THE JOURNAL OF THE LEARNING SCIENCES, *11*(2&3), 187–215

Reasoning With Tools and Inscriptions

Paul Cobb
Department of Teaching and Learning
Vanderbilt University

The unit of analysis that I use when discussing the 2 sample episodes is that of a classroom mathematical practice together with the students' diverse ways of contributing to its continual regeneration. Analyses cast in terms of this unit account for the mathematical learning of the classroom community. As I clarify, a classroom mathematical practice is itself composed of 3 interrelated types of norms: a normative purpose, normative standards of argumentation, and normative ways of reasoning with tools and inscriptions. In keeping with the theme of this special issue, I step back from the sample analysis by focusing on the last of these 3 aspects. In doing so, I introduce the notion of a chain of signification to illustrate a way of accounting for mathematical learning in semiotic terms.

The dominant view of mathematical symbols as external representations has been challenged in recent years by an alternative perspective that emphasizes the activity of symbolizing. In this newer perspective, the focus of investigations shifts away from the analysis of symbols as external supports for reasoning and moves toward students' participation in practices that involve symbolizing. Rather than describing the properties of tools such as physical devices, computer icons, and notations independently of their use, this perspective treats symbolizing as an integral aspect of mathematical reasoning (Dörfler, 1993; Kaput, 1994; Meira, 1995, 1998; van Oers, 1996). This change in analytic focus is informed by contemporary semiotics and leads to rejection of the view that the process of developing meaning for symbols (broadly defined) involves associating them with separate, self-contained referents. Instead, the ways that symbols are used and the meanings they come to have

Correspondence and requests for reprints should be sent to Paul Cobb, Vanderbilt University, Peabody College Box 300, Nashville, TN 37203. E-mail: paul.cobb@vanderbilt.edu

are seen to be mutually constitutive and to coemerge[1] (van Oers, 2000). In the nondualist approach to symbol and meaning that I will take when analyzing the *Batteries* and *AIDS* episodes, my primary concern will be to tease out relations between the ways that tools and symbols are used and the mathematical meanings that they come to have. For ease of explication, I reserve the term *tool* for the two computer minitools that the teacher and students used during the design experiment, and I speak of *inscriptions* when referring both to written notations and to the graphical displays produced by using the minitools.

Because I participated in the classroom design experiment, the overall orientation from which I will analyze the episodes is that of an instructional designer. I have argued elsewhere that this orientation gives rise to two basic challenges (P. Cobb, 2001). The first of these concerns the time frame of the analysis. The immediate task at hand is to analyze video recordings that span a few minutes. However, my interest is to contribute to the improvement of the instructional design that was enacted during the 34 classroom sessions of the design experiment by delineating major shifts in mathematical reasoning. For my purposes, therefore, the relevant time scale is one of weeks and months rather than seconds and minutes. As a consequence, I propose relatively global conjectures about the mathematical practices established by the classroom community with the understanding that their viability should be tested by analyzing other episodes. In practice, my colleagues and I typically work through the entire data set generated during a design experiment by using a variant of Glaser and Strauss's (1967) constant comparison method. We discussed and illustrated this analytical approach in some detail in a previous issue of this journal (P. Cobb, Stephan, McClain, & Gravemeijer, 2001).

The second challenge that emerges from my orientation as an instructional designer involves the choice of an appropriate unit of analysis. It is important for my purposes that analyses feed back to inform the improvement of the instructional design. As my colleagues and I have argued elsewhere, this requires that we analyze the evolution of the classroom microculture that constitutes the immediate situation of the students' learning (cf. Gravemeijer, 1997). Given this interest, an appropriate an-

[1]As Sfard (2000) observed, meaning sometimes is equated with the way that symbols are used. She noted, for example, that Wittgenstein (1953) defined the meaning of a specific word in terms of the set of rules that govern the use of that word. It is therefore important to note that a strict adherence to this view that meaning can be reduced to accustomed use leads to an epistemological behaviorist position in which understanding involves nothing more than acting in accordance with established conventions (cf. Rorty, 1979). Crucially, a position of this type that focuses exclusively on social use from the observer's perspective does not address what might be termed the experiential aspects of meaning. In the case of mathematics, this experiential aspect includes the sense of acting in a mathematical reality that is assumed to be at least partially shared with others (Bloor, 1976; Davis & Hersh, 1981; Greeno, 1991; Schoenfeld, 1998; Sfard, 1991). To differentiate the viewpoint I illustrate in this article from the position that mathematical meaning can be equated with acting in accord with conventional rules, I follow van Oers (2000) in speaking of the coemergence of meaning and ways of symbolizing.

alytical unit consists of a classroom mathematical practice together with students' diverse ways of contributing to its continual regeneration. In documenting the emergence of new practices as reorganizations of previously established practices, we in effect account for the mathematical learning of the classroom community. Furthermore, in conducting such analyses, we chart the emergence of what is traditionally called mathematical content. However, as will become apparent, the objectification inherent in the standard notion of content is displaced by a concern for the process of mathematizing in which the use of tools and inscriptions plays a central role.

In introducing this unit of analysis, I must clarify what I mean when I speak of a classroom mathematical practice. This is essential given that the term *practice* enjoys wide currency at present and has a variety of different meanings. As I illustrate when analyzing the two sample episodes, a classroom mathematical practice consists of three interrelated types of norms:

1. Normative purpose.
2. Normative standards of argumentation.
3. Normative ways of reasoning with tools and inscriptions.

These three types of norms delineate three aspects of mathematical activity on which my colleagues and I have found it useful to focus for our purposes as instructional designers. I make this observation because other more extensive schemes for parsing mathematical activity have been proposed (e.g., Kitcher, 1984). In this regard, our overriding concern has been to develop analytical tools that are adequate for our purposes rather than to be exhaustive in analyzing mathematical activity.

This pragmatic orientation acknowledged, it is worth noting that three types of norms which comprise a mathematical practice are compatible with the distinctions that Wenger (1998) made between three interrelated dimensions of a community of practice. In the case of a classroom community, the first aspect that Wenger discussed, a joint enterprise, seems to correspond closely to the normative purpose for engaging in mathematical activity established by a classroom community. The second dimension of a community of practice that Wenger identified, mutual relations, involves the constitution of normative standards of argumentation as well as more general classroom social norms. Wenger's third dimension, a well-honed repertoire, appears to relate directly to normative ways of reasoning with tools and inscriptions. As I have indicated, my colleagues and I did not attempt to apply Wenger's analysis of communities of practice to the mathematics classroom when we developed the construct of a classroom mathematical practice. However, the parallels locate our analytical approach in a broader theoretical context.

The definition that I have given of a classroom mathematical practice gives rise to one further issue that must be addressed before I analyze the two episodes: clarifying my use of the term *normative*. Briefly, norms are inferred by discerning patterns or regularities in the ongoing interactions of members of a community, such as the teacher and students in a classroom. A norm, therefore, is not an individualistic notion but instead refers to patterns in collective activity within a community. As a consequence, my colleagues and I view classroom norms as joint accomplishments that are interactively constituted by the teacher and students. Thus, we emphasize that the students necessarily have to play their part even when a teacher consciously expresses her institutionalized authority in action by attempting to initiate the renegotiation of classroom norms.

Methodologically, the process of developing and testing conjectures about classroom norms involves, in part, teasing out those aspects of activity that appear to be beyond justification. Although evidence of this type is indirect and thus less than conclusive, it can be given some weight in cases where students view it as their obligation to question peers when they do not agree with them or understand what they are saying. This was the case in the classroom in which the statistics design experiment was conducted (McClain, Cobb, & Gravemeijer, 2000). A second, more robust type of evidence, is indicated by Sfard's (2000) observation that normative ways of acting are not mere arbitrary conventions for members of a community that can be modified at will. Instead, these ways of acting are value-laden in that they are constituted within the community as legitimate or acceptable ways of acting.[2] This observation indicates the importance of searching for instances where a student appears to violate a proposed norm to check whether his or her activity is constituted in the classroom as legitimate or illegitimate. In the former case, it would be necessary to revise the conjecture, whereas in the latter case, the observation that the student's activity was constituted in the classroom as a breach of a norm provides evidence that supports the conjecture (cf. P. Cobb et al., 2001). Finally, a third and even more direct type of evidence occurs when the teacher and students talk explicitly about their respective obligations and expectations. Such exchanges typically occur when one or more of the participants perceive that a norm has been violated.

As the discussion of this methodological point makes clear, the mathematical norms that compose a classroom mathematical practice are concerned with the status that particular ways of reasoning with tools and inscriptions come to have in public classroom discourse. The notion of mathematical norms and thus of classroom mathematical practices therefore relates directly to issues of power and authority, in that a student who violates a mathematical norm is not merely viewed as

[2]This observation also bears directly on the process by which the members of a community develop distinct identities as they participate in the continual regeneration of communal norms (cf. Schulz, 1962; Wenger, 1998). In particular, members of a community do not merely act in accord with the norms of a community. They become people who consider that their rights have been infringed when they perceive that a norm has been breached.

having made an alternative interpretation. Instead, the student is positioned in classroom discourse as not understanding how the world stands mathematically.[3]

In keeping with the theme of this special issue, the analysis I present of the two episodes gives particular emphasis to normative ways of reasoning with tools and inscriptions. However, because of the interdependency of this and the other two aspects of a mathematical practice, I consider each aspect in turn for both sample episodes. Against the background of this sample analysis, I then step back to locate my discussion of normative ways of reasoning with tools and inscriptions within the context of a wider spectrum of ideas. In doing so, I introduce the notion of a chain of signification to illustrate a way of accounting for mathematical learning in semiotic terms. As will become apparent, in addressing these issues, I purposefully fail to highlight the teacher's role, because this is the focus of the analysis that Kay McClain presents in this special issue. Furthermore, in keeping with the charge to all contributors, I do not discuss the rationale for the instructional activities or the computer minitools. As a consequence of these omissions, it might appear at times that the norms I document emerge without initiation or guidance, and that the students miraculously develop relatively sophisticated forms of mathematical reasoning. To counter this interpretation, I refer the reader to McGatha's (2000) detailed description of the "design story" of this teaching experiment.

NORMATIVE PURPOSE

The instructional tasks used in the two episodes both involved comparing two data sets to decide which of two options was superior in some way. At a relatively immediate level, it would be possible to analyze the local goals that emerged for the students as they presented and critiqued analysis of the batteries data and as they judged the adequacy of inscriptions taken from their analyses of the AIDS data. However, the purposes on which I focus when attempting to delineate classroom mathematical practices are more global in nature. As a consequence, I did not conduct a detailed turn-by-turn analysis of each episode.

[3]Much and Shweder (1978) distinguished between five qualitatively distinct types of norms: regulations, conventions, morals, truths, and instructions. P. Cobb, Wood, Yackel, and McNeal (1992) built on this work to distinguish what might, in colloquial terms, be described as classrooms where mathematics is learned with understanding from those where it is learned without understanding. Cobb et al. argued that the mathematical norms established in the former type of classroom were truths, whereas those established in the latter type of classroom were procedural instructions. The assertion that students who violated a mathematical norm in the stats project classroom were positioned as not understanding how the world stood mathematically indicates that the mathematical norms constituted in this classroom had the quality of mathematical truths. If the norms had been procedural instructions, students would have been positioned as being ineffective rather than not understanding normative meanings that are endowed with ontological significance with respect to the topic under discussion.

Instead, my goal as I analyzed each episode was to tease out what this group of people was doing together mathematically. My intent in addressing this issue was to develop conjectures that were not specific to these particular episodes but that could be tested while analyzing the entire corpus of video-recordings generated during the design experiment. Given this purpose, I should clarify that I have previously conducted a longitudinal analysis of this entire data corpus (P. Cobb, 1999). My primary concern when doing so was to identify major shifts in the students' reasoning about data. In this article, I reconstruct the process of developing general conjectures by foregrounding the theoretical perspective that informs my analytical approach.

The first claim I want to make is that the normative purpose in both episodes involved analyzing data. At first glance, this might appear to be a trivial observation. However, performance assessments that we previously conducted with a group of seventh graders in the same school in which we conducted the design experiment indicated that, for those students, data analysis involved "doing something with numbers" (McGatha, Cobb, & McClain, 1999). In other words, data were not, for these students, measures of an aspect of a phenomenon that were generated to address a particular problem or issue. Instead, for these students, data were simply numbers that they were expected to manipulate. In contrast to these performance assessments, the interpretation of the horizontal bars in the first minitool as signifying the life span of batteries (Figure 1 in the Appendix, article by McClain, this issue) appeared to be normative in the first episode. Similarly, in the *AIDS* episode, the interpretation of the various inscriptions as signifying patients' T-cell counts (Figures 9–12 in the Appendix) appeared to be normative. Against the background of these general assertions, I now focus on the normative purposes for analyzing the batteries and AIDS data sets. The conjecture that I formulated when I first analyzed the *Batteries* episode was that this purpose was to identify trends and patterns in the data that were significant with respect to the question at hand. For example, Jamie stated her understanding of Caesara's argument as follows:

> She's saying that out of 10 of the batteries that lasted the longest, 7 of them are green, and that's the most number, so the *Always Ready* batteries are better, because more of those batteries lasted longest. [16]

In this case, the pattern she identified was that 7 of the 10 longest lasting batteries were *Always Ready* batteries. Blake, for his part, described a pattern that he considered significant:

> Now, see, there's still green ones behind 80, but all the *Tough Cell* is above 80. So I'd rather have a consistent battery that I know that will get me over 80 hours than one that just try to guess. [30]

The ensuing exchange between Blake and Jamie is also helpful in delineating the normative purpose for analyzing data. Jamie initiated the exchange by requesting clarification from Blake, "Um, why wouldn't the *Always Ready* batteries be consistent" [34] and appeared to be positioned in the subsequent discussion as failing to understand that the data at hand were a sample of all the batteries of each brand. It is important to stress that this claim about Jamie's positioning is concerned with the status that her contributions came to have in public discourse rather than with how she actually understood Blake's arguments. Thus, although she might have understood the distinction between sample and population, Blake's and the teacher's contributions positioned her as failing to understand [37–41]. In the process, the purpose for analyzing the data that appeared to be emerging as normative was to compare the two brands of batteries rather than these two particular batches of 10 batteries. It should be clear that this inference does not involve a claim about consensual meanings. There is, for example, no indication that Jamie was party to the apparent agreement between Blake and the teacher. Instead, the claim rests on inferences about the purposes for analyzing data that were constituted as legitimate and as illegitimate in public classroom discourse.

The process of inferring the normative purpose for doing statistics during the *AIDS* episode is necessarily less direct because the students did not present their analyses but instead assessed the adequacy of a series of inscriptions. In conducting the analysis, I therefore found it necessary to focus primarily on what was not said by attempting to tease out the implicit purpose for which the students judged the inscriptions to be adequate or inadequate. Consider, for example, Jamie's assessment of the first inscription that was discussed (Figure 9 in the Appendix):

> I think it's a pretty adequate way of sharing information because you can see where the range is starting and ending and you can see where the majority of the numbers are. [52]

Because Jamie's assessment was constituted as legitimate in the subsequent discussion, the question I sought to address was that of inferring the purpose for which an inscription that shows the ranges and the location of the majority of the numbers in both data sets might be adequate. Drawing on the conjecture I made when analyzing the *Batteries* episode, I speculate that such an inscription is adequate if the normative purpose is to delineate global patterns in data that are significant with respect to the issue under investigation.

This claim about the normative purpose remained viable as I worked through the remaining *AIDS* subepisodes. For example, Blake justified his positive assessment of the final inscription in which the data had been partitioned into four equal groups (Figure 12 in the Appendix):

> Well, it doesn't really matter where all the data is because you know from where the groups are what, what treatment is better or where the data stand on both treatments. [152]

I infer that in saying that it "doesn't really matter where all the data is," Blake was arguing that is not necessary to know the location of all the individual data points. The inscription was adequate in his view because it shows "where the data stand" or, in other words, global patterns in the data. Furthermore, he argued that these patterns were significant with respect to the issue under investigation in that "you know…what treatment is better."

The final issue I considered when inferring the normative purpose for analyzing data in the *AIDS* episode was whether data sets were samples. However, the students' contributions to the discussion are ambiguous on this point. In particular, the assessments they made of the inscriptions could be interpreted either as statements about patterns in the two specific batches of data or as conclusions about the effectiveness of the two treatments in general. That the contrast between these two interpretations did not emerge as an explicit topic of conversation could indicate that the issue had been resolved in prior discussions. However, in the absence of direct evidence, the conjecture I made when analyzing the *Batteries* episode remains to be tested by examining other episodes from the data corpus.

If we step back from the details of this sample analysis, the conjectured normative purpose of identifying global trends and patterns in data that are relevant to the question under investigation is significant in that it relates directly to the epistemology of the discipline. Statistics is concerned with situations characterized by variability and uncertainty. In the *Batteries* episode, for example, the life spans of the batteries of both brands varied and the life span of any particular battery could not be predicted with any certainty before it was tested. Statisticians seek to cope with this uncertainty by analyzing patterns that emerge at the macrolevel as data accumulate (G. Cobb & Moore, 1997). Their primary focus is on patterns in variation and change. My conjecture that the normative purpose established in the design experiment classroom was consistent with this disciplinary orientation implies that the teacher and students actually were doing statistics together. In general, an analysis of the normative purposes constituted by a classroom community is useful in delineating the overall goals of an instructional design as they are realized in the classroom. An analysis of this type frequently reveals that these purposes differ significantly from the goals the designer envisioned at the outset, thereby feeding back to inform the revision of the instructional design.

NORMATIVE STANDARDS OF ARGUMENTATION

My focus when analyzing the second aspect of a classroom mathematical practice, normative standards of argumentation, was on the norms for what counted as a legitimate data-based argument. The negotiation of these norms appeared to be explicit in the *Batteries* episode. At the beginning of the episode, Caesara developed an argument that seemed to focus on the number of batteries of each brand that were among the 10 longest lasting batteries [8–12]. In revoicing Caesara's argument, Jamie both stated an explicit conclusion ("the *Always Ready* batteries are better") and gave a warrant that explained why the data support this conclusion ("because more of those batteries lasted the longest" [16]). I draw on Toulmin's (1969) scheme of argumentation to illustrate the structure of the explanation developed jointly by Caesara and Jamie (Figure 1). In this scheme, the term *data* refers to the support that is given for a conclusion.

In the case of the batteries instructional activity, for example, one could imagine a student merely pointing to the two data sets and stating the conclusion that one of the brands is superior. In doing so, the student would treat the conclusion as a self-evident consequence of the data. If questioned, the student would be obliged to give a warrant that explains why the data support the conclusion (e.g., the conclusion that "the *Always Ready* batteries are better" is supported by the data "because more of those batteries lasted the longest").

As the episode continued, Jason challenged Caesara and Jamie's argument [18], thereby prompting the teacher to ask Caesara to justify her focus on the 10 longest-lasting batteries [19–21]. In Toulmin's (1969) terms, Caesara therefore was obliged to give a *backing* that indicates why her warrant should be accepted as having authority. She did so by explaining that a total of 20 batteries were tested and "half of the 20 is 10 so that's how I chose it" [20]. Crucially, this backing did not refer to the issue under investigation, that of deciding which of the two sets of batteries lasted longer. In contrast, when Blake explained his method of comparing the number of batteries of each brand that lasted more than 80 hr, he referred to the

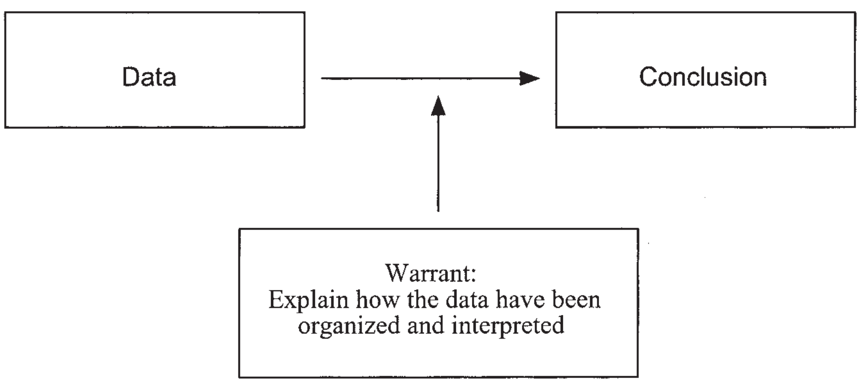

FIGURE 1 Structure of Caesara and Jamie's argument in the *Batteries* clip.

issue under investigation explicitly [30]. In doing so, he interpreted a feature of the inscription ("there's still green ones behind 80") as indicating a difference between the two brands that he considered significant. The structure of his argument is shown in Figure 2.

A key difference between Blake's explanation and the explanation developed by Caesara and Jamie was that Blake gave a backing to justify why the way in which he had organized the data was relevant to the issue of deciding which brand of battery was superior.

Later in the episode, Jessica compared Caesara's and Blake's analyses:

> I was just going to say that well, even though 7 of the 10 longest-lasting batteries are the *Always Ready* ones, the 2 lowest are also *Always Ready* and if you were using these batteries for something important, then you might end with one of the bad batteries and could [inaudible]. [42]

In making this argument, Jessica differentiated between Caesara's and Blake's analyses in terms of their practical consequences. Her argument was constituted as legitimate in the subsequent exchange, thereby contributing to the ongoing negotiation of standards of argumentation. In particular, the obligation to give a backing that justified why a particular way of organizing data (i.e., a warrant) was relevant to the issue at hand appeared to be emerging. Although it would be strong too claim

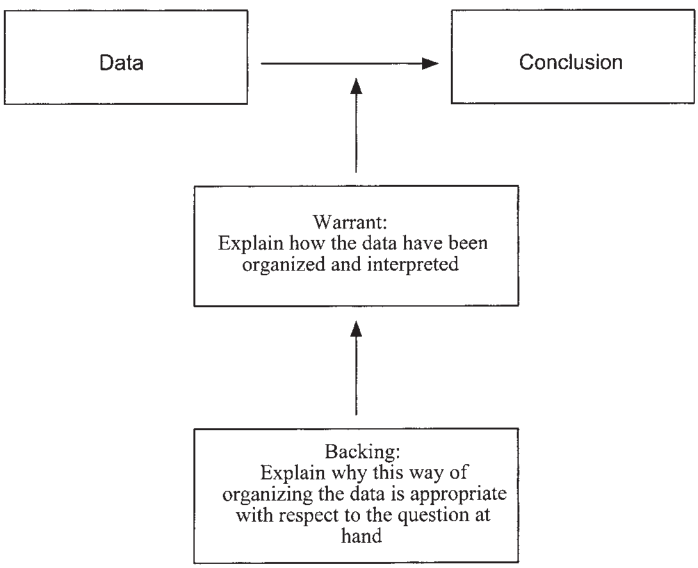

FIGURE 2 Structure of Blake's argument in the *Batteries* clip.

that this obligation was normative on the basis of this short episode, the analysis of the *AIDS* episode provides the opportunity to test this conjecture.

Because the students did not present their analyses during the discussion of the AIDS data, I attempted to infer the unspoken arguments that the students judged as adequate or inadequate when they assessed the inscriptions. An incident that is particularly helpful in this regard occurred during the discussion of the second inscription in which the two data sets had been partitioned at the T-cell value of 550 (Figure 10 in the Appendix). Vallory appeared to question the adequacy of the inscription:

> Why did they…550? I don't know? Why is 550 so important? Because the median is really 550—no, it's not, but it's not 550. [78]

In posing this question, Vallory seemed to call for a backing that justified why this way of organizing the data was relevant to the issue of determining the relative effectiveness of the two treatments. Significantly, several students indicated that they wanted to respond to her question without prompting by the teachers. This constitutes relatively strong evidence that the norms of argumentation included the obligation to give such a backing by this point in the design experiment.

I found further supporting evidence for this conjecture as I analyzed the other *AIDS* subepisodes. For example, consider again Blake's assessment of Inscription 5 (Figure 12 in the Appendix), in which the two data sets had been partitioned into four equal groups:

> Well, it doesn't really matter where the data is because you know from where the groups are what, what treatment is better or where the data stands on both treatments. [152]

In claiming that the groups show "where the data stands," Blake gave a backing to the effect that the inscription shows how the two data sets are distributed, thereby making it possible to see relevant patterns.

I had to make one modification to my conjecture about normative standards of argumentation when I analyzed the *AIDS* episode. I can best illustrate the need for this change by focusing on Jamie's assessment of Inscription 1 (Figure 9 in the Appendix), in which the location of "most of the numbers" was shown for both data sets.

> I think it's a pretty adequate way of showing information because you can see where the range is starting and ending and you can see where the majority of the numbers are. [52]

Jamie's explanation precipitated a discussion of the meaning of the term *majority*, an issue that I examine when I focus on normative ways of reasoning with tools and inscriptions. My immediate interest is with what the inscription was judged to show in an adequate way. Presumably, it was that the majority of the data in the experimental treatment were located in a higher part of the range than in the standard treatment. The underlying argument might have been that experimental treatment was more effective for three reasons:

1. The data sets have been structured by isolating the intervals in which the data were piled up.
2. This way of organizing the data shows the location of the majority of the data points in each data set.
3. It is therefore possible to see differing patterns in how the data are distributed.

The details of this inferred argument are, of course, open to question. My purpose in sketching it is to illustrate that a warrant that explains how the data were organized and interpreted can reasonably be assumed. Furthermore, there is no indication that an explicit backing was needed. Instead, the teachers and students appeared to take it for granted that this way of organizing the data resulted in the identification of patterns that were relevant to the question under investigation. In other words, the resulting patterns appeared to be ready to hand and thus beyond justification. This suggests that although an argument required a warrant, a backing was only needed when the reason for organizing the data in a particular way could not be taken as self-evident. Again, because this inference is based on the analysis of only a small fraction of the data corpus, it should be viewed as a conjecture that is subject to refutation when other episodes are examined.

In reflecting on the approach to standards of argumentation that I have illustrated, I should acknowledge that Toulmin's (1969) scheme frequently has been critiqued, not least because of its structuralist orientation. In this regard, it is readily apparent that my analysis does not capture a number of aspects of classroom discourse that are highlighted by other contributions to this special issue. I nonetheless contend that Toulmin's scheme is useful for my purposes as an instructional designer. To justify this claim, I draw on a distinction that Thompson and Thompson (1996) made between what they term *calculational* and *conceptual* orientations in teaching. My colleagues and I have found it useful to extend this distinction by talking of calculational and conceptual discourse[4] (P. Cobb, 1998; P. Cobb et al., 2001). It is important to stress that calculational discourse does not refer to conversations that focus merely on the procedural manipulation of conventional mathematical inscriptions. Instead, calculational discourse refers to discus-

[4]I use the term *discourse* in its colloquial, everyday sense to refer to the talk of members of the classroom community. This use should be distinguished from the more technical notion of a discourse as defined by Gee (1997). An analysis of normative standards of argumentation is concerned with the quality of classroom talk and serves to tease out the often implicit criteria for what counts as an acceptable mathematical argument in a particular classroom. In differentiating between calculational and conceptual discourse, I draw a distinction between two qualitatively distinct norms for argumentation.

sions in which the primary topic of conversation is any type of calculational process that is enacted to produce a result or arrive at a conclusion. These conversations can be contrasted with conceptual discourse in which the reasons for calculating in particular ways also become an explicit topic of conversation. In this latter case, conversations encompass both students' calculational processes and the interpretations that underlie those ways of calculating.

To illustrate this distinction, consider again the discussion of Inscriptions 2 and 3 of the AIDS data, in which the data sets were partitioned at the T-cell value of 550. A calculational explanation would describe the steps taken to arrive at the conclusion that the experimental treatment was more effective. Students giving such an explanation might clarify that they used the create-your-own-groups option on the minitool to partition the data sets at a particular value. They might then conclude by reporting the number of data points above and below this value in both data sets. Inscription 2 of the AIDS data appears to involve a calculational explanation in that the students who produced it provided a warrant (how the data were organized) but did not give a backing (why the data were organized in this way).

In contrast to this calculational explanation, a conceptual explanation would not merely describe the steps taken during the analysis but would include the reasons for carrying out these steps. Crucially, Vallory called for such an explanation when she asked, "Why is 550 so important?" [78]. In the ensuing discussion, the teachers and students developed a backing that related the numerical description of the data sets given in Inscription 2 to qualitative judgments about how the data were distributed. My colleagues and I consider this relation to be highly significant given the importance that we attribute to the overarching statistical idea of *distribution* (cf. P. Cobb, 1999; P. Cobb, McClain, & Gravemeijer, in press).

The value of Toulmin's scheme to us as instructional designers is precisely that it foregrounds this distinction between calculational and conceptual discourse. Our experience of conducting a series of design experiments over the past 14 years leads us to speculate that discussions that the teacher initiates and supports the development of backings can be particularly productive as settings for mathematical learning. Because a backing involves an articulation of the reasoning that lies behind the process of producing a result, students' engagement in such discussions supports their attempts to understand others' thinking. Had the discussion of the AIDS data remained calculational, the students would have had to figure out why others had used the minitool to organize data in a particular way entirely on their own. In contrast, their engagement in conceptual discourse provided them with resources for understanding how others had reasoned about the data.

This illustration is paradigmatic in that the resources for mathematical learning provided by engagement in conceptual discourse typically are not limited to what is said but include inscriptions that are pointed to and spoken about (cf. Thompson, Philipp, Thompson, & Boyd, 1994). In general, we have found that the development of conceptual discourse and development of ways of inscribing go hand in hand. This indicates the close relation between standards of argumentation and the third aspect of a classroom mathematical practice, the normative ways in which tools and inscriptions are used.

NORMATIVE WAYS OF REASONING WITH TOOLS AND SYMBOLS

As part of my analysis of the normative purpose for doing statistics, I suggested that it involved actually analyzing data rather than merely "doing something with numbers." This observation relates directly to the normative ways of reasoning with the two computer minitools that were established by the classroom community. In particular, it implies that the length of the horizontal bars in the first minitool signified measure of the life spans of batteries and that the position of the dots with respect to the axis in the second minitool signified patients' T-cell counts. As I noted, these normative ways of reasoning about data were in marked contrast to the performance assessments we conducted with another group of seventh graders in the same school.

Roth's (1996) reflections on the process of creating and inscribing data clarify that the differences between these contrasting types of reasoning are not trivial.

> To work on and think about natural phenomena (i.e., to mathematize phenomena), scientists engage in practices that ultimately produce inscriptions. Latour (1993) showed how a phenomenon and its mathematical order were produced as members of a scientific expedition first took soil samples that were subsequently placed in a two-dimensional array of boxes according to location and depth of the probe. This practice, one among many in which the members of the expedition engaged, made the phenomenon accessible to transformations into various drawings on paper (inscriptions) and to analyses wherever the drawings could be transmitted. (p. 493)

Similarly, the process of inscribing the batteries data in the first minitool involved a mathematization in which the students' initial realization that batteries have a finite life and that some last longer than others was transformed by treating the life span of a battery as a *quantity* that could be measured by, say, putting the battery in a flashlight and timing how long it gave light. I conjecture that the use of the minitool was central to this process, and, in a more extensive analysis of the use of tools and inscriptions, I would want to analyze other episodes to document the transformation of what might be termed the brute batteries situation into one that involved the analysis of quantitative data.

This mathematization appeared to have opened up new possibilities for action. For example, the discussion of Caesara's and Blake's analyses in the *Batteries* episode indicates that partitioning was a relatively routine way of organizing data sets. During the discussion of Caesara's analysis, the range tool was used to isolate the 10 longest lasting batteries [4–12]. Blake, for his part, directed the teacher to use the value tool to partition the data at 80 hr [24–30]. As the students reasoned with these tools, new meanings seemed to emerge (e.g., batteries of a particular brand are better because more of those batteries last the longest; batteries of a particular brand are consistent because all last over 80 hr). Thus, in my analytical approach, the properties that data sets came to have and the normative ways in which the computer minitools came to be used are considered to be intimately related.

The final point I want to make about the *Batteries* episode concerns the normative ways of reasoning about data that had been partitioned in some way. The argument that Caesara and Jamie developed jointly focused on the absolute number of batteries of each brand that were among the 10 longest-lasting batteries (i.e., 7 and 3). Similarly, Blake focused on the absolute number of batteries of each brand that lasted more than 80 hr. These comparisons of partitioned data sets in terms of part–whole relations were treated as legitimate, indicating that this type of reasoning was normative. Such reasoning can be contrasted with an alternative type of argument that focuses on the proportion of the batteries of each brand that are among the 10 longest-lasting batteries or that last more than 80 hr. In terms of a distinction that mathematics educators viewed as critical, these alternative arguments involve multiplicative rather than additive comparisons of data sets (cf. Harel & Confrey, 1994; Thompson & Thompson, 1996). In statistical terms, this distinction corresponds to the contrast between relative frequency and absolute frequency. I should emphasize that additive reasoning is entirely adequate when data sets are equal, as was the case in the *Batteries* episode. When I conducted the prior longitudinal analysis of all 34 classroom sessions, I therefore found it essential to examine the discussions of instructional activities in which the students used the first minitool to analyze unequal data sets.[5] The *AIDS* episode in which the students had used the second minitool is also of interest for this reason.

A significant feature of the discussion during the *AIDS* episode is the almost complete absence of references to the students' prior use of the second minitool during the small-group work that preceded the whole class discussion. In the case of Inscriptions 2 and 3, it seemed to be taken for granted that the students who produced the inscriptions had used the create-your-own-groups option to partition the data sets. Furthermore, although differences in the students' interpretations of Inscription 5 became apparent during the discussion, it seemed to be taken as self-ev-

[5]The students first used the first minitool to compare unequal data sets approximately 1 week after the *Batteries* episode. The possibility of reasoning about the data sets in proportional terms did not arise despite the teacher's efforts to support the emergence of such reasoning. Instead, the discussion focused on the issue of which data points should be omitted to equalize the data sets. Students did not begin to reason about data multiplicatively until two thirds of the way through the design experiment, shortly after the second minitool had been introduced.

ident that the data had been analyzed by using the four-equal-groups option on the minitool. These consensual understandings of the use of the minitool appeared to lie behind and to ground the discussion throughout the *AIDS* episode. In this regard, it is worth noting that a significant shift had occurred in the way that whole class discussions were organized since the *Batteries* episode. Previously, the students had used the projection of a minitool to describe how they had conducted their analyses. In contrast, they now assessed written reports of others' analyses without access to a minitool.

In considering the ways of reasoning about the data that had become normative by the time of the *AIDS* episode, I work through the subepisodes chronologically commencing with the discussion of Inscription 1. During this subepisode, the teacher and students explicitly negotiated the meaning of a term Jamie introduced, that of the majority. Sheena was the first student to offer a clarification:

> Sheena: Like, when she talks about like when she says, like when say where the majority of the numbers were, where the point is, like, you see where it goes up?
>
> Kay: I do see where it goes up.
>
> Sheena: Yeah, like right in there, that's where the majority of it is. [62–64]

I initially conjectured that Sheena's reference to "where the point is" might indicate that the majority was, for her, analogous to the mode and indicated the specific point in the range where the data were stacked the highest. However, her gestures as she spoke mitigated against this interpretation. In addition, the teacher gestured to the region beneath the upturned "V" on one of the graphs and Sheena replied, "Yeah, like right *in* there" [65] rather than, "Yeah, like right there." It therefore seems that the majority as it was constituted in this brief exchange was located in the interval spanned by the upturned "V."

In both this exchange and the remainder of the subepisode, the discussion of the term *majority* appeared to be grounded in activity with the second minitool. In particular, the teacher and students seemed to take it as self-evident that Inscription 1 showed the intervals where the data were piled up on the line plot in the second minitool. This in turn implies that data inscribed as line plots had been endowed with the property of shape. The emergence of this property seems to be intimately related to the types of inscriptions with which the students reasoned as they used the second minitool. It is doubtful, for example, that data sets would have come to have this property had the students continued to use the first computer minitool.

The exchange between Sheena and the teacher focused on the location of the majority but did not touch on the nature of the putative entity that they established was located in a particular interval. Was it a particular figural shape in the data, the piling up of data, or something else? Vallory's subsequent clarification, with which both Sheena and Jamie indicated agreement, is helpful in this regard:

> However many people were tested, that's where most of those people fitted in, in between that range. [72]

Vallory's use of the word *most* appeared to be relational. The absolute number of people tested was not central to her explanation. Instead, it appeared to be a proportional notion that signified a qualitative relative frequency.

In determining the status that this meaning of the majority had in the classroom, I found it useful to focus on Derrick, the student who first asked Jamie what she meant by "the majority of the numbers" [57]. There is no indication from the two further contributions that he made during the subepisode that he interpreted the majority as a relative frequency [58–70]. However, putting aside speculations about his personal state of mind, the crucial point to note is that he appeared to be positioned in classroom discourse as failing to understand rather than as making an alternative interpretation. This suggests that the meaning of the majority as qualitative relative frequency might have been normative.

The analysis of the remaining three *AIDS* subepisodes indicated that this conjecture was viable. The term *majority* was used in the next inscription that the teachers and students discussed:

> The new drug was better than the old. The majority of the new ones are behind 550, and the majority of the new drug was in front of 550. [75]

This use of the term *majority* was not questioned in the ensuing exchange. Instead, the discussion focused on why the students who produced the inscription had partitioned the two data sets at the T-cell value of 550. This constitutes weak confirmation that relative frequency had been established as the normative meaning of majority.

Later in the same subepisode, one of the teachers introduced Inscription 3 (Figure 10 in the Appendix), in which the absolute numbers of data points above and below the T-cell value of 525 in each data set were reported. The teacher noted similarities between this and the previous inscription in which the data sets had been partitioned at the T-cell value of 550 and then asked for questions or comments about the two inscriptions. Marissa, the first student to respond, was critical of the second of these two inscriptions.

> Marissa: I would think the second one [Inscription 3] would be more confusing because it has, since the old program has more numbers than the new program.
>
> Kay: Oh, so it looks like that there's more. They [the standard treatment protocol] had 56 that were above 525, and they only had 37?
>
> Marissa: So, it's like, I guess what I'm trying to say is it's harder to compare them. [92–94]

Marissa's objection appeared to be that it was difficult to draw a conclusion about the effectiveness of the two treatments from Inscription 3 because it stated the absolute number of data points above and below the T-cell value of 525. She viewed this as problematic because "the old program has more numbers than the new program." In contrast, the previous inscription in which the term *majority* had been used was cast in terms of relative frequency. Because Melissa's objection was constituted as legitimate in the course of the exchange, it provides reasonably strong evidence that relative frequency had been established as the normative meaning of majority.

At the beginning of the next subepisode, one of the teachers used a diagram developed from Inscription 3 in which the data sets were partitioned at T-cell counts of 525 to pose the following question:

> Couldn't you just argue, hey, this shows really convincingly that the old treatment was better, right? Because there were 56 of them, 56 scores above 525, 56 people with T-cell counts above 525 [in the standard treatment protocol], and here [in the experimental treatment protocol] there's only 37 above, so the old one just had to be better, there's more people. I mean, there's 19 more people in there, so that's the better one, surely. [99]

Several students challenged this argument during the ensuing exchange. Blake, for example, developed a rationale similar to that which Marissa had articulated when he observed that "there's more people with the old program than there is with the new program" [101]. In doing so, he seemed to indicate that the data sets should be compared in terms of relative rather than absolute frequencies. Jason, the next student to object to the teacher's argument, made such a comparison:

> Then you see that there 37 is more than half [of the experimental treatment data] over 525 and 56 is not more than half of 130… more of them on the bottom than on the top. [103]

Jason's argument appeared to be that whereas more than half of the experimental treatment data were above the T-cell value of 525, this was the case for less than half of the standard treatment data. Although it is not clear whether it was intentional, his use of half the data as a benchmark is consistent with the meaning of majority that I conjectured was normative. Significantly, both Blake's and Jason's contributions as well as Will's subsequent revoicing of Jason's argument [116] were constituted as legitimate ways of reasoning. This further substantiates the conjecture that relative frequency had been established as the normative meaning of majority. Normative ways of reasoning about data that had been partitioned by using the second minitool therefore appeared to be multiplicative in nature. This represents a significant advance when compared with the *Batteries* episode in which the normative ways of reasoning about partitioned data were additive rather than multiplicative and involved comparing data sets in terms of absolute rather than relative frequencies.

At first glance, the final *AIDS* subepisode, where the students assessed Inscription 5 in which the data sets were partitioned into four equal groups, appears to refute these inferences. Blake argued that this inscription was adequate because he could see from the differences in the groups that the experimental treatment was more effective than the standard treatment [127–133]. The teacher then asked him to justify this conclusion.

Kay: And what are you basing that comment on?
Blake: Because the three lines for the equal groups were all, what is that—525?
Kay: Yeah.
Blake: Above 525 compared to only one of them was over on the top [in the standard treatment]. [134–137]

Blake spoke of three lines on the experimental treatment graph and one line on the standard treatment graph being above the T-cell value of 525 when, in fact, four lines and two lines, respectively, were above this value. This suggests that the lines to which he referred were lower bounds of three of the groups in the experimental treatment data and one of the groups in the standard treatment data. His subsequent objection to finding how many data points were in each group [148–150] and his comment that the inscription showed where the data stood [152] indicate that each group was, for him, a proportional part of a data set. In other words, the graphs were, for him, structured multiplicatively in terms of relative frequencies.

A number of students indicated that they were not convinced by Blake's argument that the inscription was adequate. For example, Marcus suggested the following modification:

Marcus: …Anyway, so it might have helped if they put the numbers in the groups so you have a better idea of what you're seeing.
Kay: Put the numbers in here?
Marcus: In say the fourth group [of the experimental treatment graph] would be a 12 in there, and just write the number 12 in there so you know… [139–141]

Marcus's suggestion that the absolute frequencies be added to the inscriptions runs counter to Marissa's argument in an earlier subepisode, in which she indicated that absolute frequencies are confusing when the data sets are unequal. Crucially, Marcus was not positioned as failing to understand. Instead, several students indicated that they agreed with his proposal, and one of the teachers attempted to capitalize on it by asking, "…is there some way for us to know how many are in each of those groups?" [142]. I take this as a relatively strong indication that reasoning multiplicatively about data partitioned into four equal groups was not normative at this point in the design experiment.

This conclusion seems to be at odds with the conjecture I made while analyzing the prior two subepisodes, namely that normative ways of reasoning involved comparing data sets in terms of relative frequency. In accounting for this apparent conflict, I focus on the differences between the first three and the final inscriptions. The students who produced the first three inscriptions did so to report perceptually based patterns that they had identified in how the individual data points were clumped. For example, the students who produced Inscription 1 isolated the intervals where the data were stacked up. The students who produced the Inscriptions 2 and 3 also focused on the perceptual shape of the data sets as inscribed in the second minitool. They appear to have selected values at which to partition the data sets so that experimental treatment data were stacked up above the partition and the standard treatment data were stacked up below it. In contrast, the partitions in the final four-equal-groups inscription (Figure 12 in the Appendix) do not correspond to clumps or gaps in how the individual data points were distributed. Instead, the students who produced this inscription used the four-equal-groups option on the minitool to create graphs from which they could read how the data were distributed. In other words, the graphs in the minitool and the inscription were, for these students, texts from which they could read global trends in the data rather than means of documenting perceptually based patterns in how the individual data points were clumped. The discussion during the final *AIDS* subepisode indicates that reasoning of this type had not been established as normative at this point in the design experiment.

In stepping back to reflect on the claims I have made about normative ways of reasoning with tools and inscriptions, the primary distinction I have drawn is between additive and multiplicative reasoning. In the *Batteries* episode, normative reasoning involved partitioning data sets inscribed in the first minitool additively into absolute frequencies. In contrast, normative reasoning in the *AIDS* episode involved partitioning data sets inscribed as line plots multiplicatively into relative frequencies. This suggests that the classroom community's mathematical learning during the sessions between the *Batteries* and *AIDS* episodes may have been relatively significant. However, to substantiate this conjecture, it would be important to examine classroom sessions in which the students used the first minitool to analyze unequal data sets. I did in fact analyze the whole class discussions conducted during these sessions when I previously conducted a longitudinal analysis of the entire data corpus (P. Cobb, 1999). This analysis indicated that the claim about the significance of the classroom community's mathematical learning was viable. I therefore concluded that a new mathematical practice had emerged during the intervening sessions as a reorganization of the practice in which the students participated during the *Batteries* episode.

EMERGENT MODELS AND CHAINS OF SIGNIFICATION

The analysis I have presented illustrates the approach I take when testing and revising conjectures about normative ways of reasoning with tools and inscriptions as I work through a corpus of design experiment data episode by episode. At first glance, this analysis might appear to be at odds with the basic theoretical orientation that I outlined in the first part of this article. There, I indicated that my general position is one in which the meaning and ways of symbolizing are seen to co-evolve and to be mutually constitutive. To avoid a possible misunderstanding, I therefore need to stress that the interrelations between meaning and ways of symbolizing that I am interested in as an instructional designer are not located at the microlevel of moment-by-moment shifts in meaning. Instead, they concern more enduring developments in the collective mathematical learning of the classroom community. I hinted at a more global relation of this type while analyzing the *Batteries* episode when I claimed that the characteristics and properties that data sets came to have were related intimately to the normative ways in which the first computer minitool was used. A similar claim can be made about the teacher's and students' use of the second minitool (P. Cobb, 1999). Claims of this type are important from the viewpoint of instructional design in that they can immediately feed back to inform the revision of tools that were specifically designed to support students' learning. Given its pragmatic significance, I clarify this perspective on the interplay between meaning and ways of symbolizing by first introducing the notion of emergent models that is central to the design theory of Realistic Mathematics Education, to which I and my colleagues subscribe (cf. Gravemeijer, 1997). Against this background, I then discuss a possible chain of signification that might have emerged as the students used the two computer minitools to analyze data.

The overall intent of an approach to instructional design compatible with the theory of Realistic Mathematics Education is to support a so-called reinvention process in which students initially mathematize their informal reasoning in problem settings that are experimentally real to them. As the analysis I have presented illustrates, mathematizing can be viewed as a collective as well as an individual process in which whole class discussions play a crucial role. Within this general approach, the initial phase of an envisioned learning trajectory for the classroom community typically involves the emergence of *models of* informal reasoning that are tied to particular task settings and involve situation-specific imagery. In a subsequent phase of an envisioned trajectory, this initial modeling activity or its results are themselves mathematized to produce a *model for* mathematical reasoning. A proposed learning trajectory of this type therefore involves the conjecture that models of informal reasoning will gradually evolve through use into models for more general mathematical reasoning that are independent of situation-specific imagery (cf. Gravemeijer, 1994).

As I illustrate, this distinction between models-of and models-for is also useful in delineating the actual learning trajectory of a classroom community during a design experiment. In semiotic terms, the actual learning trajectory takes the form of a chain of signification that emerges as the normative ways of reasoning with tools and inscriptions evolve. This notion of a chain of signification was first proposed by Lacan (1977). However, I draw primarily on Walkerdine's (1988) illustration of its relevance to analyses of mathematical learning. A possible chain of signification that might have been constituted during the design experiment is shown in Figure 3. The links shown in Figure 3 were constituted as the classroom community established new signs that subsumed previous signs. However, given the brevity of the two sample episodes, it is important to stress that the proposed chain is necessarily highly speculative. My purpose in the following paragraphs is therefore not to argue in support of the conjectures inherent in this chain. Instead, my goal is to illustrate how my colleagues and I analyze mathematical learning in semiotic terms, in the process delineating the coevolution of meaning and ways of symbolizing.

To clarify the first conjectured link in the chain of signification, I revisit a mathematization that occurred as the teacher and students talked through the data creation process. When I analyzed the *Batteries* episode, I argued that an initial mathematization occurred as the data were inscribed as horizontal bars in the first minitool such that the life span of a battery became a quantity whose magnitude could be measured. I suggested that it was as a consequence of this mathematization that data displays in this minitool became texts about the situation from which the data were generated. This in turn opened up new possibilities for action that included using the value tool option and the range tool option to partition data sets additively. Judging from the sample episode, the normative ways of reasoning that emerged as the teacher and students used these tools involved treating the resulting data displays as models-of the situation from which the data were generated.

The second conjectured link in the chain of signification proposed in Figure 3 appeared to be constituted at some point between the *Batteries* and *AIDS* episodes.

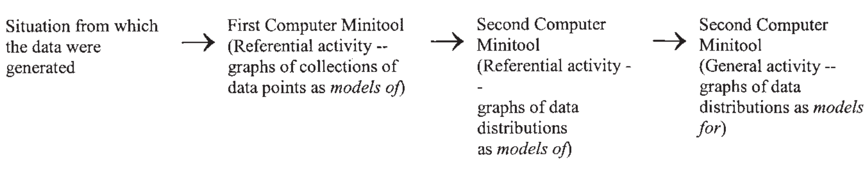

FIGURE 3 Conjectured chain of signification constituted during the design experiment.

I argued that it had become an established norm by the *AIDS* episode that the act of partitioning data sets structured them multiplicatively into relative frequencies. This was significant from the point of view of our instructional agenda, in that a concern for how data are distributed in fact assumes that those data are structured multiplicatively (P. Cobb, 1999). It is for this reason that I differentiate in Figure 3 between *data distributions* and data sets that are structured additively into *collections of data points*.

During the analysis of the *AIDS* episode, I also clarified that this norm concerning the multiplicative structuring of data sets applied only to partitionings that were based on perceptual patterns in the data. This was the case with the first three inscriptions that the teachers and students discussed, but not with the final inscription in which the data were organized into four equal groups. Given this apparent dependency on the perceptual features of particular data sets, I infer that the data displays the students produced when they used the second minitool were constituted in public classroom discourse as models-of the situations from which the data were generated (see Figure 3). The shift indicated by the second conjectured link in the chain of signification thus did not appear to involve a model-of/model-for transition but instead concerned the nature of the models-of that had become normative.

To account for this inferred shift, I follow Walkerdine (1988) in observing that although new signifiers sometimes are introduced as substitutes for preceding terms with the intention of preserving meaning, the meaning in fact evolves as previously established sign combinations slide under the new signifier. In the case at hand, the constitution of the first proposed link in the chain involved the establishment of the sign combination shown in Figure 4. Reasoning with this sign involved interpreting the situations from which data were generated additively in terms of collections of data points. The conjectures I have made about the second link in the chain imply that the line plot inscription in the second minitool did not merely displace the horizontal bar graphs as the signifier in this sign. Instead, a new sign was established, as shown in Figure 5. The signified of this sign was a situation comprised of collections of data points rather than a so-called brute situation (i.e., the situation as it might have been interpreted before the design experiment began). According to this conjecture, meaning evolved as this new sign combination was established in that it became normative to reason about the situations from which data were generated multiplicatively in terms of distributions rather than additively in terms of collections of data points.

The distinction between model-of and model-for proves useful when analyzing the discussion of the Inscription 5 in the final *AIDS* subepisode in semiotic terms. As I have noted, a model of informal reasoning is tied to particular task settings and involves situation-specific imagery, whereas a model for more general reasoning is independent of situation-specific imagery. I have already suggested that Blake's multiplicative interpretation of the four-equal-groups inscription did not involve a reliance on perceptual patterns that were specific to those particular data

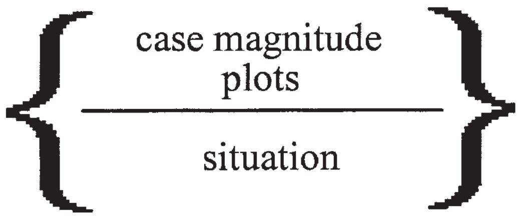

FIGURE 4 The first sign combination established as the conjectured chain of signification emerged.

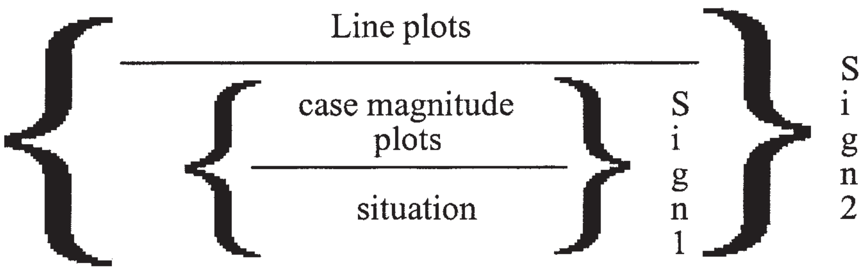

FIGURE 5 The second sign combination established as the conjectured chain of signification emerged.

sets. As part of this argument, I noted that the partitions made by the students who produced the inscription did not correspond to clumps or gaps in the data. Instead, these students used the second minitool to create data displays from which they could infer patterns in how the data were distributed. For these students, the data displays were therefore models-for reasoning about data displays independently of specific perceptual patterns. The distinction between data displays in the second minitool as models-of and as models-for clarifies why normative interpretations of the first three inscriptions, but not Inscription 5, involved structuring the data multiplicatively into relative frequencies.

As I conclude this discussion of normative ways of reasoning with tools and inscriptions, it is worth noting that several of the students' contributions to both the *Batteries* and the *AIDS* episodes reflected the relatively complex relations of signification that I have discussed. Consider, for example, Vallory's clarification of the term *majority* during the discussion of Inscription 1 of the AIDS data:

> However many people were tested, that's where most of those people fitted in, in between that range. [72]

Vallory's talk of people fitting into a range or particular interval on a graph exemplifies a phenomenon that Nemirovsky and Monk (2000) called fusion. In Vallory's clarification, references to aspects of the inscription (e.g., "in between that range") were fused with references to the signified situation (e.g., "however many people were tested"). In accordance with Walkerdine's (1988) notion of a sign combination sliding under a succeeding signifier, the people to whom Vallory referred were not merely AIDS patients who had participated in a study. They were also people who had T-cell counts that were measurable quantities and that could be inscribed as dots on a line plot. Vallory's way of talking, in which she reasoned with the inscription as a model of the situation from which the data were generated, therefore can be seen to reflect her participation in the classroom community's constitution of the sign combination shown in Figure 5.

In the introductory section of this article, I contrasted the approach I would take when analyzing relations of signification with an alternative in which the process of developing meaning for symbols is characterized as one of associating them with separate, self-contained referents. It should be clear that the approach I have followed questions whether it is productive for the purpose of instructional design to take mappings between symbols and referents at face value. Instead my focus has been on the activity of reasoning with tools and inscriptions as an integral aspect of the process of mathematizing. As a consequence, I have not treated so-called referents as immutable or fixed but instead have suggested that they evolve as new ways of using tools and inscriptions emerge. It is to this aspect of tool and inscription use that both Walkerdine and Nemirovsky and Monk orient us. Instructionally, students' use of tools and inscriptions is then framed not as a separate instructional goal but as a primary means of supporting the emergence of significant mathematical ideas. In the process, the use of tools and inscriptions becomes an aspect of the practices that students participate in rather than disembodied skills to be mastered.

CONCLUSION

In this article, I have adopted a relatively global perspective consistent with my interests in instructional design. As I clarified at the outset, the primary unit of analysis that my colleagues and I use consists of a classroom mathematical practice together with students' diverse ways of contributing to its constitution. The analysis of the *Batteries* and *AIDS* episodes illustrates that this methodological approach acknowledges the diversity of students' reasoning while simultaneously treating that diversity as socially situated. It therefore should be clear that my interest in bringing qualitative differences in students' reasoning to the fore did not signal a return to mainstream psychology. Instead, I treated students' mathematical interpretations as embodied in their acts of participating in the classroom mathematical practices that constitute the immediate social situations of their development (cf. Minick, 1987). Furthermore, I viewed the teacher and students as collectively enacting a normative world of signification rather than as representing a world (Varela, Thompson, & Rosch, 1991). This approach is consistent with Lave's (1988) call "to incorporate the active character of experience into the unit of analysis" (p. 180). Furthermore, in line with Wenger's (1998) discussion of communities of practice, this approach captures social structures that are within the scope of students' engagement with the world. In the case of mathematics instruction, these familiar structures include the relatively inconspicuous, recurrent, and taken-for-granted features of classroom life encompassed by the three aspects of a classroom mathematical practice. As a consequence, this analytic approach foregrounds students' experienced, lived-in worlds while delineating the social situations in which their mathematical learning occurs.

My primary concern when analyzing the *Batteries* and *AIDS* episodes was to illustrate the process of generating and testing conjectures about the classroom community's mathematical learning during the 12 weeks of the design experiment. In doing so, I highlighted the teacher's and students' use of designed artifacts by framing normative ways of reasoning with tools and inscriptions as a central aspect of a mathematical practice. Against this background, I subsequently developed conjectures about the interplay between changes in mathematical meaning and changes in ways of reasoning with tools and inscriptions by delineating a chain of signification. As I noted, this provided a speculative account of the classroom community's mathematical learning in semiotic terms.

In addition to illustrating the notion of a classroom mathematical practice as a unit of analysis, I also have attempted to clarify why my colleagues and I find this analytical approach useful. As the discussion of the two sample episodes indicates, this approach enables us to document what is traditionally called mathematical content as an aspect of the classroom microculture. This focus is clearly important given our concerns and interests as mathematics educators. However, in an analysis of this type, content is treated as an emergent phenomenon that is constituted and continually regenerated by the teacher and students as they use tools and other resources. For example, the central statistical idea of relative frequency emerged as a normative way of reasoning with tools and inscriptions during the classroom sessions between the *Batteries* and *AIDS* episodes. The fact that this approach takes mathematics seriously is obviously important given our purposes as instructional designers. In addition, this approach enables us to sidestep the all-to-familiar gulf between theoretical analyses and instructional practice. In the case of the design experiment from which the two episodes are taken, for example, the accounts we have developed of the classroom community's mathematical learning are tied to the teacher's and students' use of particular tools and inscriptions (Cobb, 1999; McClain et al., 2000). As a consequence, we were in a position to develop testable conjectures about how the tools used in the classroom might be improved to support the emergence of significant statistical ideas. In more general terms, the analytic approach I have illustrated in this article supports educational innovation as a process of continual iterative improvement.

ACKNOWLEDGMENTS

The analysis reported in this article was supported by the National Science Foundation under grant REC9814898 and by the Office of Educational Research and Improvement (OERI) under grant R305A60007. The opinions expressed do not necessarily reflect the views of either the foundation or OERI.

I am grateful to Anna Sfard, Rogers Hall, Ann Ryu, and James Gee for their thoughtful critiques of previous drafts of this article.

REFERENCES

Bloor, D. (1976). *Knowledge and social imagery*. London: Routledge & Kegan Paul.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly, 104,* 801–823.

Cobb, P. (1998). Theorizing about mathematical conversations and learning from practice. *For the Learning of Mathematics, 18*(1), 46–48.

Cobb, P. (1999). Individual and collective mathematical learning: The case of statistical data analysis. *Mathematical Thinking and Learning, 1,* 5–44.

Cobb, P. (2001). Supporting the improvement of learning and teaching in social and institutional context. In D. K. S. Carver (Ed.), *Cognition and instruction: 25 years of progress* (pp. 455–478). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Cobb, P., McClain, K., & Gravemeijer, K. (in press). Learning about statistical covariation. *Cognition and Instruction*.

Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences, 10,* 113–164.

Cobb, P., Wood, T., Yackel, E., & McNeal, G. (1992). Characteristics of classroom mathematics traditions: An interactional analysis. *American Educational Research Journal, 29,* 573–602.

Davis, P. J., & Hersh, R. (1981). *The mathematical experience.* Boston: Houghton Mifflin.

Dörfler, W. (1993). Computer use and views of the mind. In C. Keitel & K. Ruthven (Eds.), *Learning from computers: Mathematics education and technology* (pp. 159–186). Berlin: Springer-Verlag.

Gee, J. P. (1997). Thinking, learning, and reading: The situated sociocultural mind. In D. Kirshner & J. A. Whitson (Eds.), *Situated cognition: Social, semiotic, and psychological perspectives* (pp. 235–260). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine.

Gravemeijer, K. E. P. (1994). *Developing realistic mathematics education*. Utrecht, The Netherlands: CD-ß.

Gravemeijer, K. (1997). Solving word problems: A case of modelling? *Learning and Instruction, 7,* 389–397.

Gravemeijer, K. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning, 1,* 155–177.

Greeno, J. G. (1991). Number sense as situated knowing in a conceptual domain. *Journal for Research in Mathematics Education, 22*, 170–218.

Harel, G., & Confrey, J. (1994). *The development of multiplicative reasoning in the learning of mathematics*. Albany, NY: State University of New York Press.

Kaput, J. J. (1994). The representational roles of technology in connecting mathematics with authentic experience. In R. Biehle, R. W. Scholz, R. Strasser, & B. Winkelmann (Eds.), *Diadactics of mathematics as a scientific discipline* (pp. 379–397). Dordrecht, The Netherlands: Kluwer.

Kitcher, P. (1984). *The nature of mathematical knowledge.* New York: Oxford University Press.

Lacan, J. (1977). *Ecrits: A selection*. London: Tavistock.

Latour, B. (1993). *La clef de Berlin et autre leçons d'un amateur de Sciences* [The key to Berlin and other lessons of a science lover]. Paris: Editions La Découverte.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life*. New York: Cambridge University Press.

McClain, K., Cobb, P., & Gravemeijer, K. (2000). Supporting students' ways of reasoning about data. In M. Burke (Ed.), *Learning mathematics for a new century* (2001 Yearbook of the National Council of Teachers of Mathematics, pp. 174–187). Reston, VA: National Council of Teachers of Mathematics.

McGatha, M. (2000). *Instructional design in the context of classroom-based research: Documenting the learning of a research team as it engaged in a mathematics design experiment.* Unpublished doctoral dissertation, Vanderbilt University, Nashville, TN.

McGatha, M., Cobb, P., & McClain, K. (1999, April). *An analysis of students' statistical understandings*. Paper presented at the American Educational Research Association, Montreal, Canada.

Meira, L. (1995). The microevolution of mathematical representations in children's activity. *Cognition and Instruction, 13,* 269–313.

Meira, L. (1998). Making sense of instructional devices: The emergence of transparency in mathematical activity. *Journal for Research in Mathematics Education, 29*, 121–142.

Minick, N. (1987). The development of Vygotsky's thought: An introduction. In R. W. Rieber & A. S. Carton (Eds.), *The collected works of Vygotsky, L.S.* (Vol. 1, pp. 17–38). New York: Plenum.

Much, N. C., & Shweder, R. A. (1978). Speaking of rules: The analysis of culture in breach. *New Directions for Child Development, 2,* 19–39.

Nemirovsky, R. C., & Monk, S. (2000). If you look at it the other way… In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing, communicating. and mathematizing: Perspectives on discourse, tools, and instructional design* (pp. 177–221). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton, NJ: Princeton University Press.

Roth, W. M. (1996). Where is the context in contextual word problems? Mathematical practices and products in grade 8, students' answers to story problems. *Cognition and Instruction, 14,* 487–527.

Schoenfeld, A. H. (1998). Making mathematics and making pasta: From cookbook procedures to really cooking. In J. G. Greeno & S. V. Goldman (Eds.), *Thinking practices in mathematics and science learning* (pp. 299–320). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schutz, A. (1962). *The problems of social reality*. Hague, The Netherlands: Martinus Nijhoff.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22,* 1–36.

Sfard, A. (2000). On reform movement and the limits of mathematical discourse. *Mathematical Thinking and Learning, 2,* 157–189.

Thompson, A. G., Philipp, R. A., Thompson, P. W., & Boyd, B. (1994). Calculational and conceptual orientations in teaching mathematics. In D. B. Aichele (Ed.), *Professional development for teachers of mathematics: 1994 yearbook of the National Council of Teachers of Mathematics* (pp. 79–92). Reston, VA: National Council of Teachers of Mathematics.

Thompson, A. G., & Thompson, P. W. (1996). Talking about rates conceptually. Part II: Mathematical knowledge for teaching. *Journal for Research in Mathematics Education, 27*, 2–24.

Toulmin, S. (1969). *The uses of argument*. Cambridge, England: Cambridge University Press.
van Oers, B. (1996). Learning mathematics as meaningful activity. In P. Nesher, L. Steffe, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 91–114). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
van Oers, B. (2000). The appropriation of mathematical symbols: A psychosemiotic approach to mathematical learning. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design* (pp. 133–176). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
Walkerdine, V. (1988). *The mastery of reason*. London: Routledge.
Wenger, E. (1998). *Communities of practice*. New York: Cambridge University Press.
Wittgenstein, L. (1953). *Philosophical investigations*. Oxford, England: Blackwell.