JOURNAL OF LEARNING ANALYTICS

Volume 6(2), 70-85. http://dx.doi.org/10.18608/jla.2019.62.5

Designing in Context: Reaching Beyond Usability in Learning Analytics Dashboard Design

June Ahn[1], Fabio Campos[2], Maria Hays,[3] and Daniela DiGiacomo[4]

Abstract

Researchers and developers of learning analytics (LA) systems are increasingly adopting human-centred design (HCD) approaches, with growing need to understand how to apply design practice in different educational settings. In this paper, we present a design narrative of our experience developing dashboards to support middle school mathematics teachers' pedagogical practices, in a multi-university, multi-school district, improvement science initiative in the United States. Through documentation of our design experience, we offer ways to adapt common HCD methods — contextual design and design tensions — when developing visual analytics systems for educators. We also illuminate how adopting these design methods within the context of improvement science and research–practice partnerships fundamentally influences the design choices we make and the focal questions we undertake. The results of this design process flow naturally from the appropriation and repurposing of tools by district partners and directly inform improvement goals.

Notes for Practice

- Prior learning analytics work has used various techniques from human-centred design ranging from user-interviews to engaging practice partners in low-fidelity prototyping. A framework of design practice that is deeply embedded in partnerships with our intended users is needed.
- Designing within the context of research–practice partnerships and improvement science initiatives helps dashboard designers balance the tensions between making interface–interaction decisions while ensuring that design aligns with existing work processes, data interpretation goals, and improvement aims.
- Common data collection techniques, such as user interviews and think-alouds, can be structured and analyzed for insights into practitioners' data sensemaking needs in addition to usability analyses to inform interface or product changes.
- The purpose of learning analytics design work should not be limited only to fidelity of implementation or adoption. Rather, an indicator of success is whether productive adaptations and local needs for LA tools can be embedded in the design itself. Partnership approaches offer unique advantages to achieving these design goals.

Keywords

Human–computer interaction, human-centred design, learning dashboards, design narratives, data sensemaking, improvement science, learning sciences

Submitted: 31.10.2018 — **Accepted:** 09.03.2019 — **Published:** 05.08.2019

Corresponding author [1]Email: junea@uci.edu Address: University of California, Irvine, School of Education, 3200 Education, Irvine, CA 92697, USA

[2] Email: fabioc@nyu.edu Address: New York University, Steinhardt School of Culture, Education and Human Development, 82 Washington Square East, New York, NY 10003, USA

[3]Email: mehays@uw.edu Address: University of Washington, Seattle, College of Education, 2012 Skagit Lane, Miller Hall, Seattle, WA 98105, USA

[4] Email: daniela.digiacomo@uky.edu Address: University of Kentucky, School of Information Science, 320 Little Fine Arts Library, Lexington, KY 40506, USA

1. Introduction

The collection, processing, and use of data to improve educator practices and learner experiences are fundamental concerns in the field of learning analytics (LA). As research and development progresses in this field, LA scholars have naturally identified a set of emerging obstacles. Researchers have called for attention to the design of platforms — such as learning dashboards

— to recognize how design-decisions have fundamental impacts on the interpretation and use of analytics (Duval, 2011; Klerkx, Verbert, & Duval, 2017). LA scholars are increasingly turning to methods in human-centred design (HCD) as a way to mitigate the misalignment between dashboard designs and their intended uses with diverse stakeholders and settings (Dollinger & Lodge, 2018; Holstein, Hong, Tegene, McLaren, & Aleven, 2018). Finally, scholars in LA are also beginning to understand that the locus of design decisions must move beyond the micro-context of person and interface to consider a broad array of factors. Such factors include diverse types of users, multiple aims and data interpretation needs, theories of cognition or learning, and the institutional, leadership, micropolitical, and broader sociocultural settings within which analytics are taken up (Alhadad, 2018; Datnow, 2000; Dawson et al., 2018; Dollinger & Lodge, 2018; Jivet, Scheffel, Specht, & Drachsler, 2018).

In this paper, we contribute to these emerging concerns by documenting our design team's process of developing dashboards for a multi-university, multi-school district, improvement science initiative in the United States. We present a design narrative, which Hoadley (2002) articulates as a way to describe the history and evolution of a design over time. A goal of a design narrative is to relay important details about stakeholders, events, trade-offs, and results of design decisions through rich description. As with any narrative, one cannot describe all details in absolute high fidelity. Instead, the aim is to present key vignettes that make explicit the implicit knowledge that the design-researcher employs to create and implement new innovations. Through this process, the goal is to help other design-researchers replicate and adapt this implicit knowledge in their future work.

We make the following contributions through documentation of our design experience. First, a key challenge facing designers in LA is knowing how to translate and adapt the knowledge and wisdom of HCD practices when developing tools such as dashboards. In this paper, we highlight how our team implicitly embedded common HCD ideas — such as contextual design and design tensions — when developing dashboards for educators (Beyer & Holtzblatt, 1999; Tatar, 2007). Making these design processes explicit, and naming their intent and affordances, helps make clear the why of a design practice beyond simply deploying techniques in a rote manner. Second, as LA design-researchers move to consider broader institutional, political, and organizational contexts, we shed light on how designing within a particular framework — research–practice partnerships (RPPs) and improvement science (IS) — fundamentally influenced the design choices we made and the focal questions we undertook. We argue that this type of contextual design process — where we are embedded within the RPP and IS team — provides different insights than acting as outside observers or designers.

A key need in future design-based analytics and dashboard research will be to systematically theorize about how LA tools can be productively adapted (not adopted) to make an impact on educational practice. Making design explicit — in the ways we model here — is a core way to develop a productive "middle space" that links human-centred design approaches and learning analytics as a field (Suthers & Verbert, 2013). Here we think of the target of the middle space as the design process itself, and between the fields of HCD and LA. We are not suggesting new methodologies as the goal (although perhaps new techniques may emerge from developing this middle space in the future). Instead, in this paper, we model how HCD practices gain a richer flavour when adapted and combined with partnership approaches in education and learning analytics.

2. Theoretical Framework

2.1. Learning Dashboards and their Implementation

The design and evaluation of learning dashboards is a major area of inquiry in learning analytics research (Duval, 2011; Klerkx, Verbert, & Duval, 2017). The hope is that dashboards display learning analytics to learners or educators in ways that build awareness of practice and transform teaching and learning. Current reviews of learning dashboard research typically organize the extant research along a few thought lines. Issues of design, form, and audience remain as major areas of concern. For example, Schwendimann et al. (2017) note that most research on learning dashboards focuses on monitoring learning data for teachers and students, with a substantial portion of these studies situated within university contexts. Additionally, the majority of literature Schwendimann et al. (2017) reviewed focused on using log data of existing platforms, such as learning management systems, in which data representations such as bar charts, line graphs, and tables were used most frequently. Verbert, Duval, Klerkx, Govaerts, and Santos (2013) also summarize LA dashboard research, highlighting that data is being visualized across different platforms (e.g., mobile devices, tabletops, etc.) and types of information, ranging from performance data, resource use, student activity, and a variety of indicators. These reviews hint at particular design considerations about how to visualize data, which data to show to educators or learners, and for what interpretation aims.

Dashboard research has also focused on sensemaking around data leading to productive action. We understand data sensemaking as a motivated and continuous process of comprehending connections and intentionally creating meaning (Klein, Moon, & Hoffman, 2006). Lee et al. (2016) describe the data sensemaking processes as "conscious efforts to achieve understanding of how to interpret visual objects and underlying content in an information visualization" (p. 499). Yi, Kang, Stasko, and Jacko (2008) argue that the basic building block of sensemaking is the generation of insights, which can be fostered not only through well-designed digital interfaces but also by promoting user engagement with the data and reducing undesired cognitive load.

Scheffel, Drachsler, Toisoul, Ternier, and Specht (2017) observe that dashboard designs should support processes including data processing, awareness, reflection, and impact, while Verbert and colleagues (2013) name awareness, reflection, sensemaking, and impact as goals for dashboard designs. Past research also illuminates how particular types of visualizations, when combined with certain routines and conditions of use, may produce undesired effects, or reinforce negative beliefs about students and instruction (Gašević, Dawson, & Siemens, 2015). Furthermore, researchers in various studies note the difficulty in actually supporting teachers to move from mere awareness to changes in practice (Few, 2006; Jivet, Scheffel, Drachsler, & Specht, 2017; Jivet et al., 2018; Verbert et al., 2013; Wardrip & Shapiro, 2016). Other studies also begin to attune to implementation as the key mechanism for productive uses of dashboards. LA designers need to attend to local coordination and how users interpret different data (Wise & Vytasek, 2017), issues such as manageability of data systems (Martinez-Maldonado et al., 2015), and the organizational and political context that users are in as they come to a dashboard platform (Ferguson et al., 2014).

The extant research informs our design process by highlighting the need to design for multiple layers of educational settings such as micropolitical, sociocultural, organizational, and policy (Buckingham Shum, 2015; Datnow, 2000; Little, 2012). As we detail below, we model how to holistically undertake this complex design process by combining a deeper understanding of HCD approaches and embedding our design-research in partnerships and improvement science contexts.

2.2. Human-Centred Design Approaches

LA designers and researchers are increasingly turning to HCD methods to adopt practical design techniques and knowledge to explore the design cycle of a learning analytics project. In the current learning analytics literature, we observe calls for theory-based and participatory approaches. Alhadad (2018) notes that designs of data visualization should be rooted in learning theories of cognition. For example, paying attention to issues of cognitive load, attention, human perception, and data literacy (e.g., learned, prior knowledge about how to read visual information) have direct implications for design such as avoiding visual clutter, and chunking data into interpretable segments on the screen. This perspective aligns with what data visualization researchers understand as visual encoding and interaction design decisions (Munzner, 2014), where choices about types of charts, interface components such as text and placement of visual markers, and interactive elements align with what one knows about cognition and perception.

However, learning analytics researchers often work in applied settings with actual educators and educational settings, where one quickly observes that design decisions at the interface level alone are not sufficient to productively promote effective uses of visual dashboards (Dollinger & Lodge, 2018). To glean more information about context and user needs, a common strategy is to utilize participatory design techniques where end-users are directly invited into the design process. Participatory design for learning itself is a burgeoning field with its own history, philosophies, and corpus of techniques and methods (DiSalvo, Yip, Bonsignore, & DiSalvo, 2017). From a practical perspective, LA researchers have used a subset of techniques such as interviewing end-users about their needs ahead of time to derive design ideas (Xhakaj, Aleven, & McLaren, 2016), employing practitioner-partners as informant designers who give feedback on designs (Fiorini et al., 2018), and engaging teachers throughout the prototyping process itself to create classroom analytics (Holstein et al., 2018).

Designing from learning theory and co-design techniques is important but limits the scope of what one might learn if one only applies these practices in a rote manner. For example, dashboards should, of course, be designed with knowledge of cognition in mind, and users should be engaged to obtain feedback to validate whether a tool is usable and relevant. However, beyond being in service of research and product development, design processes can shed light on key learning analytics issues over time. In addition to a library of techniques to deploy, design practice is also a structured process of interrogating the underlying assumptions, theoretical propositions, intended solutions, and unintended consequences of a given solution (Zimmerman, Forlizzi, & Evenson, 2007).

We borrow from the HCD literature to highlight two ways to frame this process of design as inquiry. First, a core need for any design is to map out the underlying phenomena that we are designing for; the multi-dimensional, highly interdependent, array of factors that drive behaviours and outcomes (Tatar, 2007). The more ways that LA design-researchers can develop their warrants for designs, the more robust, usable, and adaptable learning dashboards may be. From this perspective, theory-informed and participatory approaches are elements of a toolkit to triangulate ideas and develop stronger warrants. We add to this toolkit the idea of contextual design (Beyer & Holtzblatt, 1999), where the designer's aim is to understand a work process and the system (personal, social, organizational, etc.) that supports this work, in order to design new solutions.

Design-researchers who utilize contextual design will often interview users to glean information about their work process and setting, in order to surface additional details and engage users in feedback on prototypes. HCD scholars are also realizing the role of observation and cultural analysis, translating methods such as ethnography in order to develop deeper, multifaceted understandings of a problem-space (Blomberg, Giacomi, Mosher, & Swenton-Wall, 2017). Recent studies in learning analytics have utilized techniques of contextual design (Holstein et al., 2018; Xhakaj et al., 2016). In the following design narrative, we add to this toolkit by highlighting how embedding the design team within an improvement science network — designers engaging in the improvement work itself, while also designing tools — deepens our ability to understand the problem space even further than typical ways of adopting contextual design.

A second design practice that often distinguishes novice from more experienced designers for education, is identifying an effective focus for design decisions to account for the broader system that surrounds a learning practice. As noted earlier, in the design of dashboards and visualizations, there are important design decisions to be made at the level of encoding and interaction with the interface (Alhadad, 2018; Munzner, 2014). However, what does it mean to also design for the complex, interdependent structures that shape educators' work, or learners' experiences, which substantially shape how they come to and interpret data (e.g., institutional pressures, leadership, organizational politics, etc.)? Tatar (2007) offers the idea of design tensions that, alongside other design practices, is essential to adequately understand and develop technologies for educational settings. She observes that "The design of technology... must inevitably involve trade-offs and center on satisficing, finding decisions that are good enough… Design tensions conceptualize design not as problem solving but as goal balancing" (p. 415). When using a design tensions framework, the designer "identifies not a problem or solution but rather a limited resource or choice across one or more criteria. Design tensions help us search the situation for channel factors, the few crucial emergent configurations that may make or break a system" (Tatar, 2017, p. 417).

In the following design narrative, we shed light on how designing for different systems can look in practice, as we work with multiple school districts within an improvement science network. We show how seemingly simple design decisions such as the choice of a data representation, or creating a tool to align usability feedback to local work processes, are not just mundane tasks that come from theory or asking end-users, but are elevated to the level of resolving design tensions which may make or break the ultimate use and impact of a dashboard tool.

2.3. Embedding Design in Context: Aligning with Partnership and Improvement Science

An important contribution of this design narrative is the opportunity we had as design-researchers to design in context. Our process was shaped as part of a broader research–practice partnership (RPP) grounded in an improvement science (IS) framework. Two central tenets of RPPs are a commitment to a jointly negotiated problem of practice and the generation of an improvement-oriented, real-world solution by both the school district and the university research team (Coburn, Penuel, & Geil, 2013). Rather than being informed only by practitioner experience, as is the case of typical HCD methods, or driven primarily by theory, as is the case of traditional academic research, RPPs are long-term collaborations between practitioners and researchers. These partnerships are fundamentally about bringing relevant and often on-demand research to bear on contemporary problems of educational practice. In RPPs, research gains its rigour from the idea of relevance-to-practice (Gutiérrez & Penuel, 2014), where the work is made consequential through examining persistent problems of practice "in their context of development, with attention to ecological resources and constraints, including why, how, and under what conditions programs and policies work" (p. 19).

Improvement science is a recent approach to educational research that encourages the field to learn fast to implement well (Bryk, Gomez, Grunow, & LeMahieu, 2015). Responding to the omnipresent problem of how to scale up research-based knowledge to inform practice (Coburn & Stein, 2010), improvement science is led by the following six core principles (Bryk et al., 2015):

1. Make the work problem specific and user centred
2. Variation in performance is the core problem to address
3. See the system that produces the current problems
4. One cannot improve at scale what one cannot measure
5. Anchor practice improvement in disciplined inquiry
6. Accelerate improvements through networked communities

In this way, an improvement science approach to inquiry aims not only to learn quickly from iterative cycles of research, design, and implementation, but also to respond rapidly for the purposes of observable improvement in practice.

Buckingham Shum (2015) was one of the first to identify the synergy between learning analytics research and improvement science methodologies. A potential innovation from building connections between these two fields is creating a design practice that can adequately take into account the complex, interconnected systems within which LA tools are taken up, and to better marry LA design with real-world impact (Dollinger & Lodge, 2018). The following design narrative begins to articulate what this design process within improvement science looks like in practice.

3. Design Context and Methods

3.1. Design Context

Our design team is embedded within a larger, multi-institution, multi-school district, improvement network called PMR2: Practical Measures, Routines, and Representations for Improving Instruction. The network began in 2016 and is currently active, funded by a grant from the U.S. National Science Foundation. The PMR2 network involves four universities — University of Washington (UW), Vanderbilt University (VU), University of California, Riverside (UCR), and University of California, Irvine (UCI) — and three partner school districts in the United States. Our design team is led by a co-principal investigator at UCI (the first author) and a group of student user-experience designers and programmers who support the entire network. Each university institution is engaging in an RPP with a local school district. While each individual RPP attends to local contexts and problems of practice, the entire improvement network is focused on two common issues:

1. Supporting middle-school, mathematics teachers (in the United States, grades 6–8) to move from rote teaching practices to facilitating more student collaboration and discussion of mathematics, allowing for deeper student learning and conceptual understanding of math (Franke, Kazemi, & Battey, 2007; Jackson, Garrison, Wilson, Gibbons, & Shahan, 2013; Jackson et al., 2015).
2. Developing practical measures of the quality of student math discussion. Practical measures are metrics of *process data* about what happens in a math classroom such as whole class discussion, small group discussion, and other instructional moves. These measures provide fast, formative data back to teachers and their instructional coaches (Yeager, Bryk, Muhich, Hausman, & Morales, 2013). A summary of PMR2's scope and activities can be found in Figure 1.

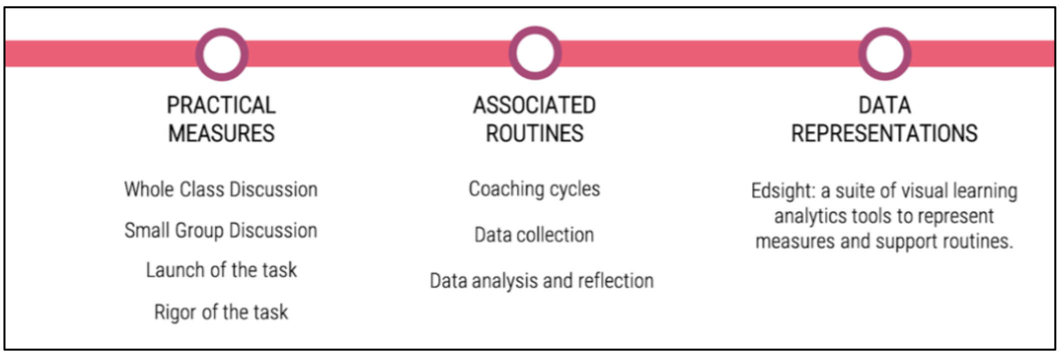

Figure 1. In the PMR2 Project, the collaborating RPP's focus on using practical measures in specific organizational routines and developing visualizations to support this work across the network.

These details about context inform our design process in a few ways. Each RPP is actively attuned to problems of practice that are unique to their local settings. For example, the UW partnership began their work in 2016 by realizing that instructional coaches wanted to focus on curriculum development first. Their theory of action continues to be that aspects of curriculum, and recommendations to math teachers about pedagogical elements such as the rigour of mathematical tasks and how they are launched, relates to student discussion of complex mathematics in the classroom. The VU partnership has taken up a slightly different problem of practice, focusing on supporting instructional coaches who are, in this context, the key bridge to supporting teachers. This focus continues to influence our future design approaches as dashboards and sensemaking for this stakeholder will take on a different flavour than a focus on teachers (which is the audience of our dashboards in the below narrative).

A common, anecdotal response we receive from other RPP teams engaged with data is that there are many tools for data collection via online methods (either automatic log data or online survey), visualization, and analysis. However, the tools often conflict with key organizational routines or intractable district policy requirements. These details may seem mundane, or filed under usability or management issues, but Tatar (2007) teaches us that these are design tensions that make or break the system. For this reason, embedding a design team within an improvement science network is advantageous because we are not designing and hyper-optimizing solutions for a single situation, as would be the case in a university course or a single MOOC platform. Rather, we are continually analyzing and balancing our design process across comparative cases (e.g., each RPP) to develop solutions that hopefully can work for a wider range of scenarios, while being directly linked to observable impacts and improvement goals.

3.2. Data Collection and Analysis

The construction of the following design narrative comes from several sources of data and information. Our design process began in January 2017 and has included weekly meetings of the design team, including PMR2 researchers, designers, web developers, and database programmers. In addition to these meetings, the broader PMR2 project team, across multiple institutions, has met weekly to coordinate project activities in an additional, cross-team meeting focused on making co-design decisions on the dashboard application. Thus, our design process has been embedded within the weekly project management routines of the initiative itself for the past two years (as of this writing).

We also employed data collection and co-design techniques that are common in HCD practice. For example, our initial corpus of data includes early, pilot interviews with partner teachers where they were presented with sample data representations of their classroom data and asked to think aloud as they interpreted and made sense of this information. Think-alouds are often described as a method suitable for tasks that involve analysis and interpretation of data. In our case, it permitted a low entry barrier for users, since they were encouraged to use their own language (Van Someren, Barnard, & Sandberg, 1994) when facing different types of charts and graphs. To obtain verbalizations that accurately reflected cognitive processes (Fonteyn, Kuipers, & Grobe, 1993), we refrained from giving detailed instructions about what to do when respondents faced a data representation for the first time (Lee et al., 2016). In addition, we regularly developed low-fidelity prototypes that were presented to RPP teams for feedback and design input (Buley, 2013).

Our team's focus on practical measures directly addresses the need for process-oriented learning analytics that can directly inform practice improvements. Recent scholars have called for more process-oriented feedback to inform practice, versus only performance-oriented data (Sedrakyan, Malmberg, Verbert, Järvelä, & Kirschner, 2018) that is limited to the level of awareness (Jivet et al., 2017). Recent LA research has taken up the design and use of practical measures to drive system-wide pedagogical improvement (Krumm et al., 2016). Our research project contributes to this emerging literature.

The broader PMR2 improvement network also convenes annual RPP meetings that bring research and practice partners across all teams together for an intensive time of information sharing, agenda setting, and joint research. During these meetings, we engaged in co-design sessions to glean more design insights for the dashboard platform. In July 2017, we conducted a series of brainstorming sessions and requirements-related consultations with our research partners to inform the initial designs of the platform. In July 2018, we conducted two co-design sessions with researchers and partner teachers, instructional coaches, and district leaders to make further design decisions on the dashboard application. We utilized a variety of techniques including interviews, sticky-noting and affinity mapping, user persona development, and journey maps. Artifacts from these sessions include audiotaped recordings of our co-design groups' discussions and their feedback on the system, and physical artifacts (such as sticky notes and paper prototypes) that speak to the ideas in the design session.

Later in our design cycle, as we developed a robust dashboard application for testing, we conducted a formative user study that involved in-depth interviews with our partner teachers and coaches. We conducted seven interviews with three teachers and four coaches from two of the RPP teams from August–October 2018. The goal of these interviews was two-fold: 1) in the more traditional UX design tradition, participants engaged in a think-aloud as they explored the features of our dashboard to provide feedback on usability issues; and 2) participants explored with the researchers how the data being represented by the dashboard were interpreted and could be used to improve practice. While the semi-structured parts of the interview allowed us to capture the context and conditions of use of our learning dashboard prototype, the think-aloud revealed several heuristics employed by practitioners to make sense of classroom data. Artifacts collected from these interviews included videotaped recordings of each video conference interview, researcher field notes, and the interview protocol.

Each of these interviews was then transcribed. To analyze our data, interviews were imported into the qualitative software Dedoose, and a subset of the research team inductively coded within and across the interviews through an open coding process that emphasized the generation and application of low-inference descriptive codes such as "teacher/coach reaction to data" (with sub-codes such as "surprise," "distrust," or "confirmation") and "pedagogical implication" (with sub-codes of "judgement," "prediction," and "strategy building"; Miles & Huberman, 1994). After primary inductive coding was complete, we collaboratively engaged in a series of conversations aimed at refining and narrowing the codes. We then identified emerging themes and patterns across the data and created analytic memos. To ensure our themes were contextualized by the real conditions of use present in each partner-school district, we analyzed our analytic memos in light of the background information we knew from our roles as part of the partnership teams. From this process, a handful of design narratives emerged, reflecting our practice partners' responses to our dashboard.

4. Vignettes of Design in Context

In the following, we present an overall description of our dashboard prototype and then highlight two cases where design decisions arose from a careful balancing of understanding our partners in context. We returned to these decisions as we gleaned more information from our user-study to shed light on issues of sensemaking with the dashboard tool, to result in new, iterative design considerations.

Our design process resulted in a web-based, dashboard application called Edsight (see Figure 2), developed with responsive web design that accommodates different device types through which educators might access the dashboard. When teachers log into Edsight (http://edsight.io), they first see an overview dashboard that displays the practical measure data they have collected thus far, in which classrooms they have the most data, and how many days have elapsed since they logged data for analysis. These orienting metrics are aimed to help teachers have a bigger picture of their overall data-driven process.

Figure 2. "Overview," "Report," and "Activity List" modes, displayed in desktop, tablet, and phone screens.

The main dashboard area is called a Report, where teachers can select dates, classrooms, and practical measures to display (Figure 3). Once a teacher filters by these selections, the dashboard displays results for different data or survey items. Up to this point, the design choices that are implicitly embedded in the screenshot artifacts reflect a wide variety of other dashboard applications. In the following vignettes, we delve more deeply to share how specific design decisions arose from our embedded process in two ways: designing for perceptual change and designing for improvement routines.

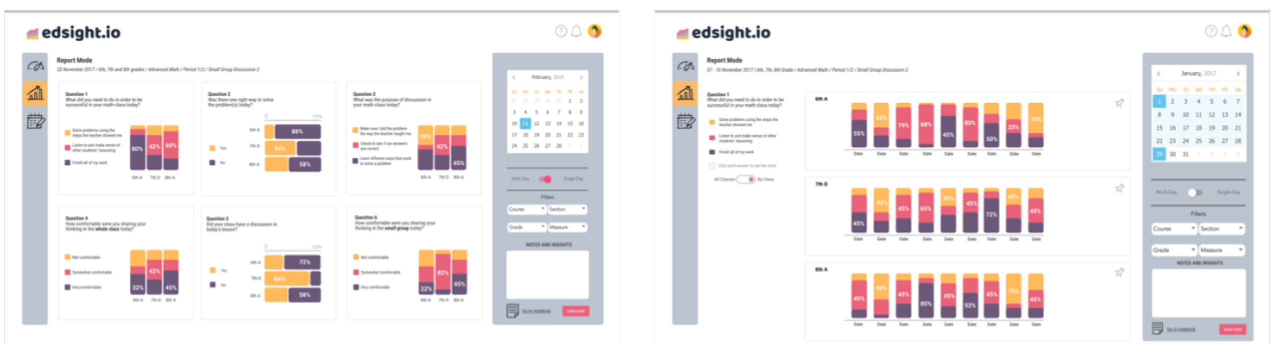

Figure 3. (L) A single-event report displayed in Edsight, with filters on the right side. (R) Multi-Day view, with ability to compare patterns across classes and sections.

4.1. Promoting Perceptual Change: Not Signalling the Right Answer

In our very first prototypes, we experimented with a variety of chart types and ways of displaying data to teachers. For example, our practice-partners requested that we display all data in pie charts, largely due to their ubiquitous availability and people's general familiarity with them. However, pie charts are often not ideal for situations where differences in proportion are less clear (e.g., 53% of one answer vs. 47% of another answer) and are also ill-suited to effectively display longitudinal patterns (Munzner, 2014; Knaflic, 2015). The nature of our practical measures data required that we communicate both proportion (which proportion of students answered in which way) and longitudinal patterns of whether improvement was occurring over time. Thus, at the data encoding level, using stacked bar charts was a natural decision (Figure 3).

The data visualization literature is replete with research and recommendations for design. However, making design decisions from theories of cognition, or best practices in data visualization design, sometimes resulted in early conflicts. In our initial prototypes and co-design process with partners, we found out that the audience and context of use we were designing for — educators working in network improvement communities — required adaptations of what are considered best practices for information visualization. One example was an idea to use techniques that more effectively direct the user's attention to salient aspects of a data visualization. In Figure 4, we present an example technique using colour in a stacked bar chart to direct attention to a specific answer choice in the data, which may orient a teacher to focus more intently on improving that answer response over time.

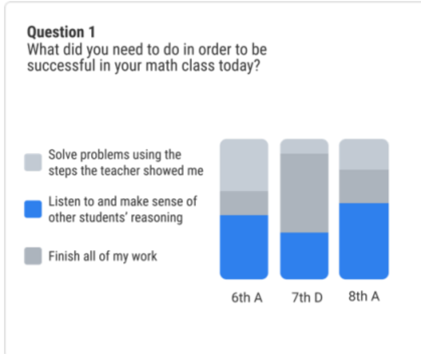

Figure 4. Using colour to direct attention to salient parts of the data.

There are many benefits to designing representations to orient users to specific elements. In a dashboard with many data representations, helping users focus on salient aspects through this sort of signalling may reduce cognitive load (Mayer & Moreno, 2003; Fiorella & Mayer, 2014; Sweller, 1988). In our case, we initially assigned visual aids, such as colour and font size, to manage the visual clutter. We wanted to prevent teachers not only from engaging with nonessential aspects of visualizations but, most importantly, to identify the response item on which to focus their improvement efforts.

We soon learned from our RPP partners that signalling preferred responses could produce undesired reactions in their work. Since we were designing dashboard tools for an improvement science process with teachers, a foundational goal was to enhance their sensemaking about their classroom processes, to provide aids to engage with instructional coaches, and to learn their way through the process. By orienting teachers to a "correct" choice, we introduced dynamics that are counter to the process. Indicating more and less "correct" choices meant interpreting results a priori for practitioners and imposing an external entity that was telling teachers they were doing something wrong and were being watched.

We also understood signalling the desired answer as a form of automation of teachers' sensemaking efforts. Echoing Merceron, Blikstein, and Siemens' (2015) conclusions, judgements made about learning data should not simply be outsourced to machines. Instead of fostering collective negotiation of meaning from the classroom data, to drive reflection and new ideas for improvement, we introduced an experience of external accountability and monitoring. Finally, a sole focus on a right answer would also promote gaming of the system, as users may be more motivated to optimize to obtain better results (e.g., telling students how to answer these practical measure surveys).

With this early feedback in mind, and through experience designing within the context of our improvement network, we were able to quickly shift to focus on balancing design tensions. We experimented with different colour schemes to allow for multiple displays of answers without overloading one's attention, and we used textual elements to orient teachers (and their coaches) to which survey answers dominated in a given classroom and day (Figure 5). In this way, teachers could easily see patterns across time and classrooms and spur collaborative discussion and sensemaking.

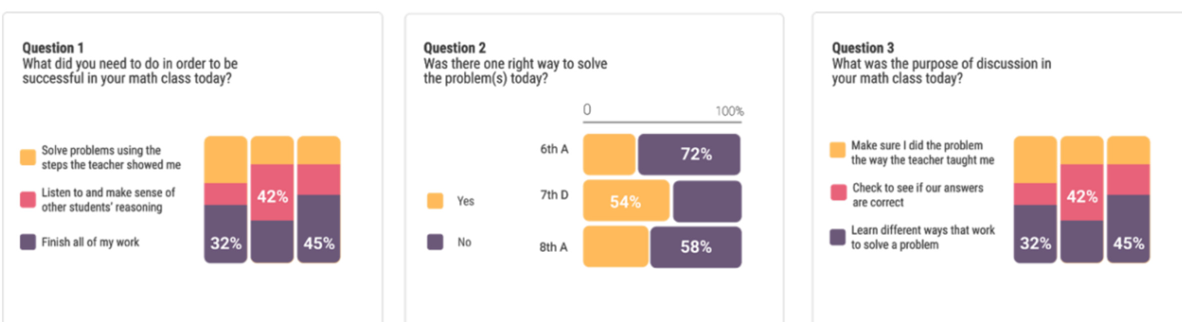

Figure 5. Examples of different questions, where the desired answer is not signalled to the user. Colour choices and textual elements focus attention to prioritize sensemaking and discussion.

We saw this example as a way that being embedded in an improvement network, working closely with partners, helped us gain a richer understanding of a work process and context than we would have by undertaking efforts such as only interviewing users. Furthermore, the vignette illustrates ways that design tensions (e.g., between standards of effective data representation vs. the needs of the practice situation) result in more grounded design decisions.

4.2. Designing for Improvement Routines

Organizational routines are key to the enactment of improvement efforts. For example, the VU partnership focused on supporting instructional coaches in the district to work with teachers on their practice in what the partnership called coaching cycles. These routines involved collaborative tasks between coach and teacher including the following:

- Negotiating the goals of a coaching cycle (between a teacher and coach)
- Selecting and scheduling classes and times for observation and data collection
- Planning lessons to try new ideas, and then collecting data about how students experienced new pedagogical practices
- Looking at the data together, and participating in coaching and professional development meetings
- Negotiating strategies to bridge potential gaps between expected and actual results

A key realization is that the data (practical measures) and the tools to support such coaching cycles (the dashboard) should amplify this process as much as possible while reducing potential obstacles. A substantial amount of the design team's initial process involved creating interfaces that would support these workflows. Making the coaching-cycle experience even slightly more burdensome would be a make-or-break issue. Designing for this workflow provided a good entry point to also engage with the organizational, social, and political factors that shape this practice in our different districts. We present two example features that illustrate seemingly small but key design decisions that support work routines: *Scheduling and Journaling.*

Scheduling: Through several iterations with our practice-partners, we ultimately decided that a key choice was to make scheduling prominent in the interface. In the coaching cycles, deciding when to collect data, selecting classrooms to collect data from, and committing to this initiation of the cycle were major milestones to meet. This particular routine is fundamental because it feeds the subsequent steps by injecting data into the system. Thus, a scheduling interface is immediately prominent on the right side of the interface when a teacher logs in (Figures 2 and 3). Further, the overview dashboard orients teachers to quickly understand what data they have, which classrooms they've examined (and which they've not), and how long ago they've engaged in a data cycle. When teachers schedule a new data collection (via the calendar interface), they also select which measures to deploy, which grades, classes, and sections to schedule for (supporting the selection process). The widget also has a field for pre-survey notes, opening an important avenue for reflection and logging goal definitions for the cycle, classroom and pedagogy notes, and shared notes between teachers and their coaches. These notes feed into another key routine around which we centred our design efforts: *journaling.*

Journaling: Bryk et al. (2015) suggest the importance of reflective strategies to support pedagogical improvement, hence journaling. An initial design conjecture for our team was that only displaying data representations would not be sufficient to support the deep, reflective cycles of sensemaking that occur between teachers and coaches. Seeing and making sense of data is one process, but keeping records of the surrounding discussions, plans, and goals from the data — in close proximity to the data representations themselves — would be important to support a sustained improvement process. Prior research on teachers' sensemaking with data suggests that it does not happen in a vacuum, but rather is highly influenced by previous knowledge, beliefs, interactions, and ancillary forms of information (Wardrip & Herman, 2018). As a result, we designed features on the dashboard that we hoped could foster this type of reflection, interpretation, and collaboration with the potential to lead to data-informed pedagogical improvement (Jivet et al., 2017).

We developed a Notebook feature that would keep a running log of notes (Figure 6). The notebook is organized as an open-ended tool so that teachers and coaches can adapt the feature for their own, local improvement goals. Teachers and coaches are prompted to include pre-class notes (i.e., notes about class objectives and expectations, taken when teachers negotiate instructional goals and schedule a data collection event), and post-class notes (i.e., notes taken when users analyze a report of classroom data). We made a conscious design decision to display timestamps to each entry, so that teachers can keep track of their thinking. To facilitate further sensemaking and reflection, pre- and post-survey notes are displayed side by side for each survey event. Finally, we included a special field to allow for notes about multiple events, as coaches and teachers often want to examine longitudinal trajectories of their work.

5. Understanding Sensemaking of Data

To this point, the prototype of the Edsight dashboard tool was the culmination of about a year of design activity (as detailed above). We then engaged in a rapid user study to validate our design decisions and collect formative data about the user experience with the dashboard. Here we want to illustrate the use of rapid, formative user studies not only to inform surface design decisions (such as changing an interface element), but to also inform broader, theoretical understanding about key issues surrounding our dashboard design.

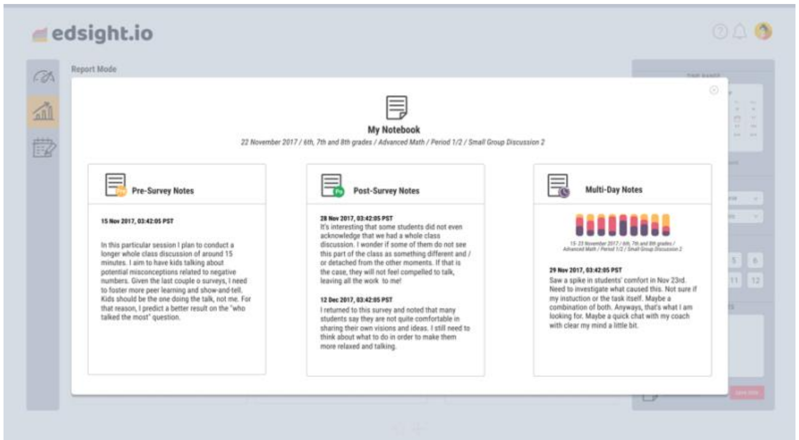

Figure 6. Notebook feature, with pre-survey, post-survey, and multi-day notes.

A key question for both our practice partners and our research team, is to understand if and how sensemaking and classroom decision-making can be supported through our designs. From our user interviews, a few themes emerged that will inform the next phase of our dashboard design-research and also suggest key challenges that are salient for future dashboard designs.

5.1. Building Supports for Remembering

One theme in particular surfaced in nearly all of our think-alouds and interviews. Practitioner partners all voiced challenges in recalling past events when looking at dashboards of classroom data from a retrospective view. Teachers consistently reported having difficulty remembering the particular lesson, curricular units, or student interactions represented by a given piece of data. As a consequence, teachers scanned the dashboard's various interface elements for supplementary information connected to each data event. For example, when asked about how they interpreted a particular data point, teacher partner Janice (all names are pseudonyms) shared, "It makes me curious. I want to go back to my plan book and find out what I was actually doing that day. And how I could have changed it because this is what I want." Janice's response is illustrative of the pattern that emerged around teachers wanting to upload or retrieve notes, as well as artifacts such as handouts or lesson plans in order to situate the figures and visual representations they encountered. We inductively coded this pattern as "teacher/coach desires ancillary information" and noted it emerged across all of the think-alouds.

The same theme emerged when we interviewed instructional coaches, with one coach, Tracy, noting, "That's why I was saying it's [the process of coaching with this data] a combination of this data with notes, with just conversational debriefing that helps us to look at the whole picture." She continued by saying: "Gosh, having an old class list in front of me would help [in] trying to put myself back in that class frame. Um, do you mind if I grab an old class list? It's right behind me." Without ancillary fragments of information that aided them in comprehending the broader classroom context in which prior instructional events occurred, coaches were often not able to form an opinion or judge teacher practices, or they were hesitant to do so.

These insights from user interviews validate some of our original design conjectures and decisions around the need to better support note taking. Such insights are also consistent with what Wardrip and Herman (2018) have found around the importance of "informal data" as requisite tools to think with when analyzing data to inform teaching. For future design and research, these findings illuminate a need to understand how to better promote the uptake of archiving this informal data. Questions that we anticipate asking include those of the sensemaking for valid interpretation action type, such as What kinds of notes or artifacts promote the most productive remembering?, How should informal data be collected, stored, and used as tools for remembering?, and Who should collect it? We contend that these questions can be most robustly understood by continuing to emphasize design in context, including ongoing and iterative feedback sessions with partner teachers, coaches, and students, in practice. From within such a position, embedded within an improvement science oriented research–practice partnership, co-design techniques can help designers move beyond just making decisions from their own mental models or theory, to focusing the locus of design on high-impact concerns for practice.

5.2. One Chart, Many Meanings: Exploring Stories of Practice

Another illuminating theme in our participant interviews was observing the wide variation in their interpretations of the same data representations. This phenomenon of one visualization eliciting numerous attributions, stories, conclusions, and subsequent actions has been extensively documented by the fields of learning analytics and information visualization. By utilizing a think-aloud protocol, we set out to understand the range of interpretations that our partners brought to the dashboard

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

tool. For a variety of reasons, nearly half of those we interviewed expressed doubt as to whether the data they were seeing was reflective of what actually occurred in a given classroom (as they remembered it). Consider Marie's example when we asked her if she thought that the data accurately reflected her practice over time:

> *Probably, yes and no. Like, again, the kids are getting tired of the survey. So they are probably just trying to hurry up and finish it, and I know that the timing also would have affected how seriously they took it. I know a lot of the times we gave it with a couple of minutes to spare because I wasn't watching the time and I had forgotten that we transition early and I thought I had more time and I so I think that also kind of affected it, yeah. (Think-aloud with Marie, October 2018)*

Marie's words surface an element of uncertainty toward the data representations regarding the reflectiveness of the data in relation to what may have been occurring in the actual classroom. We inductively coded this theme as "teacher/coach interpretation: partial reality of a moment in time." The above quote illuminates how details of practice within our RPP helps us understand the implications of our tools. Marie's hesitation seems to be explained by her own data collection strategies: with less time left for administering surveys and students taking them less seriously ("I know that the timing also would have affected their how seriously took it… I had forgotten that we transition early"). An interesting question to consider is how to design for or support these nuances of practice.

Another representative response of this type came from instructional coach Bonnie, when we asked her about why she would need more information to make sense of the dashboard data representations:

> *So, there's different angles to the truth. You have different perspectives. You have students' perspectives, and sometimes a student will answer the way they think you want them to answer. Sometimes they may be interpreting a question in a different way than the way it was intended. That's a piece there. There's the teacher's perspective. The teacher, in his or her mind, [has] preconceived notions about the class and what will happen going in, and those flavour things. And then there's an outside observer's perspective, who may or may not be truly an outside observer. For example, with the coaching, I see teachers at multiple points during the year, so I have multiple data points to use, to think about, what I might see, and what I'm seeing in the data. And in the information that's there. Where [the researcher], coming in for the first time, may not have those ideas and that background knowledge. (Think-aloud with Bonnie, September 2018)*

As Bonnie's response illustrates, interpretation of data is similarly subject to the user's perspective on the validity of the data itself, as well as how it is positioned and understood amongst other pieces of information about persons and practice.

Often users made meaning of the data representations by attributing causality to student behaviour, teacher performance, or external factors. Consider Coach Leslie's response to our question about how she would use the data to support improvement of her teacher's instruction, in particular regarding the practical measure survey question related to problem solving:

> *It looks like in February students were kind of pretty much even about solving this problem the way the teacher showed them versus making sense of other students' reasoning. In March, it seems like there was some confusion, maybe in the purpose of the lesson because some students didn't have a response, or they wanted to finish all their work. And there still seems to be a large group that want to solve the problem the way the teacher showed them. There might be a problem in the way the lesson was described or explained, or maybe the purpose. (Think-aloud with Leslie, September 2018)*

Leslie's response is illustrative of the type of varied meaning-making that coaches and teachers engaged in when viewing the data representations. In Leslie's case, she understood the data as representative of the range of students' problem-solving practice. At the same time, she expresses a number of different reasons as to why many of the students answered in a particular way to the practical measure question ("maybe a problem in the teacher's... or the way the lesson was described, or the purpose"). In the case of coach Theresa, who also understood the data as mostly reflective of actual practice, she attributed the students' response patterns to the district-wide emphasis on socioemotional learning, as well as a particular teacher's instructional behaviour (e.g., small group discussion preceding whole group discussion).

> *This [chart] is telling me [about] the comfort of students sharing out in [the] whole group. In this case, no one was not comfortable with it at all. A lot of kids were somewhat comfortable, and some were very comfortable. Which that's actually a great thing, that no one just totally felt uncomfortable… I do know that [the teacher] focused a lot on having them to discuss in small groups first before they became whole groups, that could be part of why that helped with that comfort. (Think-aloud with Theresa, August 2018)*

For Theresa, then, as well as Leslie, the data representations served as useful tools to help think about how their teachers were organizing their classrooms. At the same time, what our think-aloud analysis made visible was the wide range of stories of practice that both teachers/coaches expressed when viewing the data representations.

This feedback from user interviews, combined with our roles as RPP-embedded designers, deeply knowledgeable about our partner contexts, helps inform future design questions for our team. A critical area for exploration will be to decide where, in the system, to support these varied ways of interpreting the same data representation. For example, we may decide to scaffold particular types of data interpretation on the interface of Edsight, which is often a first option that comes to mind for designers. However, by designing within an RPP and IS team, we can also naturally see that our partners themselves might be the most effective scaffold to guide productive sensemaking through their existing routines (e.g., instructional coaching cycles etc.). The ultimate decision of which mechanisms and locus of design are most effective are likely to be dependent on the local context and actors one is partnering with, thus highlighting the affordances of embedding dashboard design in partnership and improvement work.

6. Discussion

In the above design narrative, we outlined various ways that design practice can occur when developing data dashboards for educators. We focus on a few vignettes using rich description, from a multitude of possible design moments that occurred in the project, to contribute key insights to the development of learning analytics dashboards.

A common way that learning scientists and learning analytics researchers might approach design is to create solutions from one's own mental models or theory, interviewing users to obtain requirements and contextual knowledge, or seeking feedback on analytics and interfaces. In our design experience, we observed that *designing in context* through the normal workflow of partnership research, offers unique affordances in comparison to positioning the designer as separate from the research or practice partners. In particular, design in context may be a more efficient way to deeply understand diverse users and use cases, problems of practice and systemic factors such as organizational, political, and social dynamics that ultimately impact learning analytics adoption in a school district, than those afforded by decontextualized design practices. Furthermore, we see great promise in embedding design work within emerging models of research–practice partnerships and improvement science models because being embedded as design-researchers provides rich, implicit, understanding of the local contextual factors and design tensions that may be easily lost through only doing interviews or a few observations to inform design.

Past research in learning analytics has focused on documenting how improvement metrics could be derived (Krumm et al., 2016) and scholars have articulated the synergy between LA and these RPP or IS models (Buckingham Shum, 2015). We add to this nascent literature by documenting how *learning analytics design practice* can look as scholars and practitioners work together. In addition, we articulate how established HCD practices can be embedded in a research–practice partnership between researcher-designers and education practitioners.

Designing a learning dashboard embedded in partnership work (RPP and IS approaches) revealed nuances not initially anticipated by our team. We began our process equipped with tools and techniques from existing HCD practices, such as empathy and experience mapping, participatory design, think-alouds, rapid prototyping and iterative user testing. The internal dynamics of an RPP, combined with external pressure and other forces at the district level, exposed how the standard HCD toolkit needs to be reinforced with other ideas and techniques. In other words, crafting a suite of learning analytics *for* and *with* the direct participation of teachers, district leaders and instructional coaches involves recognizing and attending to an additional layer of practice within the overall design process. We see this extra layer — the RPP/IS design loop — not as a concurrent process but one that *encases*, *guides* and *gives meaning* to regular HCD practices when designing for district partnerships (Figure 7).

One salient feature of this multilayered process concerns the time employed to generate solutions and responses to district needs. Based on past experience and initial design expectations, we anticipated shorter design sprints, with rapid cycles of prototyping, testing, and hypothesis validation. Our experience as embedded within an RPP proved that relying solely on traditional "in-the-lab" user testing techniques was not enough. We soon realized that the amount of time required to accurately comprehend practitioner's problems of practice — and how they might or might not justify new features or data representations — should be extended to real-world improvement cycles. This realization did not mean the complete disregard of quicker, in-the-lab cycles but the addition of "in-the-wild" user testing (Martinez-Maldonado et al., 2015): evaluating designs and functional prototypes within teachers' and coaches' practices and with an extended timeframe. The longer timeframe employed in our design process, and the elimination of the stimulus–response focus that is commonplace in regular lab tests, allowed our design ideas to be assimilated and appropriated by our partners.

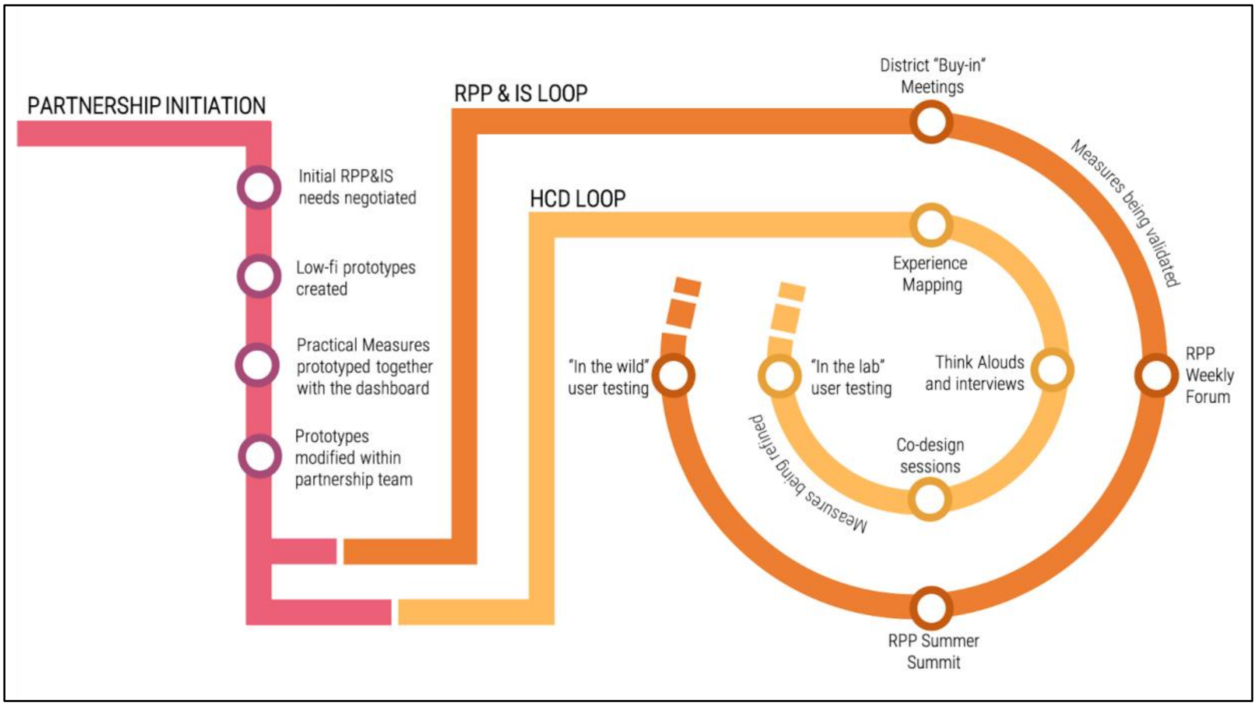

Figure 7. A model relating HCD design practices to RPP and IS routines.

The amplification of the design timeframe and the acknowledgement of an overarching RPP/IS design loop not only made researchers and practitioners more cognizant of needs and routines they seek to respond to but influenced the very nature of the problems themselves. The current theory of action mostly present in the learning analytics literature flows from the identification of a problem, to the creation of a suite of tools (such as a dashboard), to a desired change in teaching or learning (Martinez-Maldonado et al., 2015) or even from practitioner awareness to instructional impact (Verbert et al., 2013). A linear fidelity of implementation process is implied as a core outcome.

Within a research–practice partnership, or in the context of an improvement network of districts and schools, this theory of action might not be enough. In reality, we observe a less linear chain of events. Instead we see the design process unfolding with the problem itself being changed by its intended solution, and the solution being constantly repurposed by the user. In our research-design process, for instance, we noticed how key routines, such as coaching cycles and survey scheduling, were refined and even re-conceptualized by the constant interaction between the researcher, coaches, or teachers in relation to the various prototypes and the testing procedure itself. In such cases, the focus shifts away from fidelity of implementation to designing for **robust adaptation and local appropriation of tools**. With this focus in mind, partnership approaches to research and design provide an advantageous location to design for these outcomes.

The systemic design work necessary for RPP/IS contexts resonates with Judith Little's (2012) work about data use in schools. Little upholds that studies about teachers' use of data generally rely on interviews, self-reports, diaries, user logs, and surveys. The field does not account for the dynamics of practice relating to how analytics are in fact employed for instructional improvement. Through her micro-process perspective, however, "investigators delve into the ways that the members of an organization create shared meaning and activity through interaction, employing methods of observation that capture what individuals do with one another and with the relevant tools and objects of a field" (p. 6). Potvin, Kaplan, Boardman, and Polman (2017) contribute to this view by maintaining that the process of co-designing with practitioners (e.g., a curriculum or a digital tool) "in the wild" is an ever-morphing process that will attune research-designers to yet undiscovered user necessities and needs of adaptations. This is precisely the locus of design in which our RPP/IS work is situated.

Finally, while reiterating that the final design of a dashboard is not the goal of this paper, we consider the necessity of reflecting about how our team evaluates the results of our work so far. A common way that designers measure the outcomes of their practice is by observing adoption metrics such as number of users or total time logged in a platform or digital product. In the case of a partnership-embedded work, however, we have come to appreciate other metrics of success by considering adaptation (not adoption alone) as a design goal. The iterative nature of the design process will provide solutions remarkably different from the original intents. When this process is productive — when it engages school partners in a manner that generates strong, locally relevant insights and designs — we see our work as conducive to improvement in educational practice, which is the ultimate goal and hope of the learning analytics endeavor.

Acknowledgments

We thank the network of collaborators and partners in the PMR2 team. We acknowledge the invaluable contributions of our peers at University of Washington, Vanderbilt University, UC Riverside, UC Irvine, New York University, and Stanford University. Our special thanks to all of the teachers, instructional coaches, and leaders from our partner-districts who, amidst the challenges inherent to school life, found time and energy to collaborate intensely with our team of researchers.

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported by the National Science Foundation, through grants 1719744, 1620851, 1621238, and 1620863.

References

Alhadad, S. S. (2018). Visualizing data to support judgement, inference, and decision making in learning analytics: Insights from cognitive psychology and visualization science. *Journal of Learning Analytics, 5*(2), 60–85. http://dx.doi.org/10.18608/jla.2018.52.5

Beyer, H., & Holtzblatt, K. (1999). Contextual design. *Interactions*, *6*(1), 32–42.

Blomberg, J., Giacomi, J., Mosher, A., & Swenton-Wall, P. (2017). Ethnographic field methods and their relation to design. In J. Simonsen & T. Robertson (Eds.), *Routledge international handbook of participatory design* (pp. 123–155). New York: Routledge. https://doi.org/10.1201/9780203744338

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Cambridge, MA: Harvard Education Press. https://doi.org/10.1002/sce.21223

Buckingham Shum, S. (2015, May 4). Learning analytics meet improvement science [Blog Post]. Retrieved from https://medium.com/@sbskmi/learning-analytics-meet-improvement-science-66748565bcc4

Buley, L. (2013). *The user experience team of one: A research and design survival guide*. New York: Rosenfeld Media.

Coburn, C. E., & Stein, M. K. (2010). Research and practice in education: Building alliances, bridging the divide. Lanham, MD: Rowman & Littlefield.

Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013, January). *Research–practice partnerships: A strategy for leveraging research for educational improvement in school districts.* New York: William T. Grant Foundation.

Datnow, A. (2000). Power and politics in the adoption of school reform models. *Educational Evaluation and Policy Analysis 22*(4), 357–374. https://doi.org/10.3102/01623737022004357

Dawson, S., Poquet, O., Colvin, C., Rogers, T., Pardo, A., & Gašević, D. (2018). Rethinking learning analytics adoption through complexity leadership theory. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK '18), 5–9 March 2018, Sydney, NSW, Australia (pp. 236–244). New York: ACM. https://10.1145/3170358.3170375

DiSalvo, B., Yip, J., Bonsignore, E., & DiSalvo, C. (2017). *Participatory design for learning.* New York: Routledge.

Dollinger, M., & Lodge, J. M. (2018). Co-creation strategies for learning analytics. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK '18), 5–9 March 2018, Sydney, NSW, Australia (pp. 97–101). New York: ACM. https://10.1145/3170358.3170372

Duval, E. (2011). Attention please! Learning analytics for visualization and recommendation. In P. Long, G. Siemens, G. Conole, & D. Gašević (Eds.), *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (LAK '11), 27 February–1 March 2011, Banff, AB, Canada (pp. 9–17). New York: ACM. https://doi.org/10.1145/2090116.209011

Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014). Setting learning analytics in context: Overcoming the barriers to large-scale adoption. In *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (LAK '14), 24–28 March 2014, Indianapolis, IN, USA (pp. 251–253). New York: ACM. https://doi.org/10.1145/2567574.2567592

Few, S. (2006). Information dashboard design: The effective visual communication of data. Sebastopol, CA: O'Reilly Media.

Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology, 39*(2), 75–85. https://doi.org/10.1016/j.cedpsych.2014.01.001

Fiorini, S., Sewell, A., Bumbalough, M., Chauhan, P., Shepard, L., Rehrey, G., & Groth, D. (2018). An application of participatory action research in advising-focused learning analytics. In *Proceedings of the 8th International Conference* *on Learning Analytics and Knowledge* (LAK '18), 5–9 March 2018, Sydney, NSW, Australia (pp. 89–96). New York: ACM. https://doi.org/10.1145/3170358.3170387

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research, 3*(4), 430–441. https://doi.org/10.1177/104973239300300403

Franke, M. L., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225–256). Greenwich, CT: Information Age Publishers.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Gutiérrez, K. D., & Penuel, W. R. (2014). Relevance to practice as a criterion for rigor. *Educational Researcher*, *43*(1), 19–23. https://doi.org/10.3102/0013189X13520289

Hoadley, C. P. (2002). Creating context: Design-based research in creating and understanding CSCL. In *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community* (CSCL 2002), 7–11 January 2002, Boulder, CO, USA (pp. 453–462). International Society of the Learning Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates. https://dl.acm.org/citation.cfm?id=1658679

Holstein, K., Hong, G., Tegene, M., McLaren, B. M., & Aleven, V. (2018). The classroom as a dashboard: Co-designing wearable cognitive augmentation for K–12 teachers. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK '18), 5–9 March 2018, Sydney, NSW, Australia (pp. 79–88). New York: ACM. https://doi.org/10.1145/3170358.3170377

Jackson, K., Garrison, A., Wilson, J., Gibbons, L., & Shahan, E. (2013). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education*, *44*(4), 646–682. https://doi.org/10.5951/jresematheduc.44.4.0646

Jackson, K., Cobb, P., Wilson, J., Webster, M., Dunlap, C., & Appelgate, M. (2015). Investigating the development of mathematics leaders' capacity to support teachers' learning on a large scale. *ZDM: Mathematics Education*, *47*(1), 93–104. https://doi.org/10.1007/s11858-014-0652-5

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, M. Pérez-Sanagustín (Eds.), *Data Driven Approaches in Digital Education: Proceedings of the 12th European Conference on Technology Enhanced Learning* (EC-TEL 2017), 12–15 September 2017, Tallinn, Estonia (pp. 82–96). Lecture Notes in Computer Science, Springer. https://doi.org/10.1007/978-3-319-66610-5_7

Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018). License to evaluate: Preparing learning analytics dashboards for educational practice. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK '18), 5–9 March 2018, Sydney, NSW, Australia (pp. 31–40). New York: ACM. https://doi.org/10.1145/3170358.3170421

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems*, *21*(4). https://doi.org/10.1136/bmj.2.3022.599

Klerkx, J., Verbert, K., & Duval, E. (2017). Learning analytics dashboards. In C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.), *The handbook of learning analytics* (pp. 143–150). Beaumont, AB: Society for Learning Analytics Research (SoLAR). https://doi.org/10.18608/hla17

Knaflic, C. N. (2015). Storytelling with data: A data visualization guide for business professionals. Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781119055259

Krumm, A. E., Beattie, R., Takahashi, S., D'Angelo, C., Feng, M., & Cheng, B. (2016). Practical measurement and productive persistence: Strategies for using digital learning system data to drive improvement. *Journal of Learning Analytics*, *3*(2), 116–138. https://doi.org/10.18608/jla.2016.32.6

Lee, S., Kim, S. H., Hung, Y. H., Lam, H., Kang, Y. A., & Yi, J. S. (2016). How do people make sense of unfamiliar visualizations? A grounded model of novice's information visualization sensemaking. *IEEE Transactions on Visualization and Computer Graphics*, *22*(1), 499–508. https://doi.org/10.1109/TVCG.2015.2467195

Little, J. W. (2012). Understanding data use practice among teachers: The contribution of micro-process studies. *American Journal of Education*, *118*(2), 143–166. https://doi.org/10.1086/663271

Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., & Clayphan, A. (2015). Latux: An iterative workflow for designing, validating, and deploying learning analytics visualizations. *Journal of Learning Analytics*, *2*(3), 9–39. https://doi.org/10.18608/jla.2015.23.3

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Merceron, A., Blikstein, P., & Siemens, G. (2015). Learning analytics: From big data to meaningful data. *Journal of Learning Analytics*, *2*(3), 4–8. https://doi.org/10.18608/jla.2015.23.2

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook.* Thousand Oaks, CA: Sage Publications. https://doi.org/10.1016/s1098-2140(99)80125-8

Munzner, T. (2014). *Visualization analysis and design*. (A. K. Peters visualization series). Boca Raton, FL: CRC Press/Taylor & Francis Group. https://doi.org/10.1201/b17511

Potvin, A. S., Kaplan, R. G., Boardman, A. G., & Polman, J. L. (2017). Configurations in co-design: Participant structures in partnership work. In B. Bevan & W. R. Penuel (Eds.), *Connecting Research and Practice for Educational Improvement: Ethical and Equitable Approaches* (pp. 135–149). Abingdon, UK: Taylor & Francis. https://doi.org/10.4324/9781315268309-9

Scheffel, M., Drachsler, H., Toisoul, C., Ternier, S., & Specht, M. (2017). The proof of the pudding: Examining validity and reliability of the evaluation framework for learning analytics. In *Proceedings of the 12th European Conference on Technology Enhanced Learning* (EC-TEL 2017), 12–15 September 2017, Tallinn, Estonia (pp. 194–208). Lecture Notes in Computer Science, Springer. https://doi.org/10.1007/978-3-319-66610-5_15

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., ... & Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, *10*(1), 30–41. https://doi.org/10.1109/tlt.2016.2599522

Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2018). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*, in press. https://doi.org/10.1016/j.chb.2018.05.004

Suthers, D., & Verbert, K. (2013). Learning analytics as a middle space. In *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (LAK '13), 8–12 April 2013, Leuven, Belgium (pp. 1–4). ACM. https://doi.org/10.1145/2460296.2460298

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Tatar, D. (2007). The design tensions framework. *Human–Computer Interaction*, *22*(4), 413–451. https://doi.org/10.1080/07370020701638814

Van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical approach to modelling cognitive processes*. Academic Press. London. http://hdl.handle.net/11245/2.149552

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, *57*(10), 1500–1509. https://doi.org/10.1177/0002764213479363

Wardrip, P. S., & Herman, P. (2018). "We're keeping on top of the students": Making sense of test data with more informal data in a grade-level instructional team. *Teacher Development*, *22*(1), 31–50. https://doi.org/10.1080/13664530.2017.1308428

Wardrip, P. S., & Shapiro, R. B. (2016). Digital media and data: Using and designing technologies to support learning in practice. *Learning, Media and Technology*, *41*(2), 187–192. https://doi.org/10.1080/17439884.2016.1160929

Wise, A. F., & Vytasek, J. (2017). Learning analytics implementation design. In C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.), *The handbook of learning analytics* (pp. 151–160). Beaumont, AB: Society for Learning Analytics Research (SoLAR). https://doi.org/10.18608/hla17

Xhakaj, F., Aleven, V., & McLaren, B. M. (2016). How teachers use data to help students learn: Contextual inquiry for the design of a dashboard. In *Proceedings of the 11th European Conference on Technology Enhanced Learning* (EC-TEL 2016), 13–16 September 2016, Lyon, France (pp. 340–354). Lecture Notes in Computer Science, Springer. https://doi.org/10.1007/978-3-319-45153-4_26

Yeager, D., Bryk, A., Muhich, J., Hausman, H., & Morales, L. (2013). *Practical measurement*. Palo Alto, CA: Carnegie Foundation for the Advancement of Teaching.

Yi, J. S., Kang, Y. A., Stasko, J. T., & Jacko, J. A. (2008, April). Understanding and characterizing insights: How do people gain insights using information visualization? In *Proceedings of the 2008 Workshop on BEyond time and errors: Novel evaLuation methods for Information Visualization* (BELIV '08), 5 April 2008, Florence, Italy (Article No. 4). New York: ACM. https://doi.org/10.1145/1377966.1377971

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '07), 28 April–3 May 2007, San Jose, CA (pp. 493–502). New York: ACM. https://doi.org/10.1145/1240624.1240704