[8] http://centerforgamescience.org/blog/portfolio/refraction/
[9] http://www.civilization5.com/

show progression in a similar visualization across games for an accessible comparison of play. This potentially supports the development of player profiles and allows insight around common states of interaction. With a similar goal in affording clear comparison of play, Scarlatos and Scarlatos (2010) built a cross-game tool that visualizes play progress as a glyph, the shape of which (standardized across games) can be interpreted to determine desirable progression or failure. Such generalizable tools have limits, since they essentially equate win states across games, even though they may not actually be comparable in terms of difficulty or rigor; they also may not apply across genres or platforms. However, for assessing student play style across games, especially in large systems that contain multiple games designed to work together, these analytics can have value for informing iterative design and player profile formation. In early stages they may also support understanding where attrition points occur in gameplay for iterative design improvement.

Building on basic game interactions, visualization and descriptive statistics can also be used to illuminate player patterns orthogonal to the click-by-click logfiles, such as biometric trends. Eyetracking is a capability that commercial games are increasingly developing—particularly in games with a camera built into the platform device interface (e.g. the PC game *Rise of the Tomb Raider*[10]). Leveraging this potentially powerful data source, Kiili, Ketamo, and Kickmeier-Rust (2014) evaluated serious game eye-tracking data using statistical analysis and heat maps, which revealed that low performers directed too much attention to areas of little relevance compared to high performers. Other forms of biometrics, applicable at the intersection of neuroscience and game-based learning behavior (e.g. Beall et al., 2013), have looked to visualizations of brain activity for insights about learning. One such study (Baker et al., 2015)

[10] https://store.na.square-enix.com/product/294275/rise-of-the-tomb-raider-pc-download

took cortical measurements of brain activity during play of a fractions game, with heat map results revealing similar brain activity to that which results from traditional mathematical activities in the same domain. Used in conjunction with user testing and event-stream data analysis, these biometric visualizations can support early testing of cognitive engagement and iterative design choices to optimize user attention.

Inferences about player behavior and affect can also be made in conjunction with play, which can begin to be explored through visualization and descriptive statistics—for example, through distilling event-stream data into snapshots of play in the form of *text replays* (Baker, Corbett, & Wagner, 2006), which are designed to support human evaluation of player behaviors (e.g. Owen, 2014). Descriptive statistics have also been useful for representing coded instances of strategic behavior in games (cf. Berland & Lee, 2011; Steinkuehler & Duncan, 2008), elucidating favorable and unfavorable student patterns useful for consideration in early design and beyond. These kinds of descriptive visualizations can also set a foundation for building behavior models in relationship with play data for more complex analyses in later stages.

Beta Development: Structure Discovery and Relationship Mining

In more advanced phases of game development (i.e. Beta design stages) the LA methods categories of structure discovery and relationship mining can be used to understand player decisions on a deeper level—with capability to identify sequence and attrition points, as well as interaction patterns related to engagement, success, and strategy. These analytics offer opportunity to refine game design to support successful student trajectories based on organic play patterns (rather than relying on "ideal" pathways defined a priori), and can continue to offer insight throughout the final stages of design.

Recent research has utilized structure discovery methods such as cluster analysis with large amounts of event-stream game data to surface strategies and interactions related to game success. Kerr and Chung (2012) explored clustering techniques in the elementary math game *Save Patch* in order to capture the kinds of strategies used by students. Building on this work, in which fuzzy clustering was most useful, the game design was revised to minimize the ability to pass a level using incorrect mathematical strategies. Empirical testing of the new game version revealed the changes resulted in more correct strategies in fractions problem-solving used to pass levels of the game, with more positive student reception of the updated version (Kerr, 2015). In an analysis of another game in the same domain of fractions, hierarchical clustering was used to group player strategy; this analysis demonstrated that exploration of splitting (i.e. partitioning a whole into equal-sized parts) mechanics in-game significantly improved students' fraction understanding, and that splitting strategy improved from early to late gameplay (Martin et al., 2015). Other game-based analysis work has applied methods like latent class analysis (LCA) to derive emergent student groups for play profiles; in a recent study of the learning game *Physics Playground*, LCA results derived emergent player trajectories indicative of student play styles, including achievers, explorers, and disengaged players (Slater et al., 2017). Other research into the psychology of play has also mined the structural relationships within play profile attributes in an online game, using factor analysis to distill ten motivations for play grouped into achievement, social, and immersion components (Yee, 2007). In related structure discovery with game data in the *GlassLab* game *Mars Generation One*[11] (designed to build argumentation skill in middle school players), factor analysis was used to distill survey-based game measures of

[11] https://www.glasslabgames.org/games/AA-1

engagement and self-efficacy, which was then aligned with event-stream data in predicting self-reported learning (Owen et al., 2015).

Relationship mining for game analysis has been used to discover associations between play variables. Recent research has explored association patterns between player profile attributes and in-game data, finding evidence that player types and psychological attributes provides key insight into play behaviors (e.g. Canossa et al., 2015; Yee et al., 2011). Also exploring associations between game data and out-of-game behavior, Andres and team (2014) found that affect (specifically the state of being confused) is negatively related to high in-game achievement and efficiency in physics problem-solving.

Sequence mining has also been a particularly popular method, as play data can offer a rich and varied trajectory of sequential player decisions, particularly for non-linear games. Exploration of n-grams (i.e., sequences of play behavior, in the context of serious games), for example, have supported adaptive level progression tailored to the player's history of in-game behavior (e.g. Butler et al., 2015). In a serious math game for elementary school students, n-gram analysis was utilized for mining the most frequent sequential play patterns (Aghababyan et al., 2016) as an extension of understanding strategic play trajectories in a serious game (Martin et al., 2015). N-gram analysis has also been paired with other methods for increased insight into play. Owen (2014) pairs bi-gram and tri-gram counts of in-game activity with correlation mining, showing that specific productive failure trajectories are significantly associated with learning gains in a middle-school biology game. N-gram analysis has also been used in combination with logistic regression in the study of a role playing game (RPG), to show trajectories of play that differentiate high expertise players from those with lower expertise (Chen et al., 2015). Moving into probabilistic modeling, Markov models have been used to show the probability of a player transitioning from one state to another in gameplay (e.g. the likelihood of moving from one game level to the next, or to oscillate between states of success and failure). In the context of a middle-school science game, for instance, a first-order Markov model was employed to determine the stages of play in which students are most likely to quit (Owen, Shapiro, & Halverson, 2013). Hidden Markov modeling (HMM) has been used to explore latent states of student understanding during play across multiple game platforms—including computer games (e.g. Clark et al., 2012) and digitally interactive tabletop games (e.g. Tissenbaum, Berland, & Kumar, 2016). Tissenbaum and colleagues (2016) mined the sequence of player circuit-forming as unproductive or productive with an HMM, identifying productive learning trajectories of students who had started in unproductive states and moved to success (2016), within the context of a game-based museum exhibit. Overall, structure discovery and relationship mining can thus support understanding of play trajectories connected with positive game performance and learning outcomes; while these are valuable insights for understanding student behavior on their own, they can also inform iterative design to support such trajectories with adaptive leveling or enhanced scaffolding at key points in the game.

Late Beta and Final Release: Predictive Learning Analytics

In final stages of game development, including late Beta and final release, learning analytics can be used to predict in-game actions and performance most characteristic of learning. Predictive modeling can reveal a great deal about student growth during play, and mine key predictors of behavior from the game data event stream—especially in combination with ongoing insights from previous-stage analytics, including visualization, structure discovery, and relationship mining. These investigations have the potential to support field-enriching inferences about learning and behavior, as well as fuel data-driven design through real-time detection of students' pathways to inform adaptive, personalized game progression.

Various methods of prediction have been used in analyzing serious game data, from canonical statistical models (e.g. linear regression and HLM; cf. Marascuilo & Serlin, 1988) to data mining algorithms for classification and regression (Baker, 2010). Utilizing different prediction models to investigate strategy use in a real-time strategy (RTS) game, Weber and Mateas (2009) evaluated various algorithms (including linear regression, additive logistic regression, J48 classification, and M5' regression); overall, it was found that M5' overall had the smallest relative error in predicting timed player construction of key game resources. Prediction has also been leveraged in the form of HLM for evaluation of collaboration and competition in games, with recent research showing that competition increased in-game math learning compared to individual play, and both collaboration and competition elicited greater situational interest and enjoyment (Plass et al., 2013). In another math game, researchers used predictive modeling with logistic regression to show that different kinds of fraction errors are predictive of learning outcomes (Kerr & Chung, 2013)—implying that in-game scaffolding design should not treat all errors equally. In further predictive modeling, survival analysis was used to investigate the game *Quantum Spectre*, specifically pinpointing conditions of play that influenced player dropout with an accelerated failure time model (Hicks et al., 2016). Prediction has also been used to support adaptive game play, as seen in the use of reinforcement learning to predict optimal player scaffolding through narrative in the learning game *Crystal Island* (Rowe & Lester, 2015). Similarly, adaptive learning design has been explored using decision trees in game-like e-learning environments, using prediction to prescribe customized learning paths through the system (e.g. Lin et al., 2013).

Recent research in the application of LA/EDM to learning games utilizes predictive data mining to build event-stream detectors of behavior, a method first applied in the context of intelligent tutoring systems (e.g. Baker, Corbett, & Koedinger, 2004). With the increasing availability of log file data in digital learning games, event-stream detectors have been leveraged to more deeply understand and predict player behavior. In the context of a physics game, for example, detectors of affective states and off-task behaviors were built based on video logs as well as event-stream data to build a predictor of behavior and affect throughout play (Kai et al., 2015). Results showed distinct event-stream behaviors indicative of each state (e.g. boredom's predictors included number of items "lost" or moved off screen during play, and amount of time elapsed between actions). The video-based detectors were more accurate than the interaction-based detectors, but could not be used in many situations (due to occlusion of the face, for example, a joint detector using both types of data was more effective than either type alone (Bosch et al., 2015)). Also focusing on players' approaches to games, other research has created game-based detectors on behaviors related to goals and strategy. DiCerbo and Kidwai (2013) built a detector of whether players were serious about completing a game’s quests, with implications for enabling design support of players to complete game objectives. Productive failure and boundary-testing have also been modeled in recent studies, with a detector of thoughtful exploration built for a middle school biology game (Owen, Anton, & Baker, 2016). The results gave insight into emergent player pathways in which failure was a healthy part of a trajectory to ultimate game success. The implication that many pathways can lead to learning has guided related work, as seen in a detector designed to capture emergent strategy for level completion within the physics game Impulse (Asbell-Clarke, Rowe, & Sylvan, 2013).

Game-based detectors have also been used to predict learning performance based on in-game player choice. A prime example is measurement of science inquiry skill in a game-based virtual environment, in which classifiers were used to detect students' learning of the STEM content during play (Baker & Clarke-Midura, 2013). Achievement in a physics game was also the subject of a recent prediction analysis in a physics game, with detectors built to predict in-game level completion at the highest level (gold) and a moderate level (silver). The findings suggested that gold achievers tended to be more efficient with time and resources than their silver-winning counterparts (Malkiewich et al., 2016). In related work, Rowe and her colleagues (2017) leveraged detectors towards creating valid, computer-based assessment of implicit science learning, using validated in-game measures as outcome variables in event-stream prediction of learning performance in physics games. Broadly, this detector-based approach has opened learning insight beyond simply looking at a pre- or post-test and treating the game as a black box; it enables understanding the emergent, event-stream interactions that support learning outcomes and target behaviors—and in turn creates the opportunity for design refinements that can support student growth moment-by-moment in play. It also creates strong potential for process-based assessment of learning, particularly in the context of complex skills and problem solving.

Overall, in support of iterative serious game design, learning analytics can leverage multimodal data streams for insights about learning and player patterns at various stages of development. The analyses reviewed here reflect recent trends in empirical game-based learning research—including usage of learning data frameworks and visualization, structure discovery and relationship mining, as well as prediction methods—with applicability to progressive stages of design (i.e., Alpha, Beta, and final release).

Discussion and Conclusion

Learning analytics and educational data mining are a set of methods that can be used to fuel the advancement of educational games research through leveraging the rich data streams enabled by digital educational games, helping to finely-tune data-driven design for personalized, engaging game-based learning experiences. Challenges and opportunities for future work in game-based learning analytics at scale are constantly expanding, in parallel with advances in technology and increases in the sophistication of game delivery systems (e.g. 3D, Augmented Reality, and Virtual Reality), leading to compelling playful learning experiences.

Implications

Application of LA to the complex, data-rich medium of serious games is a challenging endeavor with great potential for harnessing interest-driven learning (cf. Squire, 2006; Steinkuehler, 2004). As the body of empirical work in this area grows, there is opportunity to advance theory in the context of this complex, engaging learning medium. As we explore in this chapter, empirical work modeling event-stream player patterns at scale has utilized core LA methods of visualization, structure discovery, relationship mining, and prediction. This growing base of research provides great opportunity for game-based application of a broader array of educational data mining algorithms recently explored in different contexts, including probabilistic modeling (e.g. Bayesian Knowledge Tracing; Corbett & Anderson, 1995) and advanced predictive algorithms (e.g. deep learning; Botelho et al., 2017). Experimental design and game experiences geared towards building research in learning sciences also has considerable potential—from expanding knowledge of areas like embodied cognition (Abrahamson, 2009; Gee, 2008) to apprenticeship models (e.g. National Research Council, 2000;

Steinkuehler & Oh, 2012) to learning epistemology (e.g. Hofer & Pintrich, 1997; Martinez-Garza & Clark, 2017) .

Games also offer opportunity for expanding approaches to assessment and measurement in virtual learning environments (cf. Mislevy et al., 2014). Good games—intrinsically motivating learning environments which provide just-in-time information through a series of well-ordered problems (Gee, 2003)—inherently provide occasion for players to discover the underlying rule system of games through boundary testing (e.g. Owen et al., 2016). This kind of exploration is an implicit norm in the medium of games, in which equally engaged players may interact differently with the system—often in ways designers themselves don't anticipate (cf. Squire, 2011; Salen & Zimmerman, 2004; Juul, 2013). Therefore, analysis methods well-matched to the game context, and intent on capturing the most information about learner pathways, can be best equipped to mine emergent player patterns. These kinds of methods native to EDM can be used in conjunction with more traditional assessments to expand approaches to rigorous competency measurement in complex, game-like environments (e.g. Rowe et al., 2017; Baker & Clarke-Midura, 2013).

Finally, forays into studying organic patterns of play also enable a critical application of learning analytics in serious games: data driven design for personalized learning. As detailed in this chapter, iterative design based on emergent play patterns can support game development through multiple stages. Robust data frameworks, visualizations and descriptive statistics can be helpful early on (e.g. Alpha stages) to capture basic player interactions while core mechanics, level design, and fundamental user experience are being shaped. Later, during Beta development, structure discovery and relationship mining can be leveraged to streamline the player experience across multiple levels of play through identifying play sequence and attrition points, as well as interaction patterns related to engagement, success, and strategy. These methods can build on one another, supporting final application of predictive modeling within late Beta and final release stages—and to inform user-adaptive play in highly evolved game design. For example, personalized game experiences can utilize prediction to provide different core content for players (e.g. Rowe & Lester, 2015; Liu et al., 2013), or inform game overlays for just-in-time scaffolding based on behavior detection (as proposed by DiCerbo & Kidwai, 2013). Mining organic predictive patterns of play allows for personalized learning experiences for the player, which has significant implications for moment-to-moment engagement and system efficacy. Since serious games by definition have potential to teach while sustaining engagement, game-based application of LA methods can detect for learning as well as engaged behavior and afford personalization on both of these dimensions. This analytics-fueled advancement in adaptive digital design has huge implications for serving a wide range of students—at massive scale—to support individualization and learning gains in both formal and informal learning environments.

Conclusion and Future Work

As noted above, future work in game-based learning analytics affords increased opportunity for enhancing both theory and learner experiences and outcomes. Digital data streams afford investigation of learning patterns—through data that captures student process, not just a final answer—at a scale not previously possible in educational research. Advancement of technology is only fueling this potential, enabling even larger bodies of data through the advent of innovative game genres like 3D, Augmented Reality, and Virtual Reality. As these kinds of technologies reach players globally, a challenge presents itself to harness this potential and increase the size and scope of targeted studies. This future work is one link in a chain of challenges related to learning analytics and optimized design: leverage game-based engagement to create compelling and polished games for learning using emergent game genres, sustainably distribute these games to the desired population sample, utilize the technology to reach a larger number of students, and maintain development work long enough to meaningfully implement data-driven design. Successful navigation of these challenges may be possible as the realms of commercial and learning games converge in various forms: widely used subscription-model learning games (e.g. *ABCmouse* [12]and *ST Math*[13]); the modding of commercial entertainment games for learning (e.g. *SimCityEDU*[14], *Words With Friends EDU*[15], *Plants vs. Zombies EDU*[16]); and powerful tangential learning leveraged from existing commercial games (e.g. *Minecraft*[17], *Civilization*, and even *Assassin's Creed*[18] (e.g. Berger & Staley, 2014)). In these examples, highly polished games are sustainably created and distributed to a target audience, with potential for the study of data-rich environments that foster engaged learning. Still, the barriers to entry in any one of these models (particularly the third category) are substantial, and sustainable creation, research, and ongoing refinement of quality learning games remains a challenge.

In particular, clearly structured, comprehensive learning data is key to fruitful analysis (e.g. Halverson & Owen, 2014). As discussed in this chapter, interpretable, design-aligned data is critical for analysis feature selection, understanding analysis results, and using feedback to subsequently inform design. Building in such a framework early on in development can also support best practices in learning design. However, such implementation takes planning, technological resources, and a viable event-stream data framework. Thus, building in this

[12] https://www.abcmouse.com/
[13] http://www.stmath.com/
[14] https://www.glasslabgames.org/games/SC
[15] https://wordswithfriendsedu.com/
[16] https://www.glasslabgames.org/games/PVZ
[17] https://minecraft.net/en-us/
[18] https://assassinscreed.ubisoft.com/game/en-us/home/index.aspx

framework from early stages of design, or undertaking the non-trivial task of retrofitting after game completion, can be formidable; recent efforts in learning game data architecture have expanded the options and attempted to reduce implementation logistics (e.g. Serrano-Laguna et al., 2017; Danielak, 2014; Chung, 2015), but there remains opportunity for standardization and accessibility across the field.

Lastly, future work lies in adopting best practices of commercial game development within the creation of learning games. In order to benefit from data-driven design, in other words, one has to engage in it. Even a relatively small investment of resources in an iterative, user-centric design approach, which is common in industry, can increase the quality of the learner experience: e.g. fail early and often, with both small-n qualitative playtests and larger event-stream analysis where possible. In the realm of serious games this can make for substantially better products, ones that students may voluntarily play outside of school or experimental conditions, potentially empowering interest-driven learning at unprecedented scale. Through an increase in demand, such work might also increase the viability and sustainability of serious game development models.

Overall, learning analytics in application to the complex medium of learning games can support advancement of theory in the field, adaptive game-based learning, and powerful crafting of an engaged learning experience through iterative, data driven design. As we explore in this chapter, recent research has established a growing body of empirical game-based studies in learning analytics. These methods include visualization, structure discovery and relationship mining, as well as predictive modeling—which, respectively, can support alpha, beta, and final release stages of game development. In combination with a robust data collection framework, leveraging learning analytics throughout the design process and beyond is key to supporting students in personalized, engaging play experiences optimized for learning at scale.

Acknowledgements

The authors would like to thank Laura Malkiewich and Ani Aghababyan, as well as the team originally at Games+Learning+Society at UW-Madison (including Rich Halverson, Constance Steinkuehler, Kurt Squire, Matthew Berland, and Ben Shapiro).

References

Abrahamson, D. (2009). Embodied design: constructing means for constructing meaning. *Educational Studies in Mathematics*, *70*(1), 27–47. https://doi.org/10.1007/s10649-008-9137-1

Aghababyan, A., Martin, T., Janisiewicz, P., & Close, K. (2016). Microgenetic Learning Analytics Methods: Hands-on Sequence Analysis Workshop Report. *Journal of Learning Analytics*, *3*(3), 96–114. http://dx.doi.org/10.18608/jla.2016.33.6

Aghababyan, A., Symanzik, J., & Martin, T. (2013). Visualization of "states" in online educational games. In *Proceedings of the 59th World Statistics Congress* (pp. 3382–3387). the Hague, Netherlands: International Statistical Institute.

Andersen, E., Liu, Y., Apter, E., Boucher-Genesse, F., & Popović, Z. (2010). Gameplay analysis through state projection. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games* (pp. 1–8). New York, NY: ACM Press.

Anderson, C. G., Binzak, J. V., Dalsen, J., Saucerman, J., Jordan-Douglass, B., Kumar, V., Turker, A., Berland, M., Squire, K., & Steinkuehler, C. A. (2016). Situating Deep Multimodal Data on Game-Based STEM Learning. In C.-K. Looi, J. Polman, U. Cress, & P. Reimann (Eds.), *Proceedings of the 12th International Conference of the Learning Sciences* (Vol. 2, pp. 974–977). Singapore: International Society of the Learning Sciences.

Andres, J. M. L., Rodrigo, M. M. T., Sugay, J. O., Baker, R. S., Paquette, L., Shute, V. J., Ventura, M., & Small, M. (2014). An Exploratory analysis of confusion among students using Newton's playground. In C.-C. Liu, H. Ogata, S. C. Kong, & A. Kashihara (Eds.), *Proceedings of the 22nd International Conference on Computers in Education* (pp. 65–70). Nara, Japan: Asia-Pacific Society for Computers in Education.

Asbell-Clarke, J., Rowe, E., & Sylvan, E. (2013). Assessment design for emergent game-based learning. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 679–684). New York, NY: ACM Press.

Baker, J. M., Martin, T., Aghababyan, A., Armaghanyan, A., & Gillam, R. (2015). Cortical Activations During a Computer-Based Fraction Learning Game: Preliminary Results from a Pilot Study. *Technology, Knowledge and Learning*, *20*(3), 339–355. https://doi.org/10.1007/s10758-015-9251-y

Baker, R. S. (2010). Data mining for education. In P. Peterson, E. L. Baker, & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., pp. 112–118). Oxford, UK: Elsevier.

Baker, R. S., & Clarke-Midura, J. (2013). Predicting Successful Inquiry Learning in a Virtual Performance Assessment for Science. In S. Carberry, S. Weibelzahl, A. Micarelli, & G. Semeraro (Eds.), *Proceedings of the 21st International Conference on User Modeling, Adaptation, and Personalization* (pp. 203–214). New York, NY: Springer.

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2004). Detecting student misuse of intelligent tutoring systems. In J. C. Lester, R. M. Vicari, & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems* (pp. 531–540).

Baker, R. S., Corbett, A. T., & Wagner, A. Z. (2006). Human classification of low-fidelity replays of student actions. In M. Ikeda, K. Ashlay, & T.-W. Chan (Eds.), *Intelligent Tutoring Systems* (pp. 29–36). New York, NY: Springer.

Baker, R. S., & Siemens, G. (2014). Educational data mining and learning analytics. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (2nd Edition, pp. 253–274). New York, NY: Cambridge University Press.

Baker, R. S., & 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. Retrieved from http://jedm.educationaldatamining.org/index.php/JEDM/issue/view/5

Beall, M., Farajian, R., Owen, V. E., Slater, S., Smith, A., Solis, E., Steinkuehler, C. A., & Davidson, R. (2013, June). *Games for Mindfulness and Pro-Social Behavior: the Tenacity Project Collaboration.* Presented at the 9th Annual Games+Learning+Society Conference, Madison, WI.

Berger, W., & Staley, P. (2014). Assassin's Creed III: The complete unofficial guide, a teacher's limited edition. *Well Played*, *3*(1), 1–10. Retrieved from http://press.etc.cmu.edu/index.php/product/well-played-vol-3-no-1/

Berland, M., & Lee, V. R. (2011). Collaborative Strategic Board Games as a Site for Distributed Computational Thinking: *International Journal of Game-Based Learning*, *1*(2), 65–81. https://doi.org/10.4018/ijgbl.2011040105

Bosch, N., D'Mello, S., Baker, R. S., Ocumpaugh, J., Shute, V., Ventura, M., Wang, L., & Zhao, W. (2015). Automatic Detection of Learning-Centered Affective States in the Wild. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (pp. 379–388). New York, NY: ACM Press.

Botelho, A. F., Baker, R. S., & Heffernan, N. T. (2017). Improving Sensor-Free Affect Detection Using Deep Learning. In E. Andre, R. S. Baker, X. Hu, M. M. T. Rodrigo, & B. du Boulay (Eds.), *Artificial Intelligence in Education: 18th International Conference* (pp. 40–51). Cham, Switzerland: Springer.

Butler, E., Andersen, E., Smith, A. M., Gulwani, S., & Popović, Z. (2015). Automatic Game Progression Design through Analysis of Solution Features. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2407–2416). New York, NY: ACM Press.

Butler, E., & Banerjee, R. (2014). *Visualizing progressions for education and game design* (whitepaper). University of Washington: Center for Game Science. Retrieved from https://www.ericbutler.net/assets/papers/unpub2014_visualization.pdf

Canossa, A., Badler, J. B., El-Nasr, M. S., Tignor, S., & Colvin, R. C. (2015, June). *In your face (t) impact of personality and context on gameplay behavior*. Presented at the FDG Conference (Foundations of Digital Games), Pacific Grove, CA.

Chen, Z., El-Nasr, M. S., Canossa, A., Badler, J., Tignor, S., & Colvin, R. (2015, September). *Modeling individual differences through frequent pattern mining on role-playing game actions*. Presented at the Eleventh artificial intelligence and interactive digital entertainment conference (AIIDE), Santa Cruz, CA.

Chung, G. K. W. K. (2015). Guidelines for the design, implementation, and analysis of game telemetry. In C. S. Loh, Y. Sheng, & D. Ifenthaler (Eds.), *Serious games analytics: Methodologies for performance measurement, assessment, and improvement* (pp. 59–79). New York, NY: Springer.

Clark, D. B., Martinez-Garza, M. M., Biswas, G., Luecht, R. M., & Sengupta, P. (2012). Driving Assessment of Students' Explanations in Game Dialog Using Computer-Adaptive Testing and Hidden Markov Modeling. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in Game-Based Learning* (pp. 173–199). New York: Springer.

Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popovic, Z., & Foldit players. (2010). Predicting protein structures with a multiplayer online game. *Nature*, *466*(7307), 756–760. https://doi.org/10.1038/nature09304

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, *4*(4), 253–278. Retrieved from https://link.springer.com/journal/11257/5/3/page/1

Danielak, B. (2014). *Analyzing Data with ADAGE*. Gitbooks. Retrieved from https://capbri.gitbooks.io/makescape-adage-gitbook/

DiCerbo, K. E., & Kidwai, K. (2013). Detecting Player Goals from Game Log Files. In S. K. D'Mello, R. A. Calvo, & A. Olney (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 314–316). Massachusetts, USA: International Educational Data Mining Society.

Gee, J. P. (2003). *What Video Games Have to Teach Us About Learning and Literacy*. New York, NY: Palgrave Macmillan.

Gee, J. P. (2008). Video Games and Embodiment. *Games and Culture*, *3*(3–4), 253–263. https://doi.org/10.1177/1555412008317309

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research*, *3*, 1157–1182. Retrieved from http://dl.acm.org/citation.cfm?id=944968

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. https://doi.org/10.1504/IJLT.2014.064489

Hao, J., Smith, L., Mislevy, R., von Davier, A., & Bauer, M. (2016). *Taming Log Files From Game/Simulation-Based Assessments: Data Models and Data Analysis Tools* (ETS Research Report Series) (pp. 1–17). Princeton, NJ: Educational Testing Service.

Hicks, D., Eagle, M., Rowe, E., Asbell-Clarke, J., Edwards, T., & Barnes, T. (2016). Using game analytics to evaluate puzzle design and level progression in a serious game. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 440–448). New York, NY: ACM.

Hofer, B. K., & Pintrich, P. R. (1997). The Development of Epistemological Theories: Beliefs About Knowledge and Knowing and Their Relation to Learning. *Review of Educational Research*, *67*(1), 88–140. https://doi.org/10.3102/00346543067001088

Institute of Play. (2013). *Digging Into Data with SimCityEDU* (Vol. 21). Retrieved from https://www.youtube.com/watch?v=Y5_7Y7wKE6A&feature=share

Juul, J. (2013). *The Art of Failure: An Essay on the Pain of Playing Video Games*. Cambridge, MA: MIT Press.

Kai, S., Paquette, L., Baker, R. S., Bosch, N., D'Mello, S., Ocumpaugh, J., Shute, V., & Ventura, M. (2015). A Comparison of Video-Based and Interaction-Based Affect Detectors in Physics Playground. In O. C. Santos, J. G. Boticario, C. Romero, M. Pechenizkiy, A. Merceron, P. Mitros, J. M. Luna, C. Mihaescu, P. Moreno, A. Hershkovitz, S. Ventura, & M. C. Desmarais (Eds.), *Proceedings of the 8th International Conference on Educational*

Data Mining (pp. 77–84). Massachusetts, USA: International Educational Data Mining Society.

Kerr, D. (2015). Using Data Mining Results to Improve Educational Video Game Design. *Journal of Educational Data Mining*, *7*(3), 1–17. Retrieved from https://jedm.educationaldatamining.org/index.php/JEDM/article/view/JEDM048

Kerr, D., & Chung, G. (2012). Identifying key features of student performance in educational video games and simulations through cluster analysis. *Journal of Educational Data Mining*, *4*(1), 144–182. Retrieved from https://jedm.educationaldatamining.org/index.php/JEDM/article/view/25

Kerr, D., & Chung, G. (2013). *The Effect of In-Game Errors on Learning Outcomes* (CRESST Report No. 835). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from https://www.cse.ucla.edu/products/reports/R835.pdf

Kiili, K., Ketamo, H., & Kickmeier-Rust, M. D. (2014). Evaluating the Usefulness of Eye Tracking in Game-based Learning. *International Journal of Serious Games*, *1*(2), 51–65. https://doi.org/10.17083/ijsg.v1i2.15

Kim, J. H., Gunn, D. V., Schuh, E., Phillips, B., Pagulayan, R. J., & Wixon, D. (2008). Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 443–452). New York, NY: ACM Press.

Lin, C. F., Yeh, Y. C., Hung, Y. H., & Chang, R. I. (2013). Data mining for providing a personalized learning path in creativity: An application of decision trees. *Computers & Education*, *68*, 199–210. https://doi.org/10.1016/j.compedu.2013.05.009

Liu, Y., Mandel, T., Butler, E., Andersen, E., O'Rourke, E., Brunskill, E., & Popovic, Z. (2013). Predicting player moves in an educational game: A hybrid approach. In S. K. D'Mello, R. A. Calvo, & A. Olney (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 106–113). Massachusetts, USA: International Educational Data Mining Society.

Malkiewich, L., Baker, R. S., Shute, V. J., Kai, S., & Paquette, L. (2016). Classifying behavior to elucidate elegant problem solving in an educational game. In T. Barnes, M. Chi, & M. Feng (Eds.), *Proceedings of the 9th International Conference on Educational Data Mining* (pp. 448–453). Massachusetts, USA: International Educational Data Mining Society.

Marascuilo, L. A., & Serlin, R. C. (1988). *Statistical methods for the social and behavioral sciences*. New York, NY: W. H. Freeman.

Martin, T., Petrick Smith, C., Forsgren, N., Aghababyan, A., Janisiewicz, P., & Baker, S. (2015). Learning Fractions by Splitting: Using Learning Analytics to Illuminate the Development of Mathematical Understanding. *Journal of the Learning Sciences*, *24*(4), 593–637. https://doi.org/10.1080/10508406.2015.1078244

Martinez-Garza, M. M., & Clark, D. B. (2017). Investigating Epistemic Stances in Game Play with Data Mining. *International Journal of Gaming and Computer-Mediated Simulations*, *9*(3), 1–40. https://doi.org/10.4018/IJGCMS.2017070101

Mislevy, R. J. (2011). *Evidence-Centered Design for Simulation-Based Assessment* (No. 800). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Mislevy, R., Oranje, A., Bauer, M., von Davier, A., Hao, J., Corrigan, S., Hoffman, E., DiCerbo, K., & John, M. (2014). *Psychometric Considerations in Game-based Assessment*. Redwood City, CA: GlassLab.

National Research Council. (2000). *How People Learn: Brain, Mind, Experience, and School (expanded edition)*. (J. D. Bransford, A. Brown, R. Cocking, S. M. Donovan, & J. W. Pellegrino, Eds.). Washington, DC: National Academies Press.

O'Rourke, E., Butler, E., Liu, Y.-E., Ballweber, C., & Popovic, Z. (2013). The effects of age on player behavior in educational games. In *Proceedings of the 8th International Conference on the Foundations of Digital Games* (pp. 158–165). Santa Cruz, CA: Society for the Advancement of the Science of Digital Games.

O'Rourke, E., Peach, E., Dweck, C. S., & Popovic, Z. (2016). Brain Points: A Deeper Look at a Growth Mindset Incentive Structure for an Educational Game. In *Proceedings of the 3rd ACM Conference on Learning@Scale* (pp. 41–50). New York, NY: ACM Press.

Owen, V. E. (2014). *Capturing In-Game Learner Trajectories with ADAGE (Assessment Data Aggregator for Game Environments): A Cross-Method Analysis* (Doctoral dissertation). University of Wisconsin-Madison, Madison, WI. Retrieved from https://search.proquest.com/openview/870bb9267879b7646024ca482c7bf68b/1?pq-origsite=gscholar&cbl=18750&diss=y

Owen, V. E., Anton, G., & Baker, R. S. (2016). Modeling User Exploration and Boundary Testing in Digital Learning Games. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization* (pp. 301–302). New York, NY: ACM Press.

Owen, V. E., DiCerbo, K. E., Riconscente, M. M., Kline, E., Hoffman, E., Jackson, T. G., & Bertling, M. (2015). Where the Rubber Meets the (Cross)Road: Insights into Game-

Based Learning & Assessment Design. In K. E. H. Caldwell, S. Seylor, A. Ochsner, & C. A. Steinkuehler (Eds.), *Games+Learning+Society (GLS) 11.0 Conference Proceedings* (pp. 299–305). Pittsburgh, PA: ETC Press.

Owen, V. E., Shapiro, R. B., & Halverson, R. (2013). Gameplay as Assessment: Analyzing Event-Stream Player Data and Learning Using GBA (a Game-Based Assessment Model). In *The Computer Supported Collaborative Learning (CSCL) Conference 2013* (Vol. 1, pp. 360–367). Madison, WI: International Society of the Learning Sciences.

Plass, J. L., O'Keefe, P. A., Homer, B. D., Case, J., Hayward, E. O., Stein, M., & Perlin, K. (2013). The impact of individual, competitive, and collaborative mathematics game play on learning, performance, and motivation. *Journal of Educational Psychology*, *105*(4), 1050–1066. https://doi.org/10.1037/a0032688

Ramirez, D. (2016). *How Player Movement Data Improves Educational Game Assessments.* (Doctoral dissertation). University of Wisconsin-Madison, Madison, WI. Retrieved from ProQuest Dissertations and Theses database (14118).

Rowe, E., Asbell-Clarke, J., Baker, R. S., Eagle, M., Hicks, A. G., Barnes, T. M., Brown, R. A., & Edwards, T. (2017). Assessing implicit science learning in digital games. *Computers in Human Behavior*, *76*, 617–630. https://doi.org/10.1016/j.chb.2017.03.043

Rowe, J. P., & Lester, J. C. (2015). Improving student problem solving in narrative-centered learning environments: A modular reinforcement learning framework. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Artificial Intelligence in Education: 17th International Conference* (pp. 419–428). Cham, Switzerland: Springer.

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

Sao Pedro, M. A., Baker, R. S., & Gobert, J. D. (2012). Improving construct validity yields better models of systematic inquiry, even with less information. In J. Masthoff, B. Mobasher, M. Desmarais, & R. Nkambou (Eds.), *User Modeling, Adaptation, and Personalization: 20th International Conference* (pp. 249–260). New York, NY: Springer.

Scarlatos, L. L., & Scarlatos, T. (2010, September). *Visualizations for the assessment of learning in computer games*. Presented at the 7th International Conference & Expo on Emerging Technologies for a Smarter World (CEWIT 2010), Incheon, Korea.

Schell, J. (2008). *The art of game design: A book of lenses*. Burlington, MA: Elsevier.

Serrano-Laguna, A., Martinez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjóna, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces*, *50*, 116–123. https://doi.org/10.1016/j.csi.2016.09.014

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer Games and Instruction* (pp. 503–524). Charlotte, NC: IAP.

Shute, V. J., & Kim, Y. J. (2014). Formative and Stealth Assessment. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 311–321). New York, NY: Springer.

Siemens, G., & Long, P. (2011, September). Penetrating the fog: Analytics in learning and education. *Educause Review*, *46*(5), 30–32. Retrieved from http://www.elmhurst.edu/~richs/EC/OnlineMaterials/SPS102/Teaching%20and%20Learning/Penetrating%20the%20Fog.pdf

Slater, S., Bowers, A. J., Kai, S., & Shute, V. J. (2017, July). *A Typology of Players in the Game Physics Playground*. Presented at the Digital Games Research Association (DiGRA), Melbourne, Australia.

Squire, K. (2006). From content to context: Videogames as designed experience. *Educational Researcher*, *35*(8), 19–29. https://doi.org/10.3102/0013189X035008019

Squire, K. (2011). *Video Games and Learning: Teaching and Participatory Culture in the Digital Age*. New York, NY: Teachers College Press.

Steinkuehler, C. A. (2004). Learning in massively multiplayer online games. In *Proceedings of the 6th International Conference of Learning Sciences* (pp. 521–528). Santa Monica, CA: International Society of the Learning Sciences.

Steinkuehler, C. A., & Duncan, S. (2008). Scientific Habits of Mind in Virtual Worlds. *Journal of Science Education and Technology*, *17*(6), 530–543. https://doi.org/10.1007/s10956-008-9120-8

Steinkuehler, C. A., & Oh, Y. (2012). Apprenticeship in Massively Multiplayer Online Games. In C. A. Steinkuehler, K. Squire, & S. Barab (Eds.), *Games, Learning, and Society: Learning and Meaning in the Digital Age* (pp. 154–184). Cambridge, MA: Cambridge University Press.

Tissenbaum, M., Berland, M., & Kumar, V. (2016). Modeling Visitor Behavior in a Game-Based Engineering Museum Exhibit with Hidden Markov Models. In T. Barnes, M. Chi, & M. Feng (Eds.), *Proceedings of the 9th International Conference on Educational Data Mining* (pp. 517–522). Massachusetts, USA: International Educational Data Mining Society.

Tufte, E. R. (2001). *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

U.S. Department of Education. (2012). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief*. Washington, DC: SRI International.

Wallner, G., & Kriglstein, S. (2015). Comparative visualization of player behavior for serious game analytics. In C. S. Loh, Y. Sheng, & D. Ifenthaler (Eds.), *Serious Games Analytics* (pp. 159–179). Switzerland: Springer International Publishing.

Weber, B. G., & Mateas, M. (2009). A data mining approach to strategy prediction. In *2009 IEEE Symposium on Computational Intelligence and Games* (pp. 140–147). Piscataway, NJ: IEEE.

Yee, N. (2007). Motivations for play in online games. *CyberPsychology & Behavior*, *9*(6), 772–777. https://doi.org/10.1089/cpb.2006.9.772

Yee, N., Ducheneaut, N., Nelson, L., & Likarish, P. (2011). Introverted elves & conscientious gnomes: The expression of personality in World of Warcraft. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 753–762). New York, NY: ACM Press.