A Video Game Telemetry Project: Modeling Data Science and Machine Learning in Educational Research

PG Schrader, University of Nevada, Las Vegas, United States
Mark Carroll, University of Nevada, Las Vegas, United States

The Asian Conference on Education 2025
Official Conference Proceedings

Abstract

This paper models a process for exploring questions situated in digital video games and broader educational contexts. Findings from an ongoing study are used to outline a process to extract data from the encapsulating social media ecosystem using applied data science, artificial intelligence (AI), and machine learning (ML) techniques. Previous efforts involving Delphi techniques to examine player behavior and social discourse across online game environments (Schrader et al., 2020) tend to be cumbersome and laborious. Alternatively, educational data science techniques, in conjunction with AI and ML, provide mechanisms to triangulate data from three distinct sources: player discussions about performance on social media, game telemetry data, and game outcomes. In this study, researchers applied supervised and unsupervised machine learning approaches alongside statistical and learning analytics methods to extract and analyze behavioral patterns to evaluate if players' social media assertions about success aligned with the actions they take (Baker & Siemens, 2014; Siemens, 2013). Findings reveal that players' claims often reflect personal anecdote rather than observable behaviors that are linked to success. This work carefully outlines the integrated process (e.g., ML with public APIs) as a model for educational researchers to engage in similar studies. We advocate for broader utilization of data science, AI, and machine learning in research contexts that represent large, complex, but publicly available datasets (e.g., telemetric user data). This session serves as an entry point for researchers who are new to AL and ML, especially those investigating learning in contexts with high-volume data streams.

Keywords: educational data science, artificial intelligence, machine learning, video games, research methods

Introduction

Researchers have used many and varied arguments to assert that the study of video games is a relevant, important, and useful area of inquiry, especially in the context of education and educational psychology. These arguments range from the extraordinary popularity of games and games' pervasive reach across age bands and cultures across the world. The tremendous annual global revenue of approximately $455 billion (Clement, 2024) and the massive scope of many titles that have millions of players worldwide (e.g., World of Warcraft) (Active Player, 2025) undergird these respective arguments.

When considered as a partner in cognition or context for learning, as opposed to a programming or design content, researchers also assert that the educative value of games involves them as informal learning contexts and situated environments (Barab et al., 2005; Squire, 2006), simulacra or digital petri dishes (Kafai, 1996; Schrader et al., 2016), or even tools for assessment of learning (i.e., stealth assessments, McCreery et al., 2019). The examination of games has deepened our collective understanding across a variety of topics, such as literacy (Gee, 2003), cognition and cognitive skills (Bediou et al., 2018), spatial reasoning, problem-solving, decision-making, personality, motivation, goal orientation, self-directed learning (Connolly et al., 2012), and social and collaborative learning.

Further, video games and the social platforms that encapsulate them represent a growing class of non-traditional online learning environments (OLEs) (Barab, 2014; Schrader et al., 2017; Steinkuehler & Squire, 2017). These environments are data rich, globally distributed, and characterized by continuous interaction, rapid feedback, and emergent patterns of behavior (Schrader et al., 2020; Schrader & McCreery, 2012). Unlike classrooms, learning management systems, or even massive open online courses, games generate intensive, time-stamped behavioral data as a byproduct of participation. At the same time, they are supported by expansive social ecosystems, including discussion forums, social media, and streaming platforms. As such, these tools serve as a third space, where players exchange views about their activities, discuss performance, and articulate the needs for success. Collectively, these systems present both an opportunity and a challenge for educational researchers. There are abundant learning data, but their extraction requires novel, and often complicated, methods to capture the complexity of these environments.

Despite sustained interest in games as learning contexts, much of the existing research relies on methods that are not necessarily aligned with the properties of contemporary OLEs or poised to take full advantage of their technological capabilities (McCreery et al., 2011; Schrader et al., 2017). Surveys, interviews, and expert-driven techniques such as Delphi studies have proven useful for identifying constructs, perceptions, and socially validated hypotheses, but they are labor intensive, slow to iterate, and heavily reliant on self-report. More importantly, these approaches often elevate belief about play, rather than objective data related to learning and performance within the system itself. As a result, there remains a gap in the field's understanding about concrete, observable behaviors in these systems and what players believe about performance.

Fortunately, recent advances in educational data science, learning analytics, and applied machine learning offer additional tools for researchers who are addressing questions of this ilk (Baker & Siemens, 2014; Romero & Ventura, 2020; Siemens, 2013). Public application programming interfaces (APIs) in harmony with contemporary modeling techniques make it possible to collect, process, and analyze large volumes of behavioral and social data from online platforms. However, many educational researchers lack concrete skills, methods, and models for how these tools can be integrated into praxis. Further, examples are commonly drawn from fields outside of educational research and emphasize technical novelty or isolated analytic techniques.

As a result, this study demonstrates how applied data science, artificial intelligence, and machine learning techniques can be integrated with learning analytics and educational research practices to examine learning and behaviors in a complex online environment. Specifically, the paper illustrates a replicable workflow that triangulates (a) game telemetry data, (b) social discourse from online communities, and (c) game outcomes to examine the relationship between players' claims about success and the behaviors that predict it. Rather than advancing a new theory of learning or asserting causal claims about games, this work is intended to serve as a model. In this way, it is methodological in nature and offers as a practical, worked example for researchers seeking to study learning and performance in data-intensive, non-traditional OLEs. Ultimately, this work describes the end-to-end process, which connects data acquisition, modeling, and interpretation in the context of education and educational psychology.

Literature Review

The work presented here is expanded from previous work in novel contexts. Further, it is grounded in three complementary and related perspectives: complex systems, educational data science and learning analytics, and prior mixed-methods research in online game environments.

Prior Mixed-Methods and Delphi-Based Work

The present study builds directly on prior mixed-methods research that used Delphi techniques, surveys, and factor analytic methods to identify socially validated hypotheses about successful gameplay (Carroll, 2023; Schrader et al., 2020). That work demonstrated the value of leveraging player communities to surface domain-specific knowledge and to operationalize constructs in novel contexts (i.e., the League of Legends). Although informative, Schrader et al. (2020) noted key limitations to existing methods. In the case of Delphi studies, they are arduously slow, difficult to scale, and dependent on expert judgment and consensus. Moreover, they capture *beliefs* about success rather than direct behavioral evidence of it.

Building on this, the current work acknowledges the limitations and extends them. Further, social discourse remains a valuable data source. However, it is analyzed at scale using automated methods (e.g., AI, Term Frequency-Inverse Document Frequence (TF-IDF), and KMeans). In this analytical pipeline, behavioral hypotheses are not taken at face value, but are derived using scalable approaches and evaluated against quantitative findings (i.e., telemetric results). In this sense, the present work represents an evolution from expert- and discourse-driven inquiry toward an integrated, data science-informed approach that remains compatible with educational theory and interpretation.

Games as Complex Systems

From a learning sciences perspective, digital games exhibit all the characteristics of a complex system (Hilpert & Marchand, 2018; Jacobson et al., 2019; Schrader et al., 2020;

Wilensky & Jacobson, 2014). They consist of multiple interacting components (e.g., players, rules, roles, resources, objectives, feedback mechanisms) that operate across time and scale. Micro-level actions (e.g., movement, resource collection, coordination with teammates) function dynamically to evolve into macro-level outcomes (e.g., victory, defeat, or sustained performance improvement). Most importantly, these different levels of interactions are nonlinear, emergent, and multidimensional (i.e., compound over time). Said another way, small differences in timing or coordination can produce disproportionate effects on outcomes. As such, learning and performance in games cannot be fully understood through static snapshots or isolated variables.

Ontological framing using a complex systems lens has two important methodological implications. First, it requires data that reflect interactions over time rather than punctate, momentary events. Second, this view necessitates analytic approaches and analyses that can handle high-dimensional, interdependent features without assuming simple linear relationships. Traditional educational methods tend to collapse performance into measured change at specific points in time (e.g., pretest–posttest designs or retrospective self-report). These approaches, even those that include time series, are poorly suited to address the high levels of complexity and change in these systems, especially where learning evolves continuously.

Educational Data Science and Learning Analytics

Educational data science (EDS), including the related domains of educational data mining and learning analytics, provides a complementary methodological framework (Baker & Inventado, 2014; Romero & Ventura, 2020; Siemens & Gašević, 2012). Broadly, these fields emphasize the collection, analysis, and interpretation of large-scale data about learners and their contexts with the goal of understanding and improving learning. While much EDS research has focused on institutional systems such as learning management systems and MOOCs, EDS methods are also applicable to non-traditional OLEs, like games. Further, EDS also emphasizes interpretability and human-in-the-loop analysis, particularly when models are used to support educational inference rather than prediction alone.

When applied to games, EDS techniques enable researchers to move beyond proxy measures of learning and toward direct observation of behavior. Specifically, telemetric data (i.e., data that are automatically gathered by the systems) include information about specific actions that are captured during gameplay. These typically include time-stamped, sequential data that correspond to in-game events as they unfold during play. These data offer fine-grained records of actions, sequences, and timing. When interpreted using machine learning models and data science techniques, these data can be used to identify patterns associated with success, failure, or change over time.

Methods

This study employed a multi-stage, integrated methodological approach designed to examine learning-relevant behavior in a complex online environment using large-scale, naturally occurring data (Baker & Siemens, 2014; Creswell & Plano Clark, 2017). The methods were intentionally organized to reflect an end-to-end research workflow common in educational data science and learning analytics: data acquisition, preprocessing and feature construction, modeling, and interpretation. Rather than privileging a single analytic technique, the approach triangulated complementary methods to address distinct but related questions about behavior, discourse, and outcomes. Game telemetry data were used to model observable performance patterns over time, while social discourse data were analyzed to capture players' articulated beliefs about success. These data sources were then examined through a set of analytic pipelines that balanced interpretability, predictive utility, and scalability. Collectively, the methods were selected to be transparent, replicable, and adaptable to other non-traditional online learning environments, with an emphasis on demonstrating process rather than establishing causal claims.

Data Sources and Acquisition

To demonstrate an integrated research workflow, data were collected from two primary sources: game telemetry from League of Legends and social discourse from Reddit. These sources were selected because they represent complementary perspectives on learning and performance—one behavioral and outcome-oriented, the other discursive and interpretive.

Game Telemetry Data

League of Legends provides public access to detailed gameplay data through the Riot Games API (Chen, 2015; Lee et al., 2015). Using this API and some code in Python, match-level telemetry was collected, including timestamped event logs, player statistics, team outcomes, and contextual information about each match (see Figure 1). The data consisted of 100,000 match timeline files in structured JSON format, each containing timestamped logs of every in-game event (e.g., kills, objective captures, structure destruction, and resource accumulation). There are roughly 1,000 events per match. The final dataset included more than 100 million events.

Figure 1
Riot Games API Scrape

3.0: Execute Crawl

This command implements the codeblocks above and runs the scrape.

```
[9]: crawl_ranked_solo_last_year(target_matches=TARGET_MATCHES, resume=True)
```

```
1_5366379866/timeline
Collected 4250 matches so far…
Collected 4275 matches so far…
Collected 4300 matches so far…
Collected 4325 matches so far…
[429] Backing off 81s … https://americas.api.riotgames.com/lol/match/v5/matches/NA
1_5320960488/timeline
Collected 4350 matches so far…
Collected 4375 matches so far…
Collected 4400 matches so far…
Collected 4425 matches so far…
[429] Backing off 82s … https://americas.api.riotgames.com/lol/match/v5/matches/NA
1_5271529501/timeline
Collected 4450 matches so far…
Collected 4475 matches so far…
Collected 4500 matches so far…
[PARQUET] wrote 515,875 rows -> data_riot/parquet/region=americas/events-00009.par
quet
Collected 4525 matches so far
```

Note. Scrape was executed over the course of 24 hours and included high-level players over the last year.

To make these data analytically tractable, raw events were processed through an extract–transform–load (ETL) pipeline. Events were aggregated into fixed temporal bins (e.g., three-minute intervals) at the team level, producing count-based features that captured the evolving state of each match. This approach preserved temporal structure while reducing noise and dimensionality, allowing subsequent models to focus on patterns of interaction rather than individual actions in isolation.

Social Discourse Data

Social data were collected using an API for a widely used discussion platform that hosts multiple communities dedicated to League of Legends strategy, coaching, and improvement (i.e., Reddit). Researchers have argued that these spaces function as "Third Places" and data from these contexts reflect shared activity, interest, and social functions (Haythornthwaite et al., 2018; Steinkuehler & Williams, 2006). Using the Reddit API, posts and comments were scraped from selected subreddits over defined time windows. Queries targeted advice-oriented and instructional discourse, using keywords commonly associated with improvement and gameplay strategy.

Data consisted of roughly 20,000 posts with a range of 80–250 tokens per post. A token is a basic unit of text, including words and punctuation. The resulting corpus underwent a multi-stage cleaning process, including deduplication, removal of deleted or low-information entries, normalization of text, and retention of minimal metadata for traceability. Engagement indicators such as upvotes and comment counts were preserved to provide contextual information about the visibility and social validation of claims. When completed, roughly 80–90% of the text remained usable. This supported the idea that the scrape was successful and secured usable data that was suitable for large-scale text analysis.

Data Processing Techniques

Following acquisition, telemetry data were processed through a series of extract–transform–load (ETL) procedures designed to support downstream modeling while preserving interpretability. Raw telemetry events were parsed, validated, and aggregated into fixed temporal bins (i.e., 3 minutes) at the team level, producing structured, count-based features that captured evolving gameplay dynamics. This transformation reduced noise, mitigated sparsity, and enabled comparison across matches of varying duration.

Social discourse data required parallel but distinct processing steps. Text was cleaned and normalized, duplicates and system-generated artifacts were removed, and documents were organized into structured corpora suitable for vectorization. Where appropriate, datasets from multiple sources were aligned using shared temporal windows and contextual metadata, allowing behavioral and discourse patterns to be examined in relation to one another. These processing steps were critical for ensuring that subsequent analyses reflected meaningful patterns rather than artifacts of data collection or platform-specific noise.

Analytical Pipelines

The central contribution of this paper lies in the modeling of three complementary analytical pipelines. Each pipeline addresses a different aspect of learning and performance in complex online environments, and together they illustrate how educational researchers can integrate data science techniques into coherent workflows.

Pipeline 1: Principal Components Analysis Dimension Reduction and Logistic Regression

The first pipeline employed principal component analysis (PCA) followed by logistic regression to model match outcomes based on telemetry-derived features. PCA was used to reduce the dimensionality of the binned event data while retaining the majority of variance. This step served both computational and conceptual purposes: it mitigated multicollinearity among event counts and provided a compact representation of overall patterns of play.

Logistic regression was then trained on the resulting components to predict binary outcomes (win/loss). This approach offers several advantages for educational research. It is computationally efficient, relatively interpretable, and aligns well with learning analytics traditions that prioritize transparency and replicability. While limited in its ability to capture complex nonlinear interactions, the PCA–logistic regression pipeline provides a stable baseline against which more sophisticated models can be compared.

Pipeline 2: Gradient Boosting and SHAP Interpretation

To address the limitations of the linear decision boundary in logistic regression and limited interpretability of factors, a second pipeline employed gradient-boosted decision trees (XGBoost) trained on the same temporally aggregated telemetry features. Unlike logistic regression, gradient boosting can model nonlinear relationships and interactions among features without strong distributional assumptions. This makes it particularly well suited to complex systems where the impact of an action may depend on timing, context, and co-occurring events.

Model interpretation was supported using SHapley Additive exPlanations (SHAP), which attribute contributions of individual features to model predictions. SHAP values were used not as definitive explanations, but as interpretive aids that highlight which types of events and temporal patterns were most influential in predicting outcomes. This combination of predictive power and post hoc interpretability provides a pragmatic balance between model performance and educational insight.

Pipeline 3: Social Discourse Modeling

The third pipeline focused on modeling social discourse at scale. Text from Reddit posts and comments was vectorized using term frequency–inverse document frequency (TF-IDF) representations and clustered using unsupervised methods such as k-means and non-negative matrix factorization. These techniques identify recurring themes and topics without requiring manual coding or predefined categories.

To support interpretation, representative terms and excerpts from each cluster were examined, and AI-assisted summarization was used to generate concise descriptions of dominant themes. Importantly, this step was treated as an analytic aid rather than an authoritative interpretation. Human judgment remained central in evaluating whether identified themes aligned with established gameplay constructs or reflected common misconceptions.

Together, these three pipelines illustrate how behavioral data, predictive modeling, and discourse analysis can be integrated within a single research design. The following sections use this integrated workflow to examine points of convergence and divergence between what players say about success and the behaviors that predict it.

Results

The integrated analytical pipelines yielded complementary results that illustrate how different data sources and modeling approaches surface distinct, and at times conflicting, accounts of performance and success. Rather than reporting exhaustive model metrics, the results are organized around the types of patterns each pipeline revealed and the role those patterns play in informing educational inference.

Telemetry-Based Modeling Results

Across modeling approaches, telemetry-derived features reliably supported prediction of team-level match outcomes. The PCA and logistic regression pipeline demonstrated that a reduced set of latent dimensions derived from temporally binned event data was sufficient to predict win–loss outcomes at rates well above chance. These dimensions reflected broad axes of play that combined multiple event types rather than isolated actions, suggesting that performance emerges from coordinated patterns of behavior over time rather than single metrics. When manipulating the number of factors, results also indicated that there was a sufficient amount of area under the curve (AUC) and accuracy rating (see Figure 2). For reference, an AUC is an indication of a model’s ability to classify outcomes, where .50 is random guessing and 1.0 is perfect prediction.

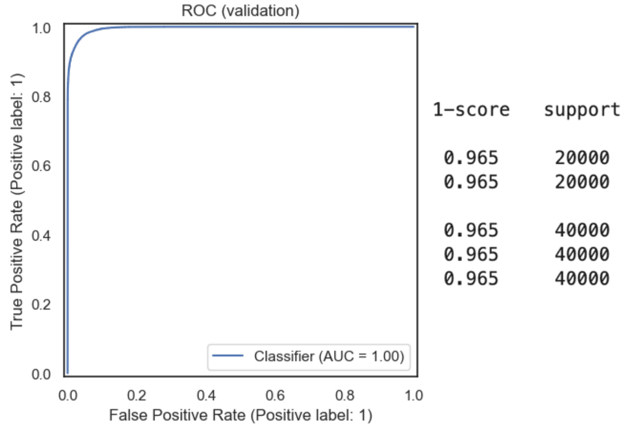
Figure 2
Principal Components Analysis Model Fit Statistics (12 Factors)

```
PCA components kept: 12 | variance: 0.3707
AUC: 0.9640453554796121
ACC: 0.90475

Report:
              precision    recall  f1-score   support

           0      0.911     0.897     0.904     19955
           1      0.899     0.913     0.906     20045

    accuracy                          0.905     40000
   macro avg      0.905     0.905     0.905     40000
weighted avg      0.905     0.905     0.905     40000
```

Note. The original model retained 102 components. This model retained 12 components and was 90% accurate.

The gradient-boosted models further improved predictive performance by capturing nonlinear interactions among event types and temporal windows (see Figure 3). SHAP analyses indicated that certain event categories exerted differential influence depending on when they occurred within a match. For example, early structural objectives and coordinated team actions were consistently associated with increased likelihood of victory, whereas other frequently discussed actions showed weaker or context-dependent relationships with outcomes. Importantly, no single feature dominated predictions across all matches; instead, influence was distributed across interacting behaviors, reinforcing a complex systems interpretation of gameplay.

Figure 3
XGBoost Model Fit Statistics and Receiver Operator Curve



Note. The original model retained 102 components. This model retained 12 components and was 90% accurate.

Social Discourse Modeling Results

Unsupervised clustering of Reddit discourse revealed recurring themes related to improvement and success, including vision control, lane management, jungle tracking, champion selection, and role-specific responsibilities. These themes were highly stable across clustering methods and aligned closely with constructs identified in prior Delphi-based research, suggesting that large-scale discourse modeling can recover socially validated knowledge structures without manual coding.

However, discourse frequency and salience did not consistently correspond to telemetry-based indicators of success. Topics that were heavily discussed or strongly emphasized within the community were not always those most strongly associated with winning behavior in the telemetry models. In many cases, discourse emphasized actions that were highly visible, role-specific, or cognitively salient, even when their aggregate relationship with outcomes was modest.

Cross-Source Contrasts

When examined together, the results highlight a systematic divergence between discourse-driven accounts of success and behaviorally grounded patterns derived from telemetry. Social data captured how players explain, justify, and narrate success, while telemetry data captured how success is enacted through coordinated behavior over time. These perspectives were not mutually exclusive, but they were not isomorphic. The contrast between them provides a methodological insight: reliance on either discourse or behavior alone yields an incomplete account of learning and performance in complex online environments.

Findings and Discussion

Collectively, these findings are highly informative relative to playing the game: League of Legends. More broadly, the findings also demonstrate the value of integrating applied data science methods into educational research on non-traditional OLEs. First, the results confirm that large-scale telemetry data can be leveraged to model performance patterns in ways that are interpretable and theoretically meaningful when paired with appropriate analytic techniques. Even relatively simple models, when applied to well-constructed features, yielded stable and informative results.

Second, the analysis of social discourse underscores the importance of distinguishing between beliefs about learning and observable learning-relevant behavior. Player communities generate rich, socially validated knowledge, but that knowledge is shaped by anecdote, role identity, and visibility rather than systematic outcome associations. From an educational perspective, this mirrors well-documented discrepancies between self-report and behavioral measures in formal learning contexts.

Third, the integration of these data sources illustrates a practical instantiation of learning analytics principles in an informal environment. Rather than treating machine learning models as endpoints, the workflow emphasized interpretation, comparison, and triangulation. In doing so, it offers a model for how educational researchers can engage with AI and ML tools without ceding epistemic authority to them.

More broadly, the findings reinforce the argument that learning in complex systems is best understood through patterns of interaction rather than isolated variables. Success in games—and by extension, in other complex OLEs—emerges from coordinated, temporally situated behavior that cannot be fully captured through surveys or single indicators. The methods demonstrated here provide one avenue for addressing this challenge at scale.

Limitations

Several limitations should be acknowledged. First, the study is situated within a single game context, and the specific behavioral patterns identified are not assumed to generalize directly to other games or learning environments. The contribution of this work lies in the process rather than the transferability of particular findings.

Second, the modeling approaches employed, while robust, are sensitive to feature construction and aggregation choices. Different temporal binning strategies or alternative representations of events may yield different patterns. Similarly, although SHAP provides useful interpretive affordances, it does not offer causal explanations and should be interpreted cautiously.

Third, social discourse data reflect participation biases and platform-specific norms. Not all players contribute equally to online discussions, and highly visible discourse may overrepresent certain roles or playstyles. These limitations mirror those of traditional qualitative methods and underscore the importance of triangulation.

Future research should explore the application of this workflow to other online learning environments, including learning management systems, collaborative platforms, and serious games. Longitudinal analyses that track change over time and designs that integrate additional data sources, such as communication logs or physiological measures, may further enhance understanding of learning in complex systems.

Recommendations

Based on the methods and findings presented, several recommendations are offered for educational researchers and methodologists seeking to study learning in complex, data-intensive online environments.

First, researchers are encouraged to move beyond single-method designs when investigating learning in non-traditional OLEs. Telemetry, discourse, and outcome data each provide partial views of learning and performance; relying on any one source risks oversimplification. Integrated workflows that triangulate behavioral and discursive data are better suited to capturing the dynamic and emergent properties of learning in complex systems.

Second, educational researchers should treat data processing and feature construction as substantive methodological decisions rather than technical preliminaries. Choices about temporal aggregation, unit of analysis, and representation fundamentally shape what patterns can be detected and interpreted. Explicit documentation of these decisions enhances transparency and supports replication, particularly when adapting workflows to new contexts.

Third, machine learning models should be deployed with interpretive intent rather than predictive emphasis alone. While predictive accuracy is valuable, educational relevance depends on the extent to which model outputs can inform understanding of learning processes. Techniques that support post hoc interpretation, such as dimensional reduction or feature attribution methods, are especially important when working with complex behavioral data.

Finally, discourse from online learning communities should be treated as a complementary analytic resource rather than a proxy for behavior. Social data are invaluable for understanding how learners make sense of their experiences, but they should be examined alongside behavioral evidence to avoid conflating salience with efficacy. This distinction is particularly important when educational interventions or design recommendations are informed by learner-generated content.

Conclusion

This paper presented a methodological exemplar for studying learning-relevant behavior in complex online environments using applied data science, learning analytics, and machine learning techniques. By modeling an end-to-end workflow that integrates game telemetry, social discourse, and outcome data, the study demonstrates how educational researchers can leverage publicly available, high-volume data sources to examine learning beyond traditional instructional settings.

From an educational perspective, the central contribution of this work lies not in the specific findings associated with a particular game, but in the process by which those findings were generated. The workflow illustrates how learning can be studied as an emergent, interactional phenomenon, and how discrepancies between what learners say and what they do can be examined systematically rather than assumed or dismissed. This approach aligns with contemporary views of learning as distributed, situated, and dynamic.

As learning increasingly occurs in digitally mediated and informal environments, educational research methods must evolve accordingly. The model presented here offers one pathway for that evolution—one that balances computational power with theoretical grounding and human interpretation. It is hoped that this work encourages broader adoption of integrated data science approaches in education, while reinforcing the importance of methodological rigor, transparency, and educational relevance.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author confirms that AI or AI-assisted technologies have been used to improve the readability of this manuscript. However, the ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

References

Active Player. (2025). *World of Warcraft: World of Warcraft Live Player Count and Statistics*. https://activeplayer.io/world-of-warcraft/

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 61–75). Springer.

Baker, R. S., & Siemens, G. (2014). Educational data mining and learning analytics. *Cambridge Handbook of the Learning Sciences*, *2*, 253–272.

Barab, S. (2014). Design-based research: A methodological toolkit for engineering change. In *Handbook of the Learning Sciences* (pp. 151–170). Routledge.

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making Learning Fun: Quest Atlantis, A Game Without Guns. *Educational Technology Research & Development, 53*(1), 86–107.

Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Green, C. S., & Bavelier, D. (2018). Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychological Bulletin, 144*(1), 77–110.

Carroll, M. C. (2023). *Toward data science for research in novel online learning environments: A mixed method, complex systems approach* [Doctoral dissertation, University of Nevada, Las Vegas].

Chen, M. (2015). *Data-driven game design*. CRC Press.

Clement, J. (2024). Video game industry – Statistics & Facts. Statista. https://www.statista.com/topics/868/video-games/?srsltid=AfmBOopyVpsEZm9QcgDaS5sCh9AfYUt3Lh9zRvXPJC7X7oJIbZIhb-0L

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2012), 661–686.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Sage.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave/St. Martin's.

Haythornthwaite, C., Andrews, R., Fransman, J., & Meyers, E. (2018). The social learning relations of MOOCs. *Distance Education*, *39*(3), 321–337.

Hilpert, J. C., & Marchand, G. C. (2018). Complex systems research in educational psychology. *Educational Psychologist*, *53*(2), 85–100.

Jacobson, M. J., Kapur, M., So, H.-J., & Lee, J. (2019). The ontologies of complexity and learning. *Learning, Culture and Social Interaction*, *21*, 1–11.

Lee, J., Wohn, D. Y., & Park, J. (2015). Understanding player behavior through telemetry. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*.

McCreery, M. P., Krach, S. K., Bacos, C. A., Laferriere, J. R., & Head, D. L. (2019). Can video games be used as a stealth assessment of aggression? A criterion-related validity study. *International Journal of Gaming & Computer Mediated Simulations*, *11*(2), 40–49. https://doi.org/10.4018/IJGCMS.2019040103

McCreery, M. P., Schrader, P. G., & Vallett, D. B. (2011). Games as assessment environments. *Educational Assessment*, *16*(4), 256–271.

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(3), e1355.

Schrader, P. G., Carroll, M. C., McCreery, M. P., & Head, D. (2020). Mixed methods for Human-Computer Interaction research: An iterative study using Reddit and social media. *Journal of Educational Computing Research, 58*(4), 818–841. https://doi.org/10.1177/0735633119878066

Schrader, P. G., & McCreery, M. P. (2012). Are all games the same? Examining three frameworks for assessing learning from, with, and in games. In D. Ifenthaler, D. Eseryel, and X. Ge (eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives* (pp. 11–28). Springer.

Schrader, P. G., McCreery, M. P., & Vallett, D. B. (2017). Games and learning: A complex systems perspective. *Educational Psychologist*, *52*(2), 1–15.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, *57*(10), 1380–1400.

Siemens, G., & Gašević, D. (2012). Guest editorial—Learning and knowledge analytics. *Educational Technology & Society*, *15*(3), 1–2.

Squire, K. D. (2006). From content to context: Video games as designed experiences. *Educational Researcher, 35*(8), 19–29.

Steinkuehler, C., & Squire, K. (2017). Videogames and learning. In *The Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 819–838). Cambridge University Press.

Steinkuehler, C., & Williams, D. (2006). Where everybody knows your screen name. *Journal of Computer-Mediated Communication*, *11*(4), 885–909.

Wilensky, U., & Jacobson, M. J. (2014). Complex systems and education. *Educational Psychologist*, *49*(2), 71–82.

Contact email: pg.schrader@unlv.edu