Making Sense of Sensemaking: Understanding How K-12 Teachers and Coaches React to Visual Analytics

5 authors, including:

Fabio Campos
New York University
15 PUBLICATIONS **61** CITATIONS

Ha Nguyen
University of California, Irvine
24 PUBLICATIONS **50** CITATIONS

Maria Hays
University of Washington Seattle
4 PUBLICATIONS **33** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   The Conference of The Birds View project

Making Sense of Sensemaking: Understanding How K–12 Teachers and Coaches React to Visual Analytics

Fabio C. Campos[1], June Ahn[2], Daniela K. DiGiacomo[3], Ha Nguyen[4], Maria Hays[5]

Abstract

With the spread of learning analytics (LA) dashboards in K–12 schools, educators are increasingly expected to make sense of data to inform instruction. However, numerous features of school settings, such as specialized vantage points of educators, may lead to different ways of looking at data. This observation motivates the need to carefully observe and account for the ways data sensemaking occurs, and how it may differ across K–12 professional roles. Our mixed-methods study reports on interviews and think-aloud sessions with middle-school mathematics teachers and instructional coaches from four districts in the United States. By exposing educators to an LA dashboard, we map their varied reactions to visual data and reveal prevalent sensemaking patterns. We find that emotional, analytical, and intentional responses inform educators' sensemaking and that different roles at the school afford unique vantage points toward data. Based on these findings, we offer a typology for representing sensemaking in a K–12 school context and reflect on how to expand visual LA process models.

Notes for Practice

- At K–12 schools, educators play different roles, such as teachers and instructional coaches. Understanding how these roles offer distinct vantage points toward data and how they shape sensemaking is fundamental to developing human-centred learning analytics.

- Training coaches to facilitate sensemaking with teachers is key for K–12 instructional improvement. Providing a common language to observe, categorize, and discuss reactions to data allows coaches and school leaders to design professional development interventions attuned to teachers' practices and needs.

- Visual analytics tools need to account for the possibility of "over-fitting"—or the normalization of new instructional information into teachers' pre-existing views—and facilitate question-asking and information-seeking behaviours to mitigate this risk.

Corresponding author [1]*Email: fabioc@nyu.edu Address: New York University, Steinhardt School of Culture, Education and Human Development, 82 Washington Square East, New York, NY, USA. ORCID ID: https://orcid.org/0000-0001-5499-5312*

[2]*Email: junea@uci.edu Address: University of California, Irvine, School of Education, 3200 Education, Irvine, CA, USA. ORCID ID: https://orcid.org/0000-0003-0987-8410*

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

[4]*Email: thicn@uci.edu Address: University of California, Irvine, School of Education, 3200 Education, Irvine, CA, USA. ORCID ID: https://orcid.org/0000-0001-7138-1427*

[5]*Email: mehays@uw.edu Address: University of Washington, College of Education, 2012 Skagit Lane, Seattle, WA, USA. ORCID ID: https://orcid.org/0000-0001-9359-6955*

1. Introduction

Using classroom data to inform instruction has become a necessary competency for K–12 school educators in the United States (Mandinach, 2012). Dashboards that display teaching and learning data are becoming ubiquitous in K–12 institutions, allowing educators to visually display and keep track of relevant metrics and events. A question persists about *how* exactly different school educators build interpretations based on such data.

Sensemaking is key for human information processing (Dervin, 1998). Learning analytics (LA) and human-computer interaction (HCI) researchers have attempted to conceptually explain how visual analytics tools influence teaching and learning outcomes for teachers and students. These models, however, have yet to account for the multiple steps taken by K–12 educators to frame an instructional situation and adequately explain their representational world. Similarly, existing theoretical models tend to group many different activities under the banner of sensemaking, such as reflection, awareness, insight generation, and interpretation, thus losing precision in detailing such key processes. Current theoretical models also do not emphasize how interpretations may vary depending on the different roles an educator plays in K–12 schools. Assuming all educators go through similar sensemaking practices—as well as taking sensemaking in the K–12 environment as a given—may conceal significant particularities involved in this process and lead to user-agnostic tools that fall short of enabling instructional change.

In this paper, we provide a nuanced view of how middle school mathematics teachers and their instructional coaches (both often called "educators" hereafter) make sense of classroom data delivered through dashboards. Our study is the result of a larger network of four research-practice partnerships (RPPs) between universities and school districts (Coburn, Penuel, & Geil, 2013). Over the course of two years, our research team co-designed a visual analytics dashboard with educators from each RPP. The dashboard displays survey data related to students' collaboration and discussion in classrooms and aims to improve educators' knowledge about students' mathematical thinking (Figure 1). A core research need in this broader endeavour is to better understand how educators make sense of these visualizations as part of their practice.
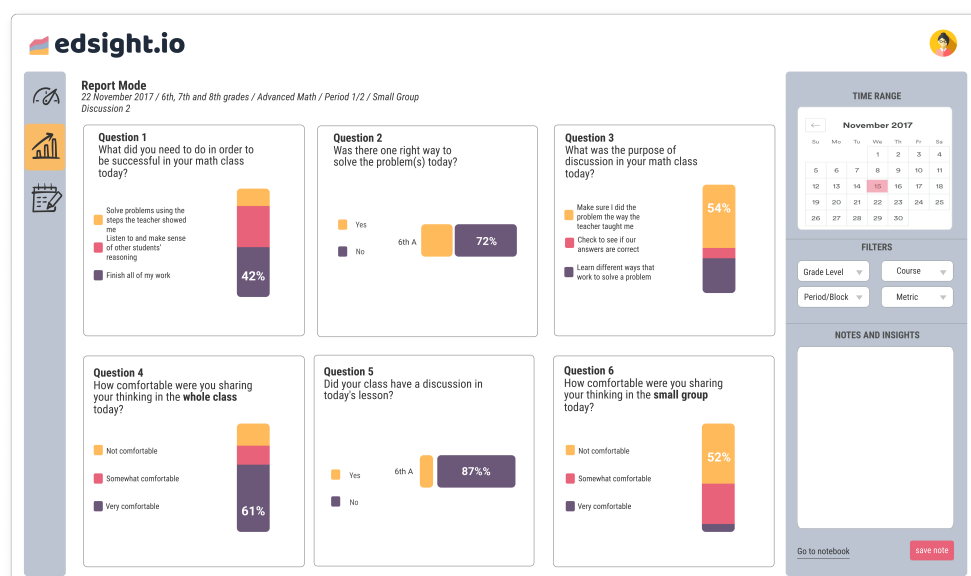

Figure 1. Edsight's report page, showing an example dashboard of a single-day classroom report.

Within the context of our RPPs, we conducted think-alouds and semi-structured interviews with teachers and instructional coaches to understand how they made sense of data delivered through a dashboard. Teachers in our sample are responsible for direct interaction with students and instruction in their mathematics classroom. Instructional coaches are experienced educators, specially trained to offer on-site, subject-specific instructional support for school teachers (Marsh, Sloan McCombs, & Martorell, 2010). The interaction between teachers and coaches—a well-documented capacity-building intervention in US K–12 schools (Marsh, Bertrand, & Huguet, 2015)—often involves continuous cycles of data analysis and the design of corresponding instructional responses (Farrell & Marsh, 2016). In this study, math teachers and their coaches were presented with practical measures (PMs)—frequent measures to support and improve teachers' daily practice (Yeager, Bryk, Hausman, Muhich, & Morales, 2013)—collected through in-class surveys. These measures represent how students experience specific aspects of instruction, such as small-group or whole-class discussion. With minimal guidance from our team, educators were prompted to make sense of reports of various measures, delivered by the dashboard, representing one or more classroom events.

Our analysis draws from areas beyond the LA and HCI fields to characterize the sensemaking process in K–12 schools. From the organization and information studies fields, we draw on descriptions of sensemaking in organizations—paying special attention to how interpretations are achieved as a *result* of sensemaking, and how taking *action* may be an independent, yet correlated, process (Weick, 1995). This paper also draws from education researchers who investigate the conversational nature of sensemaking in the K–12 context (Horn, Kane, & Wilson, 2015), particularly when coaches are involved in supporting teachers. We view this specific dynamic—teachers being "coached" to make sense of learning data to improve instruction—as

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)

2

unique to K–12 schools and an opportunity to expand current LA theories of action by exploring the different vantage points people bring to data.

By closely examining K–12 educators' data sensemaking practices, this study makes three contributions. First, we offer a *fine-grained typology* of sensemaking reactions to visual data in school settings. Although sensemaking is often described as central to planning and delivering instruction (Xhakaj, Aleven, & McLaren, 2016), descriptions of reactions to data in the LA and education literature have been limited in granularity thus far. Based on the understanding that LA research can support professional development at the school level (van Leeuwen, van Wermeskerken, Erkens, & Rummel, 2017), we argue that a detailed typology will allow researchers and school leaders to support teacher sensemaking more appropriately by providing nuanced ways to observe, identify, and collectively discuss reactions to data.

Second, we add to the growing body of work that explores how human and organizational factors inform data sensemaking. The broader HCI literature has explored experiences with data in multiple contexts, but with limited studies in K–12 school settings that are often indifferent to educators as a particular type of user. The LA literature has focused on the education domain but only recently started exploring issues related to human-centred design (Buckingham Shum, Ferguson, & Martinez-Maldonaldo, 2019; Ahn, Campos, Hays, & DiGiacomo, 2019). With this study, we aim to bridge this gap by exploring *sensemaking patterns, particular to a specific context and group of users*: teachers and instructional coaches in K–12 school environments. By revealing the most prevalent cases from our research, we foreground how different vantage points act as frames of reference, filtering and shaping the sensemaking process.

Finally, we posit that sensemaking does not happen in a vacuum, devoid of organizational influences, nor is it a purely individual endeavour but a cyclical process in which educators create authorial accounts of instructional events. Based on our findings, and leveraging previous LA models, we offer an expanded *process model* of K–12 sensemaking of visual analytics. Our findings, presented in Section 4, highlight the critical need to understand the range and structure of reactions to a given data representation and their implications for LA theory and practice.

2. Literature Review

2.1 Expanding the Notion of Sensemaking

In a world where reality changes constantly and gaps in information abound, individuals are continuously making efforts to interpret complexity, diversity, and uncertainty (Dervin, 1998; Yi, Kang, Stasko, & Jacko, 2008). Sensemaking is generally theorized as the process through which people ascribe meaning to their experience in the world (Klein & Moon, 2006). This phenomenological view implies that, no matter the nature of information, the visible world will likely be interpreted in highly variable ways (Dervin & Naumer, 2010; Dervin, 1998; Klein & Moon, 2006).

Researchers summarize sensemaking as a *process*—and thus not a single act—that involves identifying patterns, making connections, creating meaning, and anticipating trajectories effectively (Klein & Moon, 2006). The idea of meaning *making*—as opposed to meaning *comprehension*—is also supported by Weick (1995, 1993), who views sensemaking as an active authoring of reality, as opposed to the mere discovery of something hidden in numbers, text, or events. In Weick's (1995) model, *interpretations* describe the resulting product of the larger sensemaking process. Treating sensemaking and interpretation as a single act wrongly implies that meanings already exist and are merely codified in their various representations, waiting to be discovered (Brown, Colville, & Pye, 2015). In a similar way, education scholars posit that data alone cannot speak for itself, nor can it show an optimal path to educators without an active, intentional process of sensemaking (Farrell & Marsh, 2016; Shapiro & Wardrip, 2015; Mandinach, 2012).

Individuals organize their lived experiences and perceptions of reality in the form of stories (Bruner, 1991). From this process, one might convert ambiguous or equivocal events of the past into *plausible retrospective accounts* (Weick, 1995; Brown et al., 2015; Maitlis, Vogus, & Lawrence, 2013; Colville, 2007). In this sense, one major threat to sensemaking might be the over-rationalization of past events and their representations and the *normalization* of lived reality into pre-existing categories (Maitlis et al., 2013).

In our analysis, we approach sensemaking as a process not solely influenced by objectivity and rationality, but one prone to be shaped by subjectivity and emotions. First, sensemaking is highly contingent upon frames of reference, cognitive filters that organize and ascribe meaning to newly acquired information (Sherin & Russ, 2014; Klein & Moon, 2006). Dervin and Naumer (2010) put forward the idea that individuals process new information only when anchored in their own stories and frames of reference. Second, individuals' affective states might relate to filtering the content that will ultimately be analyzed (Bower & Cohen, 1982; D'Mello, 2017; You & Kang, 2014; Wise & Jung, 2019; Maitlis et al., 2013). Maitlis and colleagues (2013), for example, proposed that the valence of emotions (i.e., negative or positive) significantly fuels and affects how sensemaking is triggered (or hampered). Whereas negative emotions might ignite sensemaking more easily, positive states might lead to normalizing evidence into existing accounts of reality. Following this logic, our study also analyzed several emotional responses from teachers and coaches when faced with visual analytics.

The information visualization (InfoVis) and HCI communities have broadened the understanding of how users make sense of information delivered by data-rich tools. Studies in these fields are typically concerned with evaluating either visualizations (e.g., design decisions) or their underlying processes (e.g., user experience) (Lam et al., 2012). A common finding speaks to how specific features of a tool or visualization may affect or augment resulting interpretations (Larkin & Simon, 1987; Scaife & Rogers, 1996; Grammel, Tory, & Storey, 2010; Romat et al., 2019; Yi, Kang, Stasko, & Jacko, 2007). However, studies that delve into the nuanced practices that educators undertake when making sense of learning data are only beginning to emerge in the HCI and InfoVis fields. This gap justifies the descriptive and non-evaluative approach we adopt in this paper.

Sensemaking, by itself, may not lead to action. While actionability is the desired outcome of data-rich tools and routines (Farrell & Marsh, 2016; Mandinach, 2012), pedagogical activities—or the lack thereof—are not measures of sensemaking effectiveness. Making decisions and taking actions based on data is a complex process, likely to be affected and constrained by factors such as organizational policy, hierarchy, team dynamics, and individual capabilities (Weick, 1995; Horn et al., 2015; Mccoy & Shih, 2016). In essence, individuals might gain insights or build vivid stories from data without necessarily acting on practical decisions based on what they learned. This view resonates with that of Wise and Jung (2019), who describe sensemaking as a process that goes beyond taking pedagogical action per se. Their findings suggest that data sensemaking might result in pedagogical reflection, thus suggesting "the potential of analytics to support instructors in engaging in a deep inquiry process that reshapes their pedagogical knowledge and fosters their professional growth" (p. 63). Recognizing that sensemaking and action are two different but connected processes guides our analysis and bounds the scope of this paper.

2.2 Data Sensemaking in K–12 Schools

K–12 schools in the US (from which we draw our participants) present pedagogical, cognitive, and organizational features that influence how teachers attribute meaning to learning data (Sherin & Russ, 2014; Schildkamp, 2019). Factors such as pedagogical perspectives (Coburn & Turner, 2011), a priori knowledge (Simon, 2017), and extant views of students' capabilities (Jackson, Gibbons, & Sharpe, 2017), as well as organizational dynamics, such as control over data (Little, 2012), have been described as being critical to understanding K–12 sensemaking. How these dynamics play out is unique to the K–12 level (e.g., they will play out differently in higher education).

For example, in a K–12 setting, curricular standards and external accountability pressures create significant demands for educators, often shaping the culture of data use in schools (Frank, Kim, Salloum, Bieda, & Youngs, 2020; Farrell & Marsh, 2016). Teachers and coaches are expected to make sense of instructional data to set classroom goals, track student progress, re-design a course, or identify new pedagogical strategies, among other tasks (Farrell & Marsh, 2016). To support these complex activities, educational leaders have invested in capacity-building policies at the school and district levels. One such policy—instructional coaching—has become a cornerstone in reform policies across the US, often employed in support of instructional improvement efforts at scale (Marsh et al., 2010; Russell et al., 2020).

Instructional coaches are defined as "master teachers who offer on-site and ongoing instructional support for teachers" (Marsh et al., 2010, p. 873). Typically, coaching programs are school based, organized around a specific subject (e.g., literacy coaches, mathematics coaches), and aimed at helping teachers make sense of various forms of learning data (e.g., assessment, dashboard visualizations, classroom videos, student work) and determine adequate instructional responses (Farrell & Marsh, 2016). In the US, mathematics coaching has been associated with a focal shift from the rote demonstration of solutions to the development of students' conceptual understanding (Russell et al., 2020). This move results in specific practices and routines at the school, namely individual and group coaching, and the development of professional learning communities (PLCs), generally conceptualized as small groups of educators who gather in organized cycles of data inquiry (Russell et al., 2020; Cobb, Jackson, Henrick, & Smith, 2018). In our study, mathematics coaches propose goals, observe classes, give structured feedback, and co-design instructional strategies with teachers, both individually and organized in math PLCs.

The interactions between coaches and teachers are often dialogue intensive (Horn et al., 2015) and use learning data as a starting point. These deep and specific conversations revolve around content, pedagogy, and student thinking and occur over lesson-planning or debriefing sessions (Russell et al., 2020). To that end, both teachers and their coaches are expected to notice salient aspects of classroom data (Erickson, 2011; Van Es & Sherin, 2002). Based on what they notice, school educators may make connections to—or fill in the gaps of—external information. However, this conversation-based sensemaking at the K–12 level (Horn et al., 2015) depends on the role an educator assumes at a school (Mandinach, 2012). This implies that teachers and coaches will potentially manifest different vantage points when approaching learning data. Understanding how school roles influence sensemaking as observed in verbal responses to data is one of the major purposes of this study.

Building capacity for sensemaking requires supporting and reorganizing teachers' and coaches' *conversations* (Murnane, Sharkey, & Boudett, 2005). However, developing a lingua franca—or a shared understanding of patterns and responses to learning data—requires detail and granularity. While education scholars have studied how teachers generally respond to data, theoretical models in these areas are often restricted to top-level categorical descriptors. For example, researchers might describe response types to assessment data as being "conceptual" (i.e., data is used to enrich conversations about instruction),

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)

4

"symbolic" (i.e., data is used to justify a pre-existing stance), or "instrumental" (i.e., data is used to make instructional decisions) (Murnane et al., 2005). Despite their contributions to categorizing educators' responses, such models fall short of providing enough granularity, which we may need in order to better understand people's responses to data and design tools for finer-tuned supports.

Finally, education research has identified major cognitive structures that K–12 educators employ when planning future instruction. Borko, Livingston, and Shavelson (1990), for instance, describe teachers' thinking as organized around three main structures: scripts (i.e., information about familiar situations), scenes (i.e., experiential knowledge of classroom events), and propositional structures (i.e., factual knowledge about key components of teaching, such as pedagogies, content knowledge, and students' abilities). Similarly, Horn and colleagues (2015) find that teachers and coaches create common ground around instructional practices by collectively *replaying* past events and *rehearsing* future episodes. Both studies reveal that educators approach learning data grounded in their experiences, contexts, and frames of reference, which is a foundational concept for our analysis. In our work, we subscribe to the notion that sensemaking happens retrospectively, with users first grasping the big picture and then detecting patterns to ultimately cross the bridge between their vantage points and what is displayed by the data (Yi et al., 2008).

In summary, K–12 sensemaking is a complex process that involves making connections and filling the gaps between specific information and the larger context it refers to (Alhadad, 2018). Understanding the nuanced ways K–12 educators make sense of learning data is critical for LA design and theory. In the following section, we discuss how current LA theoretical models might need to be further detailed to better reflect practices and routines on the ground.

2.3 LA Theories of Action

Conveying a model that accounts for the granularity and variability of educators' sensemaking is still a major opportunity for LA researchers (Duval, 2011). LA process models, summarized in Figure 2, often describe steps flowing from instructional events to data visualizations (particularly through dashboards); followed by one or more stages, such as awareness, reflection, and/or interpretation; and culminating in some form of instructional decision (Bocala & Boudet, 2015; Jivet, Scheffel, Drachsler, & Specht, 2017; Lee et al., 2016; Van Es & Sherin, 2002). Scheffel, Drachsler, Toisoul, Ternier, and Specht (2017), for instance, name steps such as data processing, awareness, reflection, and impact, while Verbert, Duval, Klerkx, Govaerts, and Santos (2013) describe a similar model that involves awareness, reflection, sensemaking, and impact. Similarly, Xhakaj and colleagues (2016) hypothesized a causal chain representing how dashboards may affect how educators plan and prepare for a class session.
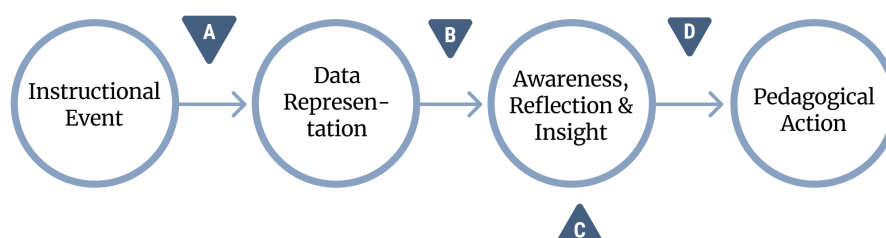

Figure 2. A summary of LA process models.

At the core of these theoretical representations is the assumption that educators will make sense of instructional situations to establish potential courses of action. However, information delivered through LA dashboards can only be used to adjust the course of instruction if educators can properly translate data displays into pedagogical insights (Molenaar & Knoop-van Campen, 2018; Xhakaj et al., 2016). Taking sensemaking for granted may hide numerous particularities involved in this process and result in a misrepresentation of how educators reason and plan future instructional moves.

A careful analysis of extant models representing LA sensemaking reveals key opportunities for further investigation (Figure 2, marked with letters A to D). First, design and computational decisions might directly affect the way instructional events are represented through graphs and charts (A). For example, allowing for more or less agency over visual representations (Peck, Ayuso, & El-Etr, 2019; Ahn, Nguyen, & Campos, 2021), understanding how specific representations generate distinct responses (Voyiatzaki & Avouris, 2014), and creating more transparent algorithms (Holstein, Mclaren, & Aleven, 2019; Sun, Mhaidli, Watel, Brooks, & Schaub, 2019; Chen & Zhu, 2019) are among the most documented design decisions by LA and HCI scholars but rarely make their way into theoretical process models. The second area of opportunity (B) lies between data representations and the sensemaking processes. LA scholars, such as Van Leeuwen, Rummel, and van Gog (2019), have made significant advances in understanding how different interpretive aids, hardwired in data representations (e.g., dashboards), might support

teacher sensemaking. However, we seek additional understanding of how different educator roles in K–12 settings might engage in sensemaking according to their specific vantage points and the responsibilities they entail.

Another area of opportunity (C) refers to the need for a deeper understanding of the steps involved in sensemaking through visual LA, which is one of the gaps this paper seeks to address. One of the few studies in LA that describe fine-grained responses to data comes from Voyiatzaki and Avouris (2014), who mapped several responses to data elicited by teachers while supervising computer-supported collaborative activities. Their study, although focused on synchronous classroom orchestration, found several data sensemaking heuristics, such as *comparing*, *monitoring*, and *exploring* information displayed by a digital tool. Likewise, Molenaar and Knoop-van Campen (2018) explored the question of how teachers make data actionable and detailed educators' pedagogical responses to data delivered through dashboards. In our work, we seek to build on this growing body of research, which seeks to open the black box of sensemaking and establish a more nuanced model of educators' responses to data representations of learning events.

Finally, D refers to understanding how resulting interpretations might facilitate pedagogical actions, which itself is an interesting opportunity for future research (e.g., see Molenaar and Knoop-van Campen (2018)). However, we noted earlier how sensemaking and action can be distinct processes. In the present study, we bound and focus our analysis on opportunities B and C. Although shown for clarity and completeness of the model, A and D are beyond the scope of this particular study.

In summary, K–12 educators are expected not to simply "comprehend" information but to notice, frame, connect, and fill in the gaps when making sense of data. Considering the critical need to *map* and understand *how* these interpretive moves are employed, this study builds upon current LA theories of action by pursuing the following research question:

> **RQ: What are K–12 educators' types and patterns of responses to visual representations of data?**

3. Methods

3.1 Context of Investigation

Our research team is part of a larger network of RPPs (Coburn et al., 2013) called PMR2: Practical Measures, Routines and Representations. The network started in 2016 and was funded by the US National Science Foundation. At the time of this study, the PMR2 network involved four RPPs, comprising four universities and four school districts in different regions of the US. The network is focused on supporting middle school math teachers (6th to 8th grades) to use data to facilitate student collaboration and discussion, allowing for a richer conceptual understanding of mathematics.

In a typical RPP, the focus of investigation is jointly negotiated between district partners and researchers (Coburn & Penuel, 2016). In our case, the RPPs focused on improving instructional practices and routines of math teachers and their coaches. These are precisely the two school-based roles that frequently engage in data sensemaking at the classroom level. Although other K–12 educators—such as district leaders and subject specialists—participate in our RPPs, their potential reactions to data are beyond the scope of our analysis.

3.1.1 PMs

One of the network's core activities is the development of PMs of classroom instruction. These formative measures return fast and actionable data to teachers and their coaches to inform instructional action, versus performance-oriented data that might be limited to the level of awareness (Yeager et al., 2013). PMs are meant not to measure learning or satisfaction but to indicate how students experience specific aspects of instruction. In our case, these aspects include the quality of whole-class and small-group discussions conducted by the teacher to improve mathematical thinking among students (Jackson, Henrick, Cobb, Kochmanski, & Nieman, 2016).

In the PMR2 research network, PMs are collected through student-facing surveys—containing around eight questions each—administered by teachers in various moments of a class period. For example, students evaluate their experience participating in a whole-class discussion by responding to a survey at the end of a session, often as an exit slip. Examples of questions involved in these measures are "Were you comfortable sharing your thinking in today's whole-class discussion?" and "Who talked the most in today's whole-class discussion?", among many others. All questions have multiple-choice answer options. In most cases, responses are collected electronically, via Google Forms; in a smaller number of classrooms, data is collected using pen and paper (data from both formats was used in this study). Student data is then anonymized, processed, and delivered to the educators in the form of a dashboard that visualizes single- and multi-day "classroom reports" (Figure 3).

3.1.2 The Visual Analytics Tool

As part of PMR2, our research group developed Edsight, a suite of visual analytics tools that not only delivers classroom reports but also allows for multiple queries, comparisons, and note taking (Figure 1). Students' identities are never disclosed to either coaches or teachers in Edsight's interface. One of the key objectives of our research group is to introduce Edsight as a

sensemaking tool to support the work of coaches and teachers, especially by providing a space for reflection about particular aspects of instruction. Other types of data beyond PMs, such as attendance and grades, are currently not represented in our tool.
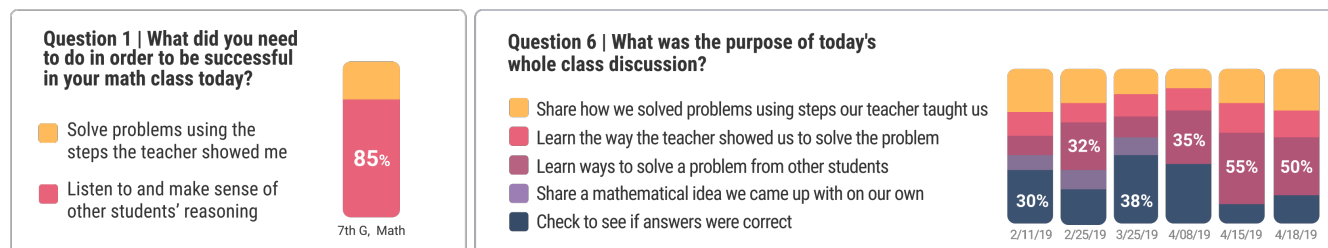

Figure 3. Left: A survey item from a single-day report. Right: A survey item from a multi-day report.

3.1.3 Study Participants

To examine sensemaking practices with our educator partners, we recruited a purposeful sample of middle school mathematics teachers ($n = 9$) and instructional coaches ($n = 9$) from a pool of 74 educators spread across the four districts that participate in our network of RPPs (Table 1). Given the exploratory nature of our research question, our sample is not proportional to the size of each local cohort of educators.

In our sample, the majority (83%) of participants were female ($n = 15$; 9 coaches, 6 teachers). Teachers had an average of 10 years of professional experience as educators (SD = 9.6) and no experience with visual analytics tools whatsoever. All coaches from our sample were experienced educators, often recruited from the districts' teacher pools, with an average of 17 years of experience (SD = 7.3). Only one instructional coach reported previous experience using LA dashboards.

Table 1. Study participants by district and gender

District	Teachers		Coaches		Total
	M	F	M	F	
South 1	1	4	0	0	5
South 2	0	0	0	5	5
Northwest	2	2	0	2	6
West	0	0	0	2	2
Total	**3**	**6**	**0**	**9**	**18**

Within our RPPs, coaches and teachers work together in different types of inquiry cycles, from large professional development meetings to one-on-one sessions. Coaches are customarily expected to support teachers to reflect upon different aspects of instruction and understand the necessary steps for improving it. The example below illuminates the type of continued support provided by coaches and their particular vantage point toward learning data:

> **Coach J.** *I coach about five to six individual teachers right now. They are [the] teachers who struggle the most. Three of the teachers are not new to the profession. (…) However, the issue is that they're working extremely hard in the classroom, but it's not showing up on test scores. So, our whole focus has been where's the disconnect and how can we fix it. (…) The issue though on the back end is being able to sustain whatever change we put in place, so I would see the change happen for about two weeks, and we will see success. And then, come back in a week and then it's like 'Okay, what happened to what we talked about?'*

3.2 Data Collection

To capture educators' full range of responses to data visualizations, we conducted think-aloud sessions and semi-structured interviews with teachers and coaches. Most of the sessions occurred through video conferencing ($n = 12$) due to budgetary and scheduling restrictions—especially considering our partner districts were located in four different states. The remaining sessions occurred in person ($n = 6$). All sessions lasted approximately one hour—including both think-aloud and interview components—and were audio-recorded and transcribed by our team.

When real classroom events were accessible (e.g., graphs based upon surveys collected from students), we presented their respective visual representations to either the teacher or the corresponding coach. When real events were not accessible— particularly due to local conditions of the RPP—we presented educators with example graphs (modelled after actual data from a classroom, but anonymized) and invited them to role-play a situation that involved making sense of data. Example graphs

were built to reflect similar reports collected from other teachers who participate in one of the RPPs. Most of the participants saw their own data ($n = 12$, 67%), and 33% ($n = 6$) saw example data based on similar classroom contexts. We established the validity of such visual representations by consulting with our network of partners (both researchers and district leaders). This feature of the research process also offered some level of external validity, as teachers in our K–12 contexts often look at their own classroom data but also make sense of others' or example data (e.g., in professional development sessions or department meetings). Each participant was exposed to approximately 15 visualizations, distributed in single- or multi-day reports.

3.2.1 Procedure

In each session, we asked participants to analyze data reports and to think aloud as they interacted with graphs. In addition, we asked semi-structured, open-ended questions that explored how the data was more or less representative of their practice. While the interviews allowed us to explore contextual factors, the think-alouds revealed the interpretive processes employed by practitioners. To obtain verbalizations that accurately reflected cognitive operations, we refrained from giving detailed instructions to participants or directly asking how the data could be used to inform instructional changes (i.e., at least until after the think-aloud section of the interview was completed). Instead, we invited practitioners to verbalize their thoughts with prompts such as "Could you tell us what you notice?" and "What does this data tell you?" This decision not only provided us with more details about how educators react to each data visualization but, most important, revealed their strategies for moving across charts and building a cohesive chronicle of classroom events in their own terms.

Our choice for employing think-aloud protocols is explained by the need for a dialogic method (Dervin & Naumer, 2010) suitable for tasks that involve building interpretations (Fonteyn, Kuipers, & Grobe, 1993; Charters, 2003). In our case, it permitted a low entry barrier, since users were encouraged to use their own language when facing different types of data visualizations. Think-aloud protocols have also been widely used by InfoVis researchers as a method of directly tracing users' thinking. Liu and Stasko (2010) argued that such protocols are key to studying the internalization and processing of new information. The efficacy of the method, however, depends on maintaining a free flow of thought. Likewise, directly stimulating responses may result in biased data that fails to reproduce how teachers create interpretive narratives outside the confines of the lab (Van Es & Sherin, 2002; Lee et al., 2016).

3.3 Data Analysis

Data analysis was organized in six stages:

Stage One. Each interview was transcribed and imported into the qualitative software Dedoose. To accurately balance the units of analysis, we divided the transcripts into smaller excerpts related to one single interpretive act (Sherin & Russ, 2014). Sometimes these units encompassed more than one dashboard item (i.e., educator was comparing across items). In other occurrences, our unit of analysis comprised reactions to a single data visualization. One example of a unit of analysis, which shows a response to "What does it tell you?", can be found below:

> ***Coach Z.*** *Well, it looks like in February some students weren't comfortable sharing their thinking. And all those students now feel more comfortable. But there's still a pretty big discrepancy between being very comfortable and somewhat comfortable. But there was growth there so that's good.*

Stage Two. We conducted an initial inductive analysis of the data, looking for how educators reacted to single-day or multi-day reports of student surveys. After this initial phase, our team started grouping some matching responses into larger categories, which we understood as dimensions of educators' sensemaking acts (Sherin & Russ, 2014; Lee et al., 2016). Our decisions were also informed by past work in the LA, education, and organization studies fields. For example, we decided to include an *emotional* dimension to capture potential affective reactions to data, resonating with the work of Wise and Jung (2019). The other two dimensions were *analytical*, aimed at capturing analytical reasoning and story building about the data, and *intentional*, where we would group instructional change plans and intentions (Maitlis et al., 2013). Since our study did not follow educators back into schools, we did not include a *pedagogical action* category but limited ourselves to capturing verbalized plans and devised activities. The way these dimensions interconnect will be explored in the Findings section (Section 4).

Stage Three. With the dimensions and units of analysis defined, we started an open coding process that emphasized the generation of low-inference descriptive codes (Miles & Huberman, 1994). Examples include *empathy for students* in the emotional dimension, *asking questions* in the analytic dimension, and *improvement plan* in the intentional dimension, among many others. In this stage, we coded all occurrences within each excerpt, without limiting the total number of codes. After analyzing the initial results, our team noted that more than one emotional or intentional code rarely occurred in the same excerpt. Conversely, analytical codes typically occurred once or twice within the same excerpt, but rarely above this mark. This observation led our team to define rules for the final round of coding, as presented below in Stage Five.

Table 2. A typology of educators' responses to visual data

Dimension	Response Type	Definition
Emotional	Positive Surprise	Educators see results superior to expectations.
	Negative Surprise	Educators see results inferior to expectations.
	Satisfaction	Educators are satisfied with what the data shows.
	Empathy	Educators consider what motivated students to provide a certain answer.
	Distrust	Educators distrust the data, the instrument of collection, or the visualization.
	Responsibility	Educators feel directly responsible for the results.
	No Response	Educators do not express any emotion related to the data.
Analytical	Re-stating	Educators re-state what is represented by the data in their own terms.
	Asking Questions	Data prompts educators to ask further questions. Includes acts of curiosity and wondering.
	Recall	Data makes educator remember or backtrack to past events.
	Attribution of Cause	Educators try to attribute the cause of the result to something external to the data.
	Confusion	Educators are confused by what the data displays.
	Gap Filling	Educators fill in the gaps of the data with opinions, beliefs, stances toward teaching, etc.
	Connecting the Dots	Educators connect multiple instances or aspects of the data.
	Comparing	Educators compare more than one feature or instance of the data.
	Part vs. Whole	Educators interpret a part of the result as representative of the whole, or vice versa.
	Prediction	Educators predict how results would happen under different conditions.
	Multiple Lenses	Educators consider the fact that data can be analyzed from multiple angles.
	Judgment	Educators formulate an opinion (e.g., about other teachers, students) based on data.
	Confirmation	Educators confirm a belief or prediction based on what they interpret.
	No Response	Educators do not convey any form of analysis of the data.
Intentional	Action Intention	Educators communicate the intention of improving their (or someone else's) instruction.
	Planning	Educators list actions they would like to take in order to improve instruction.
	Seek Information	Educators communicate the need to know more (e.g., causes) to act upon the data.
	No Response	Educators don't mention potential actions of instructional change.

Stage Four. After the initial inductive coding was finalized, the research team collaboratively engaged in refining and narrowing the codebook. A major contribution of this study is a nuanced typology of educators' sensemaking; thus the development and refinement of the codebook into the final typology was a core foundation of our analysis (Table 2). In the refinement process, codes with negligible occurrences in the data were merged into other more significant codes or eventually discarded. One example is *trust*: while *distrust* of data occurred often, almost no educators spontaneously manifested trust in the data. Other codes, such as *wondering* and *hypothesizing*, were merged into either *asking questions* or *prediction*, depending on the nature of the response.

Stage Five. With a final codebook created, we conducted a second analysis. This time, each excerpt was coded in all three dimensions, using the following structure: one code for emotional, one or two codes for analytical, and one code for intentional responses. This scheme was developed based on observations made on the open coding phase of our analysis (Stage Three). It is important to mention that, in the absence of an emotional, analytical, or intentional response, we coded *no emotional response*, *no analytic reaction*, or *no improvement plan*. These codes not only represent observed situations but also reflect extant findings in the LA literature (Wise & Jung, 2019). Three researchers participated in this phase, with at least two coding each transcript. Inter-rater reliability indicated substantial agreement (Cohen's kappa = 0.82), and any disagreements were resolved through consensus (Miles & Huberman, 1994).

Stage Six. Finally, we compared the occurrence of codes per professional role (coaches vs. teachers). To determine whether differences of medians between teachers and coaches were statistically significant, we ran a Wilcoxon rank-sum test for each response type (Table 3).

4. Findings

In this section, we explore the typologies of responses to data (Table 2), their occurrence per school role (Table 3), and code co-occurrence (Figure 4). We highlight cases with higher frequencies and those central to understanding how sensemaking of learning data may relate to instructional improvement.

4.1 Response Dimensions

Our analysis of the data yielded 25 response types, which were organized into three interpretive dimensions: emotional, analytical, and intentional. In Table 2, we list and define each response type occurring per dimension.

In the *emotional dimension*, we mapped responses that pointed to affective reactions to data. In *positive or negative surprise*, for instance, educators declare that results were either superior or inferior to their initial expectations. Consider this example: "Yeah... This is kind of a red [flag]. (...) That shouldn't... That's big. I don't like that at all." These two response types—negative and positive surprise—necessarily required a reference point (i.e., previous expectations), which makes it different from *satisfaction*. Consider the following example, in which a teacher shows satisfaction when reacting to a graph but without stating a reference point for the reaction: "Looking at the data is like, it is good to hear students do learn from one another."

Whenever an educator empathized with students while trying to build an interpretation, we coded for *empathy*. This situation is represented in the following passage, taken from a teacher: "I have one little boy in this class who thinks he is terrible at math. When I sit down with him I'm like, 'You know exactly what you're doing.' He just has no confidence. I wonder how many more of them feel that way." We used *responsibility* whenever educators declared being directly responsible for the result shown by the data. Teacher M, for instance, declared "I didn't keep a log of what we did. That's my failing. (...) I'm keeping better notes this year," which is evidence of a practitioner taking responsibility. Finally, whenever an educator showed distrust toward a data representation, we marked it as *distrust*. One example is "You can't just take one piece of information and [say] that [piece] defines a student or the teacher and so."

In the *analytical dimension*, we included reactions that pointed directly to systematic reasoning with and about the data. One of the most typical responses among coaches and teachers was *asking questions*; we utilized this code every time data visualizations prompted educators to ask further questions about what was displayed. This type of response includes utterances such as "I am curious about" and "I wonder." Similarly, we used *recall* when data made educators remember or backtrack to past events, usually to attach a cause to the displayed phenomenon. As expected, this response tended to occur alongside *attribution of cause* and was often used as a form of information triangulation (Wise & Jung, 2019). Consider how Teacher G reacted to a graph by recalling class events and suggesting a causal chain that explains the data: "There wasn't much discussion from the students. (...) They weren't paying attention to what they were saying anyway."

Educators also triangulated information by filling in the gaps of the data with beliefs, stances toward teaching, or contextual information. In our data, the *gap filling* type of response is usually associated with (but not restricted to) causal explanations. Coach W, for instance, filled in the gaps between different instances of a longitudinal report: "This looks like it was a solidifying practice and then a developing task." In identifying a potential pair of task types ("solidifying" and "development"), Coach W also utilized a propositional structure (Borko et al., 1990) based on the school's pedagogical practices. Whenever educators connected multiple pieces of information contained in a report (e.g., different graphs, different dates, different response types), we coded it *connecting the dots*. One example can be found in the following passage, in which a coach reacts to a graph by connecting it with what was seen previously on the same report: "There is a connection between these two questions. It would be interesting to explore more about question five and question one." Furthermore, in what we understand as a higher-order type of reasoning, educators declared that data should be analyzed using *multiple lenses*. Coach B's quote is an eloquent example of this type of response to data: "There are different angles to the truth: students' perspectives [and] teacher's perspective, their preconceived notions about the class."

In the *intentional dimension*, we grouped reactions that pointed to various forms of intentions for future actions. For instance, we observed an *action intention* whenever educators communicated their willingness to improve instruction but did not convey a plan to do so. Conversely, we used *planning* to indicate when an educator expressed a concrete action to improve or fix a given situation. Consider Teacher L's reaction to a single-day report: "[I will] try to do different stations. (...) Maybe do a jigsaw of those people coming together, seeing what data they can get from that." In this response, Teacher L recognizes the problem and develops a theory of action to act upon the phenomenon displayed by the data. Finally, we observed several instances of *seek information*, which refers to instances when educators manifest the intention of seeking additional information. This type of reaction is evident in the following excerpt, taken from a coach: "Wait... What? I want to talk to her [a teacher]. I have as many questions as thoughts at this point. Holy smokes!"

4.2 Patterns by Professional Role

The second level of our analysis looked for how responses to data occurred by professional role. In general, we observed a stark difference in how coaches made sense of the data when compared to teachers. Results from a Wilcoxon rank-sum test suggested that while teachers typically described situations in their own terms and attributed causes to what they saw, coaches tended to engage in information-seeking behaviours and often reasoned in terms of potential pedagogical action. Table 3 presents the total count of codes per role, as well as the median per role (the statistic used in the Wilcoxon test) and the significance of the comparison between roles. Additional descriptive statistics can be found in Table 4 (Appendix).

Table 3. Total count of codes and median per role, and results of Wilcoxon test

Dimension/Code	Coaches ($n = 9$)		Teachers ($n = 9$)		Wilcoxon
Emotional	**Count**	**Median**	**Count**	**Median**	*p*-value
Positive Surprise	5	0	10	0	0.75
Negative Surprise	12	1	11	0	0.39
Satisfaction	15	1	4	0	0.1
Empathy	1	0	4	0	0.14
Distrust	1	0	9	1	0.04 *
Responsibility	0	0	7	0	0.03 *
Total Emotional	34	3	45	5	0.28
No Response	69	7	35	4	0.04 *
Analytical	**Count**	**Median**	**Count**	**Median**	*p*-value
Re-stating	30	2	11	0	0.13
Asking Questions	34	4	13	1	0.09
Recall	4	0	21	2	0.01 *
Attrib. of Cause	21	2	47	5	0.02 *
Confusion	5	1	4	0	0.42
Gap Filling	26	3	12	1	0.13
Connecting Dots	21	2	7	0	0.01 *
Comparing	9	0	2	0	0.22
Part vs. Whole	3	0	1	0	0.3
Prediction	5	1	1	0	0.06
Multiple Lenses	8	0	4	0	0.46
Judgment	8	1	1	0	0.05
Confirmation	6	0	6	1	0.77
Total Analytics	180	20	130	15	0.12
No Response	2	0	3	0	0.65
Intentional	**Count**	**Median**	**Count**	**Median**	*p*-value
Action Intention	25	3	11	1	0.05 *
Planning	6	0	0	0	0.08
Seeking Info.	19	1	2	0	0.04 *
Total Intentional	50	6	13	1	0.002 **
No Response	53	5	66	8	0.29

Notes. Obs. totals of each dimension exclude cases of *no response*. * $p \leq 0.05$; ** $p \leq 0.01$

Coaches asked more than twice the number of questions (*asking questions*: 34 occurrences) as teachers (13 occurrences). At the same time, we observed more than twice the number of *attribution of cause* among classroom teachers (47 occurrences) as among coaches (21 occurrences). Similarly, we found a total of 26 occurrences of *gap filling*, more than twice the total observed among teachers. Consider how Coach T fills in the gaps when confronted with a teacher's data: "Thinking about our district's focus on social-emotional learning, that's a good question if we're thinking about the class culture." The framing of the phenomenon as being connected to class culture—which was not explicitly present in the data report or directly measured by our PMs—suggests an explanatory structure (Borko et al., 1990) proposed entirely by the coach as an attempt to explain longitudinal patterns. Teacher K, however, took a different direction and created concrete explanations anchored in past instructional events: "A lot of them would pick that [response]. They [are] just people-pleasers. Show me how to do it and I will do it just like you showed me. A lot of them are lacking confidence in math and I think they don't want to go off on their own."

Our analysis also revealed that all responses in the *intentional* dimension, when summed, occurred more than three times as often with coaches (50 occurrences) as with teachers (13 occurrences; $p < 0.01$). Consider how Coaches B and P approach the data by prospecting potential doorways into instructional improvement:

> **Coach B.** *I'm assuming there was a lot of show and tell (…). But that's why I [agree] with increasing purposeful student talk, not just showing and telling. For next session, [what are] ways to help students make connections? What questions could be asked to help students make connections?*

> **Coach P.** *I think this information tells me that the teacher is doing a majority of talking. Students are probably in direct instruction and they're probably not being exposed to many different ways of doing math. (…) So looking at*

those types of lessons with the teacher and then even talking about how the kids felt, I would maybe go back to these two dates and all the previous ones and really look at how the kids responded.

In the above example, Coach B develops a propositional structure (Borko et al., 1990) based on her local RPP's objectives—in this case, increasing students' purposeful talk. This framing was then followed by an implicit intention of supporting the teacher in fostering connections between students' talk and the current math topic, as evidenced by the two questions found in the excerpt. Similarly, Coach P's analysis is coupled with an action intention, framed as the need to forage for additional information to form a more comprehensive explanation.

4.3 Patterns of Response Co-occurrence

To investigate emerging sensemaking patterns, we developed a matrix of response co-occurrences. The numbers in Figure 4 represent how many times two codes occurred together within a single coded excerpt. Due to decisions described in the Methods section (Section 3), each excerpt received three of four codes: one for the emotional dimension, one or two for the analytical—which depended on the breadth and depth of the participant's analysis—and one for the intentional. In all three of these groups, a *no response* code was included. Since no excerpts received two codes for emotional and intentional, there are no possible co-occurrences within these dimensions. Thus, we do not display emotional-emotional or intentional-intentional co-occurrence tables.

	Satisfaction	Empathy	Distrust	Responsibility	Positive Surprise	Negative Surprise	No Response	Multiple Lenses	Attrib. of Cause	Judgement	Asking Questions	Comparing	Confirmation	Prediction	Confusion	Connecting Dots	Gap Filling	Part vs Whole	Recall	Re-Stating	No Response
	Emotional (one code per exerpt)							Analytical (one or two codes per excerpt)													
Multiple Lenses					2	1	9														
Attrib. of Cause	7	3	10	5	9	9	34	3													
Judgement	1					4	4		1												
Asking Questions	5	1	2		3	8	32	3	5	1											
Comparing	3			1	1	5	8		1		4										
Confirmation	1		1		2	2	7		7			2									
Prediction	2		1		1		4		2			2	1								
Confusion						2	10					3	2								
Connecting Dots	4			1	2	4	18		4			9	1		1						
Gap Filling	6		2		2	10	27	2	19	4	7	2			1	3					
Part vs Whole	2				1		1								1	1	2				
Recall		2	5	4	2	4	9	2	14		3			1	1	2	1				
Re-Stating	5		1	1	1	3	31	1	5	2	6	1			1	4	6	5			
No Response		1				1	5														
Action Intention	5	1	1	4	1	7	24	4	9	2	11	5	1	3	1	8	13	2	4	8	2
Planning						1	6	1	3		1					1	1		1	2	1
Seeking Info	2				3	1	16	2	3		17	1	1	1	1	5	3			3	2
No Response	15	4	11	4	11	18	70	5	62	7	23	12	11	4	10	15	30	2	21	29	2

Legend: ☐ Zero Co-Occurrences | 4 to 6 co-occurrences | 7 to 10 co-occurrences | 11 to 14 co-occurrences | 15+ co-occurrences

(Left axis groupings: Analytical — one or two codes per excerpt; Intentional — one code per excerpt.)

Figure 4. Absolute number of co-occurrences of all response types.

A high frequency of co-occurrences involving *no response* items was observed in all three dimensions (e.g., *attribution of cause* and *no intentional response*, in the analytical dimension, with 62 occurrences). This finding confirms previous researchers' observations that *no action* is usually the most common response observed when studying educators' instructional

responses to data (Farrell & Marsh, 2016; Molenaar & Knoop-van Campen, 2018; Wise & Jung, 2019). Other patterns with a high frequency may be explained by their inherent operational definitions. This is the case for *attribution of cause* and *gap filling* (19 occurrences), for example. In the following sections, we examine and present a few illuminating cases that associate analytical and intentional responses.

4.3.1 Asking Questions and Seeking Information (17 Occurrences)

Multiple co-occurrences happened between *asking questions* and *seeking information*. Although related, the first response indicates inquiries made exclusively about the data, while the second represents an educator's intention of seeking further information outside the dashboard, as a necessary step to inform future instruction. The following excerpt, taken from Coach D, exemplifies how these two reactions interconnect: "I wonder about what type of teaching is happening and who's doing the thinking. (…) I want to go and ask some questions around instruction. I might also want to go in and observe a class." Coach D's wondering seemed to act as a lever for developing a plan to collect more data at the school (i.e., asking questions and observing a class) in order to support the teacher. A similar pattern happened among other participants, with the great majority being coaches.

4.3.2 Gap Filling and Action Intention (13 Occurrences)

Analysis of data suggested that teachers and coaches who triangulated their explanations of displayed data with values, beliefs, or outside sources of information were more likely to utter an intention of instructional improvement. This finding can be observed in the excerpt below:

> **Coach W.** *Looking at this data makes me feel like this is a solidifying task. But if it's not, then that's a whole different conversation. (…) I would want to dig deeper into how often this kind of data would show up. And also, like maybe having the opportunity to have a teacher try a different kind of lesson. (…) It might have been a review before an assessment, you know? I would really like to dig at whether this is a common pattern or not. And then, if it is, to really help the teacher shift how they plan their lessons.*

Coach W's words provide evidence of the link between filling in the gaps with external reality and conveying potential courses of action for instruction. After triangulating the graphs with an external explanation based on her own experiences and scenes (Borko et al., 1990), Coach W considers ways of acquiring more information ("dig deeper") to ultimately inform an improvement plan ("try a different kind of lesson" and "shift how they plan their lessons").

On the teacher side, Teacher M provides a similar response when confronted with data from one of her previous classes: "I feel my students have this mentality of 'I gotta finish and get the right answer,' and the process along the way doesn't matter. That's a shift that I always talk about making." Although a plan was not at all conveyed, the teacher frames the problem by filling in the gaps with her impressions and declares a need to alter the course of instruction.

4.3.3 Asking Questions and Action Intention (11 Occurrences)

Our results indicate that inquiry-oriented statements toward dashboard data co-occur with the formulation of improvement intentions. The *asking questions* and *improvement intention* response types co-occurred more often among coaches than among teachers. For instance, when prompted to write a note about a teacher's classroom report, Coach J transitioned from wondering to indicating an intention to support the teacher:

> **Coach J.** *I wonder if this is a typical lesson [and] what the teacher believes in terms of instruction. Is this [a] sit and take notes, teacher-centered classroom? Next steps: maybe observe, get at beliefs around student discourse. Think about R4R as entry point.*

In this occurrence, Coach J starts with questions around the nature of the lesson ("typical," "sit and take notes"). Without being prompted to do so, Coach J formulates an intention to understand more about the teacher's beliefs around instruction ("get at beliefs") while coming up with a potential entry point for instructional improvement ("R4R," or Resources for Rethinking, a database of curriculum resources).

4.3.4 Attribution of Cause and Recall (14 Occurrences)

Lastly, our co-occurrence analysis revealed that educators often attributed cause to a given data visualization by recalling concrete past events. Given that our coding procedure allowed two responses at the *analytical* dimension, these codes often occurred together. Interestingly, these two particular response types showed a high co-occurrence with the absence of action intentions. The *attribution of cause* and *no intentional response* pair occurred 62 times, while *recall* and *no intentional response* occurred 21 times. Table 3 shows that this pattern is more prevalent among teachers than among coaches. The excerpt below, taken from Teacher A, illustrates how educators reacted to visual data by engaging in *attribution of cause* and *recall*:

Teacher A. There wasn't much discussion from the students. (. . .) The whole thing that I've been trying to get is 'you did ratios and proportions, think unit rate, value of one item.' I had to clue them in as to what it was. But they heard [Student X] say it, and then they . . . still didn't get it. I came back to it again, because there was another question that was along the same line. You still have to remind them what it was.

In this occurrence, Teacher A reacts to a graph showing how many students declared that they were able to share how to solve a mathematical problem—which points to a central improvement goal for that district. The teacher recalls particular aspects of an instructional event to build an assertive causal explanation for what the data revealed. No intentions to seek further information or change an aspect of instruction are present in these and other excerpts from the same interviewee.

5. Discussion and Implications

While discussing how data visualizations relate to internal explanatory structures, Liu and Stasko (2010) asked what specific operations individuals might perform when faced with visual data. In posing such a question, the authors called for the development of a typology that could describe and explain how users react to visual representations of data delivered through interactive tools. The same need applies to the domain of K–12 education. Understanding what accounts of past events are created, how they are created, and to what extent has become a vital component of effectively improving instruction and moving from a data-centric to a human-centred LA paradigm (Buckingham Shum et al., 2019). With more data-informed tools present in schools, this study creates ways to categorize and understand such operations by creating a shared understanding of the many building blocks of LA sensemaking. In this section, we discuss how our findings add to current theoretical models of LA and HCI and explore potential implications for scholars and educators.

5.1 Leveraging Different Vantage Points

The patterns that emerged from our analysis show how coaches' analytical reasoning may differ from that of classroom teachers. While the sensemaking literature has shown that users in different roles are expected to react differently to data (Dervin & Naumer, 2010; Dervin, 1998; Klein & Moon, 2006; Klein, Moon, & Hoffman, 2006), LA designers need to understand *how* these variations manifest in the K–12 school domain to understand *how* they may inform the development of visual analytics tools.

In general, while teachers attributed cause to the displayed graphs by creating detailed narratives, coaches often extrapolated potential pedagogical actions in levels that vary from simple intentions to more structured and concrete steps. Considering that devising potential action from data is a *habit of mind* (Bocala & Boudet, 2015), such dispositions toward pedagogical intention and action need to be fostered among all school educators. It is critical to understand teachers' and coaches' different *vantage points* when ascribing meaning to data and to understand how LA dashboards may supplement these heuristics.

A closer look into coaches' sensemaking revealed key insights for LA research and practice. First, coaches identified relevant aspects of instruction by making connections to contexts and domains external to the data, as seen in the multiple instances of *gap filling*. Second, coaches interpreted particular data points in light of their propositional structures (e.g., pedagogical challenges, teachers' abilities) (Borko et al., 1990). Coaches also gravitated toward devising instructional plans and actions from data visualizations. Understanding coaches' unique practices and vantage points may help bridge the gap between leveraging data to develop insights and acting upon that data.

Teachers also employed differential responses when making sense of dashboard-delivered information. Analyzing data from an insider perspective (i.e., from someone directly involved in the events represented by the dashboard), teachers often attributed cause to dashboard visualizations or engaged in recalling past events. These patterns resonate with Borko and colleagues' (1990) theories that describe teachers thinking in *scenes*: experiential knowledge of people, lessons, and other features of instruction allows for rapidly storing and retrieving information from classroom events. Considering that much of this knowledge is tacit, a question remains about how visual LA tools could provide other educators with enough context and relevant information (e.g., classroom artifacts) (Charleer, Klerkx, Duval, De Laet, & Verbert, 2016; Wardrip & Herman, 2018) to contribute to understanding instructional situations without creating more cognitive load for the users.

5.2 Applying the Typology in Education Practice and Research

Data use at K–12 schools has been described by researchers as a complex process, likely to be influenced by individual and organizational factors (Coburn & Turner, 2011; Farrell & Marsh, 2016; Schildkamp, 2019). In this context, teacher-coach collaboration has used data extensively to promote reflection and learning about past classroom events (Marsh et al., 2010). The way coaches and teachers work together may vary across sites but is frequently organized around noticing particular features revealed by data, creating narratives that account for instructional events, and ultimately establishing or re-negotiating improvement goals. Hence, training coaches to better facilitate sensemaking with teachers is key to school instructional

improvement (Marsh et al., 2010). Such professional development practices may benefit greatly from a common language, which we begin to develop in the present study. Providing nuanced ways to observe, categorize, and discuss reactions to visual analytics might allow instructional coaches and school leaders to support teacher sensemaking more appropriately.

Our findings suggest that teachers tended to develop interpretations that fit existing accounts of past events, as seen in the higher numbers of *recall* and *attribution of cause* occurrences as compared with coaches. This empirical evidence suggests that teachers might normalize data representations before asking more questions or engaging with other analytical operations, such as comparing or connecting multiple dots. The threat of promptly normalizing new information is consonant with other theoretical sensemaking models (Weick, 1995; Maitlis et al., 2013). In this respect, specific patterns of responses—such as asking questions and the intention of seeking additional information or filling in the gaps and the intention for acting—might be actively sought by coaches and school leaders. This is precisely where an understanding of patterns of responses to visual data might be of practical benefit for school pedagogical improvement efforts.

In summary, given the reality of K–12 schools, where educators are expected to create meaning based on learning data, researchers might need to further develop a lingua franca, a common language that facilitates deep conversations around data (Horn et al., 2015; Russell et al., 2020). Although current research does offer frameworks of data sensemaking, such models typically are not granular enough and disregard educators' role at the school (Murnane et al., 2005). Our experience embedded in RPPs suggests that a sensemaking typology—easily recognizable by multiple educator roles and with enough granularity to account for different problems of practice—might serve as an initial step to ground and guide deep data-driven conversations.

5.3 Toward a New LA Process Model

Our study showed that making sense of data represented through a visual LA tool is indeed a highly authorial and flexible endeavour, prone to be shaped by the user's perspectives (Yi et al., 2008), organizational factors, and team dynamics (Weick, 1995, 1993; Coburn & Turner, 2011; Farrell & Marsh, 2016; Schildkamp, 2019). In this section, we discuss how our findings, combined with previous studies from the education, LA, HCI, and organization studies fields, might render the initial elements for a revised LA process model. Before doing that, we stress that our model—presented in Figure 5—is based on visual representations of teaching and learning data and that additional cases, data types, and experimental settings are necessary for a model capable of representing all potential cases of interactions with LA tools.

Our process model focuses on explaining the centrality of the overall sensemaking process, illuminating the role of resulting interpretations as plausible retrospective accounts (Weick, 1995; Colville, 2007; Brown et al., 2015), and expanding on the many contributing factors to the process as a whole. Understanding that academic scholarship forms from a continuous theoretical discussion over time, we developed our model based on previous LA models, summarized by Figure 2 and discussed in the Literature Review section (Section 2) of this paper. The same sub-processes and links are mapped below in Figure 5, represented by the letters A to D. Given that the present study design was not intended to infer reactions to data vis-à-vis design decisions, nor to map actual pedagogical actions, we focus our analysis on B and C.

No single model should carry the burden of representing all possible cases of a phenomenon. Instead, different process models might provide distinct lenses to prime and inform the work of researchers, designers, and practitioners. For example, Wise and Jung's (2019) model highlights the centrality of aligning information structures with pedagogical concerns and illuminates how university instructors' curiosity and initial questions ignite and shape the subsequent sensemaking steps. Other models, such as that of Xhakaj, Aleven, and McLaren (2017), clarify how dashboards impact teachers' planning and learning. In our case, we offer a more granular understanding of data sensemaking operations (i.e., the typology) and possible patterns (i.e., the co-occurrences), in hopes of shedding light on how school educators take on information delivered through visual analytics tools.

5.3.1 The Model

Our process model flows from one or more instructional events into its multiple data representations. The link represented by letter A represents the many design, computational, and algorithmic decisions that will ultimately shape what is displayed and how it is represented to the user. The light blue rectangle represents the sensemaking process as a whole and separates its product (interpretations) from its generating sub-processes (emotional, analytical, and intentional dimensions), thus reflecting what Brown and colleagues (2015) propose: "sensemaking refers to processes by which people generate what they interpret" (p. 267). Although the model does not explicitly map response types (e.g., *positive surprise*, *gap filling*, *action intention*), it implies that such operations pertain to the three mapped dimensions.

Learning is a situated phenomenon (Lave & Wenger, 1991) in which sensemaking and reflexive action are both constituent parts (Wise & Jung, 2019). Interpretations are represented in Figure 5 as tenable accounts of instructional events and thus point to the past. This cyclical aspect of the model is based on research into teachers' thinking that explores how K–12 educators create cognitive structures for sensemaking based on past events (Borko et al., 1990). In this respect, even if no action per se is taken (e.g., re-planning, course correction), pedagogical knowledge and reflection are generated by the process (Figure 5,

letter E) (Wise & Jung, 2019). Similarly, the model places pedagogical decision-making as an external, yet connected process (Figure 5, letter D). In developing the present model, our team was attuned to the idea that, in order to make decisions, individuals need to accurately define what the challenge is (Weick, 1993). In school settings, problems of practice might not reveal themselves through a graph or a chart. Rather, teachers and coaches will need to make sense of such representations to construct—and not merely discover—what the upcoming challenge might be. This process, the authoring of interpretations that re-signify events (letter C in Figure 5), is therefore separated from *actions* in the model.
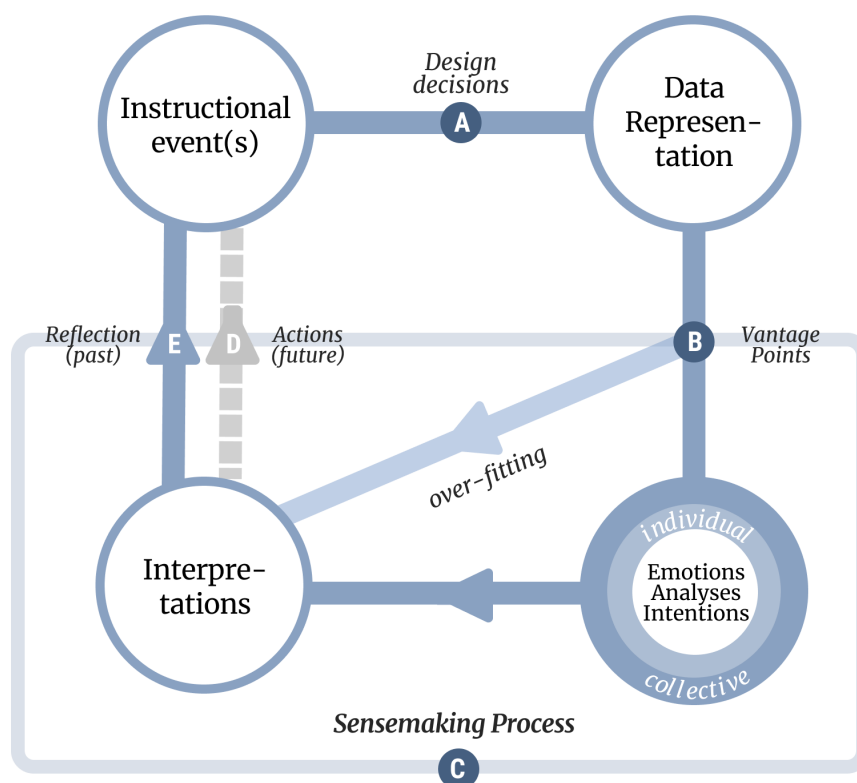

Figure 5. A proposed process model for LA sensemaking through dashboards.

5.3.2 Frames of Reference

Our findings illuminate how particular frames of reference (Van Es & Sherin, 2002; Dervin, 1998) or extant characterizations of teaching (Borko et al., 1990) play a role in shaping educators' sensemaking. This idea is supported by Klein and colleagues (2006), who argue that "when people try to make sense of events, they begin with some perspective, viewpoint, or framework—however minimal" (p. 2). In our case, we view such perspectives as the different *vantage points* afforded by the distinct roles educators might play at a school. One limitation of the model, however, is to explain how exactly such frames and viewpoints shape the subsequent steps, such as changing aspects of the instruction or actively looking for additional information, which will hopefully be explored in future studies.

5.3.3 Over-Fitting

In statistics (and, more recently, machine learning), *over-fitting* happens when a data set excessively conforms to a model, thus making the model lose its ability to accurately predict events based on new data. LA users might encounter a similar situation when faced with visual representations of learning data: over-normalizing new evidence into what is seen as plausible explanations—or over-fitting new information into existing frames—might lead to insubstantial or even repeated interpretations. This threat aligns with several education studies showing how data sensemaking in schools may open a door for simply reproducing earlier instructional practices, confirming biases, or experimenting with instruction in ineffective ways (Farrell & Marsh, 2016). Within our results, this phenomenon is evidenced by the high co-occurrence numbers between instances of *attribution of cause* and *no intentional response* ($n = 62$) and between *recall* and *no intentional response* ($n = 21$). These occurrences happened especially among teachers, who attributed potential causes to data representations by recalling concrete

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)

16

scenes and scripts (Borko et al., 1990) of the past, therefore assuming that what they see is not new. Interestingly, such educators rarely engaged in making other types of analytical inferences, such as asking additional questions or triangulating different aspects of the data, or revealed any *intentional* responses, such as seeking information or planning the next instructional move. In the *emotional* domain, responses such as *distrust*, for example—defined as "educators distrust the data, the instrument of collection, or the visualization"—led mostly to *no intentional response* ($n = 11$), thus confirming a similar over-fitting pattern.

The threat of over-fitting new events into old narratives might be especially relevant to the LA community, considering that design and computation decisions (letter A, Figure 5) might determine the ensuing steps of the process, and that over-normalizing a narrative may lead to missed opportunities for deeper insights.

5.4 Study Limitations

Our exploratory study has five major limitations. First, with a limited number of participants, we aimed at generating insights and not making any generalizable causal claims about observed patterns and relationships. For this reason, we employed purposeful sampling to reflect situations that are common for teachers (e.g., analyzing example data and looking at their own data) (Yin, 2012). Second, our study exclusively utilized measures of students' views of their teachers' instructional actions as stimuli for think-aloud sessions. Third, due to sampling decisions, our design does not balance all possible contextual elements and conditions of data use, such as different stakeholders' teaching experience and different forms of interaction between coaches and teachers in each partner district that could potentially occur. Fourth, while acknowledging the centrality of collective forms of sensemaking, our study was not designed to capture how meaning is created and negotiated between collaborating individuals. Finally, we acknowledge that other reactions could be elicited in other experimental or non-experimental settings, thus making our typology only a first step in mapping the building blocks of data sensemaking. We invite the research community to further add to this typology by studying other organizational contexts within the field of education.

6. Conclusion and Future Directions

Our study offers three major contributions to the LA and HCI communities. First, our typology of responses to visual data permits a new common language for educators, researchers, and designers when designing, studying, or using LA. Second, we broaden the understanding of how different roles within an organization—teachers and coaches—respond differently to data visualizations. For example, our analyses illuminated how coaches in our sample voiced more information-seeking behaviours, while teachers were more likely to recall past events and attribute causal explanations to the data. Third, we offer a process model to expand our scholarly community's understanding of data sensemaking through dashboards, one that combines our own findings with a review of LA, education, HCI, and organization studies scholarship. We hope this model serves as a lens for researchers and designers to further their understanding of how data representations work in tandem with other key human factors in the process of generating plausible interpretations of instructional events. While it is beyond the scope of this analysis to speak to in situ classroom practice—and recognizing that sensemaking by itself might not lead to pedagogical action—we do conjecture that by elevating and making visible the range and variability of teacher and coach sensemaking, RPPs may well be better positioned to support districts' vision for instructional change.

This study paves the way for new research directions, hopefully to be taken up by the LA and HCI communities. First, we encourage our colleagues to become attuned to the different vantage points educators may have when making sense of visual learning data, especially those unique to K–12 schools. For example, an experimental study can compare the different patterns in educators' reactions when looking at visualizations of their own classroom events versus sample data or others' data. Second, we invite the research community to think about what other instructional contexts might add to our typology of sensemaking responses. Third, we see an opportunity for analyzing what coaches' and teachers' sequences of sensemaking strategies (i.e., as seen in Figure 4) might reveal when cohorts of teachers and coaches are compared. This would provide even more clarity about how different professional vantage points shape K–12 data sensemaking. Finally, we invite researchers to explore the design implications of a sensemaking typology and process model, with special attention to reducing over-automation and over-fitting of interpretations. Considering that creating meaning is the path all humans inevitably follow (Dervin, 1998), opening the black box of sensemaking in learning environments may help us understand and imagine new ways to encourage educators to ask their own questions and tell their own stories with data.

Acknowledgements

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

References

Ahn, J., Campos, F., Hays, M., & DiGiacomo, D. (2019). Designing in context: Reaching beyond usability in learning analytics dashboard design. *Journal of Learning Analytics*, *6*(2), 70–85. https://doi.org/10.18608/jla.2019.62.5

Ahn, J., Nguyen, H., & Campos, F. (2021). From Visible to Understandable: Designing for Teacher Agency in Education Data Visualizations. *Contemporary Issues in Technology and Teacher Education.*, *21*(1).

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. https://doi.org/10.18608/jla.2018.52.5

Bocala, C., & Boudet, K. (2015). Teaching educators habits of mind for using data wisely. *Teachers College Record*, *117*(4), 1–20. Retrieved from `https://www.tcrecord.org/library/content.asp?contentid=17853`

Borko, H., Livingston, C., & Shavelson, R. J. (1990). Teachers' thinking about instruction. *Remedial and Special Education*, *11*(6), 40–49. https://doi.org/10.1177/074193259001100609

Bower, G. H., & Cohen, P. R. (1982). Emotional influences in memory and thinking: Data and theory. In M. S. Clark & S. T. Fiske (Eds.), *Affect and Cognition—17th Annual Carnegie Mellon Symposium on Cognition.* New York: Psychology Press.

Brown, A. D., Colville, I., & Pye, A. (2015). Making sense of sensemaking in *Organization Studies*. *Organization Studies*, *36*(2), 265–277. https://doi.org/10.1177/0170840614559259

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*(1), 1–21. Retrieved from `https://www.jstor.org/stable/1343711`

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonaldo, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, *6*(2), 1–9. https://doi.org/10.18608/jla.2019.62.1

Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2016). Creating effective learning analytics dashboards: Lessons learnt. In K. Verbert & M. Sharples (Eds.), *Adaptive and Adaptable Learning* (EC-TEL 2016), 13–16 September 2016, Lyon, France. *Lecture Notes in Computer Science* (pp. 42–56). Cham: Springer. https://doi.org/10.1007/978-3-319-45153-4_4

Charters, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education: A Journal of Educational Research and Practice*, *12*(2), 68–82. https://doi.org/10.26522/brocked.v12i2.38

Chen, B., & Zhu, H. (2019). Towards value-sensitive learning analytics design. In *Proceedings of the Ninth International Conference on Learning Analytics & Knowledge* (LAK 2019), 4–8 March 2019, Tempe, AZ, USA (pp. 343–352). New York: ACM. https://doi.org/10.1145/3303772.3303798

Cobb, P., Jackson, K., Henrick, E., & Smith, T. M. (2018). *Systems for Instructional Improvement: Creating Coherence from the Classroom to the District Office*. Boston, MA: Harvard Education Press.

Coburn, C. E., & Penuel, W. R. (2016). Research–practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher*, *45*(1), 48–54. https://doi.org/10.3102/0013189X16631750

Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013). Research-Practice Partnerships: A Strategy for Leveraging Research for Educational Improvement in School Districts. New York: William T. Grant Foundation.

Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement*, *9*(4), 173–206. https://doi.org/10.1080/15366367.2011.626729

Colville, I. (2007). Sensemaking. In R. Thorpe & R. Holt (Eds.), *The SAGE Dictionary of Qualitative Management Research*. SAGE Publications.

Dervin, B. (1998). Sense-making theory and practice: An overview of user interests in knowledge seeking and use. *Journal of Knowledge Management*, *2*(2), 36–46. https://doi.org/10.1108/13673279810249369

Dervin, B., & Naumer, C. M. (2010). Sense-making. In M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of Library and Information Sciences* (3rd ed., pp. 4696–4707). Boca Raton, FL, USA: CRC Press.

D'Mello, S. K. (2017). Emotional learning analytics. In C. Lang, George Siemens, A. F. Wise, & D. Gašević (Eds.), *Handbook of Learning Analytics* (1st ed., pp. 115–128). Society for Learning Analytics Research. https://doi.org/10.18608/hla17.010

Duval, E. (2011). Attention please! Learning analytics for visualization and recommendation. In *Proceedings of the First International Conference on Learning Analytics and Knowledge* (LAK 2011), 27 February–11 March 2011, Banff, AB, Canada (pp. 9–17). New York: ACM. https://doi.org/10.1145/2090116.2090118

Erickson, F. (2011). On noticing teacher noticing. In M. Sherin, V. Jacobs, & R. Philipp (Eds.), *Mathematics Teacher Noticing: Seeing through Teachers' Eyes* (pp. 17–34). Routledge.

Farrell, C. C., & Marsh, J. A. (2016, 11). Contributing conditions: A qualitative comparative analysis of teachers' instructional responses to data. *Teaching and Teacher Education*, *60*, 398–412. https://doi.org/10.1016/j.tate.2016.07.010

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

Frank, K. A., Kim, J., Salloum, S. J., Bieda, K. N., & Youngs, P. (2020). From interpretation to instructional practice: A network study of early-career teachers' sensemaking in the era of accountability pressures and Common Core state standards. *American Educational Research Journal*, *57*(6), 2293–2338. https://doi.org/10.3102/0002831220911065

Grammel, L., Tory, M., & Storey, M. (2010, 11). How information visualization novices construct visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *16*(6), 943–952. https://doi.org/10.1109/TVCG.2010.164

Holstein, K., Mclaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. *Journal of Learning Analytics*, *6*(2), 27–52. https://doi.org/10.18608/jla.2019.62.3

Horn, I. S., Kane, B. D., & Wilson, J. (2015). Making sense of student performance data: Data use logistics and mathematics teachers' learning opportunities. *American Educational Research Journal*, *52*(2), 208–242. https://doi.org/10.3102/0002831215573773

Jackson, K., Gibbons, L., & Sharpe, C. J. (2017). Teachers' views of students' mathematical capabilities: Challenges and possibilities for ambitious reform. *Teachers College Record*, *119*(7), 1–43. Retrieved from `https://www.tcrecord.org/LIBRARY/PrintContent.asp?ContentID=21791`

Jackson, K., Henrick, E., Cobb, P., Kochmanski, N., & Nieman, H. (2016). *Practical Measures to Improve the Quality of Small-Group and Whole-Class Discussions [White Paper]*. Retrieved from `https://bit.ly/30WIhys`

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* (EC-TEL 17), 12–15 September 2017, Tallinn, Estonia. *Lecture Notes in Computer Science* (pp. 82–96). Cham: Springer. https://doi.org/10.1007/978-3-319-66610-5_7

Klein, G., & Moon, B. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems*, *21*(5), 88–92. https://doi.org/10.1109/MIS.2006.100

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems*, *21*(4), 70–73. https://doi.org/10.1109/MIS.2006.75

Lam, H., Bertini, E., Isenberg, P., Plaisant, C., Lam, H., Bertini, E., . . . Lam, H. (2012). Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, *18*(9), 1520–1536. https://doi.org/10.1109/TVCG.2011.279

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*(1), 65–100. https://doi.org/10.1111/j.1551-6708.1987.tb00863.x

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.

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

Liu, Z., & Stasko, J. (2010). Mental models, visual reasoning and interaction in information visualization: A top-down perspective. *IEEE Transactions on Visualization and Computer Graphics*, *16*(6), 999–1008. https://doi.org/10.1109/TVCG.2010.177

Maitlis, S., Vogus, T. J., & Lawrence, T. B. (2013). Sensemaking and emotion in organizations. *Organizational Psychology Review*, *3*(3), 222–247. https://doi.org/10.1177/2041386613489062

Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist*, *47*(2), 71–85. https://doi.org/10.1080/00461520.2012.667064

Marsh, J. A., Bertrand, M., & Huguet, A. (2015). Using data to alter instructional practice: The mediating role of coaches and professional learning communities. *Teachers College Record*, *117*(4), 1–40. Retrieved from `https://www.tcrecord.org/content.asp?contentid=17849`

Marsh, J. A., Sloan McCombs, J., & Martorell, F. (2010). How instructional coaches support data-driven decision making: Policy implementation and effects in Florida middle schools. *Educational Policy*, *24*(6), 872–907. https://doi.org/10.1177/0895904809341467

Mccoy, C., & Shih, P. C. (2016). Teachers as producers of data analytics: A case study of a teacher-focused educational data science program. *Journal of Learning Analytics*, *3*(3), 193–214. https://doi.org/10.18608/jla.2016.33.10

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: SAGE.

Molenaar, I., & Knoop-van Campen, C. (2018). How teachers make dashboard information actionable. *IEEE Transactions on Learning Technologies*, *12*(3), 347–355. https://doi.org/https://doi.org/10.1109/TLT.2018.2851585

Murnane, R. J., Sharkey, N. S., & Boudett, K. P. (2005). Using student-assessment results to improve instruction: Lessons from a workshop. *Journal of Education for Students Placed at Risk (JESPAR)*, *10*(3), 269–280. https://doi.org/10.1207/s15327671espr1003_3

Peck, E. M., Ayuso, S. E., & El-Etr, O. (2019). Data is personal: Attitudes and perceptions of data visualization in rural Pennsylvania. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI 2019), 4–9 May 2019, Glasgow, UK (pp. 1–12). New York: ACM. https://doi.org/10.1145/3290605.3300474

Romat, H., Riche, N. H., Hinckley, K., Lee, B., Appert, C., Pietriga, E., & Collins, C. (2019). Activeink: (Th)inking with data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI 2019), 4–9 May 2019, Glasgow, UK (pp. 1–13). New York: ACM. https://doi.org/10.1145/3290605.3300272

Russell, J. L., Correnti, R., Stein, M. K., Thomas, A., Bill, V., & Speranzo, L. (2020). Mathematics coaching for conceptual understanding: Promising evidence regarding the Tennessee math coaching model. *Educational Evaluation and Policy Analysis*, *42*(3), 439–466. https://doi.org/10.3102/0162373720940699

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human Computer Studies*, *45*(2), 185–213. https://doi.org/10.1006/ijhc.1996.0048

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 É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Data Driven Approaches in Digital Education* (EC-TEL 17), 12–15 September 2017, Tallinn, Estonia. *Lecture Notes in Computer Science* (pp. 194–208). Cham: Springer. https://doi.org/10.1007/978-3-319-66610-5_15

Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational Research*, *61*(3), 257–273. https://doi.org/10.1080/00131881.2019.1625716

Shapiro, R. B., & Wardrip, P. S. (2015). Keepin' it real: Understanding analytics in classroom practice. *Technology, Instruction, Cognition and Learning*, *10*, 127–149.

Sherin, M. G., & Russ, R. S. (2014). Teacher noticing via video: The role of interpretive frames. In B. Calandra & P. J. Rich (Eds.), *Digital Video for Teacher Education: Research and Practice* (pp. 11–28). New York: Routledge.

Simon, J. (2017). A priori knowledge in learning analytics. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning* (pp. 199–227). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-52977-6_7

Sun, K., Mhaidli, A. H., Watel, S., Brooks, C. A., & Schaub, F. (2019). It's my data! Tensions among stakeholders of a learning analytics dashboard. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI 2019), 4–9 May 2019, Glasgow, UK (pp. 1–14). New York: ACM. https://doi.org/10.1145/3290605.3300824

Van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, *10*(4), 571–596. Retrieved from https://www.learntechlib.org/primary/p/9171/

van Leeuwen, A., Rummel, N., & van Gog, T. (2019). What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations? *International Journal of Computer-Supported Collaborative Learning*, *14*(3), 261–289. https://doi.org/10.1007/s11412-019-09299-x

van Leeuwen, A., van Wermeskerken, M., Erkens, G., & Rummel, N. (2017). Measuring teacher sense making strategies of learning analytics: A case study. *Learning: Research and Practice*, *3*(1), 42–58. https://doi.org/10.1080/23735082.2017.1284252

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

Voyiatzaki, E., & Avouris, N. (2014). Support for the teacher in technology-enhanced collaborative classroom. *Education and Information Technologies*, *19*(1), 129–154. https://doi.org/10.1007/s10639-012-9203-2

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. http://doi.org/10.1080/13664530.2017.1308428

Weick, K. E. (1993). The collapse of sensemaking in organizations: The Mann Gulch disaster. *Administrative Science Quarterly*, *38*(4), 628–652. https://doi.org/10.2307/2393339

Weick, K. E. (1995). *Sensemaking in Organizations*. Thousand Oaks, CA, USA: SAGE Publications.

Wise, A. F., & Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics*, *6*(2), 53–69. https://doi.org/10.18608/jla.2019.62.4

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 K. Verbert, M. Sharples, & T. Klobučar (Eds.), *Adaptive and Adaptable Learning* (EC-TEL 2016), 13–16 September 2016, Lyon, France. *Lecture Notes in Computer Science* (Vol. 9891, pp. 13–16). Cham: Springer. https://doi.org/10.1007/978-3-319-45153-4_26

Xhakaj, F., Aleven, V., & McLaren, B. M. (2017). Effects of a teacher dashboard for an intelligent tutoring system on teacher knowledge, lesson planning, lessons and student learning. *Data Driven Approaches in Digital Education* (EC-TEL 17), 12–15 September 2017, Tallinn, Estonia. *Lecture Notes in Computer Science*, *10474*, 315–329. https://doi.org/10.1007/978-3-319-66610-5_23

Yeager, D., Bryk, A., Hausman, H., Muhich, J., & Morales, L. (2013). *Practical Measurement*. Palo Alto, CA.

Yi, J. S., Kang, Y. a., Stasko, J. T., & Jacko, J. A. (2007). Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, *13*(6), 1224–1231. https://doi.org/10.1109/TVCG.2007.70515

Yi, J. S., Kang, Y.-a., Stasko, J. T., & Jacko, J. A. (2008). Understanding and characterizing insights: How do people gain insights using information visualization? In *ACM Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization* (BELIV 2008), 5 April 2008, Florence, Italy (pp. 4:1–4:6). New York: ACM. https://doi.org/10.1145/1377966.1377971

Yin, R. K. (2012). Case study methods. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA Handbook of Research Methods in Psychology: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological* (Vol. 2, pp. 141–155). American Psychological Association. https://doi.org/10.1037/13620-009

You, J. W., & Kang, M. (2014). The role of academic emotions in the relationship between perceived academic control and self-regulated learning in online learning. *Computers and Education*, *77*, 125–133. https://doi.org/10.1016/j.compedu.2014.04.018

Appendix

Table 4. Additional statistics per code and professional role, and results of the Wilcoxon test

Dimension/Code	Coaches (*n* = 9)						Teachers (*n* = 9)						Wilcoxon Test	
Emotional	**Count**	**Mdn**	**Min**	**Max**	**M**	**SD**	**Count**	**Mdn**	**Min**	**Max**	**M**	**SD**	**Z**	***p*-value**
Pos. Surprise	5	0	0	3	0.56	1.01	10	0	0	4	1.11	1.76	−0.32	0.75
Neg. Surprise	12	1	0	5	1.33	1.66	11	0	0	6	1.22	2.22	0.86	0.39
Satisfaction	15	1	0	4	1.67	1.66	4	0	0	2	0.44	0.73	1.66	0.1
Empathy	1	0	0	1	0.11	0.33	4	0	0	1	0.44	0.53	−1.48	0.14
Distrust	1	0	0	1	0.11	0.33	9	1	0	3	1	1.12	−2.01	0.04 *
Responsibility	0	0	0	0	0	0	7	0	0	2	0.78	0.97	−2.13	0.03 *
Total Emotional	34	3	1	6	3.78	1.64	45	5	2	8	5	2.18	−1.09	0.28
No Response	69	7	2	13	7.67	4.03	35	4	0	11	3.89	3.3	2.05	0.04 *
Analytical	**Count**	**Mdn**	**Min**	**Max**	**M**	**SD**	**Count**	**Mdn**	**Min**	**Max**	**M**	**SD**	**Z**	***p*-value**
Re-stating	30	2	0	8	3.33	3.04	11	0	0	4	1.22	1.56	1.51	0.13
Asking questions	34	4	0	11	3.78	3.38	13	1	0	3	1.44	1.33	1.70	0.09
Recalling	4	0	0	3	0.44	1.01	21	2	0	4	2.33	1.5	−2.69	0.01 *
Attrib. of Cause	21	2	0	7	2.33	2.35	47	5	2	8	5.22	2.59	−2.27	0.02 *
Confusion	5	1	0	1	0.56	0.53	4	0	0	2	0.44	0.88	0.81	0.42
Gap Filling	26	3	0	10	2.89	2.89	12	1	0	3	1.33	0.87	1.52	0.13
Connecting Dots	21	2	1	5	2.33	1.5	7	0	0	4	0.78	1.3	2.44	0.01 *
Comparing	9	0	0	4	1	1.41	2	0	0	1	0.22	0.44	1.21	0.22
Part vs. Whole	3	0	0	1	0.33	0.5	1	0	0	1	0.11	0.33	1.04	0.3
Prediction	5	1	0	1	0.56	0.53	1	0	0	1	0.11	0.33	1.89	0.06
Multiple Lenses	8	0	0	2	0.89	1.05	4	0	0	1	0.44	0.53	0.74	0.46
Judgment	8	1	0	4	0.89	1.27	1	0	0	1	0.11	0.33	1.92	0.05
Confirmation	6	0	0	3	0.67	1	6	1	0	2	0.67	0.71	0.29	0.77
Total Analytical	180	20	9	32	20	7.38	130	15	10	22	14.44	3.84	1.55	0.12
No Response	2	0	0	1	0.22	0.44	3	0	0	1	0.33	0.5	0.45	0.65
Intentional	**Count**	**Mdn**	**Min**	**Max**	**M**	**SD**	**Count**	**Mdn**	**Min**	**Max**	**M**	**SD**	**Z**	***p*-value**
Action Intention	25	3	1	5	2.78	1.79	11	1	0	4	1.22	1.39	2.00	0.05 *
Planning	6	0	0	3	0.67	1.12	0	0	0	0	0	0	1.77	0.08
Seeking Info.	19	1	0	6	2.11	2.42	2	0	0	1	0.22	0.44	2.10	0.04 *
Total Intentional	50	6	2	11	5.56	2.83	13	1	0	5	1.44	1.67	2.98	0.002 **
No Response	53	5	0	13	5.89	4.62	66	8	3	12	7.33	3.04	−1.07	0.29

Notes. Obs. totals of each dimension exclude the number of cases of *no response*. * $p \leq 0.05$; ** $p \leq 0.01$.