

Paper Title: Learning from Assessment Data: Epistemic Foundations of Data Use

Author(s): Brette Garner and Ilana Horn

Session Title: IPS-3

Session Type: Interactive Paper Session

Presentation Date: April 12, 2016

Presentation Location: San Francisco, California

**Authors/presenters retain copyright of the full-text paper. Permission to use content from this must be sought from the copyright holder.**

## Learning from Assessment Data: Epistemic Foundations of Data Use

Brette Garner and Ilana Horn, Vanderbilt University

Since the enactment of the No Child Left Behind Act of 2001, educators have been held accountable to increasing students' scores on standardized assessments, particularly in mathematics and English language arts. Teachers are encouraged to use data to "drive" instruction, but the details of this process vary across contexts (Datnow & Hubbard, 2015): What counts as data? Which data are emphasized, and for which students? How do educators draw conclusions from data? The answers to these questions have serious implications for equity in students' learning opportunities.

Data-driven decision-making (DDDM) models (e.g., Mandinach, 2012) outline processes by which educators can organize, analyze, and interpret data to inform instructional decision-making. While DDDM models provide a rational path to organizational improvement, these frameworks assume teachers have shared understandings about what data represent, what can be learned from data, and how to best respond to data. Yet as social studies of scientists have found, analyzing data to make evidence-based decisions is more complicated than this rhetoric suggests (Pickering, 2010; Goldstein & Hall, in press). We conjecture that the ways in which educators analyze, learn from, and respond to data are similarly complex and varied.

Meanwhile, there is a lack of theory and specification about what practices are central to productive data use for educators, which has serious implications for educational equity. Using techno-rational rhetoric, data can be used to justify the reification of systemic racism and classism at the district and school levels (Khalifa, Jennings, Briscoe, Oleszweski, & Abdi, 2013; Horn, under review). This is especially relevant in mathematics departments and in urban schools, where accountability pressures are high.

We address the following: How do mathematics educators learn from student test data to inform instruction? Teachers approach assessment data in many ways: some with more nuanced attention to sources of data and what they indicate about student thinking. Their approaches indicate epistemic stances about what data are and how they can be used to inform instruction. Teachers make *epistemic assumptions* about what data represent. These assumptions give rise to *epistemic practices* as teachers analyze data to draw conclusions. The practices consequentially privilege different *instructional responses*. By understanding the epistemic foundations underlying data use, we can inform the design of professional development, which will lead to more effective data use practices for teachers and more equitable learning opportunities for students.

### Conceptual framework

Drawing on the work of Horn and colleagues investigating professional learning opportunities in teacher workgroups, we focus on mathematics teachers' discussions in collaborative meetings, often under the facilitation of a principal or instructional coach. We take a situative view of teacher learning, assuming learning happens in interaction (Engeström & Sannino, 2010). To study teachers' professional learning opportunities, we examine how interactions (a) marshal conceptual resources for teachers and (b) mobilize teachers for future work (Hall & Horn, 2012; Horn, Kane & Wilson, 2015). In meetings with rich learning opportunities, teachers collectively

develop concepts about pedagogical issues, and then connect the concepts to their future instruction. These sorts of conversations provide learning opportunities insofar as they prepare teachers to change their professional practices. Horn and colleagues (2015) identified four key elements of teachers' conversations that shape learning opportunities: activity structures, frames, epistemic stances, and representations of instructional practice. In the present inquiry, we build on this literature by analyzing the epistemic foundations undergirding mathematics educators' data use.

The issue of epistemics is particularly salient in teachers' data use conversations, as data use is an inherently interpretive activity. When analyzing assessment data, teachers make *epistemic assumptions* about what data represent, what can be known from it, and why it is of value. Often, such assumptions remain tacit throughout conversation, but they manifest in the ways that teachers analyze data to draw conclusions (i.e., their *epistemic practices*). Teachers' epistemic assumptions and practices (or collectively, *epistemic foundations*) around data intersect with broader epistemic stances on mathematics teaching as they plan for future instruction.

In many teacher workgroup settings, teachers' data use is further influenced by the expectations of principals, coaches, and others in positions of authority. Instructional leaders provide teachers with various material resources (e.g., tables of numerical data, copies of the assessment, student work) and establish goals for data use activities (e.g., identify students for an intervention, plan next week's instruction). In this way, instructional leaders often embed their own epistemic foundations into the design of activities. Such designs are then taken up differentially, based on teachers' epistemic foundations.

## **Data and Methods**

The data for this analysis come from a larger design-based research study of instructional improvement in middle-school mathematics. From a sample of schools in two large, urban districts, we selected focal teacher workgroups based on the presence of catalysts for teachers' learning (e.g., an experienced instructional coach). We selected approximately eight groups in each of four years. During each year, we recorded approximately five meetings during which teachers planned to discuss data. Approximately 110 hours of recorded meetings from the primary corpus of data. From the larger study, we draw on secondary data sources that provide supplementary information about the school contexts and participants' backgrounds.

Across our corpus, data analysis was one of the most common activities for teacher workgroups; approximately 20% of meetings were focused on data. For the present inquiry, we analyzed the recorded meetings in which teachers focused on assessment data (approximately 22 out of 110 hours).

We use Episodes of Data Reasoning (EDRs) as our primary unit of analysis. EDRs are topically bounded segments of teachers' conversation in which they reason about or make sense of student data. For instance, an EDR might begin when teachers might ask, "What did students do on question #27? Why did only 42% of students answer it correctly?" or "What standards were the lowest on this assessment? Why?" The ensuing conversation could return to issues of instructional practice (e.g., "Question 27 was asked in a way different from how I had taught it"). We coded EDRs using a constant comparative method (Glaser & Strauss, 1965), paying special

attention to what participants in activity consider relevant and consequential for sensemaking, and treating learning as a members' phenomenon (Sacks 1967/1992; Stevens, 2010).

### **Preliminary findings**

There are multiple dimensions across which educators' use of data varies. In our preliminary analysis, we find that an epistemic assumption that is consequential for teachers' sensemaking is that of the *ontological status of data*--that is, what is the nature of data? What do data represent? Teachers' assumptions about the ontological status of data drive different epistemic practices around data use by necessitating varying types and amounts of information to draw conclusions about what students know and are able to do, as well as what future instructional responses are most appropriate.

The primary distinction that we draw is between data that is used as a *measurement* or as an *indicator* (Figure 1). Teachers who use data as a *measurement* of student learning treat like a car's fuel tank light. When the light comes on, the problem is straightforward: there is not enough gasoline in the tank. There is little room for error or interpretation; drivers typically trust that the reading is accurate. The solution is similarly clear: the driver must add more gasoline to the tank. The driver must select an appropriate station to purchase gas, but the essential response is the same regardless of their choice.

In an educational context, this approach arises when teachers take data as a straightforward measurement of what students know and are able to do. For instance, if students do poorly on an assessment item, teachers using a data-as-measurement approach may assume that students simply do not understand the content being assessed. Since the measurement of students' knowledge is straightforward, these teachers typically do not coordinate multiple data points to draw a conclusion; re-teaching is the appropriate response. Often, though, teachers using this approach emphasize either (a) which *content* needs to be re-taught or (b) which *students* need the intervention, with one almost to the exclusion of the other. Without considering the ways in which students interact with content, there are few opportunities to plan for more equitable instructional practices.

Teachers who use data as an *indicator* of student learning treat data more like a check engine light. Though a check engine light suggests that there is a problem with one's car, it does not point to a specific problem. To identify the problem, a driver (or mechanic) will need to triangulate different sources of evidence to gain a clearer picture of what is happening inside the car. Different diagnoses, then, require different solutions. Furthermore, it's common knowledge that the check engine light can come on because of a non-issue (e.g., a faulty wire); it may be that there is not a problem at all.

In an educational context, this approach arises when teachers assume that there is some uncertainty in the assessment of students' knowledge and skills: poor performance on an assessment item could be caused by any number of factors. Teachers using this approach must then coordinate multiple sources of information (e.g., student work, questions on related content, the wording of the item, etc.) in order to draw a more nuanced--and perhaps tentative--conclusion of what students know and are able to do. By considering the connections between students *and* content, teachers can open up discussions of access and equity.



Figure 1: Epistemic assumptions about the ontological status of data.

To illustrate each of these approaches, we present three EDRs. The episodes are drawn from meetings within six weeks of the state test, so accountability pressures are high. Teachers and their principals are using assessment from recent district benchmark assessments to plan instruction and interventions for the weeks remaining before the test.

Episodes 1 and 3 (data-as-measurement emphasizing content and data-as-indicator) come from Riverview Middle School, a diverse school in a large urban district. The principal of this school, Vera Cardwell, organized a full-day professional development session for teachers to analyze data from a recent district benchmark test. In Episode 1, we detail the approaches taken by the Riverview 6th-grade math team: Rachel, Crystal, and Devon. In Episode 3, we detail the approach taken by Ms. Cardwell.

Episode 2 (data-as-measurement, emphasizing students) comes from an hour-long meeting with the 7th-grade math teachers at Creekside Middle School, another diverse middle school in the same urban district. The Creekside principal, Mr. Russell, guides the teachers in using benchmark data to identify students for an intervention.

*Episode 1: Data-as-measurement--emphasizing content*

Throughout the Riverview teachers' data analysis session, the 6th-grade teachers frequently began EDRs by focusing on content--either assessment items or standards--that seemed problematic.

As this episode begins, Rachel looks at a list of assessment items by percentage correct. She points to one of the lowest-scoring items on the list and notes that the item assessed standard 6.2E, which addresses the order of operations. After identifying a problem area, Devon launches into a possible instructional strategy--using a checklist for students to go through the order of operations (Turn 6). After a brief discussion about Devon's strategy, Rachel offers a similar strategy involving a set of sticky notes that function as a checklist (Turn 19).

1.	Rachel:	Let's talk about what we're gonna do with 6.2E. Are we going to do the thing where they number each part?
----	---------	---

2.	Devon:	Where, what's E again? The adding fractions?
3.	Rachel:	6.2E is order of operations
4.	Devon:	I'd say a checklist--like you go down a list and you check off answers
5.	Rachel:	You do what now?
6.	Devon:	<i>[Gesturing in the air as he speaks]</i> 6.2E, order of operations? I'd make a PEMDAS checklist, and you check it as you, you know, do each one. Does it have this? No. Does it have this? Yes, okay, then they do it line by line. Maybe we should have them do it line by line. Maybe we should emphasize doing it line by line, just so <i>[unintel]</i> the problem each time, so that it makes a little V shape.
<i>[9 lines omitted: Teachers discuss getting a snack]</i>		
15.	Rachel:	<i>[Laughs]</i> Okay. So. You said you go line by line. Apparently what we've been doing doesn't work, but <i>[laughs]</i>
16.	Devon:	<i>[Unintel]</i> redo it. When I tried just doing it the way I did, <i>[unintel]</i>
17.	Rachel:	They didn't get it
18.	Devon:	When you do one problem each time, and draw it line by line, then students follow it.
19.	Rachel:	<i>[Nodding]</i> Mmhmm, that's good. I'm trying to remember, Ms. Jone--Ms. Jones showed me a way--it had to do with sticky notes. I think what she did was like, pretty much what you're saying. She--they had sticky notes for each part of order of operations, and they would put the sticky note. Oh, they would move over the sticky note that's in this problem. K, in this problem, there's no grouping, there's no--so I'm just going to move over the ones that are in this problem <i>[gesturing moving things over]</i> . And then like I have them stuck on my desk and I move over the ones. Now, I have to put these that I moved over in order, according to the order of operations, and I look at that while I solve the problem. That was her idea, I think that's pretty good.

During this episode, the Riverview teachers identify a problem in the data (few students answered a question about the order of operations correctly) and interpret it in a straightforward manner (students don't remember how to use the order of operations). Students are brought into

the conversation insofar as they “didn’t get it” (Turn 17) and should “follow” the checklist (Turn 18). There is no consideration of the content of the item (e.g., a word problem or an expression to simplify) or students’ sensemaking around the question.

Working under the assumption that a singular data point measures student learning about this particular content, the teachers use a logical approach: students performed poorly on the question, which means they don’t know how to use the order of operations, which means the teachers need to re-teach using a slightly different method. Yet this is problematic from an equity perspective, in that there is little room for student’s ideas, experiences, or funds of knowledge. There is no consideration of students’ sensemaking. And so this approach to data use is unlikely to result in more equitable instruction or in improved learning outcomes for students.

### *Episode 2: Data-as-measurement--emphasizing students*

At Creekside, the principal and 7th-grade teachers used a somewhat different version of data-as-measurement. For this meeting, Principal Russell pulled data from a recent district math assessment for the black students. In the previous year, Creekside failed to make Adequate Yearly Progress (AYP), in part because the black students’ math scores were lower than those of students in other subgroups. During this episode, Mr. Russell refers to the black students as “the most difficult sub-pop,” suggesting that they are the students who struggle the most at Creekside.

Mr. Russell provides teachers with a list of the black students in the grade and each students’ overall score on the benchmark assessment. He asks the teachers to categorize students according to their scores--as commended, passing, bubble, or growth--in order to allocate intervention resources. *Commended* and *passing* are categories set by the state; students who receive adequate scores pass, while those who receive excellent scores are identified as commended. *Bubble* students (c.f., Booher-Jennings, 2005) are those on the cusp of passing. *Growth* is a category used euphemistically, referring to the students with the lowest scores.

In his introduction to the activity, Mr. Russell emphasizes that the teachers are identifying students for interventions like Saturday school and math camps: “The activity for us is to get kind of an idea of where you’re at...which kids we need to be pushing into Saturday school.” He later clarifies that some students are marked with an exclamation point, which “means a kid that’s in that borderline area. He could be a good candidate for interventions.” Other students, however, are marked with an X; “the X’s would mean growth for sure, not a bubble kid.” And so the students who are being offered additional resources are not the students who are struggling the most (growth students, marked with an X), but rather those who are close to meeting the passing rate (bubble students, marked with an exclamation point). The teachers go on to categorize students according to this system.

In this meeting, Mr. Russell and the teachers use test scores as a single, unidimensional measurement of students’ mathematical knowledge. Students’ identities are similarly reduced to a binary variable (i.e., black or non-black). These variables--which are gross reductions of complicated constructs--are then being used to distribute resources inequitably. The students who most need additional instructional interventions--that is, the “growth” students--are systematically denied help. Furthermore, students in other subgroups--e.g., English Language Learners, special education students, etc.--are not considered for additional interventions.

Like the approach used by the Riverview teachers, there is a certain logic to the activity Mr. Russell organizes for the math teachers. Assuming that students’ test scores are a measurement of their mathematical knowledge--and that the black students’ passing rate is most critical for meeting AYP--identifying students who need a small push to pass and targeting them for interventions is a reasonable response. Yet again, there is little consideration for students’ sensemaking about mathematics, how future instruction might meet their learning needs, or what assets and strengths students can build on.

*Episode 3: Data-as-indicator--connecting students and content*

The Riverview principal, Ms. Cardwell, takes a different approach to data use. During her introduction to the data analysis session, she tells teachers to “look at the item, study the SE [standard], what part of the SE was addressed, what did the kids struggle with—BAM, that’s your finding.” She describes an approach to data use that requires the coordination of many pieces of information (the item, the standard, and student’s sensemaking around the content). Throughout the day, she works with different groups of teachers for short periods of time. In this episode, Ms. Cardwell works with the 6th-grade teachers to analyze a problem assessing rational number conversion (Figure 2).

<b>31</b>	Jeremy bought a skateboard on sale for \$28, which was 12.5% off the original price. What was the discount as a fraction of the original price?
<b>A</b>	$\frac{1}{4}$
<b>B</b>	$\frac{12}{5}$
<b>C</b>	$\frac{5}{7}$
<b>D</b>	$\frac{1}{8}$

Figure 2: A question from a 6th-grade district benchmark assessment.

1.	Ms. Cardwell:	<i>[Ms. Cardwell leans over SE between Devon and Crystal, reading to herself] Where am I at? Okay. Uhh, the student represents and uses rational numbers in a variety of equivalent forms. Doot doot do. Okay. So how have y'all approached that? [Stands]</i>
2.	Rachel:	Which one are you talking about talking about? 6.1B?
3.	Devon:	Um, we've done actual physical representations of, um, transformations and things. And examples of real life percentages and decimals and fractions and things.
4.	Crystal:	And we actually had to do a real world project on that
5.	Ms. Cardwell:	Okay



6.	Crystal:	On, um, actually it was [unintel] converting fractions to decimals to percents.
7.	Devon:	[unintel]
8.	Ms. Cardwell:	So what, what item is the one that's low?
9.	Devon:	It's 31 [ <i>Points to the problem in the test booklet</i> ]
10.	Ms. Cardwell:	Let me see it.
<i>[Devon hands booklet to Ms. Cardwell]</i>		
11.	Crystal:	And they had to convert twelve and a half percent.
12.	Rachel:	It's because the percent had a decimal in it. I really believe that.
13.	Devon:	But it's every time [unintel]
14.	Rachel:	Well that's the reason. Because the decimal had a percent in it. I mean, the percent had a decimal in it. I said it backwards.
15.	Crystal:	We even had them--we showed them a different way--we even had them put it on a number line
16.	Ms. Cardwell:	Mm-hmm
17.	Crystal:	And they had the fraction, and they had to match the [unintel] find the one that matched it, they had to go up to it and find the percent, and the cards were all pre-done, so it was intentional. They didn't have student-created cards. Then we had them take that hands-on activity and then create conversions on paper. And um

Upon joining the teachers, Ms. Cardwell immediately identifies a number of pieces of information: the text of the standard, the item in question, and the teachers' instructional approaches. Though Rachel offers an initial conclusion (that there was a percent with a decimal in it--i.e., 12.5%), Ms. Cardwell seeks further information:

18.	Ms. Cardwell:	Let me see the percent. What percent did they have on this one?
19.	Crystal:	Um, on forty--no, thirty-one?

20.	Devon:	Twenty-six
21.	Crystal:	Twenty-seven percent. It's 27%
<i>[Ms. Cardwell leans over to look at Devon's papers]</i>		
22.	Ms. Cardwell:	So was it a guess? Was it about twenty-something all the way across the board?
23.	Devon:	No
24.	Ms. Cardwell:	No
25.	Devon:	They actually chose B <i>[12/5]</i> more often

After identifying the percentage of students who answered the question correctly (27%), Ms. Cardwell asks for the most common answer choice (B--12/5). She also asks to see other questions on the assessment that addressed the same standard, though the teachers do not find any. Throughout the beginning of this episode, Ms. Cardwell collects multiple sources of information before coordinating them to draw tentative conclusions about what students know and are able to do with rational numbers:

27.	Ms. Cardwell:	Okay, so I would say, I would say, then, this is probably--I mean, my gut--it could be a combination of the .5, however, I--it's probably rooted in the question.
28.	Crystal:	I thought, that's what I thought
29.	Ms. Cardwell:	You know what I mean? They didn't realize that you were finding the equivalent form of a number. They might even try to--
30.	Crystal:	--do something with the 28, they don't know how to omit the non-necessary information
31.	Devon:	But we've seen that exact, um--I've given a, I've written a for sure question on a practice sheet that's in just the same format
32.	Rachel:	But it did not have a decimal in the percent. That's what goofed them up, I really believe that. They have not dealt with decimals in percents enough.

33.	Ms. Cardwell:	So what I would ask them, then, if you want to deduce that right? And you want, you know, to determine that that definitely is, then the best way is to have every kid take okay, 12.5%, represent it as a fraction. Take all the words out.
34.	Rachel:	Mmhmm. And they don't know how to do that.
35.	Ms. Cardwell:	And that, and that will tell you immediately whether or not that was one of the layers of problem. But I know, when we we talked about, um--an-an-and those are common ones that we usually go over. Like, okay, when I was a 6th grade teacher, at this point in the time, I don't know if I would say that that was overwhelmingly all of my kids' issues. Because I had a chart they filled out every single day, like, starting in January, that showed--12.5, I want it as a fraction, I want a decimal, you know, I want it as uh, uh, a percent, or whatever, my conversions. So we would include those. They do get heebie-jeebies about it, I agree, and most of them probably just thought [ <i>pause, points and looks down at problem</i> ]. Maybe they don't understand out of a hundred.
36.	Rachel:	I think they know out of a hundred
37.	Ms. Cardwell:	I mean, there's a lot of—Whoa::, not if they chose 12 over 5, they didn't, because you would have seen 12.5 over 100, at least, and then they don't know to move it over, over a thousand, and ugh, there's just a lot going on there.

In identifying multiple “layers of problem,” Ms. Cardwell hypothesizes about various ways that students may have interacted with the item. She gives consideration to Rachel’s that students may have been confused because the question used 12.5% instead of a whole percentage, like 12% (Turn 27). But she also considers the wording of the item (Turn 29) and later clarifies that prepositional phrases are particularly tricky for students to parse. She even calls students’ understanding of percent into question (Turns 35, 37), citing the most commonly selected answer choice.

Throughout this episode, Ms. Cardwell assumes that data is an indicator of what students know and are able to do: though students performed poorly on this item, there could be many sources of trouble. Only by soliciting and interpreting other sources of information--including the teachers’ previous instruction--can she hypothesize about students’ sensemaking around the item. She later proposes a scaffolded instructional response that addresses each of the various “layers of problem.”

The data-as-indicator approach, then, has greater potential to promote equity in instruction. Ms. Cardwell is able to develop a richer, more nuanced interpretation of the data and

connect it to students' thinking and their learning needs. She designs responses to assessment data that allow for richer considerations of students, their perspectives, and their ways of thinking.

### **Conclusion**

Assuming that data is a measurement of student knowledge is a rational approach under NCLB, particularly given the priority given to students' overall assessment scores. However, this approach is unlikely to lead to improved instruction or more equitable educational outcomes. Assuming, instead, that data is an indicator of student knowledge is a more complicated approach, requiring a more sophisticated consideration of students and their mathematical thinking. Yet this consideration can also open up a space to consider issues of access and equity.

As we further our analysis, we plan to examine the ways in which expertise and experience shape the ways educators "see" data and the resources that they bring to bear in analysis, as well as what supports are necessary for teachers to learn more nuanced data use practices. We also instances across our corpus in which more experienced facilitators (e.g., Ms. Cardwell) model more sophisticated data practices for others. Though others (e.g., Rachel) are able to follow the model, they do not replicate similar practices when the facilitator is absent. We conjecture that attending to teachers' perspectives on data is necessary in order to facilitate teachers' development of more sophisticated and equitable uses of data.

### **References**

- Booher-Jennings, J. (2005). Below the bubble: Educational triage and the Texas accountability system. *American Educational Research Journal*, 42(2), 231-268.
- Datnow, A., & Hubbard, L. (2015). Teachers' use of assessment data to inform instruction: Lessons from the past and prospects for the future. *Teachers College Record*, 117(4).
- Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review*, 5(1), 1-24.
- Goldstein, B.E., & Hall, R. (in press). Modeling without end: Conflict across organizational and disciplinary boundaries in habitat conservation planning. In J. Kaput, E. Hamilton, S. Zawojewski, and R. Lesh (Eds.), *Foundations for the future*, Erlbaum.
- Hall, R. & Horn, I.S. (2012). Talk and conceptual change at work: Adequate representation and epistemic stance in a comparative analysis of statistical consulting and teacher workgroups. *Mind, Culture, and Activity*, 19(2), 240-258.
- Horn, I.S., (under review). Accountability as a Design for Teacher Learning: Sensemaking about Mathematics and Equity in the NCLB Era.

- Horn, I. S., Kane, B. D., & Wilson, J. (2015). Making Sense of Student Performance Data Data Use Logics and Mathematics Teachers' Learning Opportunities. *American Educational Research Journal*, 0002831215573773.
- Khalifa, M. A., Jennings, M. E., Briscoe, F., Oleszweski, A. M., & Abdi, N. (2013). Racism? Administrative and community perspectives in data-driven decision making: Systemic perspectives versus technical-rational perspectives. *Urban Education*, 0042085913475635.
- Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist*, 47(2), 71-85.
- Pickering, A. (2010). *The mangle of practice: Time, agency, and science*. University of Chicago Press.
- Sacks, H. (1992). Omnirelevant devices; settinged activities; "indicator terms" (February 16, 1967). In G. Jefferson (Ed.), *Lectures on conversation: Volumes I and II* (pp. 515–522). Oxford, England: Blackwell. (Original work published 1967)
- Stevens, R. (2010). Learning as a Members' Phenomenon: Toward an Ethnographically Adequate Science of Learning. *Yearbook of the National Society for the Study of Education*, 109(1), 82-97.