Journal for Research in Mathematics Education
2005, Vol. 36, No. 5, 412–446

Characterizing a Classroom Practice That Promotes Algebraic Reasoning

Maria L. Blanton and James J. Kaput
University of Massachusetts Dartmouth

We present here results of a case study examining the classroom practice of one third-grade teacher as she participated in a long-term professional development project led by the authors. Our goal was to explore in what ways and to what extent the teacher was able to build a classroom that supported the development of students' algebraic reasoning skills. We analyzed 1 year of her classroom instruction to determine the robustness with which she integrated algebraic reasoning into the regular course of daily instruction and its subsequent impact on students' ability to reason algebraically. We took the *diversity* of types of algebraic reasoning, their *frequency* and *form* of integration, and *techniques of instructional practice* that supported students' algebraic reasoning as a measure of the robustness of her capacity to build algebraic reasoning. Results indicate that the teacher was able to integrate algebraic reasoning into instruction in planned and spontaneous ways that led to positive shifts in students' algebraic reasoning skills.

Key words: Algebra; Elementary, K–8; Professional development; Teacher education; Teachers (characteristics of); Teaching practice

Historically, the mathematical experience (and hence classroom practice) of most elementary teachers has focused on arithmetic and computational fluency. However, it is now widely accepted that preparing elementary students for the increasingly complex mathematics of the new century will require a different type of school experience, specifically, one that cultivates habits of mind that attend to the deeper underlying structure of mathematics (Kaput, 1999; Romberg & Kaput, 1999). Rethinking the type of curriculum and instruction in elementary grades that could effect this has led to a growing recognition that algebraic reasoning can simul-

The research reported here was supported in part by a grant from the U.S. Department of Education, Office of Educational Research and Improvement, to the National Center for Improving Student Learning and Achievement in Mathematics and Science (R305A60007-98). The opinions expressed herein do not necessarily reflect the position, policy, or endorsement of the supporting agencies.

Jim Kaput was killed in a tragic accident in late July. Although his reputation was international in scope, Jim worked tirelessly on local fronts to bring educational innovations to disadvantaged school districts. The study reported here reflects a small piece of that work. For researchers of early algebra, he helped us to imagine a more powerful mathematics for young children. We will deeply miss his vision, his insights, and his passion for thinking beyond boundaries.—*Maria Blanton* taneously emerge from and enhance elementary school mathematics (National Council of Teachers of Mathematics [NCTM], 2000). Indeed, traditional instructional and curricular elementary school practices centered on teaching arithmetic procedures, followed by a largely procedural approach to algebra from middle grades onward, have been unsuccessful in terms of student achievement (U.S. Department of Education & National Center for Education Statistics, 1998a, 1998b, 1998c). The integration of algebraic reasoning into primary grades offers an alternative that builds the conceptual development of deeper and more complex mathematics into students' experiences from the very beginning.

A MULTIDIMENSIONAL VIEW OF ALGEBRAIC REASONING

We take algebraic reasoning to be a process in which students generalize mathematical ideas from a set of particular instances, establish those generalizations through the discourse of argumentation, and express them in increasingly formal and age-appropriate ways (Kaput, 1995; 1999). For example, students are engaged in algebraic reasoning when they first describe the total number of handshakes for a group of specific size, where each person in the group shakes everyone's hand once, and then proceed to develop and express a generalization that describes the total number of handshakes for an arbitrary sized group. Depending on the experience level of the student, the generalization might be expressed in words or in symbols and could be based on the student's observation of a recursive pattern describing how to get the next total of handshakes from the current one, or a functional relationship between the number of people in the group and the total amount of handshakes (e.g., "the total number of handshakes is the sum of the numbers from '1' up to one less than the number of people in the group," or "the total number of handshakes for a group with n people is n times n - 1 divided by 2"). Similarly, students are using algebraic reasoning when, based on systematic analyses of specific cases, they generalize about the parity of the sum of arbitrary even and odd numbers, or when they recognize and express properties of a number system, such as the commutativity of addition of whole numbers.

As the foregoing examples illustrate, and as has been described elsewhere (Kaput 1998; 1999), algebraic reasoning can take various forms, including (a) the use of arithmetic as a domain for expressing and formalizing generalizations (generalized arithmetic); (b) generalizing numerical patterns to describe functional relationships (functional thinking); (c) modeling as a domain for expressing and formalizing generalizations; and (d) generalizing about mathematical systems abstracted from computations and relations. By (a), we mean reasoning about operations and properties associated with numbers, such as generalizing about the commutative property of multiplication or properties of zero, or understanding equality as a relation between quantities. Generalizing numerical patterns involves exploring and expressing regularities in numbers, such as describing growth patterns or generalizations about sums of consecutive numbers. In a similar way, modeling as a form of algebraic reasoning also involves generalizing regularities but from mathematized situations or phenomena where the regularity itself is secondary to the larger modeling task. Finally, although less common to the elementary school curriculum, generalizing with abstract objects and systems involves operations on classes of objects and is more traditionally described as "abstract algebra." Of these four forms, generalized arithmetic and functional thinking are the more common forms of algebraic reasoning in the elementary grades.

There is a growing body of research that seeks to understand how teachers and students reason algebraically and to identify the kinds of classroom practice that foster algebraic reasoning. In particular, numerous scholars have documented the capacity of students from a diversity of socioeconomic and educational backgrounds to engage in algebraic reasoning, focusing largely on generalized arithmetic and functional thinking, in ways that dispel developmental constraints previously imposed on young learners (e.g., Carpenter & Franke, 2001; Carpenter, Franke, & Levi, 2003; Carraher & Earnest, 2003; Carraher, Brizuela, & Schliemann, 2000; Falkner, Levi, & Carpenter, 1999; Kieran, 1992; Mason, 1996; Schifter, 1999). Within these forms of algebraic reasoning, our own work (see, e.g., Blanton & Kaput, 2003, 2004, 2005) focuses on functional thinking through a process in which arithmetic tasks are transformed into opportunities for generalizing mathematical patterns and relationships by varying a single task parameter (e.g., the number of people in a group, or the number of t-shirts purchased).

Although the research cited here generally takes an approach to algebraic reasoning in which students reason from particular quantities to build mathematical generalizations, there is also new research emerging from the Davydoviian (1975a, 1975b) tradition that uses the exploration of mathematical generality itself (rather than the particulars of number) as a springboard for building students' understanding of mathematical structures (Dougherty, 2003). In this approach to algebraic reasoning, students begin by comparing abstract quantities of physical measures (e.g., length, area, volume), absent of quantification, in order to develop general relationships about these measures (e.g., the transitive property of equality).

IDENTIFYING THE ALGEBRAIC NATURE OF A THIRD-GRADE TEACHER'S PRACTICE

The increasing emphasis on algebraic reasoning places elementary teachers in the critical path to mathematics reform and, in fact, the degree to which they are capable of developing children's algebraic reasoning may determine the depth of that reform (Kaput, 1999; Schifter, 1999). However, most elementary teachers have little experience with the rich and connected aspects of algebraic reasoning that need to become the norm in schools and, instead, are often products of the type of school mathematics instruction that we need to replace. Thus, if we are to build classrooms that promote algebraic reasoning, we must provide the appropriate forms of professional support that will effect change in instructional and curricular practices. In part, this requires us to understand what it means for a teacher's practice to support a culture of algebraic activity in the classroom. Our purpose in the study described here was to explore this issue. In particular, our first goal was to examine one third-grade classroom in order to identify ways in which the teacher integrated algebraic reasoning into instruction and evidence that the integration was robust and sustained. As such, we see this study primarily as making a methodological contribution by developing a framework that describes how one teacher integrated algebraic reasoning in the classroom and the frequency and diversity with which she did so. Our second goal was to examine whether that instruction affected students' capacity for algebraic reasoning.

METHODOLOGY

Context for the Study: GEAAR and June's Third-Grade Classroom

At the time of this study, the participating teacher, June,[1] was in her 2nd year of the "Generalizing to Extend Arithmetic to Algebraic Reasoning" (GEAAR) project. GEAAR, a 5-year professional development project conducted in an urban school district, was designed to develop teachers' abilities to identify and strategically build upon students' attempts to reason algebraically and to use existing and supplemental instructional resources to engineer viable classroom instructional activities to support this (Kaput & Blanton, 1999). The project, which was in its 2nd year at the time this study was conducted, included a cohort of 20 grades K–5 teachers.

The strategy of GEAAR was to embed teachers' growth within the constraints of their daily practices, resources, and capacities to grow mathematically and pedagogically. Our structure for achieving this was based on increasing teachers' capacity to transform instructional materials in order to shift the focus of their practice from arithmetic to opportunities for pattern building, conjecturing, generalizing and justifying mathematical facts and relationships. Our approach was to group teachers across grade level and engage them in solving authentic mathematical tasks and reflecting on the algebraic character of these tasks and how they might play out mathematically and pedagogically in the classroom. Teachers then adapted these tasks to their particular grade levels and implemented them in their own classrooms, focusing on student thinking and their own classroom practice. They were encouraged to think about whether a culture of inquiry was developing, what the classroom norms for argumentation were, whether students questioned each other and came to expect justification of mathematical statements, whether there were differences across classrooms, and how they themselves perceived the evolution of their practice. Teachers contributed students' work and oral and written classroom stories to ongoing biweekly teacher seminars.

June, one of the project participants, initially insisted that she was "not a math person," and she described this project as her first exposure to the ideas of alge-

[1] June is a third-grade classroom teacher in the Fall River School District.

braic reasoning. She was hesitant about participating and about her ability to do the mathematics. However, during the 1st year of the project, she conveyed a willingness to incorporate algebraic reasoning tasks in her classroom and to try new ideas suggested by the leaders and teacher participants. Moreover, she made efforts to locate classroom activities that might support algebraic reasoning and to contribute illustrations of students' work and explanations of students' thinking to seminar discussions with teachers. It was from our observations of June during the 1st year of the GEAAR project that we decided to examine her practice more closely in order to understand those types of instructional practices that indicated a generative and self-sustaining capacity to build students' algebraic reasoning.

The urban school district in which June taught was one of the lowest achieving school districts in the state based on student performance on a mandatory, statewide assessment. All of June's 18 students, representing diverse socioeconomic and ethnic backgrounds, participated in our classroom study. The SES of this class and that of the school was lower than average for the district, with 75% on free lunch and 15% on reduced lunch, 65% with parents for whom English was a second language, and 25% with no parent living at home.

Data Collection and Analysis

We used classroom observations to document June's practice. Specifically, we observed June's 90-minute mathematics class approximately twice per week for 1 academic year, observing and recording the conversations between June and her students. The data, which consisted of classroom field notes, audio recordings, June's reflections, students' written work, and classroom activities, were collected from 38 classroom visits by the researchers. Nineteen additional 90-minute mathematics classes, documented by June through written reflections and students' work, were also included in our data corpus. June's reflections included descriptions of classroom activity, such as the tasks used and specific conversations with students, as well as her perceptions on how the lessons flowed and what she thought was noteworthy about students' thinking.

Our analysis began informally in the field by first identifying classroom instances of algebraic reasoning and examining June's instructional role as they occurred. As the data were being collected and transcribed, we focused additionally on identifying those characteristics of June's practice that indicated she was beginning to think algebraically, independent of our interventions through professional development. After data collection, we analyzed field notes, audio transcriptions, and reflective writings for evidence of these characteristics and to establish how algebraic reasoning occurred throughout the academic year. In particular, we conducted a formal analysis from which we could build a profile of the robustness of June's practice. Our goal was to characterize instances of algebraic reasoning and determine how they were embedded in instruction, thereby instantiating her capacity to foster algebraic reasoning. We took the following as a measure of the robustness with which she integrated algebraic reasoning: (1) the *diversity* of types of algebraic reasoning, (2) their *frequency* and *form*[2] of integration, and (3) *techniques of instructional practice* by which algebraic reasoning could thrive.

We first coded the field notes, audio transcriptions, and reflective writings by identifying instances of *spontaneous algebraic reasoning* (SAR) and *planned algebraic reasoning* (PAR). We defined SAR as those instances that occurred without prior planning on June's part but that arose naturally in the context of the mathematics in which students were engaged and that were exploited by June's instruction. For example, June exhibited SAR when, during the course of reviewing homework of simple addition tasks, she spontaneously shifted the focus from computing sums to determining if the sum of two numbers would be even or odd. When students responded by first computing the sum to determine if it was even or odd, June began to use numbers that were sufficiently large so that students could not compute. Instead, they were forced to attend to the structure in the inscriptions themselves. (This episode is detailed later in protocol lines 1–6.)

In contrast, PAR referred to algebraic reasoning that resulted from classroom activities that June planned, in advance of class, to use because of their inherent algebraic features. We identified each type based on teacher materials that June provided us at each visit and from classroom conversations documented through field notes and audiotape. (For the 19 lessons we did not observe, we used June's reflective writings to determine if episodes were PAR or SAR.) Episodes of algebraic reasoning based directly on those tasks included in June's class materials were coded as PAR. For example, June selected the Trapezoid Problem as a preplanned instructional task. The Trapezoid Problem (detailed further in the section *"Category H: Finding functional relationships";* see also Figure 2) asks students to find a functional relationship describing the number of people who could be seated at an arbitrary number of trapezoidal-shaped tables placed end to end. We coded the episodes of algebraic reasoning related directly to solving this task as PAR.

Each episode of PAR or SAR was further categorized based on the nature of algebraic reasoning that occurred. These categories of algebraic reasoning, 13 in all, were refined throughout the initial analysis and the results were used as a framework for recoding the entire set of field notes, reflective writings, and audio transcriptions. Finally, these categories of algebraic reasoning, as well as instances of SAR and PAR, were analyzed to determine how they occurred chronologically and the frequency with which they occurred during the academic year. Activities involving instances of PAR were also differentiated between those that June culled from her own resources and those from our professional development materials.

Episodes in which algebraic reasoning occurred ranged from brief, 2- to 3-minute intervals to conversations that took 30 minutes or more. An episode was defined as a unit of conversation in which a category of algebraic reasoning occurred. In defining an episode, it was not required that students reach a final stage of generalization, since the complexity of some ideas made them more suitable to

[2] By *form* we mean whether algebraic reasoning was based on preplanned tasks or embedded spontaneously in instruction.

assimilate over time. Units of conversation were typically demarcated by arithmetic conversations or shifts in instructional focus, which subsequently served as the boundaries defining an algebraic episode. For example, June often had both spontaneous and planned conversations with students about properties of even and odd sums, conversations that were brief or extended depending on her instructional purpose. In this case, an episode was defined as the excerpt of conversation about properties of even and odd sums. In those lessons not observed by the authors but written about by June, data consisted of descriptions of conversations and not actual dialogue. For these, episodes of algebraic reasoning were defined by chunks of narrative describing classroom events that involved a particular type of algebraic reasoning.

FINDINGS: THE ALGEBRAIC CHARACTER OF JUNE'S CLASSROOM

We used the method of analysis described previously to get a sense of the robustness with which June integrated algebraic reasoning into instruction. In what follows, we discuss the coding categories that emerged from our analysis, how they occurred in instruction, and what they indicate about June's practice. Coding the data occurred along three dimensions: (1) a characterization of whether June's use of algebraic reasoning was PAR or SAR; (2) the types of algebraic reasoning that occurred and the frequency with which they occurred; and (3) the types of representational tools and processes that supported students' algebraic reasoning and that June integrated into the regular life of the classroom. Finally, we looked for aspects of June's instructional practice that supported the long-term development of students' algebraic reasoning skills. We discuss each of these in turn.

Planned and Spontaneous Episodes of Algebraic Reasoning

Out of 57 classroom instructional periods analyzed (38 based on our classroom visits; 19 contributed by June), we identified 204 episodes of algebraic reasoning. Of these episodes, 132 (65%) were characterized as SAR. That is, these were instances in which, in response to students' thinking, June spontaneously crafted instruction that required students to reason algebraically. At times, multiple instances of SAR occurred during the course of a single task. We find it significant that 65% of the episodes occurred spontaneously in instruction. The frequency of SAR not only indicates June's flexibility to spot opportunities for algebraic reasoning, it suggests growth in her content knowledge by which she could reason algebraically and see how algebraic reasoning could be integrated into a mathematically distinct context (e.g., arithmetic) so that it did not occur as a type of "contained enrichment" separate from regular instruction.

June also frequently included in her instruction preplanned algebraic tasks taken from both the GEAAR project and her own resources. Episodes of algebraic reasoning resulting from these activities were designated as PAR. Of the 204 total episodes of algebraic reasoning, 72 instances (or 35%) of PAR occurred, with multiple instances of PAR sometimes occurring for a single mathematical task. That is, each distinct category of algebraic reasoning that occurred within and as part of a planned activity was designated as PAR.

An analysis of how PAR and SAR occurred chronologically did not suggest, as one might expect, that June initially concentrated on planned activities and gradually incorporated spontaneous acts of algebraic reasoning as she became more familiar with the content. Instead, PAR and SAR were integrated throughout the year. This is perhaps because June had participated in GEAAR the year prior to this study and, thus, had some level of experience with it and a growing flexibility to spontaneously integrate algebraic reasoning in instruction.

Categorizing Algebraic Reasoning in June's Classroom

In the second phase of coding, we characterized types of algebraic reasoning and the frequency with which they occurred. We present here an explanation of the categories that emerged from this portion of the analysis, along with instances of student thinking as a way to flesh out their meaning. The categories, which are organized in two sections based on algebraic reasoning as either generalized arithmetic or functional thinking, were not necessarily disjoint and were at times intricately related in a given context. A third section presents categories that reflect processes central to algebraic reasoning that also occurred during instruction.

An analysis of how each of the categories occurred chronologically in instruction throughout the year did not indicate a particular pattern of use or frequency. It is more likely that particular categories occurred as they related to the curriculum at hand. Moreover, it was not always the case that a particular conversation or episode reached a complete level of generalization or justification of a generalization. Some conversations occurred as preliminary efforts to a more developed level of algebraic reasoning and were temporarily suspended because of the complexity of the ideas. As we will describe later, these ideas were often revisited throughout the year, allowing students to build algebraic reasoning skills.

Categories A–E: Algebraic Reasoning as Generalized Arithmetic

Instances where arithmetic was used as a domain for expressing and formalizing generalizations were coded as categories A-E. We took these instances broadly to include arithmetic processes that involved generalized quantities, not necessarily those processes that had generalization as an end result. This included instances where students were engaged in whole number operations on abstracted forms (e.g., missing number sentences), or used number in a generalized way.

Category A: Exploring properties and relationships of whole numbers. Category A describes those episodes of algebraic reasoning in which students explored various properties of, and relationships among, whole numbers. It occurred in 20 of the 204 episodes (about 10%), with 19 of those 20 episodes identified as SAR.

In particular, we found instances where students—

- generalized about sums and products of even and odd numbers;
- generalized about properties such as the result of subtracting a number from itself, expressed as the formalization $a - a = 0$;
- decomposed whole numbers into possible sums and examined the structure of those sums; and
- generalized about place-value properties.

We include as an example of Category A the following excerpt from a classroom conversation described by June (see also Blanton & Kaput, 2000):

> I asked the class what would happen if I added 2 even numbers together. Most of them said that I would get an even number. When I asked what would happen if I added 2 odd numbers together, most of them said that I would get an odd number. When asked about odd and even together, the answers were mixed. In the past I would have told them the answers by giving them some examples (e.g., 5 + 5 = 10). But . . . I wanted them to see how it really works, so that they could see that it would [generalize to all cases]. We did [an] activity combining (square) grid-paper shapes to model adding even and odd numbers. I asked the same questions again. This time they answered with more certainty. One student explained that 'the sum of any two odd numbers is even' using the idea of adding square shapes: "If you have two odd numbers it makes it even because if you have leftovers the two leftovers go together."
>
> The only confusion came when [Sarah[3]] said that odd + even was odd and even + odd was even. [Stephen] responded that that couldn't be. He used numbers in place of odd and even and said that it (using "odd" and "even") was the same as using letters instead of numbers. Sarah explained to the class, "I thought that all the time when odd is the first one it was supposed to be odd and when even was first it was going to be even. [But then I saw that that wasn't correct] because once you start turning them around, then it's the same thing. It doesn't make a difference."

As students participated in classroom instruction and peer argumentation, their generalizations about even and odd sums evolved to more sophisticated and mathematically grounded notions. Sarah's perception that the position of a term in a number sentence determined the parity of a sum ("I thought that all the time when odd is the first one it was supposed to be odd.") was challenged by Stephen, who seemed to interpret *even* and *odd* algebraically, as placeholders or variables. Sarah was eventually able to construct a commutativity argument that disproved her initial generalization ("[But then I saw that that wasn't correct] because once you start turning them around, then it's the same thing. It doesn't make a difference.").

The frequency with which this category occurred (10% of all episodes) and the mostly spontaneous way June used it in instruction (95% of Category A episodes were SAR) suggests that it was strongly embedded in her mathematical and pedagogical understanding of algebraic reasoning. Its accessibility is perhaps because it leverages the familiarity of arithmetic in ways that do not require the implementation of lengthy tasks or the typically more unfamiliar functional thinking.

---

[3] All student names are pseudonyms.

Thus, exploring properties and relationships of whole numbers may offer more familiar content for teachers to use and may suggest an appropriate starting point for professional development.

Category B: Exploring properties of operations on whole numbers. Category B describes those episodes of algebraic reasoning in which students explored the structure of mathematical operations. We identified 21 instances (about 10% of all episodes) of Category B, 20 of which were characterized as SAR. Category B includes looking for generalities in operations such as subtracting negative numbers, as well as exploring relationships between operations, such as commutativity of addition and multiplication or the distributive property of multiplication over addition. For example, as students were looking for patterns in the hundreds chart, June asked them to describe the types of operational actions required to move between various numbers. She asked, "What if I'm at 75 and I go to 65? What did I do?" The representation itself (i.e., the hundreds chart) encodes multiple ways of thinking about possible operations on 75. Students could move directly up one row, which amounted to subtracting once by 10. Or they could move left by single units and subtract 1 from 75 ten times. They began to look at more complex moves, such as the result of 25↓↓→, where each arrow represented a directional move to the adjacent number.

In fact, as students worked with these types of problems, they began to spontaneously generalize about commutative characteristics in the operational symbols and in certain sequences of operations. For example, they quickly saw that 64→↓ was equivalent to 64↓→ and that the result of 46↓↓↑↑ was 46 because "when you add 20 and subtract 20, you're adding nothing." We maintain that these examples reflect algebraic reasoning because of the emphasis on relationships between operations on numbers, not on the results of specific computations.

Category C: Exploring equality as expressing a relationship between quantities. There were 8 accounts of exploring the algebraic role of "=," 7 of which were categorized as SAR. June spent time developing the notion of equality as a relationship between quantities using a balance scale and problems such as $8 + 4 = \square + 5$ (Falkner, Levi, & Carpenter, 1999). When modeling problems with the scale, students worked on both sides of the equality to counter their experiences with computational exercises that could lead them to interpret "=" as an action object (Behr, Erlwanger, & Nichols, 1980; Kieran, 1981). As a result, they began to treat equations as objects expressing quantitative relationships. In one episode, June asked students to solve $(3 \times n) + 2 = 14$. Her description of one student's response illustrates how he had come to view equality:

> Sam said that we could take the 2 away. He said that if we take the 2 from one side we have to take it from the other side. This was to make it balance. After we [took] the 2 away, he said to take the 12 tiles and put them in groups of 3. There were 4 groups, so the answer had to be 4. We tried replacing the *n* with 4 and it worked.

We infer from Sam's explanation that he viewed "=" as expressing a relationship between quantities. He interpreted "=" as signaling a balance (as opposed to a call for action) in that any action on the equation (i.e., subtracting 2) required equivalent operations on both quantities. In Sam's strategy for solving the resulting equation $3 \times n = 12$, it seemed that the process of "dividing both sides by 3" was not an operation available to him (division was not yet a common operation for this class). Thus, he could not perform equivalent operations on both quantities as he had previously done with subtraction. However, he was flexible enough in his thinking to compensate by finding the number of groups of 3 in 12. We maintain that his alternative strategy was not only quite sophisticated, it also implicitly required Sam to see $3 \times n$ as equivalent to 12 in order to simultaneously reorganize one quantity in terms of another. That is, 12 was not the *result* of the action $3 \times n$ but an equivalent way of interpreting the quantity $3 \times n$.

Category D: Algebraic treatment of number. Category D describes episodes in which June treated numbers in an algebraic way, that is, as a placeholder that required students to attend to structure rather than rely on the computation of specific numbers. This is illustrated in the following interchange, where June challenged a student's use of an arithmetic strategy to deduce that 5 + 7 was even:

1 *June:*	How did you get that?
2 *Tony:*	I added 5 and 7 and then I looked over there [pointing to a visible list of even and odd numbers on the wall] and saw that it was even.
3 *June:*	What about 45678 + 85631? Odd or even?
4 *Jenna:*	Odd.
5 *June:*	Why?
6 *Jenna:*	Because 8 and 1 is even and odd, and even and odd is odd.

By using numbers large enough that students could not compute their sum, June required students to think in terms of even and odd properties to determine parity. In doing so, we maintain that June used numbers as placeholders, or variables, for *any* odd or even numbers. Moreover, in using numbers algebraically she was able to avoid the semiotic complications of using literals (e.g., $2n + 1$ for some integer n) to represent arbitrary even and odd numbers. This illustrates how the abstractness of numbers gets built as children work with particular quantities and how a teacher can set the stage for the next move, the formal expression of the generalization (which had not yet occurred in this case). There were 4 episodes of Category D, all of which were SAR.

Category E: Solving missing number sentences. This category involved not only solving simple single variable equations (e.g., $(3 \times n) + 2 = 14$) but also sets of equations as well as single equations with multiple or repeated unknowns. The complexity of these tasks and students' capacity to both symbolize (see Category F) and solve such equations developed throughout the year. Students were asked to solve problems such as "If $V + V = 4$, what is $V + V + 6$?" Missing number sentences were generated by students, sometimes spontaneously, in the course of other tasks such as operations on a number line. One particularly compelling episode involved Zolan's solution to a triangle puzzle. The puzzle was a triangle subdivided into regions, some containing numbers and others empty, where the

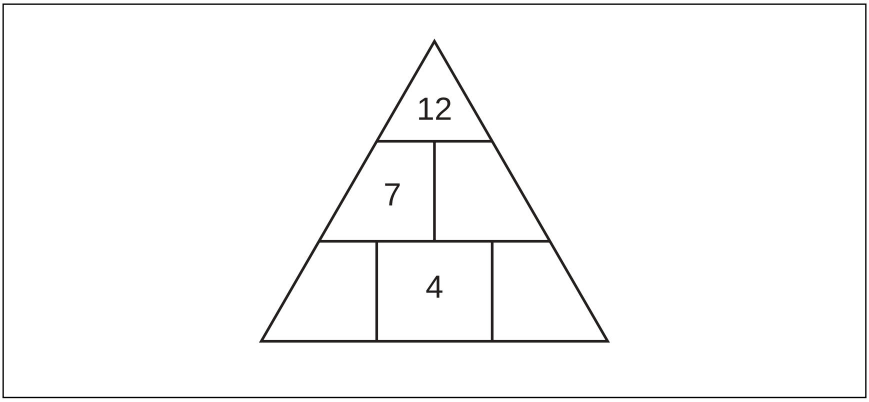


Figure 1. Zolan's triangle puzzle

regions had additive relationships between them and the goal was to "complete the triangle" by finding all the missing numbers (see Figure 1). The missing numbers were found by adding two side-by-side entries to determine the entry above them. For example, in Figure 1, 7 plus the unknown number to its right would be 12, the entry above these two numbers. Zolan spontaneously symbolized the problem by generating a set of equations ($7 + a = 12$, $e + 4 = 5$, $4 + d = 7$) to solve for the missing numbers in the triangle. Not only could he symbolize unknowns, he understood that different symbols were needed for different unknown quantities. He was able to solve each equation for its unknown and use that information in subsequent equations, finding that $a = 5$, $e = 1$, and $d = 3$. Although June had not asked students to set up such equations, that Zolan chose this process without prompting suggests that he—and perhaps others—were seeing mathematics in new, algebraic ways. Solving missing number sentences occurred in 19 instances, 14 (or 74%) of which were identified as SAR.

Categories F–J: Algebraic Reasoning as Functional Thinking

Categories F–J were coded as instances where students were engaged in generalizing numerical and geometric patterns to describe functional relationships. We took this to include attendant processes, such as symbolizing quantities or making predictions about data, which ultimately contributed to the task of describing functional relationships in June's class.

Category F: Symbolizing quantities and operating with symbolized expressions. Category F involves those instances in which students symbolized quantities or operated on symbolized expressions. Although this category is connected to Category E in its emphasis on symbolized equations or expressions (symbolizing mathematical ideas also occurs in other forms of algebraic reasoning such as generalized arithmetic), the focus here is on children's use of symbols to model problems or to operate on symbolized expressions, not the solving of an equation for an unknown or missing quantity or the symbolization of general properties of arithmetic. For example, the class developed what they described as "secret messages," which were symbolic codes for unit conversions. Secret messages were constructed by analyzing a set of numerical expressions: *3 ft* 5 *in* became 3(12) + 5 and 4 *ft* 5 *in* became 4(12) + 5. Encoding a secret message was then a process of symbolizing quantities, such as the number of feet, as variable amounts. What was being symbolized was the quantity of a particular unit. Then, to convert feet to inches, students used the secret message $F(12) + I$, where F represented the number of feet and I represented the number of inches. Here, for example, F was treated as a variable that could take on a range of values. The "secret message" acted as a function to convert a measurement from feet to inches.

Students also computed expressions such as $R + G$, where R and G represented given amounts of a particular item whose value students found by counting a set of objects. Students then used these amounts to evaluate the expressions. Note that, in this case, alphabet characters were used quite differently than that for secret messages. In particular, R and G did not represent continuous quantities but placeholders for specific, to-be-determined quantities. Although there is an important mathematical difference in whether students are working with variable as an object that can take on a range of values versus an object that functions as an unknown representation of a specific quantity, the point of this category was to characterize *both* types of situations. That is, we were interested in all cases where students were abstracting in some way from number to symbol. From this perspective, we also included in this category instances where students spontaneously symbolized mathematical relationships, such as that described previously in Zolan's solution of the triangle puzzle. We coded 17 instances of Category F, 15 (88%) of which were identified as SAR.

Category G: Representing data graphically. Category G refers to the more traditional algebra activity of plotting ordered pairs and was identified in only one episode (categorized as PAR). Although graphing is not inherently algebraic reasoning as we define it, it is included because it represents a way to encode information (graphically) that allows for the analysis of functional relationships. In this sense, it plays a supporting role in algebraic reasoning.

The fact that only one instance of representing data graphically occurred might raise questions regarding its importance, as perceived by the teacher, in developing students' mathematical ideas. As it turned out, the teacher (and students) did occasionally use visual representations, such as frequency diagrams, to represent data in statistical activities. We chose not to include these examples in this category because they served a broader representational purpose not restricted to algebraic thinking. (The broad use of graphical representations is addressed in a subsequent section concerning tools that support algebraic reasoning.) We note further that this was a third-grade class, so graphing two-variable data sets, although not impossible, was not a common activity at this point.

Category H: Finding functional relationships. Category H describes instances in which students were asked to explore correspondence among quantities or recursive relationships and develop a rule that described the relationship. We identified 13 instances of this category, 7 of which were identified as SAR. In the earlier part of the school year, Category H typically involved the use of In/Out charts to find simple additive relationships. For example, students would often examine values in the "Out" column to determine patterns such as "add two every time." However, as June's mathematical understanding about patterns and relations evolved, the complexity of functional thinking tasks evolved as well. In the following reflection from the latter part of the year, June described how students solved the aforementioned Trapezoid Problem (see Figure 2). As the excerpt opens, students have just found the number of people who can sit at 12 adjoined desks by finding a pattern in the data set describing the total number of people. The data were organized by t charts and In/Out charts.

> The strangest thing happened! I saw another pattern. I asked the class to look at the chart in another way. I wanted them to look at the relationship between the desks and the number of people. Find the rule. No luck. I then gave them the hint to see if there was a way to multiply and then add some numbers to have it always work. Jon suggested that we try and find a "secret message." After a few minutes, believe it or not, Anthony and Alicia started to multiply the number of desks with different numbers starting with one. The two children arrived at multiplying by 3 and then they would have to add 2. We tried many examples from the chart and it worked all of the time. We even tried some big numbers like 100. We then tried to make a "secret message." Anthony said that the 3 stays the same so use a *d* for desk. This is what he came up with:
>
> $$3(d) + 2 = \text{number of people}$$
>
> They realized the 3 came from the people that could sit "on the top and the bottom" and the 2 came from the two sides. This is not where I thought this was going to go!

The steps through which Anthony and Alicia established a functional relationship reflected what we think had become a fairly routine practice in June's classroom: Students used a systematic strategy to test what pattern would produce the desired

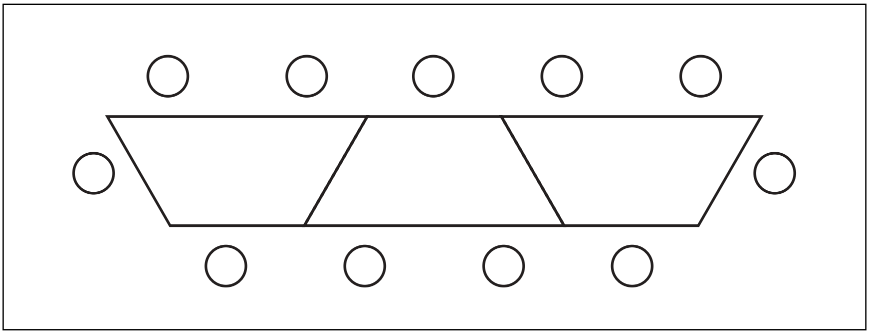


Figure 2. A 3-desk configuration for the Trapezoid Problem

data; a conjectured relationship was identified and described in everyday language ("multiplying by 3 and then . . . add 2"), then tested on a diverse domain of numbers; students symbolized the relationship by noting which quantities varied (the number of desks) and which remained constant; and students described how the physical situation was represented by their model ("They realized the 3 came from the people that could sit 'on the top and the bottom' and the 2 came from the two sides.").

The algebraic reasoning embedded in finding, describing, justifying, and symbolizing mathematical relationships between quantities that vary is crucial to elementary school mathematics because it creates conceptual underpinnings for the more formalized functional thinking that occurs in later grades. In particular, it brings to the fore relationships and structure in data that allow students to model the physical world and think about abstractions beyond the concrete constraints of particular numbers (e.g., If you wanted to seat everyone at school, how many desks would you need?).

Category I: Predicting unknown states using known data. There were 13 episodes (6%) of Category I, with one identified as SAR. Category I describes those instances in which students made conjectures about what would happen for some unknown state, given what they knew from analyzing data for functional relationships. For example, with the Handshake Problem, June asked the class to write a number sentence that would give the amount of handshakes in a group of 12 people, without enacting the handshakes. Students had already determined the amount of handshakes in groups of size 6, 7, and 8 and were looking for patterns in the unexecuted sums that resulted. An excerpt of their conversation is recorded below:

7 *June:*	If there were 12 people here and they were going to shake hands, what would you do?
8 *Ben:*	You could only shake 11 people's hands. [Based on prior conversation we infer that the student meant the first round of handshakes would involve 11 shakes.]
9 *June:*	Why?
10 *Karen:*	Because he can't shake his own hand [therefore the number sentence begins with 11 as opposed to 12].
11 *June:*	So how would your number sentence change if there were 12 people?
12 *Karen:*	Eleven, ten, nine, eight, seven, six, five, four, three, two, one.

Since thoughtful prediction about quantities has as a prerequisite the analysis of relationships between numbers, not simply operations on them, June's query about an unknown state (7)[4] became a point of entry into algebraic reasoning. Asking students to predict unknown states prompts the need for a generalization that characterizes data. From students' analysis of number sentences representing the total handshakes for various group sizes, they had conjectured what the number of handshakes for a group of any size would be and used this to make a prediction about a group of 12 people (and later a group of 20 people) without constructing a diagram or enacting the handshakes. Instruction that limits students to arithmetic

[4] Numbers refer to lines in the protocol.

skills as a solution strategy (e.g., drawing a diagram and counting out the number of handshakes) restricts students' capacity to think about extreme cases, often the very ones that are scientifically interesting.

Category J: Identifying and describing numerical and geometric patterns. Category J most frequently involved identifying patterns in numbers, where the numbers were sometimes generated geometrically. But additionally, June asked students to identify patterns in sequences of geometric shapes and in sets of number sentences. One activity June used with students was the Handshake Problem, in which students were asked to think about the number of handshakes in a group *of any size* if each person in the group shook everyone's hand once (see also Blanton & Kaput, 2003). In this activity, students created number sentences that described the number of handshakes for groups of a particular size (for example, 0 + 1 + 2 + 3 + 4 + 5 described the number of handshakes for a group of size 6), then analyzed the sums in their unexecuted form in order to identify a pattern in the set of number sentences. In this way, students were able to determine that the number of handshakes for a group of any size was given by the sum of the numbers from 0 through one less than the size of the group. We identified 55 episodes (27% of all algebraic episodes) of Category J, with 23 (or 42%) characterized as SAR.

Categories K–M: More About Generalization and Justification

In Categories K–M, we identified instances of algebraic reasoning that were not specific to the forms outlined in the definition of algebraic reasoning provided earlier. That is, they represented either acts of generalizing abstracted from a particular mathematical content or processes that we view as central for viable algebraic reasoning to occur. We conjecture that these categories reflect students' more evolved ability to reason algebraically and, because of their complexity, could indicate that algebraic reasoning was becoming a habit of mind for students.

Category K: Using generalizations to solve algebraic tasks. Category K defines those instances in which students used generalizations to build other generalizations, in all a rather sophisticated level of algebraic thinking. There were 4 instances (2% of all algebraic episodes) of Category K, one of which was SAR. One particularly compelling episode concerned generalizing about sums of even and odd numbers. In it, June had asked students to determine the result of adding three odd numbers: "If we added odd plus odd plus odd, what would the sum be?" Students argued that "the sum would have to be odd because 2 odds make an even and when you add odd plus even, you get odd." We identify this as an example of Category K because students invoked previously established generalizations ("odd + odd = even"; "odd + even = odd") to build their argument. Additionally, they were able to reason with the generalized referent "odd" in the expression "odd plus odd plus odd" and avoid the use of specific odd numbers. In this, we maintain that students were able to achieve a level of abstraction in which they could reason *with a generalization* to produce a generalization (Blanton & Kaput, 2000).

Category L: Justification, proof, and testing conjectures. Category L describes those processes that we see as essential for a culture in which algebraic reasoning can thrive but that are not unique to algebraic reasoning. They are processes that have the explication of student thinking at their core and hence provide a public, oral context in which students can engage their peers in thoughtful debate as a conjecture is established or found to be invalid.

There seemed to be a strong expectation of explanation in June's classroom; students routinely described or justified their thinking or tested their generalizations. In one episode, June spontaneously shared with students a conversation with her peers in which she had argued that zero is an even number, while others had argued that it was simply a "special number" but not even. She invited students to share their thinking. After a protracted conversation with a variety of perspectives (too lengthy to include here), the view that emerged was that zero is even because it belonged to a sequence of numbers, all of which were even, and it belonged to the sequence because it could be reached by skip counting by two ("Zero is an even number because it goes 0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, . . ."). What seemed to be, or at least was becoming, the norm for participation in mathematical conversation was that students justify their perspectives at some level of mathematical sophistication. As we saw earlier, finding a functional relationship in the Trapezoid Problem involved conjecturing and testing a relationship as well as justifying the components of a symbolized relationship ("They realized the 3 came from the people that could sit on the top and the bottom and the 2 came from the two sides."). We identified 22 episodes (11% of all episodes) of Category L, 15 (68%) of which were characterized as SAR.

Category M: Generalizing a mathematical process. Category M refers to conversations in which students built a concept that resulted in generalizing a mathematical process or formula. We see these instances as larger than finding a functional relationship (although they are certainly connected) because the generalizations addressed broad concepts in mathematics. For example, in teaching the concept of area, June developed an activity in which students determined how many two-colored counters were needed to cover a large purple square she had given each student. After discussing the limitations of the circular-shaped counters, she passed out color tiles as an alternative medium. Students immediately noticed that the color tiles covered "the whole space" of the large square. Upon introducing the notion of a unit, June had the following conversation with her students:

13 *June:* Do I know how big the square is?

14 *Mari:* No.

15 *June:* What do you see [referring to the large square which students had covered with tiles]?

16 *Zolan:* Four columns, four rows.

17 *June:* [June then covered a desk with the large purple squares like those she had given students.] So, what would the area of the desk be?

18 *Stephanie:* Twenty-four big squares.

19 *June:* What if we found the area of this table [pointing to a large table in the room]?

One student suggested using a ruler. Kevin suggested that they see how many purple squares are in a row and in a column.

20 *Zolan:* Count how many [purple squares] are in the bottom row.

21 *June:* One, two, three, four, . . . eighteen. [June counts out the number of purple squares in the bottom row.] And how many 18s do I need?

22 *Kevin:* Seven!

23 *June:* What's the best way to find area?

24 *Kevin:* You measure this way and that way [indicating length and width] and multiply.

25 *June:* What do you call "this way" and "that way"?

26 *Student:* Length and width!

We chose to separate this type of activity from other generalizations about relationships between quantities (such as the total number of handshakes in a group of any size) because it addresses a broader concept or idea in mathematics (area). We identified 7 instances of Category M, 6 of which were SAR.

Summary of Categories of Algebraic Reasoning as SAR or PAR Episodes

To summarize, Figure 3 depicts the frequency for the various categories of alge-

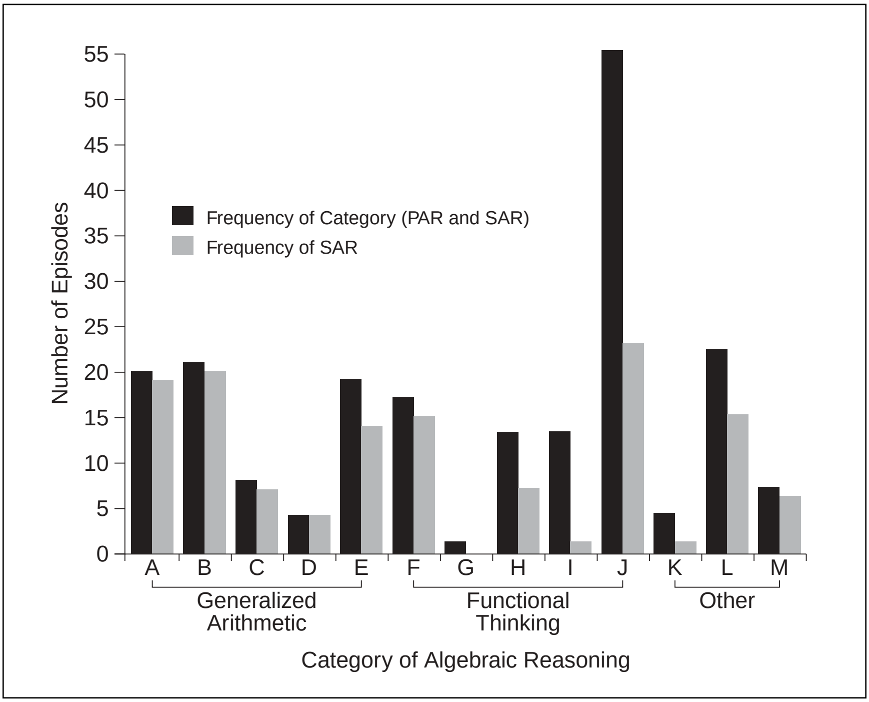


Figure 3. Frequency of categories of algebraic reasoning

braic reasoning compared with the frequency for those episodes additionally characterized as SAR. This depiction more succinctly conveys what June emphasized in her instruction and how she worked flexibly and spontaneously within classroom events. For example, Category J (Identifying and Describing Numerical and Geometric Patterns) occurred most frequently (55 episodes, or 27%) in comparison to the other categories. This is likely a reflection of the emphasis in GEAAR, which included a number of pattern-eliciting tasks such as the modified Handshake Problem (Blanton & Kaput, 2004). A secondary content emphasis of GEAAR was algebraic reasoning in the form of generalized arithmetic. In fact, with the exception of Categories K–M, which were not specifically linked to the forms of algebraic reasoning described earlier, the categories of algebraic reasoning identified in June's practice were based on either the use of arithmetic as a domain for expressing and formalizing generalizations (generalized arithmetic) or generalizing numerical patterns to describe functional relationships (functional thinking). Moreover, within these two forms, the four most frequently occurring categories were Categories A (Exploring Properties and Relationships of Whole Numbers), B (Exploring Properties of Operations on Whole Numbers), E (Solving Missing Number Sentences), and J (Identifying and Describing Numerical and Geometric Patterns), with 115 episodes combined (about 56% of all 204 episodes). Again, these frequencies seemed to reflect the content emphasis of GEAAR, not a limitation of the teacher to look at other forms of algebraic reasoning. Additionally, we found that GEAAR teachers, including June, seemed to most readily identify solving missing number sentences (Category E) as "algebraic reasoning" (perhaps because these tasks most resembled their memories of high school algebra). Since these were fairly easy to solve, they tended to be used frequently.

It could also be that teachers who are beginning to study algebraic reasoning view generalized arithmetic and functional thinking as a more natural instructional fit in an elementary mathematics classroom than the other two forms of algebraic reasoning identified here (modeling and generalizing about mathematical systems). This seems particularly true when arithmetic is used as a domain for algebraic reasoning, since arithmetic typically dominates elementary grades and represents an area where teachers might be more mathematically certain. On the other hand, teachers might connect more readily to the mathematics of modeling in a science lesson, and generalizing about mathematical systems abstracted from computations and relations is often the domain of higher mathematics.

Of the 204 episodes of algebraic reasoning, 132 (65%) were identified as SAR. We take this as an important indicator of a level of autonomy and flexibility in June's capacity to identify and support opportunities for algebraic reasoning as they occurred in instruction. Moreover, if we compare instances of SAR across categories, we notice that opportunities for SAR occurred more frequently for algebraic reasoning in the form of generalized arithmetic rather than functional thinking. Out of 132 SAR episodes, 64 episodes were in Categories A–E (generalized arithmetic) and 46 were in Categories F–J (functional thinking). We conjecture that this difference is due in part to the nature of the content; in June's class, the types of tasks involving functional thinking often took more time to implement. For example, the frequency of PAR in Category I (92%) is perhaps due to the fact that the activity of prediction requires previous steps such as generating data, identifying patterns, and describing functional relationships. Thus, it was not as likely to occur spontaneously in instruction but would more likely be the final stage of an elaborated, planned activity such as the Trapezoid Problem.

Also, GEAAR teachers, including June, seemed more familiar with arithmetic—the core of much of elementary grades mathematics—than with the complexities of covariation, correspondence, and symbolizing that occur in functional thinking. This may have led to a greater capacity to spontaneously generalize arithmetic. Additionally, functional thinking tasks tended to require more instructional time and planning and, thus, seemed less likely to emerge spontaneously. Consequently, although Figure 3 suggests that generalized arithmetic and functional thinking are both accessible topics of inquiry for elementary teachers, it also indicates that teachers might be more likely to engage in spontaneous acts of algebraic reasoning in the domain of generalized arithmetic. This suggests that generalized arithmetic might offer the most feasible point of entry for professional development designed to help teachers identify opportunities for algebraic reasoning in their everyday instruction.

We see the categories described here not as a discrete set of attributes or as simply a way to encode other teaching practices if one desired but as a statement about the robust and diverse ways in which a teacher's practice *can* integrate algebraic thinking into instruction. In other words, what we gain in part from this analysis is an existence proof that algebraic thinking can *permeate* instruction. Although classroom stories of teachers engaged in particular episodes of algebraic thinking are extremely helpful in building insights about practice, our purpose instead was to construct a detailed accounting of the long-term enactment of algebraic thinking in instruction. June showed a great deal of diversity in the forms of algebraic reasoning that she used in instruction. This, along with the frequency with which she integrated these forms as well as the flexibility with which she exploited on-the-spot opportunities for algebraic reasoning, suggests that her ability to create opportunities for algebraic reasoning was robust. Although it is difficult to claim that some particular quantification of data equates to a practice that integrates algebraic thinking in robust ways, we do conjecture that the sustained integration of planned and spontaneous opportunities for various forms of algebraic reasoning, as seen in June's practice, can lead to a habit of mind that produces algebraic thinkers.

Tools of Algebraic Reasoning

During analysis, we observed that there were certain reoccurring structures or processes that seemed to scaffold students' mathematical thinking, although the context in which they were embedded did not necessarily involve algebraic reasoning. Although these do not represent forms of algebraic reasoning per se, we came to describe them as *tools that support algebraic reasoning* (TSAR) and defined them to be objects, structures, or processes that facilitated students' mathematical reasoning and that could be used to support their algebraic reasoning. In particular, we defined it to include objects such as In/Out charts (t charts) for organizing data and concrete or visual artifacts such as number lines, diagrams, and line graphs for building and making written and oral arguments. In other words, these objects became referents around which students reasoned mathematically. TSAR was also defined to include mathematical processes, such as recording, collecting, representing, and organizing data, that may have occurred in contexts that did not explicitly involve algebraic reasoning (e.g., statistics).

Throughout the school year, June worked to scaffold students' use of tools. She often initially suggested the use of particular objects or processes (for example, the use of a t chart to record data, or the use of students' initials to track the handshakes in a group) and modeled how to use them. However, as the school year progressed, we observed that students seemed to adopt these tools for their own purposes and would use them without prompting from June. June also encouraged students to choose the tools that best suited their purposes and, as the class solved various tasks, she would often try to make explicit in the conversation the variety of tools that different students used.

From our analysis of the data, we identified 71 instances of TSAR. Given that 57 class periods were considered, we maintain that an average of 1.25 instances of TSAR per class period suggests a frequency whereby students could build a repertoire of tools for reasoning algebraically. We observed that students increasingly chose to use these tools without prompting from June. Thus, we conjecture that the significance of TSAR is that June's inclusion of these tools as a regular part of instruction helped to build a habit of mind that supported algebraic reasoning when it did occur and thereby contributed to the classroom culture of algebraic thinking.

Characteristics of Instructional Practice That Support the Integration of Algebraic Reasoning

The third measure we used to characterize the robustness of June's capacity to integrate algebraic reasoning in instruction involved the identification of techniques of practice that supported the development of students' algebraic reasoning skills. Although these techniques do not form an exhaustive list, we see them as part of an emerging profile of the type of practice that teachers skilled in algebraic reasoning might exhibit in instruction. In what follows, we describe each of these characteristics.

Seamless and Spontaneous Integration of Algebraic Conversations in the Classroom

June was able to create what we describe as an "algebraic conversation," that is, a conversation that called on students to engage in some form of generalizing or formalizing or to reason with generalizations. Moreover, she was able to do this in a manner that was spontaneous and seamless. In other words, she was able to instinctively transform what seemed to be a routine arithmetic task into one that required algebraic reasoning. In contrast to implementing only predesigned algebraic tasks as stand-alone activities separate from regular instruction, June was able to incorporate opportunities for algebraic reasoning into regularly planned, and perhaps what might have been more arithmetically focused, instruction. The protocol given in (1–6) illustrates this characteristic. In it, June and her students were discussing the parity of a sum when June transformed the discussion into an algebraic conversation by using large numbers algebraically as placeholders. We include SAR here as a characteristic of instructional practice that supports algebraic reasoning because the seamless and spontaneous integration of algebraic reasoning requires a depth of skill above and beyond a knowledge level that is confined to the use of prearranged activities in instruction. Thus, we take SAR as one of the indicators that a teacher's practice is expanding to attend to algebraic reasoning opportunities.

The Spiraling of Algebraic Themes Over Significant Periods of Time

June spiraled certain algebraic themes into her conversations with students over sustained periods of time, revisiting ideas throughout the year in deeper and more compelling ways. Whether this spiraling was planned or a result of her own developing mathematical knowledge (we think the latter), the result was to build the complexity of algebraic activity in the classroom. For example, at the beginning of the year, students generated In/Out charts and identified simple additive recursive relationships for data in the "Out" column. This progressed to analyzing the relationship between data in input and output columns and describing more complex relationships that involved both addition and multiplication, such as in the Trapezoid Problem. June also continued to revisit themes addressing number properties (such as the commutativity of addition and generalizations about even and odd number sums); symbolizing to represent unknown quantities; or varying task parameters in order to generate, identify, and describe numerical patterns. The result was that students were able to reason about these ideas in increasingly complex ways (Blanton & Kaput, 2000).

We suggest that June's ability to spiral these ideas, often spontaneously, showed that they were not isolated in June's thinking or in her implementation of them but could be brought to bear on a diversity of classroom experiences. In fact, we take as part of the evidence of robustness that she did not use these ideas in isolation or as a one-time activity, and she often integrated them by her own initiative and creativity, not by specific instruction from GEAAR.

Integration of Multiple and Independently Valid Algebraic Processes

Although doing what might be considered a stand-alone algebraic activity, June sometimes pulled into this another algebraic process that altered the complexity of the original task and, in essence, pushed the "algebra envelope." For example, during one class June asked students to use base-10 blocks to solve missing number sentences, a task that is itself algebraic in nature. After a discussion with students in which they shared the different strategies they used to solve the problem, she focused on the sentence $14 = 6 + n$ and expanded this problem in the following way: First, she asked students to solve $140 = 60 + n$, then $1400 = 600 + n$. After students shared how they had arrived at their solutions, June turned solving this family of equations into a pattern-finding activity, thus superimposing a separate algebraic process on the missing-number activity. Although this could be described as a more primitive form of pattern finding in that it did not involve identifying relationships between two quantities, it did occur early in the year as June was beginning to experiment with these types of problems. What is of note to us is June's flexibility and spontaneity in integrating these tasks and being able to transform an ongoing problem (in this case, one that has algebraic characteristics) to exploit its algebraic potential. That she was able to coordinate two separate algebraic processes, which is mathematically a different task than transforming arithmetic tasks, indicates growth in her mathematical knowledge as well.

Activity Engineering

One of June's strengths—and a characteristic that we tried to cultivate in GEAAR—was adapting or developing mathematical tasks to include algebraic reasoning. We see autonomy in task development as a critical component of teacher growth because it shows a capacity to generate resources beyond the finite resource base provided by professional development. From our analysis, we found that June became increasingly skilled at finding and adapting or developing resources that brought algebraic reasoning to the fore in instruction. For example, a few weeks after we introduced the Handshake Problem to GEAAR participants, June adapted an activity in which students explored the number of gifts received, based on the words of the song "The Twelve Days of Christmas." (Her students were learning the song for a school production.) In this song, one's paramour receives an accumulation of gifts over a 12-day period. The task was as follows:

> How many gifts did your true love receive on each day? If the song was titled "The Twenty-Five Days of Christmas," how many gifts would your true love receive on the twenty-fifth day? How many total gifts did she or he receive on the first 2 days? The first 3 days? The first 4 days? How many gifts did she or he receive on all twelve days?

As it turned out, this problem was mathematically similar to other problems, such as the Handshake Problem, that we featured in our seminars. In particular, they were all pattern-eliciting tasks that relied on the use of unexecuted sums to find and describe general functional relationships (see Blanton & Kaput, 2004). The significance here was in June's capacity to bring creative resources into her lessons without relying solely on the materials we provided. As another example, the selection of the Trapezoid Problem was inspired by June's committee work to select new school furniture. Of the 72 episodes that were identified as PAR, 63% involved tasks that June had created or selected from her own resources. We take this as further evidence of the robustness in her practice of integrating algebraic reasoning. June was developing her own algebra "eyes and ears" and was learning to plan instruction independently of the resources provided in the seminars. We summarize the findings on the algebraic character of June's classroom in Table 1.

EFFECTS OF PRACTICE: STUDENT PERFORMANCE ON ALGEBRAIC REASONING TEST ITEMS

Ultimately, any claims one makes about the effectiveness of a teacher's practice will be measured against student performance. In this study, it was reasonable to expect that the extent to which June integrated algebraic thinking in her classroom in robust and flexible ways would improve students' ability to reason algebraically. Our purpose in this section is to examine some of the evidence for this. As a caveat, we note that we see a distinction in the data presented in this section regarding what it says (or does not say) about June's practice, namely, *that* June's practice affected student achievement versus *how* it affected student achievement. In particular, our intent here is not to address *how* June's practice affected student achievement; this would require a more detailed look at how her actions played out in the classroom and how students were involved in this. Even so, it is difficult to divorce the study of characteristics of practice that promote algebraic reasoning from evidence that algebraic reasoning has been promoted. In this sense, although this section does not intend to address *how*, it does help establish *that* June's teaching had an impact on student achievement. We have already described classroom vignettes that suggest students were engaged in various levels of algebraic reasoning, and we have quantified the frequency with which algebraic reasoning was incorporated into instruction. However, more evidence concerning whether students were beginning to reason algebraically and how widespread this was in the classroom is needed. We turn our attention here to this task.

At the end of the academic year in which this study occurred, we conducted a quantitative analysis of students' performance on selected items of the fourth-grade Massachusetts Comprehensive Assessment System (MCAS), a state-wide, mandatory, standardized exam. In particular, we administered 14 items[5] (10 multiple-choice items and 4 open-response items) to June's students and to students in another third-grade control class that was from the same school and had a comparable SES. Unlike June, the teacher of the control group used a traditional arithmetic curriculum and had not participated in GEAAR. Items were selected to represent a mixture of traditional items that were of interest to the teacher (e.g., concepts of measurement, time) and algebraic reasoning items that reflected generalized arithmetic or functional thinking. Fourteen items were selected so that the test could be administered within a 1-hour episode.

[5] See Appendix for assessment items.

Table 1
Summary of Findings on the Algebraic Character of June's Classroom

Planned and Spontaneous Episodes of Algebraic Reasoning (PAR/SAR)

Out of 204 episodes of algebraic reasoning identified, 132 (65%) were identified as SAR. The remaining 72 episodes (35%) were characterized as PAR.

We did not observe a pattern by which PAR and SAR were integrated in instruction during the school year.

Types of Algebraic Reasoning and Their Classroom Frequency

Category	Frequency of Occurrence	SAR Frequency per Category
Generalized Arithmetic		
A: Exploring properties and relationships of whole numbers	10	95
B: Exploring properties of operations on whole numbers	10	95
C: Exploring equality as expressing a relationship between quantities	4	88
D: Algebraic treatment of number	2	100
E: Solving missing number sentences	10	74
Functional Thinking		
F: Symbolizing quantities and operating with symbolized expressions	8	88
G: Representing data graphically	1	0
H: Finding functional relationships	6	54
I: Predicting unknown states using known data	6	8
J: Identifying and describing numerical and geometric patterns	27	42
More About Generalization and Justification		
K: Using generalizations to solve algebraic tasks	2	50
L: Justification, proof and testing conjectures	11	68
M: Generalizing a mathematical process	3	86

Representational Tools That Supported Algebraic Reasoning (TSAR)

Objects, structures, or processes that facilitated students' mathematical reasoning and could be used to support their algebraic reasoning

Included representational tools such as In/Out charts (t charts), number lines, diagrams, and graphs, as well as processes such as recording, collecting, representing, and organizing data that may also have occurred in contexts outside of algebraic reasoning (e.g., statistics)

Out of 57 class periods analyzed, 71 instances of TSAR were identified.

Characteristics of Instructional Practice That Supported Algebraic Reasoning

Seamless and spontaneous integration of algebraic conversations in classroom instruction

Spiraling of algebraic themes over significant periods of time

Integration of multiple and independently valid algebraic processes

Activity engineering

Note: All table values are rounded percentages. Frequency of Occurrence denotes the percentage of occurrence for each category out of 204 total episodes. SAR Frequency per Category denotes the percentage of episodes within a given category that were characterized as spontaneous episodes of algebraic reasoning (as opposed to planned episodes).

MCAS was chosen as a performance measure for several reasons. First, it is analyzed in a manner that allowed for item comparisons of the treatment group (June's class) with other students across the school, district, and state. We considered this comparison important because of the literacy demands of the fourth-grade MCAS and the fact that the majority of students in June's third-grade class were from ESL homes. More to the point, MCAS includes algebraic reasoning tasks both as a separate, distinct strand and integrated across multiple content strands, and it emphasizes problem-solving skills that are consistent with algebraic reasoning, such as organizing and analyzing data, explaining, and justifying. Thus, we viewed it as a good measure of students' capacity for algebraic reasoning. Finally, because of the accountability that MCAS imposes on teachers, having a measure that persuaded teachers that algebraic reasoning was relevant to their daily practice was important for our professional development.

Item-Based Comparison of Control and Experimental Groups

State-wide achievement levels for the spring 1999 MCAS[6] were determined[7] to be (a) "advanced" for students scoring at or above 81%; (b) "proficient" for students scoring at least 67% and less than 81%; and (c) "needs improvement" for students scoring at least 41% and less than 67%. An item analysis of student responses from the control and experimental groups showed that the experimental class outperformed the control class on 11 of the 14 test items (see Figure 4). In addition, June's students performed significantly better ($\alpha = .05$) than the control group on 4 of these 11 items. Overall, the experimental group performed at the level of "proficient" on 36% of the items and at the level of "needs improvement" on 43% of the items, whereas the control group did not score "proficient" on any of the items and performed at the level of "needs improvement" on 43% of the items. The experimental group scored below "needs improvement" on only 21% of the items, whereas the control group scored below "needs improvement" on 57% of the items. Although the control group outperformed the experimental group on 3 of the 14 items, it was not at a statistically significant level.

Of the 14 items on the assessment, we identified 7 as "algebraic" (items 2, 3, 6, 7, 8, 11, and 14) because they addressed aspects of generalized arithmetic and functional thinking, requiring students to find patterns generated numerically and geometrically, understand and use whole-number properties, and identify unknown quantities in number sentences. The experimental class outperformed the control class on 6 out of 7 of these items, although not at a statistically significant ($\alpha = .05$) level. Moreover, the experimental class performed at the level of "proficient" or "needs improvement" on 72% of the algebra items (which tended to be harder than the more traditional arithmetic problems), whereas the control class performed at these levels on only 43% of the items. In particular, the experimental group scored

[6] Ten assessment items were selected from the 1999 MCAS; the remaining four items were selected from other MCAS resources.

[7] Yearly levels are variable.

"proficient" on 29% of the items and "needs improvement" on 43% of the items. The control class, however, performed at the level of "needs improvement" on 43% of the algebra items and did not score at the proficient level on any of the items.

Comparison With State and District Performance

Figure 5 provides an item analysis of those MCAS items used in our assessment that were selected from the 1999 MCAS. In particular, it shows a comparison of performances by the experimental and control groups at the state, district, and school levels.[8] Figure 5 suggests that June's students' performance was comparable to that at the state, district, and school levels for seven items (items 1, 2, 5, 6, 10, 12, 14) and in two of these (items 10, 12) exceeded fourth-grade results overall. Although items 10 and 12 had not been identified as "algebraic," they did involve interpreting graphs and counting arrangements, both of which had been included in algebraic reasoning tasks in June's classroom.

A comparison of the state and district performance on the *entire* test (not merely the algebra strand items) with the experimental group's performance on the selected 1999 MCAS items by percent per category suggests that the experi-

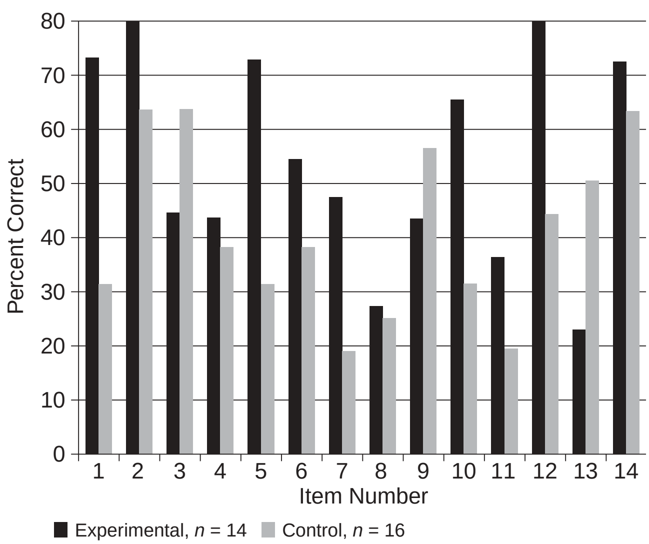


Figure 4. Students' individual performance on selected MCAS items, experimental versus control

[8] State, district, and school data for the other 4 items used in our assessment were not available. Thus, Figure 5 compares only those items taken from the 1999 MCAS.

mental third-grade class performed approximately as well as the fourth graders state-wide and significantly better than the district fourth graders (see Figure 6). We regard these as strong results, given the considerable advantage of an addi-

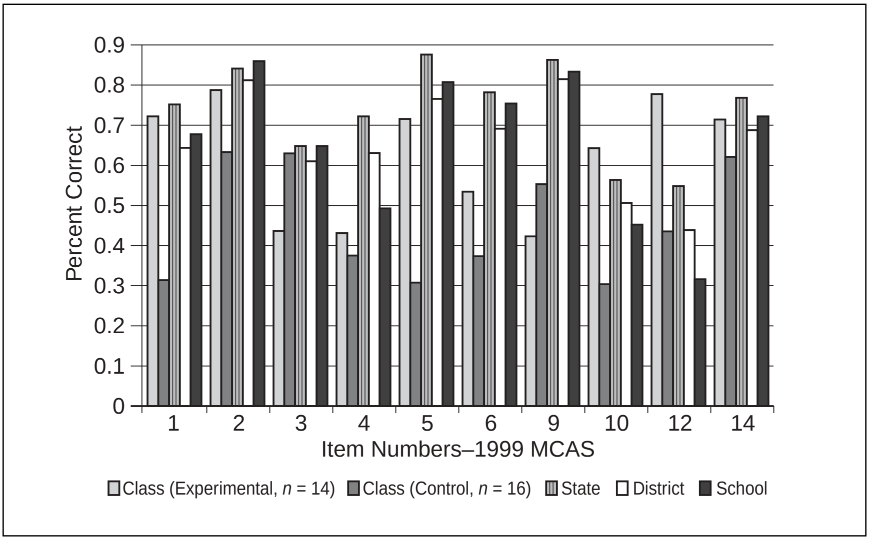


Figure 5. Percentage of correct responses for each comparison group by item

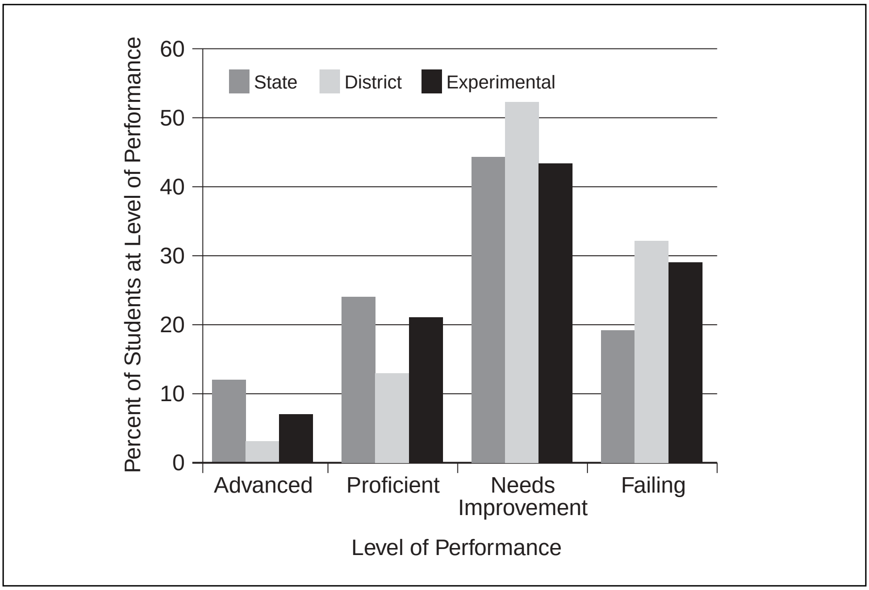


Figure 6. State and district MCAS scores compared to experimental scores

tional year's instruction for the fourth-grade comparison groups at the state and district levels, especially the significant development in the necessary verbal skills during the intervening year relevant to several of the items, and given the low SES factors for the experimental class. We caution, however, that these data provide us with only one piece of evidence. As in any quantitative analysis of student achievement, these data are subject to the complexity of how to account for all design variables in the context of teaching and learning. Thus, they are only one indicator of student performance.

CONCLUSION

In this article, we have explored how June integrated algebraic reasoning into the classroom by examining the diversity and frequency with which she did so, as well as some of the aspects of her instruction that supported this. We have also described evidence that June's instruction shaped students' ability to reason algebraically. As a result, we have a developing picture of what it looks like for a teacher's practice to cultivate students' algebraic reasoning skills in robust ways. In particular, we infer from June's practice that robustness is measured by a teacher's ability to flexibly transform a broad range of arithmetic content so that multiple domains of algebraic reasoning (e.g., generalized arithmetic, functional thinking) are woven into instruction over sustained periods of time in ways that allow the complexity of ideas to be deeply developed. Moreover, robustness is captured by the teacher's ability to either identify, modify, and adapt resources for planned instructional purposes or to spontaneously transform arithmetic conversations into those that require algebraic thinking. The frequency with which algebraic tasks are integrated in instruction is also a factor of robustness. Although it is difficult to quantify this robustness, the instructional goal is to build habits of mind whereby students naturally engage in algebraic reasoning and to use the tools (objects, structures, and processes) that support it. Marginalizing tasks as enrichment activities that occur in isolation prevents this development. In June's class, algebraic reasoning tasks were not mathematical "field trips" but were woven into the daily fabric of instruction.

Building a practice that develops children's algebraic reasoning requires a significant process of change for elementary teachers, who are often schooled in different, arithmetic ways of doing mathematics. In short, elementary teachers must develop algebra "eyes and ears" as a new way of both looking at the mathematics they are teaching and listening to students' thinking about it. This study suggests that generalized arithmetic and functional thinking offer rich (and accessible) entry points for teachers to study algebraic reasoning. Moreover, it suggests that teachers can learn to think spontaneously about these forms of algebraic reasoning and that generalized arithmetic may be particularly fruitful as an initial context for building teachers' ability to bring algebraic reasoning into classroom conversations.

Although this study illustrates one teacher's effort to integrate algebraic reasoning into her practice, more research is needed to understand the trajectories through which teachers develop as they participate in teacher communities such as those fostered by GEAAR. For example, how does teachers' knowledge of the various forms of algebraic reasoning evolve and how do they use this knowledge in instruction? Why might teachers choose one form of algebraic reasoning over another? Or, with regards to students, how might integrating algebraic thinking in elementary grades change the way that children view arithmetic concepts such as number? How do students reason algebraically with various representational tools and how does the use of these tools in instruction become a habit of mind for students? Moreover, if what occurred in June's third-grade class had occurred in grades K–5 for these students as regular, daily instruction, what would be the nature of students' understanding of elementary school mathematics and what would the implications be for their success in later grades? Furthermore, how can middle grades onward leverage algebraic reasoning in the elementary grades so that students develop a deeper understanding of more advanced mathematics? Finally, although the approach we took in analysis required us to think about classroom events largely independent of the teacher's interpretation, studies of teaching practice that incorporate both researcher and teacher interpretations of events would strengthen how we understand algebraic thinking as a classroom process.

We have used June's practice to examine the nature of algebraic reasoning that can occur in elementary classrooms. Like many elementary teachers, June had learned mathematics and how to teach mathematics in ways that were quite different from instruction that develops students' algebraic reasoning skills. Developing a cadre of elementary teachers who understand the complexity of algebraic reasoning and how to integrate it in viable ways will require long-term, sustained professional development that is sensitive to the needs of this unique population of teachers. We offer June's case as evidence that elementary teachers can engage in practices of teaching that support the development of students' ability to reason algebraically.

REFERENCES

Behr, M., Erlwanger, S., & Nichols, E. (1980). How children view the equal sign. *Mathematics Teaching, 92*, 13–15.

Blanton, M., & Kaput, J. (2000). Generalizing and progressively formalizing in a third-grade mathematics classroom: Conversations about even and odd numbers. In M. Fernández (Ed.), *Proceedings of the Twenty-Second Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 115–119), Columbus, OH: ERIC Clearinghouse.

Blanton, M., & Kaput, J. (2003). Developing elementary teachers' algebra eyes and ears, *Teaching Children Mathematics, 10*, 70–77.

Blanton, M., & Kaput, J. (2004). Design principles for instructional contexts that support students' transition from arithmetic to algebraic reasoning: Elements of task and culture. In R. Nemirovsky, B. Warren, A. Rosebery, & J. Solomon (Eds.), *Everyday matters in science and mathematics* (pp. 211–234). Mahwah, NJ: Lawrence Erlbaum.

Blanton, M., & Kaput, J. (2005). Helping elementary teachers build mathematical generality into curriculum and instruction. In J. Cai & E. Knuth (Eds.), *Zentralblatt für Didaktik der Mathematik (International Reviews on Mathematical Education*), Special Issue on Algebraic Thinking, *37*, 34–42.

Carpenter, T. P., & Franke, M. (2001). Developing algebraic reasoning in the elementary grades: Generalization and proof. In H. Chick, K. Stacey, J. Vincent, & J. Vincent (Eds.), *Proceedings of the Twelfth ICMI Study Conference* (Vol. 1, pp. 155–162). Melbourne, Australia: University of Melbourne Press.

Carpenter, T. P., Franke, M., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school.* Portsmouth, NH: Heinemann.

Carraher, D., Brizuela, B., & Schliemann, A. (2000). Bringing out the algebraic character of arithmetic: Instantiating variables in addition and subtraction. In T. Nakahara & M. Koyama (Eds.), *Proceedings of the Twenty-Fourth International Conference for the Psychology of Mathematics Education* (Vol. 2, pp. 145–152). Hiroshima, Japan: PME Program Committee.

Carraher, D., & Earnest, D. (2003). Guess my rule revisited. In N. Pateman, B. Dougherty, & J. Zilliox (Eds.), *Proceedings of the Twenty-Seventh International Conference for the Psychology of Mathematics Education: Vol. 2* (pp. 173–180). Honolulu, Hawaii: University of Hawaii.

Davydov, V. V. (1975a). Logical and psychological problems of elementary mathematics as an academic subject. In J. Kilpatrick, I. Wirszup, E. G. Begle, & J. W. Wilson (Series Eds.), & L. P. Steffe (Vol. Ed.), *Soviet studies in the psychology of learning and teaching mathematics: Vol. 7. Children's capacity for learning mathematics* (pp. 55–107). Chicago: University of Chicago.

Davydov, V. V. (1975b). The psychological characteristics of the "prenumeral" period of mathematics instruction. In J. Kilpatrick, I. Wirszup, E. G. Begle, & J. W. Wilson (Series Eds.), & L. P. Steffe (Vol. Ed.), *Soviet studies in the psychology of learning and teaching mathematics: Vol. 7. Children's capacity for learning mathematics* (pp. 109–205). Chicago: University of Chicago.

Dougherty, B. (2003). Voyaging from theory to practice in learning: Measure up. In N. Pateman, B. Dougherty, & J. Zilliox (Eds.), *Proceedings of the Twenty-Seventh International Conference for the Psychology of Mathematics Education* (Vol. 1, pp. 17–23). Honolulu, Hawaii: University of Hawaii.

Falkner, K. P., Levi, L., & Carpenter, T. P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching Children Mathematics, 6*, 232–236.

Kaput, J. (1998). Transforming algebra from an engine of inequity to an engine of mathematical power by "algebrafying" the K–12 curriculum. In S. Fennell (Ed.), *The nature and role of algebra in the K–14 curriculum: Proceedings of a national symposium* (pp. 25–26). Washington, DC: National Research Council, National Academy Press.

Kaput, J. (1999). Teaching and learning a new algebra. In E. Fennema & T. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 133–155). Mahwah, NJ: Erlbaum.

Kaput, J., & Blanton, M. (1999, April). *Algebraic reasoning in the context of elementary mathematics: Making it implementable on a massive scale.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12,* 317–326.

Kieran, C. (1992). The learning and teaching of school algebra. In D. A. Grouws (Ed.), *The handbook of research on mathematics teaching and learning* (pp. 390–419). New York: Macmillan.

Mason, J. (1996). Expressing generality and roots of algebra. In N. Bednarz, C. Kieran, & L. Lee (Eds.), *Approaches to algebra: Perspectives for research and teaching* (pp. 65–86). Dordrecht, Netherlands: Kluwer.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

Romberg, T., & Kaput, J. (1999). Mathematics worth teaching, mathematics worth understanding. In E. Fennema & T. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 3–32). Mahwah, NJ: Erlbaum.

Schifter, D. (1999). Reasoning about operations: Early algebraic thinking, grade K through 6. In L. Stiff & F. Curio (Eds.), *Mathematical Reasoning, K–12: 1999 NCTM Yearbook,* (pp. 62–81). Reston, VA: National Council of Teachers of Mathematics.

U.S. Department of Education & National Center for Education Statistics. (1998a). *Pursuing excellence: A study of U.S. fourth-grade mathematics and science achievement in international context.* Washington, DC: U.S. Government Printing Office.

U.S. Department of Education & National Center for Education Statistics. (1998b). *Pursuing excellence: A study of U.S. eighth-grade mathematics and science achievement in international context.* Washington, DC: U.S. Government Printing Office.

U.S. Department of Education & National Center for Education Statistics. (1998c). *Pursuing excellence: A study of U.S. twelfth-grade mathematics and science achievement in international context.* Washington, DC: U.S. Government Printing Office.

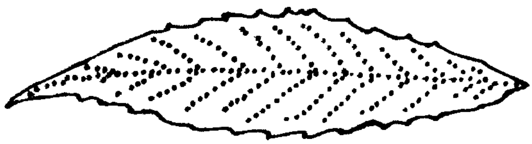
Authors

Maria L. Blanton, Department of Mathematics, University of Massachusetts Dartmouth, 285 Old Westport Road, No. Dartmouth, MA 02747-2300; mblanton@umassd.edu.

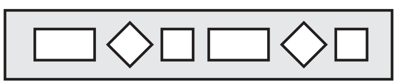
James J. Kaput, formerly at the Department of Mathematics, University of Massachusetts Dartmouth.

APPENDIX

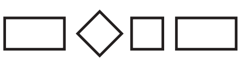
1. How many CENTIMETERS long is the leaf?



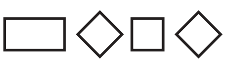
Use the pattern in the box below to answer the next question.



2. What are the next four figures in the pattern above?

A. 

C. 

B. 

D. 

3. What number does n stand for in the sentence below?

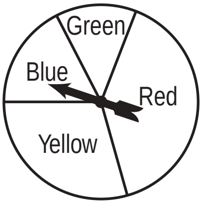
$$(8 + 2) + 6 = 8 + (n + 6)$$

A. 2 B. 6 C. 8 D. 16

4. Your lunch time begins at 12:40 P.M. If your lunch time is 35 minutes long, what time does it end?

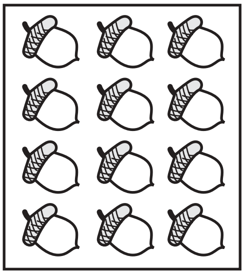
A. 12:05 P.M. B. 1:10 P.M. C. 1:15 P.M. D. 1:30 P.M.

5. This is a spinner for a game. Which color are you most likely to spin?

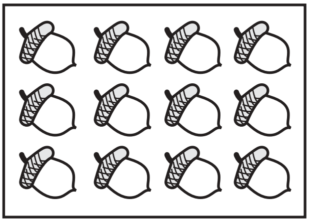


A. blue B. green C. yellow D. red

6. Melvin collected acorns from the yard. First he placed them like this:



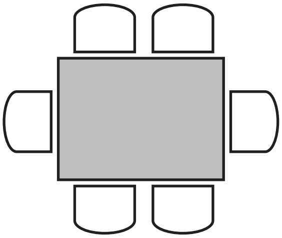
Then he placed them like this:



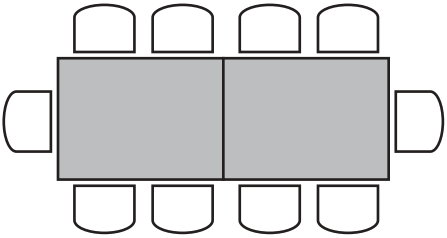
Which number sentence shows the TWO ways Melvin placed his acorns?

A. $3 \times 4 = 4 \times 3$ B. 3×4 C. $3 \times 4 > 4 \times 3$ D. $3 \times 4 \geq 4 \times 3$

7. Andrew is setting up tables for a birthday party. He knows that six people can sit about this table:



When he puts two of these tables together end to end, he can seat ten people.



How many people can Andrew seat if he puts three tables together end to end?

8. Write the RULE to find the next number in this pattern.

87, 81, 75, 69, _____

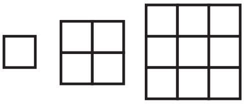
9. There are 60 pieces of art paper and 42 children. If each child gets one piece of art paper, how many pieces will be left for another project?

A. 9 B. 18 C. 27 D. 42

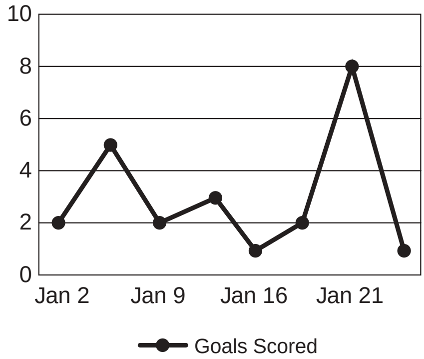
10. What is the GREATEST number of different outfits you can make with 2 pairs of pants and 5 shirts? (Each outfit must have exactly one pair of pants and one shirt.)

A. 5 B. 7 C. 10 D. 25

11. How many of the smallest squares will be in Figure 6 if this pattern continues?



12. How many goals did the Boston Bruins score in their game on January 16, 1999?

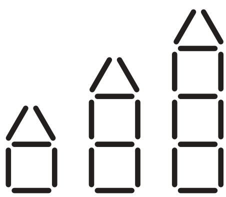


(a) 1 goal (b) 2 goals (c) 4 goals (d) 3 goals

13. Mr. Gillman wants to give apple slices to his 13 soccer players during their game. Each player will receive 3 slices. He plans to cut each apple into 4 slices. How many apples will Mr. Gillman need?

A. 8 B. 10 C. 7 D. 9

14. Donna made this pattern using sticks. Draw the next figure in the pattern.



Explain how you got your answer.