Co-Designing Assessment Dashboards with Teachers for Educational Games: A Case of Persistence

Yoon Jeon Kim [0000-0002-3467-1238], Mariah A. Knowles[0000-0002-1798-4830], and Grace Lin [0000-0001-7552-2880]

University of Wisconsin—Madison, Madison WI 53706, USA
yj.kim@wisc.edu, mariah.knowles@wisc.edu

Massachusetts Institute of Technology, Cambridge MA 02139, USA
gcl@mit.edu

Abstract. The co-creation of educational technologies is dynamic in nature and poses challenges for analyzing data generated during rapid and time-constrained co-design sessions. Thus, it's challenging to closely investigate how the co-design process played out over time, and how it led to certain technological innovations. Applying Quantitative Ethnography methods, we aim to investigate how teachers and researchers collaboratively designed data visualizations for an educational math game to measure their students' "persistence." In this poster, we report preliminary findings based on the thematic analysis codebook that the research team iteratively established.

Keywords: co-design, assessment literacy, data visualization, human-centered learning analytics, game-based learning

Introduction

Learning Analytics (LA) can add value to teaching and learning practices in classrooms, for example, by providing teachers with greater insight into students' learning so they can provide timely feedback [1]. There are three challenges for integrating LA with educational technology in classrooms: (1) teachers must understand how data is processed, (2) teachers must place appropriate value in LA's suggestions, and (3) LA's data visualization tools must be usable by teachers in real contexts. We aim to understand the interactions between teachers' assessment literacy, data visualization, and LA, within the context of game-based assessment. Our team engaged teachers as co-designers of teacher-facing, interactive, LA data dashboards, designed for enabling decision-making based on data gathered from an educational game, Shadowspect. Shadowspect is a 3D puzzle game for assessing Common Core Geometry standards, student persistence, and spatial reasoning [2]. To increase teachers' rigorous use of game-based assessment data and support pedagogical decisions based on LA, the project investigates how to co-design LA and data visualizations that elicits teachers' active participation without imposing technical barriers. The end goal of this research is to develop interactive data visualizations that embody generalizable, theoretically articulated, and empirically grounded design principles.

To those ends, among other questions, our project asks: RQ1: How do teachers and researchers come to understand each other's values and desires during the co-design process? RQ2: How do teachers' existing assessment literacy and practices influence the process identified in RQ1?

In this poster, we share preliminary findings of our thematic analysis of one case study from that project. This analysis was aimed at answering RQ1 and RQ2, and it centers on teacher/researcher discourse around the measurement and visualization of students' "persistence," a salient desire of the teachers that emerged early in the co-design process.

Methods

Context & Data Collection

We selected 8 math teachers as design fellows from 16 secondary school teachers who applied for an open call for participation. The teachers were selected on their interests in educational values of games, data use in classrooms, and co-designing processes. The team and teachers met monthly during development iteration cycles for 12 months. A typical co-design session lasted 2 hours. Due to COVID-19, all design sessions were conducted and recorded remotely via Zoom. The team collected several sources of data: design session discussions, teacher interviews, teachers' individual think-alouds, artifacts generated by the fellows, and the team's field notes. The focus of co-design activities varied, such as from generative ideation with prompts to identify indicators for a metric, to trying out a prototype and creating a user journey map to describe how they would use it in real classroom contexts [3].

Student "persistence" emerged early in the co-design process as a trait that participants wanted to measure, visualize, understand, and respond to. For example, the research team asked the fellows to create the kinds of stories that they wish to tell about students using analytics from the game, and the general consensus was that, because the game encourages students to try challenging problems, teachers wanted to see measurements related to growth, efforts, and progress. The fellows then defined, selected, and grouped a set of metrics that they wished to see in the teacher-facing game data dashboard, and their top three choices were persistence, math standards, and common errors and misconceptions. For this poster, we focus on the first of these: persistence.

There are varying definitions of persistence in the literature, so participants and the research team eventually operationalized persistence as the ability to maintain an action or complete a task, regardless of one's inclination towards the task, with the active choice to continue in a course of action, even in the face of obstacles, difficulties, or discouragement [4,5]. While game-based learning literature has previously defined and measured persistence [6,7], to date there has been no discussion on how teachers can use this information in classrooms. Therefore, for this case study, we analyzed all sessions and individual think-aloud activities that included explicit discussion of persistence in order to learn how the teachers' assessment literacies, design requirements, and use of data visualizations arose within the context of that one concept.

Thematic Analysis

Transcripts of Zoom recordings where persistence was a topic of discussion were inductively coded, where common themes were identified by a close reading of the data and iterative comparing, sorting, and defining similar data [8,9]. Themes were agreed upon through social moderation within the research team and triangulated with artifacts generated by the fellows and the team's field notes [10]. The identified themes are defined and illustrated in Table 1.

Table 1. Codebook

Label	Code	Example
MOTIVATION	Discourse invoking participants' motivations for use of the design, such as extrinsic pressures to incorporate more data in decision making and intrinsic motivations to see subjects succeed.	"if I'm talking to parents...this is really good data to have" or "We're teaching these kids not just math, but we're trying to teach them to survive in the real world."
CONCEPTION	Invoking or describing possible conceptual definitions for some bigger picture concern, such as defining a concept by its flavors or by what it requires.	"A prerequisite to some of our definitions of persistence is failure."
OPERATION-ALIZATION	Invoking or describing implementations of a metric for measuring some concept, such as tracking moves spent, tracking duration, developing composite scores, and aligning multiple metrics.	"I would want it to be...like, 'students who have made a real attempt'...defining time-wise or move-wise" or "I'm wondering if...just like the number of reattempts would be more interesting."
DESIGN	Invoking the design itself, such as proposals for features, discussion of (imagined or real) features, desires for affordances, the limits of those affordances, and evaluations of a design prototype.	"It'd be important to me that...the students along the X axis are the same, graph-by-graph" or "I'm wondering if...that visualization with the line and the bubbles seems confusing."
CONCERN	Invoking or implying concerns, critiques, or confusions of a metric, a design, or the design	"There's like kids who sit and think, and there's kids who sit and do nothing, so it's kind of hard to tell what's going on, whether it's like active time or inactive time" or "I've been

	process as a whole, such as pointing out conflating factors and confusion over what a metric represents.	asking this question kind of all along. I'm not, I don't really understand."
DESIGN TEAM	References to the design team's power in relation to the participants, such as participants trusting a metric because it was created by this team.	"We have some folks from [institution] who are letting us test out a game"
DIGGING INTO	Invoking or processes of participants' seeking richer qualitative understandings of their subjects, design elements, or concepts through engagement (imagined or real) with a design, such as comparing two subjects qualitatively and trying to uncover subjects' thinking.	"It felt like this third person...like, they kept trying. Like, they eventually checked, like they changed the perspective again. And then they eventually got rid of that triangular prism, checked their perspective again. It feels like they were more invested, and just like, couldn't quite get there" or "In all honesty...I feel like I don't have time to sit and really dive deep into the types of data that I'm getting."
ABDUCTION	Invoking or describing ventured explanations prompted by information reported by a (imagined or real) design, such as subjects' traits, motivations, skills, histories, and behaviors.	" 'Time between failure and exit.' It looks like they're solving every puzzle" or "most obvious ones are like the students that are not persisting, probably like logging out right away, and those that are, like, getting in it, are really good problem-solvers and are really good with their spatial skills, which are the rapid solvers."
OBSERVATION	Invoking participants' (imagined or real) observations or assessments from outside of the design, put into relation with (imagined or real) information reported by a design, such as subjects' in-person behaviors or assessments of subjects made prior to the design.	"students that are very different types of students that you've learned from other views in class...to see that they've spent the same time, but to want to know how they did it differently."
SELECTION	Invoking or processes of participants' seeking to identify subjects, design elements, or concepts for the purpose of taking action, such as identifying those with the highest or lowest scores on some metric.	"I don't really care about the kids kind of in the middle, 'cause they're doing what they're supposed to be doing, they're making progress...it's those outlier kids that need the extra attention."
ACTION	Describing participants' strategies for actions to be taken with respect to one or more of their subjects, in relation to (imagined or real) information reported by the design, such as praising successes and intervening in struggles.	"I would wanna sit down with that student and help them to develop strategies" or "this will be easy for me just to grab this and...I know I'm gonna put one rapid solver with one non-persistent."
TRANSPOSITION	Invoking or processes of (real or imagined) transpositions of the data collected by the design in order to focus on one dimension, such as on students, on design elements, or on concepts.	"I think I'd want, like, an alert of like, 'This misconception was shown by, you know, 70% of your students in the past week.' Like, 'Here's a puzzle you should look at.'"

Preliminary Findings

We found that although teachers agreed on the reason for persistence's central importance—persistence is a life skill—, translating that value into a single conceptual definition was difficult [11]. Multiple competing definitions arose, metrics were proposed, and it was the design team, not the teachers, that ultimately selected which to implement according to technical feasibility [12]. Throughout the co-design process, teachers demonstrated prior data literacy through critiques of these (proposed or implemented) metrics. And teachers interacted, or imagined interactions, with these metrics through the dashboard design. This design lent affordances and limitations to three major tasks for teachers' assessment and response practices: (1) digging deeper into the data on one data point in order to come away with a richer qualitative understanding; (2) “transposing” the data to focus on students, puzzles, or competencies; and (3) selecting stand-out data points for further action [13,14]. Teachers advanced several ways to incorporate the design into those actions and remarked on a number of motivating factors behind their interest in data dashboards. And as teachers thought through how they might "dig" to gain rich understanding, we noticed three patterns: (1) the design prompted teachers to reason abductively by venturing student-level explanations for why they might see the data that they do; (2) iterating between abductive reasoning and rich understandings appeared to lend confidence to both, a richer understanding resulting in better-held guesses, and vice versa; and (3) those rich understandings were used to both explain and be explained by imagined observations of their students in the broader classroom context [15]. Finally, it would be remiss to ignore the design team's own power in the process: we interpreted teachers' conceptions and motivations, we implemented the design, we were sometimes looked in to "check" the teachers' understandings, and teachers' may have limited their critiques in conversations with us [16].

Conclusion & Next Steps

Our next steps are to share our general findings with the teacher community to elicit emic feedback, automate coding of all session transcripts, and employ quantitative methods to elucidate relationships between teacher's assessment literacy, design requirements, use of visualization in classrooms, and the overall trajectory of that discourse [9,17]. Our hope for this analysis is that the application of Quantitative Ethnography (QE) methods to this case study will (1) provide the QE community a rich demonstration of the temporal dynamics of the values that emerge in co-design processes with teachers, (2) allow us to theoretically relate those dynamics to teachers' existing assessment literacy and practices, and (3) combine those results with our other research to develop interactive data visualizations that embody generalizable, theoretically-articulated, and empirically-grounded design principles.

Acknowledgements. This material is based upon work supported by the National Science Foundation under grant No. EAGER 1935450.

References.

1. Macfadyen, L. P., Dawson, S., Pardo, A., & Gaševic, D., Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge. Research & Practice in Assessment **9**, 17–28 (2014).

2. Ruipérez-Valiente, J. A., & Kim, Y. J.: Effects of solo vs. collaborative play in a digital learning game on geometry: Results from a K12 experiment. Computers & Education **159**, 104008 (2020).

3. Kumar, V.: 101 design methods: A structured approach for driving innovation in your organization. John Wiley & Sons (2012).

4. Cloninger, C. R., Svrakic, D. M., & Przybeck, T. R.: A psychobiological model of temperament and character. Archives of general psychiatry **50**(12), 975-990. (1993).

5. Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R.: Grit: perseverance and passion for long-term goals. Journal of personality and social psychology **92**(6), 1087. (2007).

6. DiCerbo, K. E.: Game-based assessment of persistence. Journal of Educational Technology & Society **17**(1), 17-28. (2014).

7. Ventura, M., & Shute, V. (2013). The validity of a game-based assessment of persistence. Computers in Human Behavior, 29(6), 2568-2572.

8. Glesne, C.: Becoming Qualitative Researchers: An Introduction. Pearson, Boston (2015).

9. Shaffer, D. W., Ruis, A. R.: How We Code. ICQE, CCIS **1312**, 62–77 (2021).

10. McDonald, N. et al: Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. Proceedings of the ACM on Human-Computer Interaction (2019).

11. Xu, Y., & Brown, G. T.: Teacher assessment literacy in practice: A reconceptualization. Teaching and Teacher Education **58**, 149-162 (2016).

12. Vezzoli, Y., Mavrikis, M., & Vasalou, A.: Inspiration cards workshops with primary teachers in the early co-design stages of learning analytics. In Proceedings of the Tenth international conference on learning analytics & knowledge 73–82 (2020).

13. Gibson, J. J.: The ecological approach to visual perception. Psychology Press (2014).

14. Schatzki, T. R.: The site of the social: A philosophical account of the constitution of social life and change. Penn State Press (2002).

15. Michos, K., et al: Involving teachers in learning analytics design: Lessons learned from two case studies. In Proceedings of the Tenth international conference on learning analytics & knowledge, 94–99 (2020).

16. Vasalou, A., et al: On power and participation: Reflections from design with developmentally diverse children. International Journal of Child-Computer Interaction **27**, 100241 (2021).

17. Brohinsky, J., et al: Trajectories in Epistemic Network Analysis. ICQE, CCIS **1312**, 62–77 (2021).