22. Understanding Youth's Personal Connection to Data in a Gamelike Assessment System Through Learning Analytics and Qualitative Analysis

SAWAROS THANAPORNSANGSUTH, YIPU ZHENG, AND NATHAN HOLBERT

Abstract: In this paper, we explore middle school students' understanding of data in a gamelike assessment system called "*Beats Empire*." *Beats Empire* situates students as music studio managers. Students use their data-analysis skills to sign artists, record songs, and determine song-release strategies to win the game. We examine how gameplay can engage students in playful construction tasks that elicit computational thinking about data. To this end, we draw qualitative data from students' play-aloud interviews and log data from classroom implementation of the game. The complementarity of learning analytics and qualitative analysis leads us to better understand how students make data-oriented gameplay decisions while actively connecting data from the insight screen with their personal experiences in music when recording new songs. The study provides methodological input on how qualitative analysis and log-data learning analytics can build upon each other to explore students' behaviors in an open-ended learning environment. It showcases the importance of integrating both quantitative and qualitative data. On the one hand, learning analytics is a useful tool for capturing insights from the students' gameplay experience. On the other hand, qualitative analysis provides space for students to explain their actions and what is meaningful to them.

Introduction

Beats Empire (Holbert et al., 2019) is a playful gamelike assessment system designed for middle school students that puts players in the role of a music studio manager. The design of *Beats Empire* is influenced by the constructionist design paradigm and culturally relevant learning. Students examine data about listeners' music preferences and trends, which vary by mood, genre, and geography. Using the data set, they make decisions about which artists to sign, what songs to record, and where to target their song release. Ultimately, students choose what kind of studio they want to create and define their own goals of success. All students' actions are logged and data about how players are interacting with the game are analyzed to offer useful information about students' understanding of data.

This paper takes an initial step in exploring methodological input on how qualitative analysis and log-data learning analytics can build upon each other to explore students' behaviors in open-ended learning. We aim to understand how students use data and what data means to them. Through students' log data and their play-aloud interviews, we observe how students use the game's insight screen in making meaningful gameplay decisions and strategic choices. Students used the data informed by the insight screen to connect with their prior knowledge in pop culture to make rational and personal gameplay decisions. They also create in-game artifacts that reflect upon their personal tastes in music. We find that by investigating the gameplay solely through log data, we miss an opportunity to understand the challenges and complexities faced by the students when comprehending data. This paper reports that students' lack of line graph usage may hinder them from making long-term predictions and analyses on the music trends.

Backgrounds

Born in the mid-1990s through the late 2010s, youth today are digital natives as they are immersed in data-driven technology since birth (Hicks, 2011; Prensky, 2001). Common Sense Media (2019) reported that youth today spend more than seven hours per day on-screen. Smartphone ownership among 13-year-olds has grown from 50% in 2015 to 72% in 2019. Youth are accustomed to advanced technology but do not necessarily have the ability to interpret or understand data. Despite the saturation of technology in youth's lives, standardized curricula, methods, and resources on data science that drives the technology we interact with every day rarely exist or are absent in schools throughout the United States (Chan & Ensari, 2019; Philip, Olivares-Pasillas, & Rocha, 2016). The ability to interpret data is a form of data literacy that has become increasingly important for youth to be able to navigate the world in the digital age (Philip et al., 2016). There is a need for educating and researching how youth make sense of data.

There are several initiatives designed to develop students' computational and statistical thinking skills so that they can access and analyze data. For instance, the Mobilize project (Phillips et al., 2016) aims to teach data science concepts across a large school district. Mobilize has received high interest from schools and students; however, it has encountered challenges in its assessment methods to capture what it had hoped to cultivate (Gould et al., 2016). *Beats Empire* seeks to assess middle school students' data knowledge in a playful, gamelike environment. In our pilot play tests, we found that assessing students' data knowledge in the gamelike environment can heighten their engagement. Students see the importance of data collection because they will later use the analysis to increase their artist's followers. In a conventional assessment, it is hard to engage students with the long-term arc of a set of related constructs (Pellicone et al., 2019).

In addition, there are limited studies on understanding how students make sense of data related to their real-life experience. Previous research on data comprehension mainly focused on students' cognitive capability in graph interpretation in formal math education. For example, Curcio (1987) defined three levels of graph reading: Read with the data, read between the data, and read beyond the data. Watson and Moritz (1999) examined students' statistical thinking under the setting of different sample sizes. Friel, Curcio, and Bright (2001) defined "graph sense" and identified factors influencing graph comprehension. Díaz-Levicoy, Batanero, Arteaga, and Gea (2019) compared the performance of Chilean children when reading different types of graphs included in the school curriculum. In this study, we hope to supplement the discussion by understanding students' concepts of data in the context of their daily life experience and reasoning.

Methodology

Thirty-five seventh graders in an urban middle school in the northeastern United States engaged in one hour of gameplay. Students ranged in age from 12 to 14 years old. All names used in this paper are pseudonyms. We collected data through automatic log files of the gameplay of all 35 students. Seven students were randomly selected to participate in pre- and postinterviews and one-on-one "play-aloud" sessions with field researchers (in which students narrated their reasoning behind choices in the game). At the end of the gameplay on the same day, the entire class participated in a focus group. For this analysis, we used the automatic log data to investigate our guiding question and to understand general patterns in how participants approached the game. Interviews and think-aloud recordings were transcribed. Time-stamped log data were exported from the back end to JSON format and then converted to comma-separated files and imported into RStudio.

The seven interviewed students were asked questions about their (a) understanding of trends and how data are used in informing real-life decisions, (b) their familiarity with line graphs, and (c) the gameplay experience. These seven are not

necessarily meant to be illustrative of the entire sample but do provide an interesting set of cases. Once the interviews were transcribed, we looked closely at segments of the interview related to the students' (a) definitions of trend, (b) usage of insight screens to inform their gameplay decisions, and (c) song-title generating habits. The data were coded "bottoms-up," in which themes and coding categories emerged from patterns in the data (Miles, Huberman, & Saldaña, 1994). For example, one code categorizes students' perceptions of themselves in relation to the understanding of trends. Example phrases from the interviews included: "What's the craze," "What is popular at the moment," and "The thing that is popular."

Findings

Phase 1: Students' Song-Title Generating Habits

We started our analysis by taking an expansive view of the data. We looked through interview transcripts, highlighting interesting gameplay behaviors or generating questions about potential gameplay patterns. We observed that the majority of students spent a lot of time in their gameplay cycling through auto-generated titles for their new song release. They would meticulously think about the song title even though it did not lead to any in-game rewards. We were interested in verifying the prevalence of this observation across subjects and exploring the reasons behind their actions. Thus, we processed and analyzed log data to help answer the question.

To our surprise, out of 35 students, 24 students generated more than 50 song titles per song release (see Figure 1). Two students even generated more than 175 song titles! What is even more surprising is that the auto-generated song title button is the second most-clicked button in the game (see Figure 2). It was used almost as much as the "back" button! (Note that a few students initially thought that the song titles they had previously seen would cycle back around so they kept going through the list.) It is clear that the auto-generated song title feature is still crucial in the students' engagement in the game. This phenomenon led us to question why the feature was so prominent in the students' gameplay experience. We also wondered if the feature would facilitate or hinder the students from exploring other meaningful features in the game that would better cultivate students' understanding and applications of data.
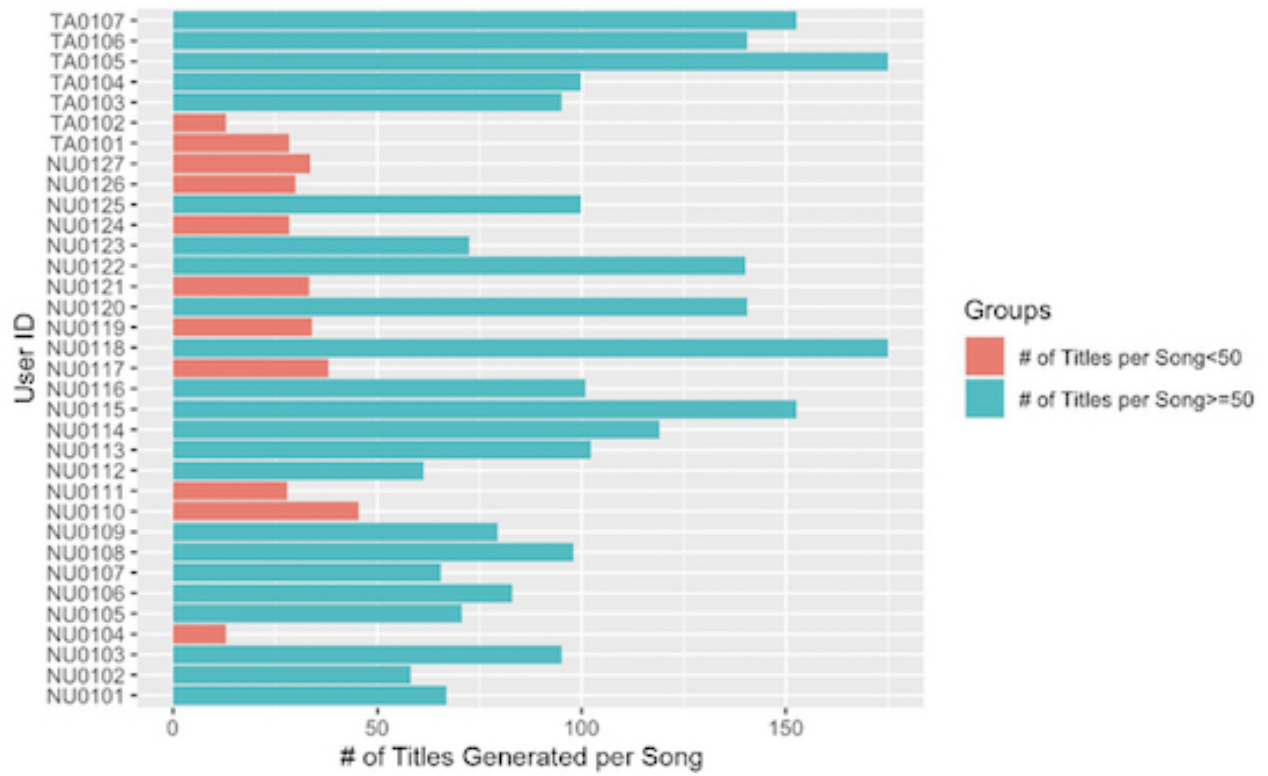
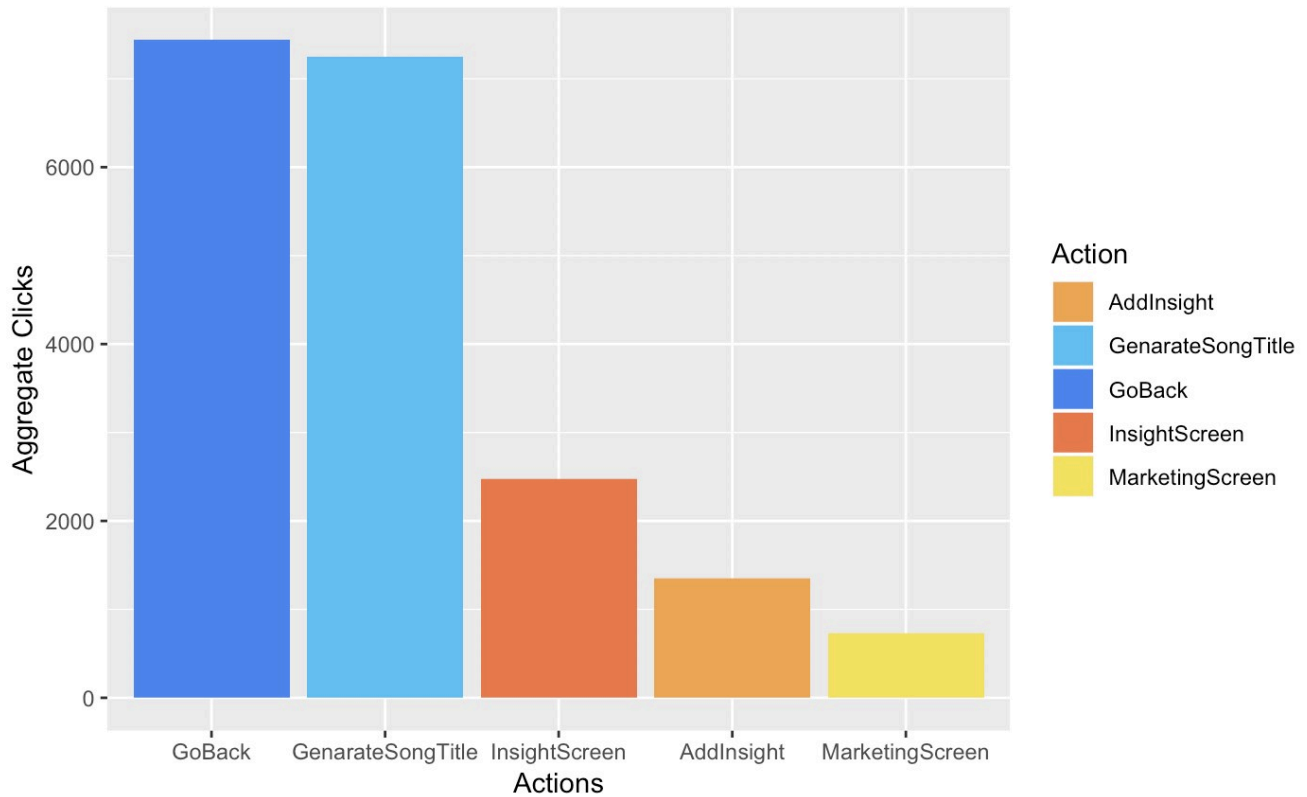Figure 1. Most students went through more than 50 song titles in order to record a song.

Figure 2. Students' aggregated clicks ranked by functions. The auto-generated song title feature is ranked second in the whole gameplay.

Phase 2: Investigating Students' Song-Title Generation Habits Through Qualitative Analysis

In this study, we looked into how students make sense of data, which went beyond managing a music studio in the gameplay and instead focused on their broader experiences in music and pop culture. The qualitative findings in our study highlight how students not only used data from the insight screen to determine which song to release but also used their intuition and prior experiences in making choices in the gameplay.

In the second phase, we were interested in why the students spent a lot of time clicking through the auto-generated song titles when this action does not lead to any in-game reward. Out of 35 students, 24 generated more than 50 song titles per song. Note that we did not explicitly inform the students that their choices on song titles had no in-game reward. Our goal was to discover the implications behind the action as it is one of the indicators of high engagement in the gameplay. Across the play-aloud interview data, our findings suggest that students who spent a lot of time scrolling through the song titles were actually looking for song titles that were suitable for the most popular moods indicated by the bar graphs in the insight screen. In other words, an otherwise unnoticed aspect of gameplay (choosing song titles) ended up revealing a very meaningful pattern relating to students' experiences. Throughout all seven play-aloud interviews, the students used their intuition and personal experience to select the final title based on the mood. For instance, Derrick looked at the insight screen and learned that the song with a nostalgic mood was popular among his fans. He said, "I always go with the highest [bar graph]." When he was recording his song, he browsed through the song titles and chose "All-American." He reasoned that the title was nostalgic because "back in the '60s there were

people who were considered all-Americans." Similarly, Alex checked his insight screen and learned that the borough of Morris enjoyed listening to the songs that signify hope. Alex spent a moment clicking through the song titles and chose "Uncrushable Joy." He said, "Because it sounds like hope." Joe commented on his song title selection habit that "I normally just pick a lot of the song titles, the one that speaks to me and that could relate to one of the topics."

Phase 3: Investigating Students' Usage of the Data Insight Screen: Bar Graph Versus Line Graph

In the third phase of our data analysis, we were interested in how the students were using the "Insight screen" in making decisions in the gameplay. The insight screen is the data-analysis tool provided within the game to help students understand the trends within each borough so that they can release songs that satisfy their fans (see Figure 3). Students could look through the data by mood, topic, and genre. They could see the trends in bar graphs, line graphs, or with geographical locations. The students must identify and act on the data and trends that they can see forming within the data visualizations provided to them.
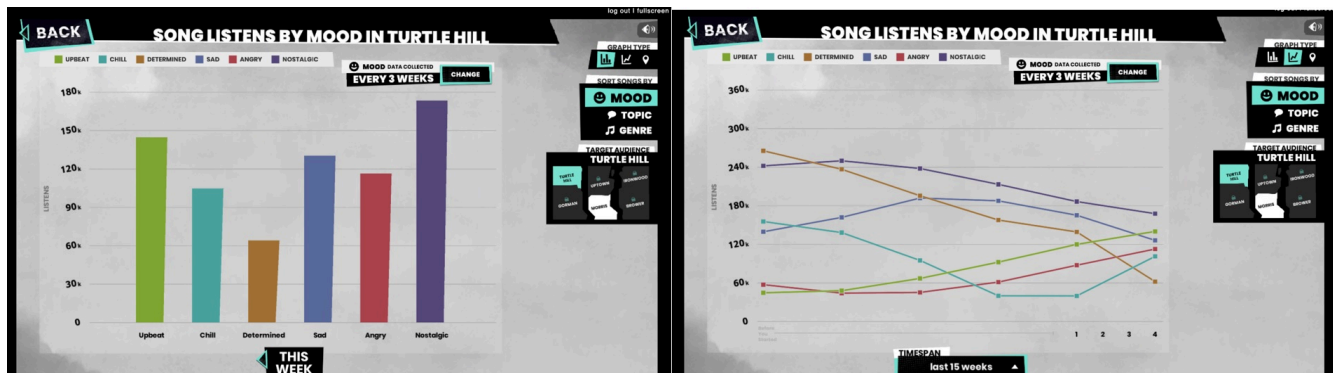

Figure 3. The screen shots of the insight screen on Turtle Hill by mood in bar graph (left) and by mood in line graph (right).

Through the log data, we found that students also extensively spent time on the insight screen (ranked eighth out of 233 unique gameplay interactions in total). Targeting each borough to release their songs, the students often employed the insight screen by looking at the borough's preference through mood and topic. From the log data, we saw how the students were frequently switching from sorting the data by mood and topic to make decisions of what to record. However, they rarely shifted from bar graph to line graph (see Figure 4). We found that when choosing what to record, the students identified a song to be "most popular" by using mainly the bar graph. The log data showed how students clicked onto the line graph but switched back to the bar graph throughout the gameplay. Less than half of the class self-reported that they were using the line graph during the post-gameplay focus group.

The play-aloud interviews revealed why they chose the bar graph over the line graph. For instance, Alex said that he would just use the bar graph because "it makes more sense." The bar graph provided sufficient data for the students to navigate the gameplay. It informed the students which mood, topic, or genre was the highest at the moment and that information was enough for them to excel. Moreover, we found that students may partially understand the benefits of the line graph. For example, Joe thought that line graphs offered better visualization. Sam said that the line graph gave her a better overview of the data while providing her more details on how they fluctuated: "Every week it went up and down, not just in summary." For her, the bar graph showed only a summary of the data, while the line graph showed how

they increased and decreased. Students seemed to omit the key point of a line graph: analyzing data in relation to time. Using the line graph, the students could make a better long-term prediction and analysis of the music preference trends based on its past and future projections.
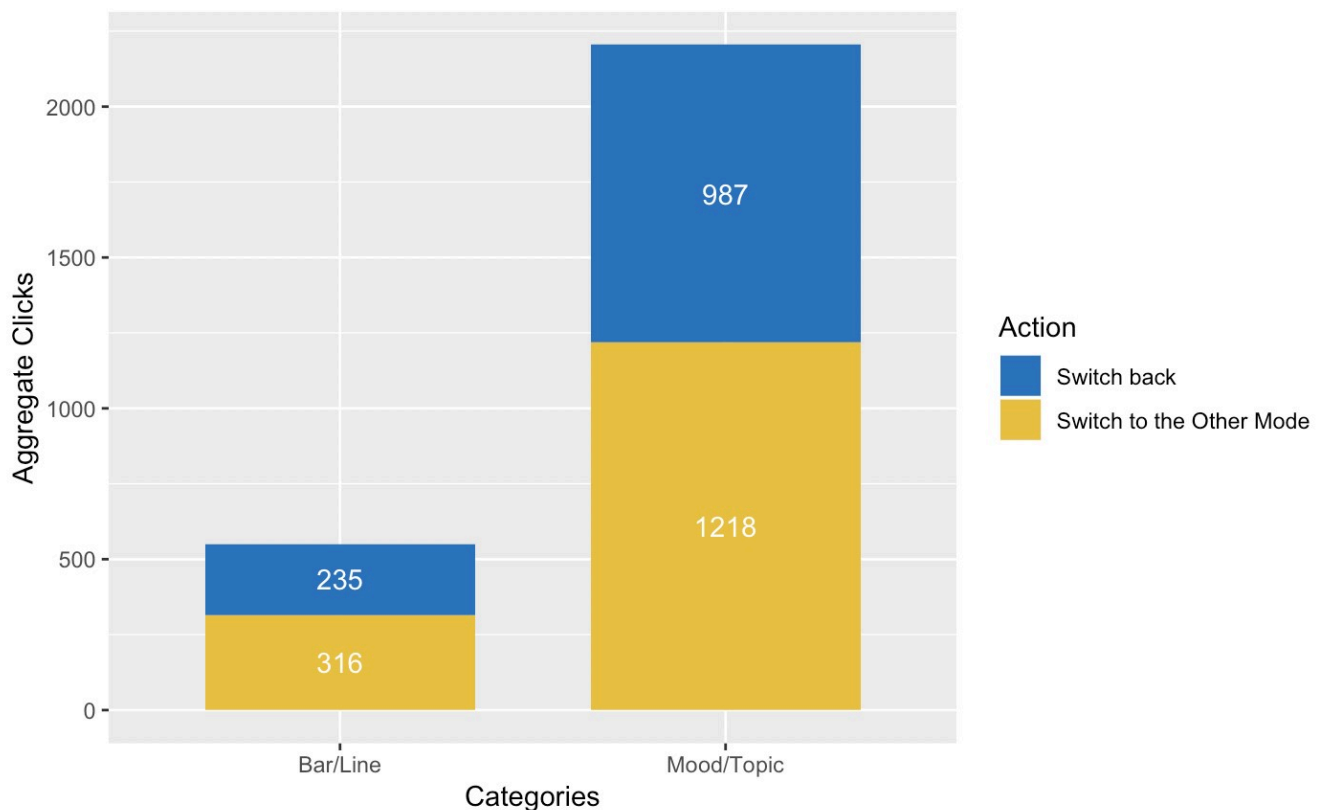

Figure 4. Students often switched from mood and topic. They rarely switched from bar graph to line graph.

Phase 4: Understanding the Students' Meaning of Trends

In the third phase, we investigated the absence of line graph usage. We found that students may overlook the benefit of a line graph in tracking changes and making predictions over short and long periods. The final phase of the finding focuses on what students thought "trend" was. Our goal is to take a step back and look into students' fundamental understanding of trends. We aim to see if the students took the aspect of "time" into account.

We asked the students, "What do you think the trend means?" Six out of seven students we interviewed thought of trends as something that is currently popular. Kendrick answered, "Trend means that a lot of people do it and then it becomes that everybody else wants to do it, so it becomes a trend, and it's really popular." Alex referred to the trend as "What's the craze. What's the thing that you like right now." The students particularly situated their understanding about trends in the context of pop culture and fashion. For instance, Alex talked about trends as "top music and clothes." Eric also mentioned that trend was how a song became popular. Sam, one of the students who used the line graph, explained that trend was "a thing or video that a lot of people are viewing and sharing because they think it's funny, entertaining, or they just like it." All in all, the students were more inclined to the nonstatistical definition of trend. They thought of trends as popularity rather than a method for understanding how and why things have changed or will change through time.

Implication and Conclusion

Game Design Implication

The goal of *Beats Empire* is to create an environment where students' understanding could be formatively assessed in a playful environment (Holbert et al., 2019). In this paper, we investigate seven grade students' understanding of data through *Beats Empire* gameplay. We demonstrate how students make sense of data and the concept of *trend* in their daily context. Trend is a core concept in data science. Data scientists use the word to indicate the change or development in a general direction over a certain period of time. However, the students in this paper thought of trends as popularity at the current moment. The discrepancy in defining trends between the data scientists and the students allows us to reflect upon design suggestions on both assessment and learning games.

For assessment games such as *Beats Empire*, designers should consider closely when selecting the context of the game. Different contexts may significantly impact students' interpretations of certain concepts. In the case of *Beats Empire*, students' interpretation of trends may derive from *trendy*, a word that is commonly used in music or fashion. When they say "trendy," they often emphasize the fact that a lot of people know or like certain things at the current moment in such a context. However, if the assessment is situated in a different context, such as climate change, their interpretations might change accordingly. Thus, conclusions about the students' understanding should not be made without carefully examining the context of the assessment game itself.

For learning games, whose goal is to teach students about data, students' interpretation of "trendy" can be leveraged as an interesting entry point to understand the statically inclined concept of trend. In addition, designers need to pay attention to design features that help students discern "at the current moment" and "over a period of time." For example, the current design of *Beats Empire* does not intentionally push students to use the line graph, since the purpose of the game is not to instruct but to assess. However, an alternative design for a learning game (if *Beats Empire* were to be a learning game) could add a specific prompt to ask students to compare and contrast the bar graph and the line graph.

Methods Implication

We showcase how integrating log-data learning analytics and qualitative interviews could strengthen our understanding of students' gameplay experience in ways that a single method could not. Throughout the analysis, we went back and forth between the log data and the play-aloud interviews. Combining both methods, the log data tell us "What happened?" while the interview data give us possible explanations of "How and why did it happen?" Qualitative data are useful in helping researchers relate naturally and holistically with students in the field. They also allow us to immerse in the students' interactions and matters that provide perspectives to others' ways of life (Emerson, Fretz, & Shaw, 2011). However, the proximity with students may engender subjectivity and biases when researchers analyze their data (Peshkin, 1988). Learning analytics methods fill this gap by supporting qualitative research in capturing transient actions and identifying overall trends in the sample population. Learning analytics also quantifies a qualitative phenomenon using combinations of log data attributes and verifies a qualitative phenomenon across a larger population (Berland, Baker, & Blikstein, 2014; Fields, Quirke, Amely, & Maughan, 2016; Zheng, Blikstein, & Holbert, 2020).

This paper uses the multimethod structure by taking an expansive view of qualitative data and spotted students' tendency to gravitate toward the auto-generated song titles function. Not only did the log data confirm our qualitative findings, but they also allowed us to further ask how students were using their personal experience to interpret the data in the game. This then led to another round of log data analysis in examining students' insight screen usage and possible

explanations in the lack of line graph usage. Last, we pull in students' interview data to further investigate students' perceptions of trends in relation to data and time.

References

Berland, M., Baker, R. S., & Blikstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning*, 19(1–2), 205–220.

Chan, M., & Ensari, I. (2019, September). *Using open-sourced data to nurture a civically engaged and computationally fluent generation.* Retrieved from https://cra.org/ccc/wp-content/uploads/sites/2/2019/09/Using-Data-to-Nurture.pdf

Common Sense Media. (2019, October). *The Common Sense census: Media use by tweens and teens, 2019.* Retrieved from https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2019

Curcio, F. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18(5), 382–393.

Díaz-Levicoy, D. A., Batanero, C., Arteaga, P., & Gea, M. M. (2019). Chilean children's reading levels of statistical graphs. *International Electronic Journal of Mathematics Education*, 14(3), 689–700.

Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). *Writing ethnographic fieldnotes.* Chicago, IL: University of Chicago Press.

Fields, D. A., Quirke, L., Amely, J., & Maughan, J. (2016). Combining big data and thick data analyses for understanding youth learning trajectories in a summer coding camp. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (pp. 150–155). https://doi.org/10.1145/2839509.2844631

Friel, S., Curcio, F., & Bright, G. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2), 124–158.

Gould, R., Machado, S., Ong, C., Johnson, T., Molyneux, J., Nolen, S., … Zanontian, L. (2016). *Teaching data science to secondary students: The Mobilize introduction to data science curriculum.* Presented at The International Association for Statistical Education (IASE) 2016 Conference, Berlin, Germany.

Hicks, S. D. (2011). Technology in today's classroom: Are you a tech-savvy teacher? *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 84(5), 188–191.

Holbert, N., Berland, M., Rutstein, D., Basu, S., Disalvo, E., Rochelle, J., … Kumar, V. (2019). *Designing constructionist formative assessment games.* Presented at the annual meeting of the American Educational Research (AERA) 2019, Toronto, Canada.

Miles, M., Huberman, A.,& Saldaña, J. (1994). *Qualitative data analysis: An expanded sourcebook.* Thousand Oaks, Ca: Sage.

Pellicone, A. J., Holbert, N. R., DiSalvo, B., Berland, M., Kumar, V., & Zhao, Y. (2019). Who played the game correctly? Data signatures of interaction in playful assessment. Presented at the Connected Learning Summit (CLS) 2019, Irvine, CA.

Peshkin, A. (1988). In search of subjectivity—one's own. *Educational Researcher*, 17(7), 17–21.

Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction*, 34(4), 361–388.

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, 9(5), 1–6.

Watson, J., & Moritz, J. (1999). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics*, 37, 145–168.

Zheng, Y., Blikstein, P., & Holbert, N. (2020). *Combining learning analytics and qualitative analysis for the exploration of open-ended learning environments.* Presented at the Constructionism 2020 Conference, Dublin, Ireland.